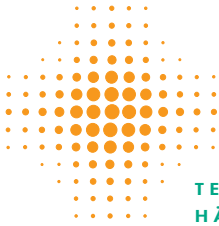
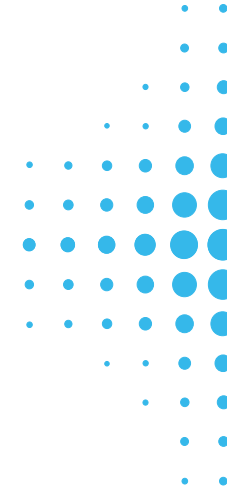
Kansanterveyslaitos
Folkhälsoinstitutet
National Public Health Institute

Kansanterveyslaitoksen julkaisuja B 26 / 2008
Publications of the National Public Health Institute



TERVEYS
HÄLSA
HEALTH
2000



METHODOLOGY REPORT

Health 2000 Survey

Helsinki
2008

Kansanterveyslaitos
Folkhälsoinstitutet
National Public Health Institute

Kansanterveyslaitoksen julkaisuja B 26/2008

Publications of the National Public Health Institute

METHODOLOGY REPORT

Health 2000 Survey

Sami Heistaro, ed.

Kansanterveyslaitos
Terveyden ja toimintakyvyn osasto

KTL-National Public Health Institute, Finland
Department of Health and Functional Capacity

Helsinki 2008

Julkaisija – Utgivare – Publisher

Kansanterveyslaitos (KTL)
Mannerheimintie 166
00300 Helsinki
Puh. vaihde (09) 47 441, faksi (09) 4744 8408

Folkhälsoinstitutet
Mannerheimvägen 166
00300 Helsingfors
Tel. växel (09) 47 441, telefax (09) 4744 8408

National Public Health Institute
Mannerheimintie 166
FI-00300 Helsinki, Finland
Telephone + 358 9 47 441, telefax + 358 9 4744 8408

Publications of the National Public Health Institute, KTL B 26/2008

ISBN 978-951-740-859-2 (print)
ISBN 978-951-740-860-8 (pdf)
ISSN 0359-3576
http://www.terveys2000.fi

Layout: Riitta Nieminen

Hakapaino Oy, Helsinki 2008

ABSTRACT

This book describes the implementation and methods of the Health 2000 Survey. It also contains an account of the sampling and the recommended methods for statistical analysis.

A random sample of the Finnish population was examined in 2000–2001 by comprehensive methods. After the fielfwork sub-populations were invited to attend several in-depth studies. This major co-operative effort was led by the National Public Health Institute of Finland (KTL).

The two-stage stratified sample comprised 10 000 persons aged 18 or over of whom 8028 were aged 30 or over. The survey was carried out in 80 study areas all over the country. The interviewers of Statistics Finland carried the home interviews. The five study teams of KTL executed the health examinations. The main objects of the study were the major public health problems, functioning and their determinants. Special emphasis was placed on cardiovascular and respiratory diseases, musculoskeletal and mental disorders and oral health. The study also concerned use of and need for care, rehabilitation and help. The methods were interviews, questionnaires, measurements, determinations from blood samples and clinical examinations.

Of persons aged 30 and over 89% participated in the home interview and 85% in the health examination (including home health examinations). Of young adults (18–29) 80% were interviewed. The study material is of high quality and it represents the whole population unusually well. Thus, it provides exceptionally good opportunities for public health research and for health monitoring. In order to enhance the use of the data KTL has created a facility to order sub-sets of the data via the Internet (www.terveys2000.fi).

An Internet-version is published simultaneously with this book. In addition to the current material it comprises links to questionnaires and other forms as well as to original operating instructions and method descriptions.

CONTENTS

1. INTRODUCTION

Arpo Aromaa

The Health 2000 Survey was conducted in 2000 and 2001. The data for the survey were collected in interviews with almost 10,000 and in comprehensive health examinations of 8,000 people. The purpose of the survey was to provide an up-to-date account of major public health problems in Finland, their causes and treatment, as well as functional capacity and work ability in the population. Furthermore, comparisons were made with the results of the Mini-Finland Health Survey 20 years earlier (Aromaa et al. 1989a) to gain a reliable picture of the development of health and functional capacity.

In order to achieve these goals it was necessary to collect a nationally representative population sample, to have a high rate of participation in the survey and to apply as rigorous and valid methods as possible. Many of the methods were chosen with a specific view to ensuring comparability with the Mini-Finland Survey. In addition, earlier domestic national population studies were consulted (Purola et al. 1967, Arinen et al. 1998, Vartiainen et al. 1998). The conceptualisation, design and methods of the research were developed and refined in several stages, beginning with the Finriski 1997 Senior Survey and then later in connection with two pilot studies in 2000. The equipment were calibrated and training was provided to the field staff so as to maximise the reliability of the results, minimise inter-rater differences and ensure comparability with earlier research.

This report and its appendices describe the data, materials and methods used in the Health 2000 Survey as well as its implementation. One important use of the report will be to serve as a benchmark for descriptions of the survey's methods in various other publications. In addition to the material published in this printed report, many of its appendices such as various forms and guidelines are available on the Health 2000 website (www.terveys2000.fi).

2. PROJECT ORGANISATION

Arpo Aromaa

The project organisation involved a wide range of research and funding agencies. Overall responsibility for project planning and implementation was assumed by the National Public Health Institute (KTL). Other agencies and organisations involved from the earliest planning stages were the Finnish Centre for Pensions, the Social Insurance Institution of Finland, the Local Government Pensions Institution, the National Research and Development Centre for Welfare and Health, the Finnish Dental Association and the Finnish Dental Society, Statistics Finland, the Finnish Institute of Occupational Health, the Finnish Work Environment Fund, the UKK Institute for Health Promotion Research and the Occupational Safety and Health Fund of the State sector.

Project operations were overseen by a steering group under the chairmanship of the Permanent Secretary for the Ministry of Social Affairs and Health. Much of the preparatory work was done within the line organisations of the various agencies and bodies involved. Overall project management was undertaken by a managing group at the National Public Health Institute (KTL), with the support of a project group and its executive committee.

The project planning organisation involved some 130 researchers and other experts who worked in planning teams for different subjects areas and who later took charge of training and the supervision of the field examinations. Upon completion of these examinations the planning teams were reorganised into research teams involving researchers both from the research institutes involved in the project and from a number of universities and hospitals. Among the themes covered by these teams were cardiovascular diseases and diabetes, mental health, musculoskeletal disorders, oral health, living habits and psychosocial factors as well as functional capacity.

The actual fieldwork was done by the project field organisation under the supervision of the planning teams.

3. SAMPLING DESIGN

Johanna Laiho, Kari Djerf and Risto Lehtonen

3.1. Definition of target population

The target population of the Health 2000 Survey consisted of individuals aged 18 or over and living in mainland Finland. In addition to the household population, people living in institutions were included. Geographically, the Autonomous Territory of Åland Islands was excluded, as were people living on islands not accessible by road.

3.2. Sampling design

The sampling design for the Health 2000 Survey was developed jointly by Statistics Finland experts and the KTL research team. The basic design was a stratified two-stage cluster sampling design. Essentially, the aim was to select a sample that properly reflected the main demographic distributions of the Finnish population and allowed data collection at reasonable cost.

In the first stage of clustering the most crucial consideration was to ensure that it was not too complicated for the participants to reach the clinic where the examination took place. Furthermore, the workload of interviewers and clinic staff had to be optimally allocated. Within these general conditions, sample selection was carried out using standard probability sampling routines.

However, there are certain features that make the study design quite complex. Firstly, Health 2000 actually involved two separate surveys: the main survey was carried out in the population aged 30 or over, and the study of young adults was focused on people aged 18–29. Furthermore, in order to obtain a sufficient number of observations from the oldest age cohorts in the main survey, people aged 80 or over were oversampled with a double sampling fraction. In other respects the sampling design was similar in both studies.

The sample was drawn at the Social Insurance Institution (SII) of Finland (Kela) using the population-wide insurance database as a sampling frame.

3.3. Stratification and clustering

Mainland Finland was divided into five geographical strata, which were the university hospital districts of Helsinki, Turku, Tampere, Kuopio and Oulu. It was agreed that the ultimate sample size for each stratum was to be proportional to the population base.

In the first stage of sampling, 80 health centre districts (clusters) were selected out of the total of 249 districts in mainland Finland. The second stage involved sampling individual persons from those districts. The maximum number of clusters was included as determined by the cost limitations. Equal allocation across strata was used, i.e. 16 clusters were selected from each stratum. However, to improve precision, it was agreed that the strata be further divided into two sub-strata as follows:

1) The 15 largest towns (their health centres) were selected with probability 1, and the sample size for each town/health centre was proportional to its proportion of the population. Simple random sampling was used to select the individuals to be examined. The method may thus be called a stratified one-stage element-level design;

2) The remaining 65 health centres were selected using a systematic probabilities proportional to size (PPS-SYS) design. As a measure of size, we used the population count of people aged 18 or over (excluding the 15 largest towns).

3.4. Selection of persons

The second stage of selection involved selecting the ultimate sampling units, i.e. target persons. The population aged 18 or over in each selected health centre district was sorted by age, and selection was carried out using systematic random sampling. The sampling thus involves implicit stratification by age. The overall sample size was set at about 10,000, with the main survey (age 30+ years) accounting for about 8,000 and the survey of young adults (age 18-29 years) for about 2,000 target persons.

In the main survey the smallest cluster-specific sample size was 50 and the largest 100 persons. In order to ensure a high precision of estimates even for the oldest generations, the decision was made to oversample people aged 80 or over. Oversampling was carried out within clusters (including the 15 largest towns) using double inclusion probabilities. In the study of young adults, the smallest and largest cluster-specific sample sizes were 10 and 25, respectively.

The sampling design used generally yields almost an equal probability of selection (EPSEM) sample when oversampling is not taken into account.

The sample sizes of both the one-stage and two-stage clusters are presented in Table 3.1, which shows that the overall sample size was 10,492. Technical checks before and during the fieldwork reduced the final sample size to 9,922 (main survey 8,028 and study of young adults 1,894). The main reason for exclusions was temporary residence abroad.

Table 3.1. Expected and actual sample size by sub-study and university hospital stratum.

University hospital district	Target population	Number of clusters	Expected sample size for young adults (18–29 yrs)	Actual sample size for young adults (18–29 yrs)	Expected sample size for main study (30+ yrs)	Actual sample size for main study (30+ yrs)	Expected total sample size	Actual total samples size	Selection probability of cluster
1. Helsinki	**1,305,804**	**16**	**658**	**625**	**2,620**	**2,811**	**3,278**	**3,436**	
1A. Clusters 1–5	822,181	5	414	445	1,650	1,716	2,064	2,161	
Helsinki	447,064	1	222	255	900	918	1,122	1,173	1
Espoo	153,597	1	77	89	309	316	386	405	1
Vantaa	131,252	1	70	63	260	283	330	346	1
Kotka	44,688	1	22	15	90	101	112	116	1
Lappeenranta	45,580	1	24	23	90	98	114	121	1
1B. Clusters 6–16	483,513	11	244	180	970	1,095	1,214	1,275	PPS
expected sample size of a cluster			*22*		*88*		*110*		
2. Turku	**533,310**	**16**	**269**	**234**	**1,070**	**1,178**	**1,339**	**1,412**	
2A. Clusters 17–18	201,053	2	105	106	400	428	505	534	
Turku	140,217	1	72	85	280	288	352	373	1
Pori	60,836	1	33	21	120	140	153	161	1
2B. Clusters 19–32	332,257	14	164	128	670	750	834	878	PPS
expected sample size of a cluster			*10*		*50*		*60*		
3. Tampere	**925,769**	**16**	**464**	**395**	**1,860**	**2,046**	**2,324**	**2,441**	
3A. Clusters 33–36	326,945	4	165	142	656	719	821	861	
Tampere	154,467	1	78	72	310	334	388	406	1
Lahti	77,087	1	43	24	150	178	193	202	1
Vaasa	44,832	1	23	28	90	90	113	118	1
Hämeenlinna	50,559	1	26	18	101	117	127	135	1
3B. Clusters 37–48	598,824	12	303	253	1,200	1,327	1,503	1,580	PPS
expected sample size of a cluster			*25*		*100*		*125*		
4. Kuopio	**675,381**	**16**	**335**	**296**	**1,360**	**1,481**	**1,695**	**1,777**	
4A. Clusters 49–51	168,790	3	84	115	340	330	424	445	
Kuopio	67,354	1	29	44	140	131	169	175	1
Jyväskylä	61,160	1	34	49	120	115	154	164	1
Joensuu	40,276	1	21	22	80	84	101	106	1
4B. Clusters 52–64	506,591	13	251	181	1,020	1,151	1,271	1,332	PPS
expected sample size of a cluster			*18*		*80*		*98*		
5. Oulu	**543,676**	**16**	**275**	**273**	**1,090**	**1,153**	**1,365**	**1,426**	
5A. Cluster 65, Oulu	89,537	1	45	59	180	181	225	240	1
5B. Clusters 66–80	454,139	15	230	214	910	972	1,140	1,186	PPS
expected sample size of a cluster			16		60		76		

Selection within stratum

A. Self-representing strata

B. Other clusters, selected using Probability Proportional to Size sampling

4. THE STAGES OF THE RESEARCH PROCESS

Arpo Aromaa

Planning for the research started in 1999. Following the initial planning stage which lasted about one year, work was started in preparation of the first pilot study. During this period the sampling plan was finalised and the first questionnaire versions were drafted. The paper and electronic versions of these questionnaires were completed after the second pilot study in summer 2000. The staff were recruited and the necessary equipment acquired during the spring and summer of 2000. Training sessions for Statistics Finland interviewers and a three-week training course for the health examination personnel were organised in August 2000.

The field examination consisted of 1) a home-visit interview and 2) a health examination, which took place a few weeks after the interview. The interviews were started on 15 August 2000 and the health examinations on 18 September 2000. If the person was unable to attend the health examination proper, they were visited at home (or in an institution) and a less extensive health examination was carried out. The people who dropped out of these stages of the study were first approached by phone and then by letter in order to retrieve key health information. Data were thus obtained from the majority of respondents through both the health interview and health examinations. At least some data collected via telephone interviews or postal questionnaires are thus available for a very large majority of the respondents.

After the field examination, data have been collected from various registers among other things on hospital treatment, disability pensions and special refund medications. The collection of register data is described separately in Chapter 17. Several more in-depth analyses were carried out on selected subsamples both while the field examination was still underway and after it was completed (Chapter 21).

4.1. Interviews and questionnaires

Much of the data collection at the various stages of the survey was based on interviews and questionnaires. The electronic and paper questionnaires are available in Finnish, Swedish and English on the project website (www.terveys2000.fi). The most important electronic forms were the home-visit interview (see Chapter 6), the symptom interview, clinical examination, and the psychiatric CIDI interview (see Chapter 8). Ordinary self-administered questionnaires were

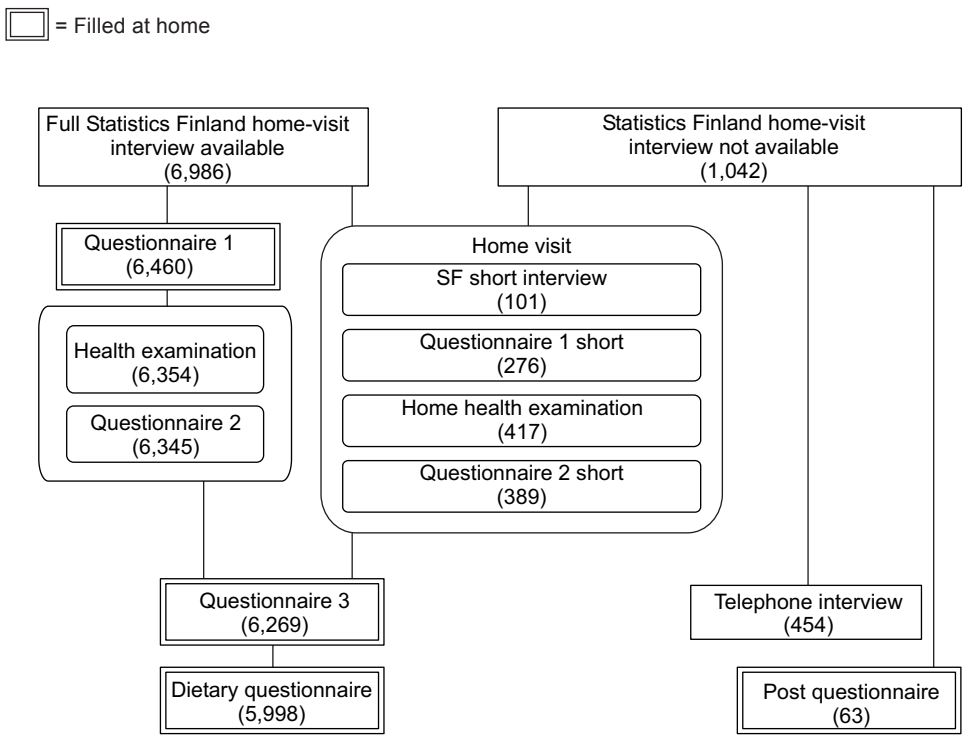
the basic questionnaire (questionnaire 1, see chapter 7), infection questionnaire (questionnaire 2), complementary questionnaire (questionnaire 3) and dietary questionnaire (see chapter 8). Telephone interviews and questionnaires for dropouts were also conducted using the same standard schedules.

The questions were partly based on questions used in earlier national surveys, especially in the Mini-Finland Survey. Partly they were based on other Finnish and international surveys and international recommendations. The aim was to achieve comparability with other Finnish surveys and also be able to compare some of the results internationally.

4.2. Health examination

The contents of the health examination are described in detail in chapter 8. To those who were not able to participate in the health examination, a shorter home health examination was carried out (please see chapter 9).

Figure 4.1. The stages of the research process.



5. PREPARATIONS FOR FIELDWORK

Sami Heistaro

5.1. General

5.1.1. Timetable and facilities

Each of the 80 municipal health centres involved in the project were contacted well ahead of the first preparations for the fieldwork. Chief physicians were contacted by phone and letter from the National Public Health Institute (KTL), with the request that facilities be made available at a later date for the health examination. The response and reaction from the local health centres was very positive. However, for various reasons not all of them were in the position to provide the necessary facilities, and in these cases the project had to content itself with rented facilities.

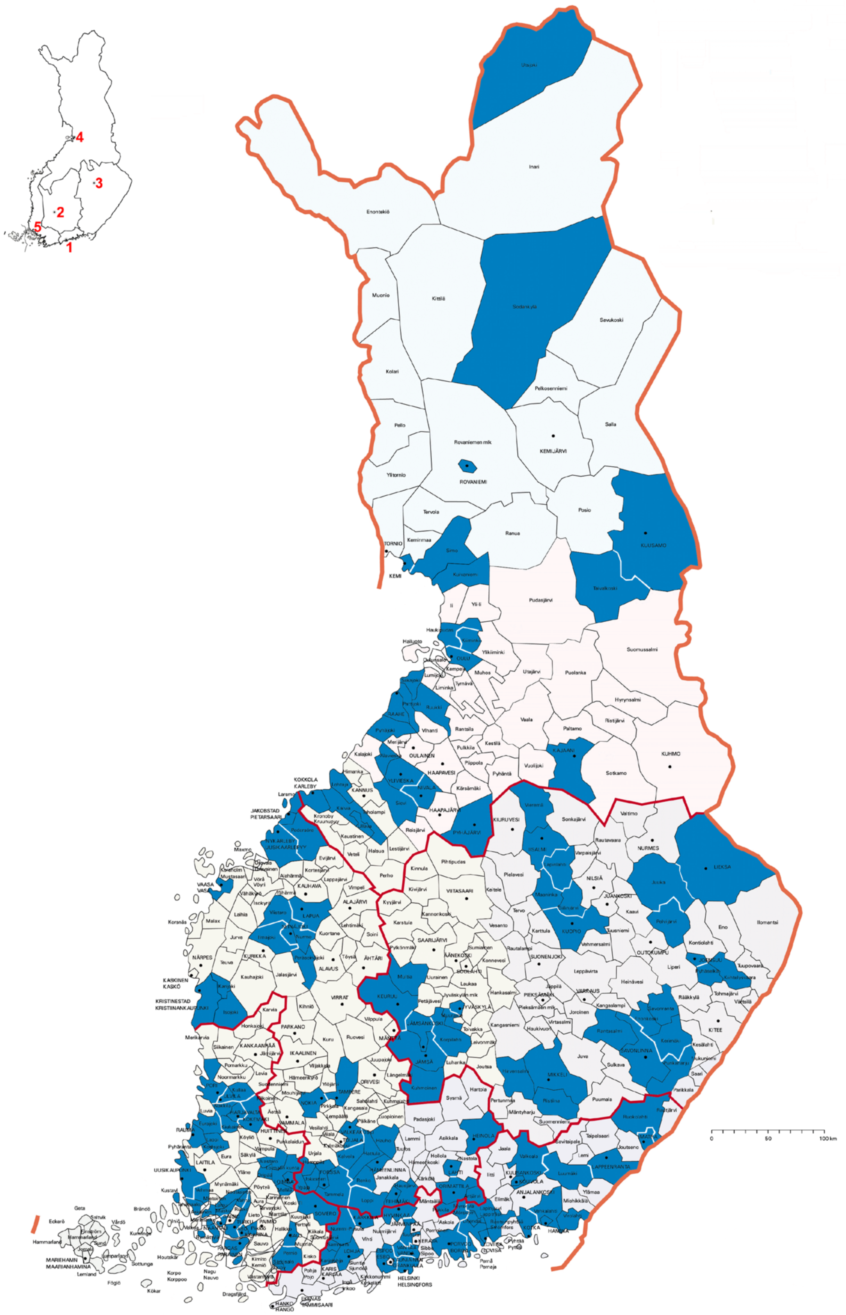
Taking account of local preferences, timetables were agreed for the health examinations. Each of the five field teams were assigned to their own geographic areas, which varied widely in size according to population density. The northernmost team was responsible for almost half of Finland, and distances between the examination sites were accordingly considerable. The field teams' equipment and materials were transported with hired trucks. Depending on the number of examinees, the teams spent anything from a few days up to a couple of months at any given location. Staff accommodation was arranged in hotels or other lodgings.

5.1.2. Implementation of health examination

One health examination lasted from four to five and a half hours. A more detailed description of the examination and its different stages is given in Chapter 8. It took the field team 3–4 hours to set up the examination stations. If possible, team members would scout the facilities in advance.

Each health examination team had 17 members. The nurse in charge of the final examination point (head nurse) also had a host of other duties, including transport for moves from one location to another, staff accommodation, communication and team spirit. The head nurses visited project headquarters about once a month to have discussions with the head of field operations about any practical problems that had surfaced in the course of the examinations.

Figure 5.1. Survey areas of Health 2000.



5.1.3. Preparation of health examination material

Tarja Kiesi

The material needed for recording the data collected and for informing the subjects was prepared at project headquarters in Helsinki about one month ahead of the start of the health examination at each location. A health examination file carrying the subject's name and examination code was prepared for each subject: this file included all the forms that the would be needed during the examination to collect data in the order of the examination points as well as relevant information for the subject. This certainly helped to speed up the examination. Separate files were prepared for home visits.

The health examination file included a personalised examination schedule, an information letter for the subject, a laboratory sample form, questionnaire 2 (infection questionnaire), a feedback form for completion by the subject, two informed consent forms, the dietary questionnaire and questionnaire 3 (complementary questionnaire), a summary of the results of the oral examination as well as sheets of sticker labels with the subject's name and code number for use on the paper forms needed in connection with home visits, for instance. Every other file also had an opinion questionnaire to collect the subject's experiences of the health examination.

The health examination files were packed into cardboard boxes bearing the name and number of the research team, the site code, the start and end dates of the examination period and examination numbers. In addition, to facilitate transportation, stickers were attached to the boxes with the name of the research team and the examination site.

5.2. Staff recruitment, training and induction

Pirkko Alha and Sami Heistaro

5.2.1. Recruitment

Field staff for the Health 2000 Survey were recruited through job advertisements placed in the Finnish Medical Journal and the magazine published by the Union of Health and Social Care Professionals ("Tehy"). The vacancies announced in Tehy were for nurses, laboratory technicians, physiotherapists and dental nurses. Following a round of interviews, ten physicians and five dentists were appointed, two plus one for each team, respectively.

More than 1,000 applications were submitted for the nursing positions announced. Large numbers applied for the positions of head nurse and examination nurse. Project headquarters appointed five head nurses, i.e. one for each team. Working closely with the appointed physicians, these five nurses were to interview the shortlisted candidates and select the staff for their teams. This was motivated by the aim of idea of creating teams with good synergy and collaboration. Eleven examination nurses and additionally two home-visit nurses were appointed to each team. The composition of the field teams is shown in Table 5.1.

Table 5.1. Composition of field teams

Nurses 1 and 2	registration and measurement point 2
Nurses 3, 4 and 5	measurement point 1 and laboratory
Dentist	clinical examination of mouth
Nurse 6	dental nurse, oral examination
Nurses 7, 8, 9 and 10	functional capacity and mental health interview
Physician 1 and 2	clinical medical examination
Nurse 11	final check point
Nurses 12 and 13	home visits and stand-ins

5.2.2. Training for Statistics Finland interviewers

One-day training sessions were organised for Statistics Finland interviewers (a total of 158 persons, for more details see Chapter 6) on the 16th, 21st and 24th of August 2000. Lectures were held at Statistics Finland among others on the following subjects: diseases, medications, classifications of surgical procedures, blood pressure measurements at home, appointments systems, contents of health examination, health questionnaires, information for the subjects, subjects' informed consent and maintaining contact with the National Public Health Institute (KTL) in different situations.

5.2.3. Training for field health examination staff

Training for the health examination staff was organised in Helsinki from 21 August to 8 September 2000. Following a general information meeting, some 100 enthusiastic field staff convened at the University of Helsinki Department of Dentistry. Training was tailored to the content of the respective examination points. A three-week timetable was prepared for the whole training period. The physicians had their own timetable. Training for the dentists and dental nurses was organised at the Department of Dentistry. Training for nurses working at the registration point, measurement point 2 and the final check point was the same, while nurses appointed to measurement point 1 and laboratory technicians also had

the same training. Nurses at the functional capacity measurement point and the CIDI (mental health) interviewers had their own training schedules.

As the nurses received training for different examination points, they were able to rotate jobs usually at one-week, sometimes at two-week intervals. This helped not only to create more variability in their work, but also to maintain consistent quality of measurements. Training for home-visit nurses covered the examination points for which the measurements were taken at home. In addition, all staff received general IT training. The contents of the examination points and their guidelines are explained in more detail under the descriptions of the respective examination points.

5.2.4. Induction training

Practical induction training was organised during the last week of the training sessions, with each field team conducting health examinations on one another. The teams moved out to their assigned areas the very next week, but at this point examinations were still carried out on specially recruited volunteers. The first health examinations proper were carried out on 18 September 2000.

5.2.5. Ongoing training

Once the health examinations got underway, the persons in charge of the different examination points made regular visits to monitor operations in the field. As expected, there were initially a whole range of questions and difficulties that were addressed as quickly as possible. This meant that minor adjustments were made to the guidelines for fieldwork even while the examinations were in progress. Staff from project headquarters continued to make field visits throughout the stage of fieldwork. As well as helping to resolve substantive issues, these visits were an important source of encouragement through the dark winter months. Additional training and refresher courses were arranged on different themes if and when necessary; one example is provided by the training day for the oral health team on 18 November 2000.

The head nurses, for their part, made visits to Helsinki about once a month to report back on progress out in the field. At the same time they were given information to take back about current and upcoming events. In connection with these visits five charge nurses took part in a separate job supervision programme organised by Sami Pirkola.

During the course of the health examinations there was some staff changeover in all the field teams. In November 2000 the decision was taken to recruit one further

examination nurse on each team. This afforded greater freedom of movement to the nurse in charge of the final check point, for instance, when they needed to move on and make the necessary preparations at the next examination site. The quality of the examinations was monitored by means of various quality measurements, which are discussed in more detail later on (see Chapter 14).

5.3. IT environment

Jaason Haapakoski, Vesa Tanskanen, Matti Sarjakoski and Esa Virtala (software)

Mikko Nissinen, Mikko Pekkarinen, Päivi Markkanen and Jonas Sundman (hardware)

The decision was made to collect as much of the data as possible directly onto computers. The following targets were set for the data collection process: 1) all data should be recorded in its final form so that only one round of data entry is needed; 2) the data shall be recorded immediately to prevent it from being lost or forgotten; 3) the data shall be recorded by the person who has produced that data, i.e. the field examiner, so as to avoid any ambiguity about responsibility; 4) data security is fully guaranteed; 5) the dataset produced is sound, has passed preliminary checks and is easy to process for further analysis; 6) the equipment and methods can be re-used in other situations; and 7) there is no need for IT staff out in the field.

The participants in the survey were divided into different subsamples. The examination and sampling protocols varied to some extent between these different samples.

IT training and support

Training for field staff consisted of three parts: 1) each member of the field staff received a brief introduction to the use of the data recording equipment; 2) each member of the staff received instruction on the use of the software at their respective examination point (1–4 points); and 3) two persons were selected in each field team to take charge of setting up and operating the local network of the field office. Among the areas specially highlighted in training were the accuracy of the data recorded and data security.

The National Public Health Institute's (KTL's) IT support services provided guidance and support by phone as well as by visiting the field teams. Programs were usually updated from the KTL mainframe as described below, but the ISDN connections available were simply too slow for copying some large programs and databases. In these cases, it was necessary to resort to other means.

The IT operations manager at the KTL was able to link up directly from her own workstation to the field team's server, working in coordination with field team's head of IT operations. 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 that software.

5.3.1. Communication

The Statistics Finland interviewers contacted the subjects in their respective areas and conducted their interviews at the subjects' home using electronic (Blaise) questionnaires. At the end of each day, the interviewer sent the data compiled to Statistics Finland using the Kermit file transfer program. Statistics Finland encrypted the interview material and forwarded it daily to a dedicated e-mail box at KTL, from which the data were automatically downloaded into the survey database. The health examination appointments made in connection with the interviews were as well downloaded into the database.

Using this database, letters reminding the subjects of their health examination were compiled and sent out to them about two weeks ahead of time. The field teams had access to the appointments made for each day, and they could take up-to-date printouts directly from the server at headquarters using a web browser (Netscape).

Once the field examination team arrived at a new examination site, a "camp" was set up that consisted of "camp days". Certain examination times had been allocated to each camp day in advance. These times had been created on the basis of a set schedule entered in a database ("model days", see below), but they could be changed as necessary on a case by case basis.

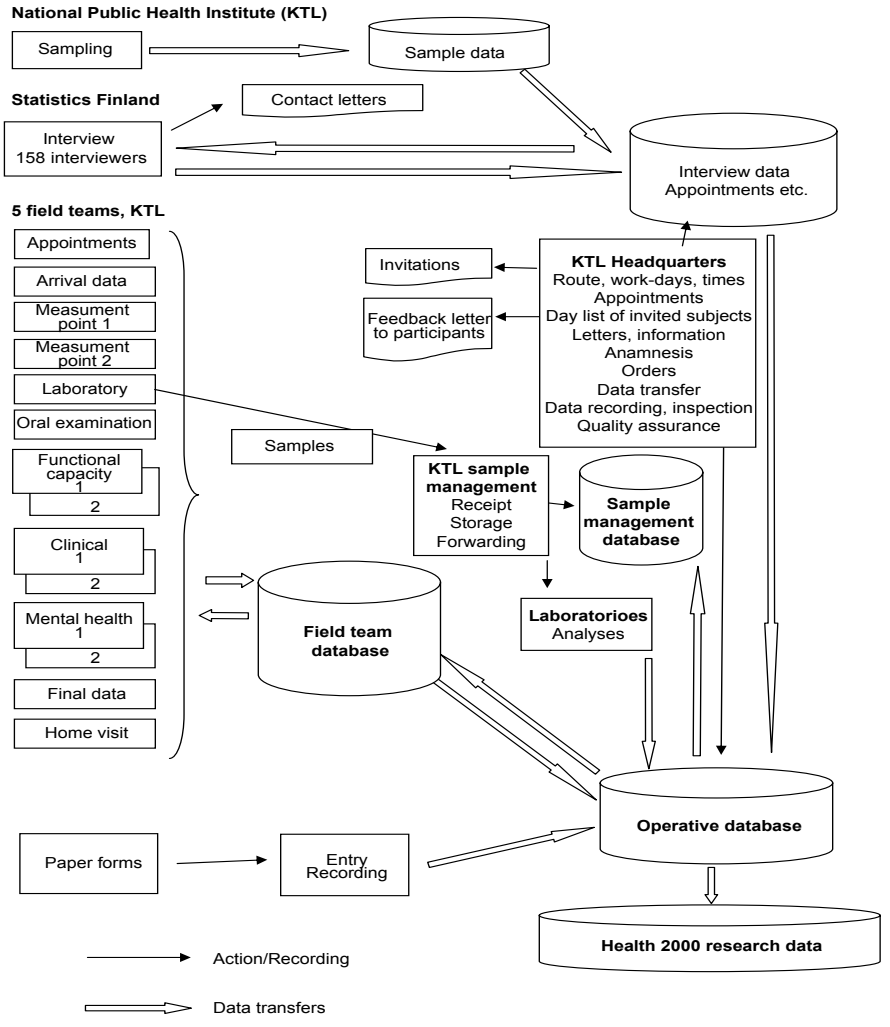
An e-mail box was allocated for each field team on the KTL e-mail server so that the team could easily communicate with project headquarters via e-mail. In addition, the field teams could use a web browser to read and print out pages from the Health 2000 website, for instance.

Each day the person in charge of the field team's IT operations started a routine to copy and transfer the new material accumulated from the field team's server to the mainframe server at headquarters. All transfer files were zipped into batches identified by the creation date and the workstation name. Batches contained only the files not transferred earlier.

Sometimes files were transferred less frequently than once a day. In this case the batch transferred consisted of all the new files created after the previous transfer. New files were distinguished from old ones on the basis of their date and, in the

case of Blaise files, on the basis of the time stamp on one individual data row (interview material for one person). The time stamp indicated the date and time when the row had last been modified. This meant that the same interview material may have been included in several transfer batches if the material had been revised since an earlier transfer. The transfer batches were copied to the KTL server using the scp (Secure Copy) command of the SSH program, which encrypts all data in transmission.

Figure 5.2. Diagram of IT-based data collection in the Health 2000 Survey,



5.3.2. Equipment

Local area networks and hardware used by field teams

The field teams' local area networks were built using two wireless Orinoco WLAN stations (IEEE 802.11b) and a CISCO 800 (LAN ISDN) router. Power supply to the stations and router was backed up with UPS. The WLAN stations and servers were connected to the router by cable, as were the two stations to each other. The work stations in the network thus had access to all the radio channels available. The network was an autonomous TCP/IP network using the addresses of the KTL internal TCP/IP address space.

The network servers were laptop computers running on Windows NT4 Server operating system. In each team the network was used by 13–15 workstations. All the servers and workstations were identical laptop computers (Toshiba Satellite Pro4270). The IT solutions adopted at each health examination point are explained in closer detail under the respective sections.

Equipment at project headquarters

The materials sent in by the field teams were received at the KTL's IT Services Unit by an AlphaServer/Tru64 Unix server. The main database (Oracle Rdb) was on the AlphaServer/OpenVMS. In addition, Windows NT servers were used as file servers. All of this equipment was already in place at the KTL.

Data transfer channels

Data communications between the field teams and project headquarters was by means of Sonera's DNRA service (DataNet Remote Access), which supported access to the KTL network by a maximum of five simultaneous users at a speed of 256 kbit/s. Primarily, communications from the field team to the DNRA service was over Sonera's or local telephone operators' ISDN network (speed 128 kbit/s). Backup was via an analogue connection from the field team's server (56 kbit/s) or GSM modem (Sonera GSM High Speed Data, 48 kbit/s). This was needed only on one occasion.

Both Statistics Finland and the Social Insurance Institution's (Kela) Turku laboratory sent their materials to e-mail boxes specially set up for these purposes on the KTL e-mail server (PMDF e-mail system on AlphaServer/OpenVMS).

5.3.3. Software

Data collection was primarily by means of electronic forms created with the Blaise system; some paper forms were also used. Blaise is a computer-assisted interview (CAI) system developed by Statistics Netherlands. Blaise version 4.3 was used both in Statistics Finland's home-visit interviews and in electronic data capture at the various health examination points.

The Blaise programs had many built-in checks. Acceptable values and the logic of the series of questions were checked at the point of data entry. Identification numbers included a check code (Gumm 1985), which was checked at data entry. Each Blaise file had certain standard variables generated by the program: these included the name of the workstation, the code of the operator, start and end time of data entry, and time when the data row had been last modified.

Data mirroring and backups

At each health examination point, the data were first recorded onto the workstation's hard disk. The Blaise program automatically backed up the form that was being completed at one-minute intervals. In connection with starting up the program, a separate Perl program was also launched to mirror, i.e. copy changed files to the network server. If the internal network connection was not available at this time, these files remained on the local hard disk until the program was restarted and the connection was open again.

The files copied to the primary server were mirrored in the same manner to the field team's secondary server. This was to ensure that both servers were all the time in an identical state. This was particularly important in the event of malfunctions on an individual server.

Software updates

The field team's server included a function that checked whether the mainframe computer had any software updates. If there were, the program automatically downloaded new versions onto the local servers. The programs running on workstations, for their part, automatically checked on start-up for the existence of a new program version on the local server. If a new version did exist, it was copied to the local workstation before starting up the program. All workstations were identical and self-sufficient with respect to their programs. Thus data entry programs could be used even if the field team's local area network was not up and running.

The programs were started with a browser. Each workstation had a local HTML home page, which contained the menu of programs as well as any news downloaded from the headquarters. It was updated from project headquarters following the same principles as in the case of the data entry programs.

5.3.4. Headquarters database

All the data collected in the field were transferred in batches to project headquarters at the KTL. The main database was Oracle Rdb relational database version 7.0 on an Open VMS system. The database was primarily used for logistics control in the following functions:

- Appointments listings to Statistics Finland interviewers
- KTL appointments system
- Invitation lists and field teams' calendars
- Recording and reporting of accountancy data (number of completed electronic forms etc.)
- Stock accounting of samples
- Creating feedback

Programming of the Oracle Rdb database was primarily based on stored SQL procedures. Various listings and reports were produced as HTML pages by using the procedures. The pages could be accessed and printed out using standard browsers both at project headquarters and in the field office. The appointments program was also based on a browser interface.

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

Following the health examinations all subjects were sent a feedback letter. The feedback was generated from the database into rtf format and printed using Microsoft Word (see Chapter 8.15).

5.3.5. SAS software

The database was constructed using version 8.02 of SAS software for statistical analysis (SAS Institute). The database consisted of SAS files created out of individual forms. The Blaise files in each transfer batch were downloaded at project

headquarters into their own directories, and then converted into text files using Blaise Manipula programs, which had been automatically created from the Blaise data dictionary. The same principles were used to generate the SAS programs with which the converted files were collated into one SAS file.

Errors detected later on in the SAS database were corrected using separate SAS programs, which included corresponding data editing commands. SAS software was also the main tool used in subsequent data processing (checks, creation of additional variables, etc.). More on this is described in later chapters.

5.3.6. Data security and privacy protection

The laptops used by the field teams had PGP encrypted disk sections for safe storage of the data collected in the health examination: even if a laptop had gone missing or had been stolen, those data would have remained safe. All communications over the wireless local area network used encryption (128 bit WEP).

Communications between project headquarters and the field teams took place over Sonera's Datanet network (DNRA), not over the public Internet. Nonetheless all electronic materials dispatched by the field teams were encrypted ("secure copy"), as was all HTTP communication from headquarters to the field teams (so-called https protocol was used).

Statistics Finland interviewers used a telephone modem to send back their interview material to Statistics Finland, which in turn forwarded this material to KTL as PGP encrypted e-mail messages. Laboratory results from SII Turku were e-mailed without encryption. However these messages included only the codes for the sample and analysis type as well as the result of the laboratory analysis. This information was not considered sensitive.

All hardware at KTL headquarters is protected by a firewall, and access is restricted to authorised users with usernames and passwords. Access to the health examination material is restricted to authorised KTL staff only. In the main database, transfer files and analysis database, subjects are identified solely on the basis of a serial code exclusive to the health examination. Personal data are recorded in a separate dataset table so that the access to these data can easily be restricted only to few people. Personal data were accessible to field staff from a separate Blaise file, which was saved on an encrypted disk portion. Paper forms with personal data were transported and kept in locked boxes and filing cabinets.

Questions of data protection are also discussed in Chapter 13.

5.3.7. Evaluation of the data collection process

Problems in data collection

Perhaps the biggest IT problems in the project were caused by late planning and preparation for key aspects of the data collection process; indeed in some instances this only got underway as the examinations were in full swing. Examples include the definition of the quality assurance studies and subsamples, the identification of the anamnesis data needed in the clinical examination, and inadequate testing of certain key programme items and functions, such as the examination entry point. Not enough time remained to analyse the experiences from the pilot studies, nor indeed was the material available in these studies broad and diverse enough; for instance the number of subjects in poor health remained rather limited. In the end the necessary hardware equipment was only fully operational once training for field staff was underway, which to some extent complicated the training process.

Some problems were experienced with the WLAN, which would not always reach all the workstations at the same time. Incorrect boot-up order or occasional malfunctions could also cause the network to drop. Even though the network was "wireless", large numbers of cables and wires had to be installed; this was sometimes a problem as well.

Hardware crashes caused some data losses. Blaise always saved its data in a backup file when it was running. If a laptop crashed, the Blaise datafile would be erased. If the unit crashed twice before data transfer, the untransferred data on the form would be lost altogether. This caused the loss of around 50 forms on functional capacity measurements and some from other health examination points.

Successes in data collection

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

The appointments system provided a sound and useful tool for organising the health examinations and the collaboration among the various parties involved. The existence of completed forms was easily recorded with a dedicated program, especially when the forms were bar coded. Bar codes and appropriate programs also helped to ensure speedy and accurate recording of the sample tubes. The same solutions and procedures have been applied later in other KTL surveys.

5.4. Invitations and appointments

Tarja Kiesi and Sami Heistaro

5.4.1. Appointments for health examination

Appointments for the health examination were made by Statistics Finland interviewers in connection with the home-visit interviews. Each interviewer had their own appointments quota. If the interviewer was unable to offer a suitable appointment time, they would call the appointments nurse while still at the subject's home to arrange a suitable appointment.

To prevent backlogs from developing at the health examination points for instance as a result of elderly subjects queuing up, successive appointments were given as far as possible to people of different ages. The last appointments of the day were only given when that was the only way that the subject could participate in the health examination. However these appointments were not given to older people. The field teams' daily timetables were built around so-called "model days" (Table 5.2).

Appointments not used from the interviewer's quota were returned to KTL appointments centre 10 days before the health examinations were due to end at the location concerned. If the interviewer conducted interviews after this and needed to make an appointment, they would phone the appointments nurse and arrange an appointment. If the interviewer did not manage to reach the appointments nurse, no appointment was made in connection with the home-visit interview.

Table 5.2. "Model days" for appointments.

Name	Code	No. of subjects invited (appointments)	Subjects' time of arrival
Base day 1	P1	23	11.00–17.45
Base day 2	P2	24	8.00–15.00
Day of arrival*	T	16	13.00–17.15
Day of departure**	L	15	8.00–11.45
Quality day***	LA	10	8.00–13.00
Long day	PP	30	9.00–17.45

* = day when field team arrived at location

** = day when field team left location

*** = for more details see Chapter 14

The Statistics Finland interviewers used an electronic form to pass on to project headquarters the names of the subjects for whom they had made appointments, the dates and times of those appointments, and the names of those who remained without an appointment. The messages could also contain additional information, such as the subject's right to use a taxi to travel to the health examination or their preference to attend the health examination at another location. The subjects themselves received a separate written note indicating the date and time of their appointment.

The subjects could also contact the appointments nurse directly at projects headquarters at KTL in order to make or change an appointment. The nurse on call kept the only appointments database up to date. She was also in the position to give appointment times that were not in the interviewers' quota, and even appointments to locations outside the subject's own home municipality. Furthermore, the appointments nurse had the authority to create new appointment times and to book double-length appointments.

The field teams were informed of the appointments for the following day in the form of invitations lists. The appointments centre delivered the list of invitations to the field team during the afternoon of the preceding day. Any changes notified to the appointments nurses after midday were conveyed to the field teams by phone or other direct contact. The invitations list included all the subjects' phone numbers so that field team members could contact them directly if necessary.

5.4.2. Cases in which subjects did not receive an appointment for the health examination

The lists of subjects who had not received an appointment for the health examination were reviewed at project headquarters about 10 days before the examinations were due to end at each location. At this stage those people had been deleted from the non-response lists who had deceased since the sample was collected, those who had permanently moved abroad or who had declined to participate for reasons of principle.

As discussed, the aim was to make appointments for all persons in the sample. Subject codes provided easy access to the electronic messages filed by the Statistics Finland interviewers, which might help to shed light on why it had not been possible to make an appointment for the subject.

When the interviewer had informed headquarters that they were unable to contact the subject or that the subject was reluctant to attend the health examination, project headquarters contacted the subject by phone and/or sent an invitation letter

stating that an appointment had been made on their behalf. During the phonecall it was stressed to the subjects that their participation in the examination was very important. A separate list was provided to the field teams on those appointments that were made without the subjects' knowledge. This was also done in cases where it was clear that the examination should have to be done at the subject's home.

Every effort was made to remove and overcome obstacles to participation in the health examination. People with physical handicaps were offered a free taxi ride if that was necessary, or it was agreed that the examinations would be done by home-visit nurses. If the subjects had said they would attend the health examination at another location, their name was added to the list of those who would be "contacted later"; these subjects were phoned when the appointments for that location became available. Interpreters or childminders were provided for the duration of the health examination if that was necessary for participation. This was always agreed upon separately with the field team. Blindness and deafness did not prevent participation, either.

5.4.3. Invitation letters

Invitations to the health examination were sent out about two weeks ahead of time. The invitation letter gave the time of the examination, instructions on how to get to the examination and other relevant information. The covering letter attached to the invitation varied depending on whether the subjects had been contacted by a Statistics Finland interviewer or an appointments nurse, or whether the subject had declined to participate in the examination.

6. HOME-VISIT INTERVIEW

6.1. Interviewers

Tarja Nieminen

The home-visit interviews were conducted by Statistics Finland: the national statistical office has a network of permanent interview staff that covers the whole country. All of the interviewers are professionals who have a basic training for the job. The allocation of assignments, job supervision, and the receipt of completed interviews is centrally coordinated via Statistics Finland's Interview Services in Helsinki.

A total of 158 Statistics Finland interviewers took part in data collection for the Health 2000 Survey. Their average age was 49 years, and most of them were women. On average, they had 11 years of experience of professional interviewing; most of them also had prior experience of health interviews.

For the Health 2000 project the interviewers attended a one-day training courses provided by experts from National Public Health Institute (KTL) and other participating organisations. In addition, the interviewers received written instructions before data collection was started, explaining the purpose of the study, its content and organisation and also providing guidance on certain individual items. Further guidelines and instructions were provided as needed during the course of data collection. Some instructions on individual items were also given on the electronic interview form. This helps to speed up the interview process and ensures that the questions are consistently interpreted. In the event of any problems during the fieldwork, the interviewers could consult Statistics Finland and other experts on different subjects; their contact information was also given in the project guidelines.

During the fieldwork, the interviewers convened regularly in regional meetings to exchange views and experiences of how the interviews were progressing as well as on any problems they had encountered. Where necessary, they could also exchange interviewees if this was considered beneficial to securing the interview.

6.2. How the interview was conducted

Tarja Nieminen

Before data collection proper, two pilot studies were conducted to test the interview questions and the fieldwork process. The first pilot was carried out in January-February and the second in April-May 2000. On the basis of the experiences from these pilot studies, a few adjustments were made to the instructions and to the interview questionnaire.

The interviews for the Health 2000 Survey were conducted between 15 August 2000 and 28 February 2001. Table 6.1 shows the monthly number of interviews carried out in each university hospital district. The number of interviews was the lowest at the beginning and at the end of the fieldwork stage. Almost 60 per cent of the interviews took place in September-October 2000. At each location, the interviews took place before the health examinations, which were conducted consecutively in different health centre districts (see Chapter 5). On average, the health examination took place one month after the interview. The interviews took place in the interviewee's home or, if they so requested, in some other suitable place outside their home. If the interviewee lived in an institution, the interview was conducted there whenever possible. Nine per cent of all interviews were done at institutions. Where necessary, the interviewees' next of kin were asked for help; this ensured that at least the most important data were obtained from as many respondents as possible.

Table 6.1. Monthly number of home-visit interviews in different university hospital districts.

	UNIVERSITY HOSPITAL DISTRICTS					
Month	Helsinki	Turku	Tampere	Kuopio	Oulu	Total
08/2000	7	25	40	11	11	94
%	0.3	2.5	2.5	0.9	1.1	1.4
09/2000	591	348	639	319	410	2,307
%	27.0	35.3	40.4	25.9	41.0	33.0
10/2000	659	268	416	271	213	1,827
%	30.1	27.2	26.3	22.0	21.3	26.2
11/2000	398	190	183	285	98	1,154
%	18.2	19.3	11.6	23.1	9.8	16.5
12/2000	207	68	95	113	76	559
%	9.5	6.9	6.0	9.2	7.6	8.0
01/2001	284	76	179	192	167	898
%	13.0	7.7	11.3	15.6	16.7	12.9
02/2001	44	10	28	41	24	147
%	2.0	1.0	1.8	3.3	2.4	2.1
Total	**2,190**	**985**	**1,580**	**1,232**	**999**	**6,986**
%	**100**	**100**	**100**	**100**	**100**	**100**

The initial contact with the individuals in the sample was by letter. The letter included the interviewer's contact information as well as an introductory brochure on the survey. Next, the interviewer phoned the recipient of the letter to make an appointment. The interview was conducted in the form of a computer assisted personal interview (CAPI). In connection with the interview, an appointment was also made with the interviewee for the health examination. The interviewer entered the date, time and place of the health examination at the end of the interview form, specified the reasons if the interviewee was unable to attend and entered any other relevant information for the health examination teams. Finally, the interviewees were given the questionnaire 1 (see Chapter 7) to complete at home, and they were asked to give their written informed consent to link certain personal register data to those collected during the fieldwork. Some interviewees were given a sphygmomanometer and a follow-up form (see Chapter 6.6).

The interview questionnaire (see www.terveys2000.fi) was divided into ten parts covering different aspects of health status, illnesses, use of health care services, and functional capacity. There were also questions on living conditions, living habits, rehabilitation, employment and work ability. Planning and design of the questionnaire involved a large number of experts from the National Public Health Institute as well as other research institutes involved in the project. The content of the interview questionnaire is described in more detailed later.

The planning of the interview questionnaire was started well ahead of time, but in the end there was a rush to the deadline following prolonged debate on the choice and exact wording of the items. Some work also went into rearranging the items. Questions also needed to be reformulated for electronic data collection, and the whole questionnaire had to be programmed in the order that the component parts were finalised. As changes were made to the questions to the very end of programming, that required special care and accuracy. The interview questionnaire was long and involved complex go-to paths, which had to be adjusted following the changes made to the questions. These changes also complicated the task of testing the questionnaire, which could not be done in one go. In view of the tight deadline and the changes and revisions that were made to the very end, the process overall was highly successful.

Several items in the questionnaire used extensive code systems that the interviewer could use directly to enter the responses to open-ended questions during the course of the interview (e.g. those on occupations, illnesses, medications). If no suitable code was provided, the response options had to be coded after the interview. Occupations were classified on the basis of Statistics Finland's 1997 classification. That otherwise served its purpose very well, but it did not have some of the job titles that would have been needed to code the occupations of the older

respondents. These had to be separately coded one by one after the interview. Most of the classifications worked very well, but the classification of surgical operations would have required some development. The interviewers were particularly pleased with the coding system that had been adapted at Statistics Finland on local municipalities: this included also some former municipalities that were lost to Russia in the Second World War.

Technically, too, the interview questionnaire worked very well. The average duration of the interview was 95 minutes. The questionnaire was prepared in both of Finland's official languages (Finnish and Swedish); if the interviewee did not speak either of these languages, an interpreter was used. However, there were only a few interviews where an interpreter was needed.

Some of the interviewees (2%) were unable to answer the questions by themselves; in these cases proxies were used. Proxy refers here to a person who is familiar with the respondent and who can, where necessary, give responses on their behalf. Most typically, proxy respondents were either the respondent's spouse or children. Some elderly respondents or those who were not well enough to answer themselves, responded to the questions together with the proxy respondents. In the interviewers' assessment, most of the responses obtained were reliable. Questions that required recollection did, however, cause some problems (Nieminen and Kuusela 2004).

Various different methods were used to try and get in touch with the respondents who could not be immediately reached, even at different times of the day. If the respondent could not be contacted by phone, even after several attempts, they would be paid a visit at their home address. Where necessary, the interviewer would visit the respondents at home at least three times. No restrictions were placed on the number of phone calls made. If the person had moved, the interviewer would find out their new address and either carry out interview himself or herself, or transfer the respondent to an interviewer who lived closer to that person.

Ten days before the health examinations were due to finish at each location, the interviewer gave up trying to contact the respondent and delegated the task to the National Public Health Institute's appointments centre; they would send out a written invitation to a health examination (see also Chapter 5). If a person who had not previously been contacted did manage to make it to the health examination, he or she would be interviewed, if possible, in connection with the examination. If attempts to contact the person were still unsuccessful, Statistics Finland's interviewer continued to try and get in touch until the end of the field stage. Interviews were conducted over the phone only in exceptional circumstances.

Most of the interviewees took part in both the interview and the health examination. In the interviewers' assessment, most respondents were quite happy to take part

in the survey. A major factor in this positive attitude was no doubt the extensive information effort ahead of the project. In spite of some early difficulties, the interviewers were also left with a positive impression of the survey (Nieminen 2003).

6.3. Electronic questionnaire and telecommunications

Vesa Kuusela

The health interview was conducted as computer-assisted: the interviewer read the questions from the laptop screen and entered the data directly into its memory. In other words, paper questionnaires were not used at all.

The electronic questionnaire differs considerably from its hardcopy equivalent, not only in terms of its physical appearance. The electronic questionnaire is actually a kind of computer program, which allows for the use of a number of features that are not available when paper questionnaires are used. For instance, the interview program defines the order in which the questions are asked. The order of questions is either strictly defined, or that order may be conditional.

One of the advantages of electronic questionnaires is that the data often contain far fewer errors than the data entered from paper forms. There is no need for any major checks on the data, because many of the procedures that are normally done with paper questionnaires after data entry are actually done during electronic data collection. The interview data collected by means of electronic questionnaires are more readily and quickly available for use than data compiled and collected using paper forms because they have already been checked and there is no need for separate data entry.

The electronic interview questionnaire differs so significantly from the traditional paper questionnaire that it is impossible to accurately describe it as such. Indeed, the printouts of the electronic versions only provide a rather crude overview of the content of the interview (see www.terveys2000.fi). The only way to gain a true picture of how the interview proceeds and of the question routings is via the electronic questionnaire itself.

Information system

The information system supporting Statistics Finland's field interviews took care of the distribution of the interviewees' contact information and questionnaires to the interviewers.

The data collected were also received and assembled by the information system. Data transfer was done using telecommunications software that forms an integral part of the information system and that uses modems to establish connections (see Kuusela and Parviainen 1997, Kuusela 2001).

In the Health 2000 Survey the interviewers returned the data they collected each day back to Statistics Finland, where the data from different interviewers were assembled. From Statistics Finland, these data were forwarded daily or every other day to the National Public Health Institute (KTL) in encrypted e-mail attachments with a time stamp, and KTL acknowledged receipt of the mail.

The interview data were thus at KTL's disposal within a few days of the actual interviews. Delivery of the necessary data to the five clinical field teams was the responsibility of KTL. An overview of the Health 2000 information system is provided in a separate article (Kuusela et al. 2000). Generally it was felt that the data collection went very well and that it was completed on schedule.

6.4. Content of home-visit interview

Arpo Aromaa

The structure of the home-visit interview is illustrated in Table 6.2.

Table 6.2. Themes covered in the home-visit interview.

Background information (marital status and relationship, household, education, occupation and job, working hours and wages, current secondary job, unemployment, information about spouse, income)
State of health and illnesses (perceived health and chronic illness, long-term diseases and injuries, treatment of illness, questions for men, questions for women)
Questions concerning parents and siblings (father and mother, brothers and sisters, childhood living itions)
Health services (availability and accessibility of services, ambulatory visits, mental health services, health examinations and preventive health services, physiotherapy and alternative treatments, medicines, oral health, self-care of the mouth, service use, dental care customer)
Living habits (eating habits, smoking)
Living environment (residential history, housing, services in the neighbourhood)
Functional capacity (usual activities, mobility and moving capacity, sensory functions, need and receipt of assistance, aids, cognitive capacity)
Work and work ability (working conditions, working capacity, working skills, pension attitudes, employment history)
Rehabilitation (use of services, need for rehabilitation)
Interviewer's assessments

6.4.1. State of health and illnesses

Information on perceived health was obtained by a standard item (de Bruin et al. 1996) with five preset responses from good to poor.

The question about long-term illness was formulated in the same way as in earlier Finnish studies (Aromaa et al. 1989a, Purola et al. 1967), covering illnesses and those disabilities and injuries that affect functional capacity and work ability. Next, the respondents were asked to identify these illnesses or disabilities. A maximum of 10 were recorded using a truncated classification that has previously been used in SII health interviews (Arinen et al. 1998).

The next questions concerned occupational diseases, changes of job or occupation due to illnesses as well as current disabilities at work due to illnesses or injuries. These were similar to the items used previously in the Mini-Finland Health Survey (Aromaa et al. 1989b).

Next was a list of common diseases. For these, the subjects were asked first of all whether a doctor had diagnosed them with that disease, and if so when, whether they had received treatment in hospital, treatment by a doctor, medical treatment and what symptoms they had had. The subsequent questions varied to some extent depending on the disease. All in all there were 43 diseases on the list, including asthma, chronic bronchitis, coronary thrombosis, heart failure, hypertension, rheumatoid arthritis, degenerative arthritis, back illness, bone fracture, osteoporosis, permanent injury caused by accident (with aetiology) mental health problem, grey cataract, macular degeneration, hearing defect, diabetes, allergy, skin disease, intestinal disease, cancer, Parkinson's disease, severe headache, incontinence and hypertrophy of the prostate. A similar but shorter list was used in the Mini-Finland Survey, and the same kind of approach has recently been recommended in the WHO Eurohis project (Nosikov and Gudex 2003).

Questions on the treatment of illnesses were organised into one cluster, with separate items for the perceived need for medical treatment and for the assessment of unmet needs. The respondents were also asked what treatments they needed more of. A corresponding battery of questions was included in the Mini-Finland Survey, and it has also been incorporated in earlier stages of health security population studies (Purola et al. 1967).

The next questions concerned waiting times for hospital admission, hospital treatments and surgical operations.

6.4.2. Questions for men and women

Men were asked about the number of children they had and about infertility and related examinations. Women's questions concerned menstruation, pregnancies and childbirths, breastfeeding, miscarriages and infertility as well as infertility treatments. Finally, women were asked about contraception and hormone replacement treatments.

6.4.3. Living conditions in childhood

After questions on parents' date of birth and causes of death, the respondents were asked how many siblings they had and their ages at death. Questions concerning childhood living conditions included items on father's and mother's occupation and education at the time that the respondents first went to school, i.e. at about age 7. Earlier, the corresponding data were collected in the repeat of the Mobile Clinic Survey in 1973–76.

6.4.4. Health services

Health services questions concerned the respondents' family doctor, primary doctor and primary nurse. Next, they were asked about ambulatory visits to the doctor due to illnesses during the past 12 months. A separate item was included on telephone and e-mail contacts. Corresponding questions were presented on visits and other contacts with the public health nurse.

A series of questions was included on the use of mental health services during the past 12 months. In the final item the respondents were asked to give their assessment of how helpful they thought the treatment had been.

The next set of questions concerned health examinations and preventive health services; only some of these items are ordinarily included in national health interviews. First, the interviewees were asked about their participation in various examinations of a general nature during the past five years (and 12 months); examples include the examination for the driving licence medical certificate, medical examination upon entering new job, and health examination for war veterans. The respondents were then asked about specific examinations such as those for eyesight, hearing, cholesterol, bone density, mammography and PSA screening.

Only limited data are available on physical treatments, and therefore questions were included on the number of visits over the past 12 months and on the provider of that treatment. No information is available at all on the use of alternative therapies, even though it is known that that is quite common. Therefore the subjects were asked about the use of these services (e.g. visits to a chiropractor, masseur or homeopath) during the past 12 months.

6.4.5. Medicines

Many European national health interviews (Aromaa et al. 2003b) have inquired about the use of medicines by presenting respondents with a list of illnesses, asking them whether they take any medication for that illness. The Eurohis project has also recommended the use of this method, the argument being that any other approach is too complex and complicated. This, however, does certainly not provide a comprehensive picture of the use of medicines, and of course it is very rarely that the reasons for medicine use are clear and unambiguous. Furthermore, people often are well informed of the medicines they are using and have complete documentation. Ever since the 1960s (Purola et al. 1967, Heinonen 1966) respondents have been asked in 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 present survey.

First, the subject's SII cards were checked to establish their entitlements to special reimbursements for medicine costs. Once these had been recorded, the respondents were asked to identify the prescription drugs they had used during the past 12 months, the names of those medicines and their use during the past seven days.

Next, the corresponding questions were asked concerning non-prescription drugs, including vitamins, natural medicines, and homeopathic products. Medicines were classified in accordance with the Pharmaca Fennica prescription medicine compendium (Pharmaca Fennica 1999, 2000, 2001), other products were grouped on the basis of a classification updated in collaboration with the National Agency for Medicines.

6.4.6. Oral health

Questions on oral health were designed to collect basic information on the condition of teeth, dentures, self-care of the mouth, self-care instruments used and particularly on the use of dental services. Subjects were asked about their visits to the dentist, their primary dentist and in detail about their latest treatments. Finally,

a question was included on the costs of dental treatment. Some of the corresponding questions were used in the Mini-Finland Survey (Vehkalahti et al. 1991), and some in a postal questionnaire designed to assess the impacts of the dental care reform.

6.4.7. Living habits

The living habits items included questions about diet, such as meals, the type of fat used on bread, type of fat used in cooking, and the amount of bread consumed as well as the consumption of fruit and vegetables. Secondly, respondents were asked about their smoking habits using a series of questions recommended by the WHO; this same battery has been used in earlier health behaviour and risk factor studies in Finland (Helakorpi et al. 2000, Vartiainen et al. 1998).

6.4.8. Living environment

The questions concerned place of residence, dwelling, any disadvantages in housing conditions, services in the immediate neighbourhood, distances to services, need for assistance to get to the shops, to see friends, relatives or neighbours and to the health centre.

6.4.9. Functional capacity and need for assistance

One of the major of concerns of this survey was with functional capacity, and therefore a large proportion of the interview items concerned ordinary activities (ADL and IADL) as well as difficulties experienced in these activities. Many of these questions were based on items developed by Katz et al. (Katz et al. 1963, Katz et al. 1970, Lawton and Brody 1969) and on OECD recommendations (McWhinnie 1981). In addition, the interview schedule included a number of questions on various other activities such as shopping, going to the bank and doing heavy cleaning jobs. The next questions concerned mobility and sensory functions. The majority of the items were similar to those used in the Mini-Finland Survey, or the same as those in the health security population survey (Aromaa et al. 1989a).

Next, the respondents were asked to say in which ordinary activities they required and received help, and further to identify from whom, how often and what kind of help they received. Finally, they were asked whether they themselves regularly provided help to other people that helped those people manage at home.

6.4.10. Aids

The next section of the interview dealt with the use of aids. The first question concerned the use of spectacles or other vision aids, the use of a hearing aid or other listening devices, the use of a walking stick and other walking aids such as a walker, wheelchair or power-operated aids. Finally, the respondents were asked to mention any other aids they used.

6.4.11. Cognitive capacity

The home-visit interview also included a brief battery of tests assessing cognitive capacity: this was a short version of the Mini-Mental State Examination (Folstein et al. 1975), which is used in clinical practice for the detection of dementia. The tests included questions to measure orientation and verbal memory, a mental arithmetic task and a drawing task to measure cognitive perception. Finally, the respondents were asked how they rated their own health, whether they had memory difficulties and what they thought might be the cause for those difficulties.

6.4.12. Work and work ability

The items concerning work and work ability were a combination of questions used in the Mini-Finland survey (Aromaa et al. 1989a) and in several studies by the Finnish Institute of Occupational Health (Tuomi et al. 1992, Piirainen et al. 2000).

The working conditions items included questions on job satisfaction, the physical load of work, appreciation of the job as well as physically strenuous job demands, such as lifting heavy objects, working on one's knees or in other awkward positions, and repetitive hand movements. Next, respondents were asked about perceived hazards such as noise, dust or vibration.

A separate set of questions was included on the use of PCs and other computer associated equipment in studies or at work. Respondents were asked to say how much time they worked with this equipment on an average day and how many years they had been doing so.

The interviewees were then asked to assess their current work ability, next to describe any restrictions in their work ability and finally to say when they had become incapacitated for work. They were then asked to compare their current work ability to their best ever work ability on a scale from 1 to 10 and to predict the future development of their work ability. If the interviewee reported restricted

work ability, they were asked about factors adversely affecting their well-being and coping at work or preventing their participation in work.

The next set of questions concerned the respondents' own perceptions of their physical and mental work ability, illnesses or injuries affecting their work ability and their impacts on work, sickness absences and working while sick. The last item in this set of questions was whether the respondents thought that as far as their health was concerned, would they be able to continue in their current job in two years' time.

As regards the skills necessary for their work ability, the respondents were asked about their need for training and how they saw their prospects of re-employment if they were made redundant.

Finally, a few items were included to measure attitudes towards retirement. First, the respondents were asked whether they had considered retiring either full-time or part-time. The next cluster of questions concerned the person's detailed employment history: respondents were asked to list all the occupations in which they had spent more than one year, and for each to describe the level of physical strain, working positions and job tasks.

6.4.13. Rehabilitation

The final battery of questions concerned rehabilitation: the use of rehabilitation services, who pays for the services, institutional rehabilitation, satisfaction with rehabilitation and aids. These were followed by items concerning work and ability programmes in the workplace. Finally, the respondents were asked about the need for rehabilitation, their preferences with respect to the type of rehabilitation and the need for aids.

6.4.14. Interviewer's assessments

Once the interview was completed, the interviewer made some technical notes concerning the interview and the appointment for the health examination. Furthermore, the interviewer was to provide an assessment of the interviewee's mobility, need for help, eyesight, hearing, speech and understanding of speech and instructions. The final notes concerned the presence of a temporary disability and the place of the interview.

6.5. Experiences of home-visit interviews

Arpo Aromaa

The computer assisted home-visit interviews were extremely long: on average they lasted 95 minutes. The programming team had been very much pressed for time to complete their job, and consequently there remained some bugs in the software that caused the loss of some individual data items. The only remedy would have been to give the programmers more time to complete their work and test the program.

The interviewers were experienced in their job, even though some of them had not worked with this subject before. It would certainly have been beneficial if instead of just one day, the interviewers could have been in training for a few days. Nonetheless, the questions were well tested and there were no difficulties in that respect. The assisted classification made it easier for the interviewers to code the responses for occupations, diseases, medicines and operations. Because of the limited amount of time available for training, there was some difficulty with supervising the questions concerning cognitive capacity, especially as it also involved a drawing task. Indeed these had to be reclassified after the field work. Another component that caused some difficulty to both interviewers and interviewees was that concerning rehabilitation: indeed it is difficult for anyone to establish the true need for and proper use of different types of rehabilitation.

Interviewers also had much difficulty in assessing the respondents' living conditions from the point of view of functional capacity and need for help. Indeed this was largely dropped from the survey after the pilot phases. It should certainly be possible to reliably collect this kind of information in a home-visit interview setting, but that would require more extensive development and training than was possible in this instance.

In contrast to ordinary home-visit interviews, these interviews also included people living in institutions. Inevitably, some of these interviews were complicated by the respondents' functional disabilities. In some individual cases the interviewers had difficulty obtaining permission to interview an institutionalised person. In this case they would contact project headquarters. As a rule a phone call by survey management to staff at the institution was enough to resolve the matter and get the go-ahead.

The survey was successful in achieving its targets of securing comparability with the Mini-Finland Survey, health security population survey, other national health surveys, several studies conducted by the the Finnish Institute of Occupational Health and also with a number of European national surveys. Furthermore, the

interview covered a whole range of issues that have not been addressed in any of these earlier studies and a number of questions that have never been used extensively before. The utility and validity of these data will be assessed upon analysing the results.

6.6. Blood pressure measurements at home

Antti Reunanen and Antti Jula

It is known that blood pressure varies according to the situation. More often than not, people arriving for a health examination will have a much higher blood pressure than when they are in the relaxed and familiar environment of their home, for instance. In order to explore this variation in more detail, some of the respondents in the home-visit interview were given a sphygmomanometer to measure their blood pressure before they came to the health examination. The measurements were carried out in line with generally accepted recommendations for the measurement of blood pressure at home (Reims et al. 2001, O´Brien et al. 2003).

The choice was made to focus on the age group 45–74: it was thought that these subjects would be the most appropriate for this analysis as some of them would be taking further examinations concerning circulatory diseases, and for these purposes the blood pressure reading from home would provide a useful complement to the measurement results.

Blood pressure measurements at home were taken with the oscillometric OMRON M4 blood pressure measuring device (Omron Matsusaka Co, Japan, OMRON Healthcare Europe B.V., Hoofddorp, The Netherlands). The units were made available to the Health 2000 survey by the Finnish Omron representative Normomedical Oy.

Statistics Finland interviewers received a brief training on the technique of blood pressure measurements at home in connection with their other training. When the interview was completed, the interviewer asked the respondents whether they would be willing to have their blood pressure measured at home. However, especially in the early stages of the survey logistical problems meant that meters could not be provided to every one of the subjects aged 45–74. During the survey the importer delivered around 1,000 sphygmomanometers for use by the survey, and in the end 2,069 subjects of a total of some 3,500 were given a meter. The meters were delivered randomly immediately when a meter was available without consideration on the health status of the recipient. All of the meters were new when

delivered by the importers. During the course of the survey one meter could be used by 2–3 subjects: meters returned when the subjects arrived for the health examination were given back to the interviewers for reuse.

The manufacturer's recommendations were closely followed in the measurement instructions. Blood pressure was measured from the upper arm of the non-dominant hand, i.e. in right-handed subjects from the left arm. Ordinarily, a cuff measuring 14 x 48 cm was used (rubber nag 13 x 23 cm). In subjects with an upper arm circumference exceeding 35 cm, a long 16 x 65 cm cuff was used (rubber bag 15 x 29 cm). The interviewer specially drew attention to the careful and correct application of the cuff, demonstrating the whole process by first taking a measurement from their own arm and then allowing the subject to follow their example and take their own reading.

Subjects were asked to take their blood pressure readings on seven consecutive days, twice a day. Measurements in the morning were to be done between 6–9 am, after washing up and before breakfast. The recommended time frame for evening measurements was between 6–9 pm. For one hour before taking their measurements, the subjects were instructed to avoid smoking, heavy physical exercise, and eating. Prior to the actual measurement, they were instructed to sit in a straight chair by a table for at least 10 minutes, with the cuff wrapped around the upper arm for at least five minutes.

For the actual measurement, the subjects were to set the maximum pressure reading at 200 mmHg, unless the subject suspected that their systolic pressure might be higher than 179 mmHg, in which case the maximum level was to be set at 240 mmHg. Once it was started up, the manometer automatically raised the pressure to the selected maximum level and then began slowly to reduce the pressure. Once the pressure had come down to base level, the readings for systolic and diastolic pressure were shown on the display unit. The subjects were instructed to record these readings on the forms supplied. In addition to the dates of the measurements, the subjects were to note whether they had used the normal or long cuff and to indicate from which arms the measurements were taken. If they had any problems with using the manometer, the subjects were instructed to study more closely the user instructions that followed with the unit.

The subjects brought along both the manometer and their results to the health examination, where the forms were checked at registration point.

No comparisons were made between the OMRON unit that was used for the blood pressure measurements at home and the results from ordinary mercury manometers. However, in the field examinations measurements were taken from the same subjects using both of these two different meters. Based on the reliability

coefficients (0.87– 0.89), there were no significant differences between the systolic pressure readings obtained with the different meters. The OMRON manometer seemed to yield slightly higher diastolic pressure readings than a mercury manometer, which is a relatively common observation in comparisons between automatic measurements and those based on hearing. For diastolic pressure, the reliability coefficients were in the range of 0.72–0.75, and the differences between the measured averages was statistically significant ($p < 0.01$). So even though the different methods of measurement were found to differ in the case of diastolic pressure, the differences in the results (1–2 mmHg in the average readings) were hardly significant at the individual level.

7. BASIC QUESTIONNAIRE (QUESTIONNAIRE 1)

Arpo Aromaa

In connection with the home-visit interviews, all subjects were given a basic questionnaire that they were asked to complete and bring along to the health examination. The questionnaire was also given to those people who were not going to attend the health examination. These people were requested to return the completed form by mail to the National Public Health Institute, using the SAE envelope provided. If the health interview itself was conducted by phone, the interviewer delivered the questionnaire to the interviewee, who was asked to return the completed form in connection with the health examination or by mail using the SAE envelope. Shorter versions of the basic and infection questionnaires were completed in connection with the home health examinations: this secured some key information without putting too much strain on these respondents who often were in poor health.

If the interviewee had some disability (such as poor eyesight) that made it difficult for them to complete the questionnaire, they would usually be assisted by a family member or staff on the health examination team. In most cases these people who required assistance did undertake to complete the questionnaire. More than half of the proxy respondents agreed to complete the questionnaire.

7.1. Content of basic questionnaire

The main content covered in the questionnaire can be categorised as follows: functional capacity, quality of life and income, common symptoms, weight and height, time use and leisure activities, physical exercise, alcohol use, health promotion, living environment, psychological experiences, mood and feelings, as well as job perception and job strain. There was a large number of items in the questionnaire that are better suited for inclusion in a questionnaire rather than in an interview. Typical examples include several items concerning psychological symptoms. Many of these items were inquired using standardised scales such as SCL90, GHQ, BDI, and MBI.

7.1.1. Functional capacity, quality of life and income

The first set of items was the EUROQOL, which is mainly designed to measure functional capacity (EuroQol Group 1990) and health status at the time of the investigation. These items were complemented with a question concerning quality of life. Next, the subjects were asked about the difficulty or inconvenience caused by an illness at work, in household chores and leisure activities. These were followed by items on income and the impact of sickness expenses.

7.1.2. Common symptoms

The next battery of questions concerned the symptoms that are listed in the SCL-90 scale (Derogatis et al. 1973). These were followed by a number of other symptom items that are frequently included in other surveys.

7.1.3. Weight and height

Weight and height questions included items on putting on weight, losing weight and slimming during the past 12 months, current weight, and weight at ages 20, 30, 40 and 50 as well as height at age 20.

7.1.4. Time use and leisure activities

A long list of items was included on time use and leisure activities, particularly concerning the frequency of participation in different activities. Most of these 16 items were the same as those in the Mini-Finland Health Survey, but a few new items were added. Several questions were asked about computer use, some of which concerned the retrieval of information about health and illnesses.

7.1.5. Physical exercise

The physical exercise component started with questions on physical exercise during leisure time based on the so-called Gothenburg scale (Wilhelmsen et al. 1972), which was also used in the Mini-Finland Survey. Next, the current version of the IPAQ scale (Craig et al. 2003) was used to measure physical exercise during leisure time and in household chores as well as walking and sitting. Physical exercise in commuting was asked in the same manner as in the Mini-Finland Survey.

7.1.6. Alcohol use

Alcohol use was estimated using a scale developed for the Mini-Finland Survey and more detailed questions that had been used in earlier Finnish and foreign studies on drinking (Simpura 2003). The first item concerned the frequency of alcohol consumption, followed by a question on the frequency of drinking beer, cider or premixed drinks and the amount ordinarily consumed per day. The last item was the Mini-Finland question regarding average alcohol consumption during the past month in bottles per week. Corresponding questions on frequency and quantities were asked for wine consumption and for spirits.

The final items on alcohol consumption were designed to assess maximum quantities consumed on any one drinking occasion as measured in terms of standard drink portions. The last question concerned the frequency of hangovers, followed by an item on service use and treatment received for alcohol problems.

7.1.7. Eating sweets

Items on eating and drinking sweets were designed to measure the frequency of sugar and xylitol consumption.

7.1.8. Health promotion

Health promotion was assessed with a question on participation in various groups and courses; among the examples listed were slimming, quit smoking, neck and back and self-care groups.

7.1.9. Living environment

The living environment questions concerned the perceived safety of one's immediate environment: some of the items were the same that have been used in earlier Finnish studies on safety and insecurity (Heiskanen et al. 2000). The next questions concerned the respondents' living environment in their formative years, followed by items concerning experiences of trust and confidence.

7.1.10. Psychological experiences, symptoms and depression

Concerned with psychological experiences and symptoms, the next component of the questionnaire was a Finnish version of the 12-item GHQ scale (Goldberg 1972, Pevalin 2000), which has been widely used as a screening tool. The 36-item version of the GHQ was also used in the Mini-Finland Survey (Lehtinen et al. 1991). Under the heading of mood and feelings, the questionnaire also included Raitasalo's (Raitasalo 1977) modification of Beck's depression inventory (Beck et al. 1961).

7.1.11. Job perception and job strain

The items concerning job perception and job strain were only presented to subjects who had been gainfully employed during the past 12 months. First, they were asked about the physical strenuousness of the job, which was followed by a series of items on job strain. This, the Maslach Burn Out Inventory (Kalimo and Toppinen 1997, Maslach and Jackson 1981, Maslach et al. 1996) has also been used in previous surveys in Finland. The next set of questions concerned job requirements and working conditions and was based on items used in studies by the Finnish Institute of Occupational Health (Piirainen et al. 2000). Other items concerned perceived threats and uncertainties, the climate in the workplace community as well as opportunities for training.

7.2. Experiences of basic questionnaire

The questionnaire schedule worked as intended, and almost all subjects returned completed forms at the health examination. The questionnaires were checked on arrival, and any missing items were completed jointly by the registration nurse and the respondent. The subjects had significant difficulties completing the IPAQ series and in assessing the maximum amounts of alcohol consumed on one drinking occasion. Looking back at the awkward structure of these questions, this was hardly surprising. At the same time as the Health 2000 Survey was underway, the IPAQ questions were used in two other national interview studies in Belgium and Denmark, and both of those projects reported problems with the current version of IPAQ.

Web link: http://www.terveys2000.fi/lomakkeet/en/t2002en.pdf

8. HEALTH EXAMINATION

8.1. Registration

Sirkka Rinne

The health examination started with registration. Staff at the reception point were instructed to record the arrival of each subject, verify their identity, check their personal file and hand out the forms they were to complete at the examination.

At the reception point, subjects were also given general information about the health examination, what it involved and what would happen later, what kind of feedback they would receive, etc. If possible, an interview was conducted to assess the presence of symptoms (symptoms interview). Each field team had at least two trained reception nurses who alternated between registration and the second measurement point (see below). In connection with home visits, registration was completed by the home-visit nurse.

Daily appointment lists were printed by the reception nurse from files that had been produced for each field group by the National Public Health Institute's (KTL) Appointments Office. The nurse delivered these lists to each health examination station, making sure that any changes or cancellations were duly recorded. If there were subjects on the list who according to the Statistics Finland interviewer or KTL Appointments Office needed personal assistance or an interpreter, that was organised by the health examination nurse together with the home-visit nurse.

Health examination files with sets of printed forms and tailored health examination schedules had been prepared in advance for each subject in the sample. The files were organised by location and health examination number. Some of the forms had prefilled identification data, others had sticker labels. At the point of registration these files were assembled in the order of the appointments list on the day before the health examination. Files were also prepared for any subjects who were not included in the sample.

The subjects were called to the point of registration by their name. After introductions and initial information, the subject's examination number was checked from the appointments list or examination file and entered into computer. The subject's personal data, contact information and subsample data were brought onto the screen by the Blaise program. The subject's identity was verified by asking them to produce their SII card, driving licence, or other similar form of identification. These data were compared with the sample data. Contact details were updated where necessary, including any imminent changes of address. The

subject was given their health examination number and information on how that number would be used during the examination and later in processing the data collected. Other information recorded at this stage included the date of the health examination, time of arrival and the nurses' user code.

Staff at the registration point were also instructed to provide information to the subjects about the Health 2000 Survey in general and to try and foster a positive attitude towards the project. All subjects were handed an information sheet and given oral information about the various stages of the survey.

Participants were asked for written informed consent. The consent form made it clear that all information collected would be handled confidentially and used for medical research purposes. It also explained that samples taken and data retrieved at a later stage from various registers would be used in research projects exploring the causes and prevention of major public health problems, functional disabilities, etc. By signing the consent form, the subjects gave their express permission that the information gathered in the health examinations may be used and linked with other data needed for medical research.

Before the subjects signed the consent form they were asked whether they felt they had received sufficient information and that they understood the purpose of the document they were about to sign. The informed consent form was prefilled with the subject's date of birth, address and printed name under the signature line; these were also checked by the subject before signing. Proxy consent was obtained in cases where for reasons of poor health the subject was unable to fill the consent form.

Sheets of sticker labels were provided to each field team for the collection of laboratory samples; one sheet was reserved for one subject. Each sheet had its own numerical ID code, which was entered into the Blaise program. At the registration point one of the labels was also affixed to the health examination schedule. In the laboratory the labels were used to identify the samples analysed.

The subjects were divided into two lines with a view to ensuring a balanced load between the assessment of functional capacity and the clinical examination. Staff at the registration point were also responsible for different questionnaires and for participant allocation to quality control and other sub-samples.

Home-visit interviews were almost always conducted prior to the health examination. On this occasion, the respondents were handed the basic questionnaire (questionnaire 1, see Chapter 7) that they were asked to complete and return at the health examination. Staff at the registration point checked whether the subject had been interviewed and whether they had filled in the questionnaire. On the day of the health examination the subjects were also asked to fill in an infection questionnaire

(questionnaire 2) as well as any missing data from the basic questionnaire. If the subject had not been interviewed in connection with the home visit, the interview was rescheduled to be done in connection with the health examination.

In connection with the home-visit interviews some of the subjects had received a sphygmomanometer and a form for recording blood pressure readings. This form was checked at the point of registration; at the same time the subjects either returned the meter, agreed on arrangements for returning it later, or filled in the necessary papers so that they could purchase the meter. Some subjects also received a form requesting their views and opinions on the Health 2000 Survey.

Finally, the subjects were asked when and what they had eaten last. The data on type of meal were fed into the Blaise program using a seven-category classification. Women under 45 were asked whether they were pregnant and if so, which week of pregnancy they had reached.

The subjects were asked to proceed and advised not to eat anything until they had been to the oral examination. Diabetic subjects, however, were advised to eat any snacks and take any medications that they would normally. A labelled receptacle was given to the subject for the collection of a urine sample.

Files and appointments lists for the next day were assembled at the end of each day. The nurse at the registration point also compiled statistics on the people who had arrived and those who had not shown up. A separate form was completed for every person who did not attend. It was also the reception nurse's job to make sure that every effort was made to contact as many non-participants as possible so that they could be seen either at the health examination or by a home-visit nurse. Most of these contacts were left to the home-visit nurse. Some of the people who failed to turn up had made an appointment in another field district; the files of these subjects were delivered to the field team in question.

8.2. Symptoms interview

Antti Reunanen and Markku Heliövaara

Following registration, the subjects were given an interview in which standard questions were used to determine the presence of symptoms of circulatory diseases, respiratory diseases and musculoskeletal disorders and to collect information on allergies and hand eczema. A major point of emphasis in interviewer training was to stress that the questions must be asked exactly as worded in the form or the computer program.

One of the key design concepts in planning the Health 2000 Survey was to create a system of collecting responses to the symptoms interview that would give doctors responsible for the clinical examinations immediate access to those responses. Unfortunately, for reasons of time constraints and technical difficulties, this never materialised. In future health surveys it would be important for clinical doctors to have access to these preliminary data. The interviewer entered the responses in the Blaise program in the ordinary fashion, and the data collected are available for other research purposes.

8.2.1. Symptoms indicative of respiratory diseases

Diseases of the respiratory system were inquired by questions about daily coughing for at least three months, the production of sputum with cough, and shortness of breath associated with exertion. The questions on habitual coughing lasting at least three consecutive months year and sputum production were originally formulated by the British Medical Research Council Committee (1965). The WHO has recommended that this set of questions be included in population studies (Rose and Blackburn 1968). The item on chronic coughing with sputum production was aimed at identifying subjects with chronic bronchitis causing these kinds of symptoms.

Shortness of breath upon physical exertion was queried with four questions, which again are in line with the original recommendations of the British MRC (Fletcher et al. 1959) and with WHO recommendations for population studies (Rose and Blackburn 1968). Shortness of breath in connection with physical exertion may be due to respiratory or circulatory diseases, or simply to poor physical fitness. Rather than aiming to establish the underlying cause of the symptom, the purpose was simply to determine its presence.

8.2.2. Symptoms of circulatory diseases

Exertion-associated chest pain symptoms indicative of coronary heart disease, intense attacks of chest pain and symptoms of intermittent claudication indicative of arterial occlusion of the lower extremities were queried 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).

8.2.3. Atopy, allergy and hand eczema

The allergy questions in the symptoms interview concerned hay fever, allergic rhinoconjunctivitis and atopic eczema. In addition, the subjects were asked whether they had ever had hand eczema. The main purpose of these questions was to determine the presence of skin symptoms with an allergic basis. For reasons of sensitivity, the questions were not limited to allergies of which the subjects were aware, but they were asked to mention any instances of hand eczema they recollected.

The allergy questions were based on a series of items recommended by Finnish dermatologists for determining the prevalence of allergies and hand eczema (Susitaival and Husman 1996). The dermatologists' battery of questions, for its part, was designed against the background of international diagnostic criteria for atopic eczema (Williams et al. 1994 a,b).

8.2.4. Symptoms of the musculoskeletal system

The rest of the items in the symptoms interview concerned symptoms of the musculoskeletal system. The first questions addressed pain symptoms in the back, neck, shoulders, joints in the limbs as well as the frequency of associated complaints and ailments.

Symptoms indicative of carpal tunnel syndrome were queried with a series of four questions.

Finally, the subjects who reported pain symptoms indicative of musculoskeletal disorders were presented with follow-up questions in which they were asked to assess the level of subjective disability caused by the pain on a VAS.

8.2.5. Quality control

The quality of the symptoms interview was monitored throughout the research process by conducting repeat interviews with the same subjects, either by the same interviewer or another interviewer on the team. The κ-coefficient describing the repeatability of responses to the question concerning exertion-related chest pain was 0.78, whereas for the sputum production question it was only 0.55. The occurrence of intense chest pain or claudication was so rare that it was not possible to obtain reliable results for repeatability. The κ-coefficients for questions concerning skin symptoms ranged between 0.69 and 0.83, for musculoskeletal disorders and balance disturbances between 0.75 and 0.97. All in all the coefficients indicated relatively good repeatability.

Long-term variation in symptom occurrence was assessed in a quality control material in which the symptoms interview was repeated six months after the base examination. The κ-coefficient for the question concerning exertion-related chest pain had dropped to 0.51, and for the bronchitis question to 0.31. The coefficients for pain symptoms indicative of musculoskeletal disorders were in the range of 0.36–0.65, for the balance disturbances question the coefficient was only 0.23.

The symptoms questions showed relatively good consistency in the field examination when the questionnaire was repeated during the same day. There was much more variation in symptoms reported by the subjects six months later. This is quite understandable in that the responses obtained are influenced not only by factors of repeatability, but also by natural biological variation.

8.2.6. Comparability with the Mini-Finland Survey

The Mini-Finland Survey (Aromaa et al. 1985) had more questions on sputum symptoms than the Health 2000 Survey. The Mini-Finland Survey sought to distinguish between cough symptoms appearing at different times of the day as well as seasonal variation in symptom occurrence. However, since the most essential symptom of chronic bronchitis is long-term productive cough, the decision was taken that the series of questions should be simplified for the Health 2000 Survey.

Questions concerning exertion-related shortness of breath were the same as in the Mini-Finland Survey.

The series of questions in Health 2000 on angina pectoris, chest pain indicative of myocardial infarction and claudication symptoms is unchanged both from earlier Mobile Clinic Surveys (Reunanen et al. 1983) and from the Mini-Finland Survey (Aromaa et al. 1985).

The Mini-Finland Ssurvey included no questions concerning allergies and hand eczema.

The questions concerning pain associated with musculoskeletal symptoms were largely the same as those asked in the interview on joint symptoms in the Mini-Finland Survey (Sievers et al. 1985). The Mini-Finland Survey had no questions on symptoms indicative of carpal tunnel syndrome or VAS questions to assess the subjective intensity of pain.

There are good grounds to assume that the symptoms interview in the Health 2000 survey is well comparable with the Mini-Finland Health Survey.

8.3. Measurement point 1

Antti Reunanen

Following the entry and symptoms interviews, the subjects were directed to measurement point 1. Here, their blood pressure and heart rate were measured and rest ECG recorded to obtain important information on circulatory diseases. In addition, basic anthropometric measurements were taken, including height, body circumference measurements and sagittal body measure. Initially measurements were also taken of neck circumference, but this had to be abandoned because of technical difficulties and time constraints.

8.3.1. Blood pressure and heart rate

The invitation letter sent out to the subjects had instructed them to abstain from smoking for at least one hour before the health examination. They had also been asked not to eat for at least four hours and to avoid physical exertion before the examination. Although this was the first point of measurement immediately after the entry and symptoms interview, many of the subjects had to sit down and wait for 5–10 minutes before they could be seen in the measurement room.

As soon as they entered the room, the subjects were asked to undress their upper body and take off their shoes and socks. Women did not have to take off their bras. Temperature control was an important consideration when the facilities were set up for the health examination. If the subjects felt the room was not warm enough, they were initially asked to remove their clothing from the upper right arm only.

Blood pressure was always measured from the right arm if possible. 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. Measurements were taken with a standard mercury manometer (Mercuro 300; Speidel & Keller, Jungingen, Germany). All the manometers used were brand new and had been calibrated before use. The width of the rubber cuff was 12 cm and its length 35 cm. If the proximal circumference of the upper arm measured at a height of 5 cm from the crook of the aim was in excess of 35 cm, a larger cuff (width 15, length 43 cm) was used.

Blood pressure was measured after the subjects had been seated quietly in the measurement room for at least five minutes. Current instructions (Rose et al. 1982, Finnish Hypertension Society working group 2002) were followed in wrapping the cuff around the upper arm, positioning the bend of the elbow at the level of the

heart and in listening to the Korotkoff sounds. Prior to the measurement proper the level of systolic pressure was determined by palpating the wrist artery. At the same time the heart rate was measured by counting the number of pulses from the artery in the wrist during 30 seconds.

For the blood pressure measurements proper, the mercury column was raised either to the level of 180 mmHg or 30 mmHg above that level at which the pulse disappeared, if the systolic pressure measured at the wrist was higher than 150 mmHg. The pressure was then steadily released at 2–3 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. The same instructions were followed when a second set of readings was taken two minutes after the first measurement.

The quality of blood pressure and heart rate measurements was constantly monitored during the field examination and in connection with separate quality control checks during which measurements were taken from people who were not included in the study population. According to the results of continuous quality control, the reliability coefficient between different people taking the measurements or between different times of the day for the same people taking measurements was 0.85 for the measurement of systolic blood pressure; 0.79 for the measurement of diastolic blood pressure; and 0.64 for the measurement of heart rate. On quality control days, the reliability coefficients between people taking measurements in different field teams ranged from 0.81 to 0.90 for systolic blood pressure; from 0.51 to 0.80 for diastolic blood pressure; and from 0.71 to 0.80 for heart rate. Based on these reliability coefficients the quality of the blood pressure and heart rate measurements may be seen as satisfactory.

As expected, systolic measurements showed much better repeatability than diastolic measurements: this is due to differences in the characteristic accuracies of the targets of measurement. However it is important to bear in mind both with blood pressure and heart rate measurements that repeatability is affected not only by factors that have to do with the actual measurement process, but also by biological variation. No matter how rigorous the quality controls, the latter cannot be taken into account with sufficient accuracy. This is reflected in measurements taken from the same subjects on average six months later, when the reliability coefficient for systolic blood pressure was recorded at 0.72, for diastolic blood pressure at 0.61 and for heart rate at 0.68.

8.3.2. Height

Height was measured using a wall-mounted stadiometer (Person-Check, Medizintechnik, KaWe, Kirchner & Wilhelm, Germany). The subjects stood upright with the feet together, head up and back against the wall. Height was preferably measured without socks, but thin socks were allowed depending on the circumstances. Height was recorded to an accuracy of 0.5 cm.

During the course of the survey it transpired that damage might be caused to the stadiometer when it was removed from the wall and moved to the next examination site. It was not always possible to get an immediate replacement. Therefore on some days height was recorded on the basis of self-report by the subjects.

8.3.3. Body circumference measurements

Body circumference measurements were taken at waist and hip level in line with recommendations for anthropometric measurements in population studies (WHO 2000, Seidell et al. 2001). By the time of these measurements all subjects had undressed the upper body. For body circumference measurements the subjects also undressed the top layer of clothing on their lower body. Measurements were taken using a regular, flexible tailor's measuring tape. Body circumference measurements were not taken if the subjects were unable to stand upright or if they were pregnant and the pregnancy was beyond the 20th week.

Waist circumference was measured in standing position, with the legs slightly apart. Horizontal waist position was determined as the mid-point between the lowest rib bones and the high point of the iliac crest, i.e. strictly on the basis of bony points of reference: shape of stomach or navel location had no bearing on the measurements. The person taking the measurements was seated in front of the subjects. Measurements were performed during light expiration. Circumference measurements were recorded to an accuracy of 0.5 cm.

Hip circumference was measured at the point of maximum girth. The determination of that point on the basis of bony reference points alone is not possible on account of individual structural variation. Most typically, the widest hip circumference was measured slightly below the iliac crest, often more or less at the upper end of the symphysis pubis. For the measurement of hip circumference the subject stood erect with legs slightly apart, in the same way as for the measurement for waist circumference. Measurements were again performed during light expiration and recorded to an accuracy of 0.5 cm.

The quality of circumference measurements was controlled almost as closely as the quality of blood pressure measurements. According to the results of continuous quality control the reliability coefficient for measurements of waist circumference was 0.95 and for hip circumference 0.97. There were only minor differences between the measurements taken by staff in different field teams: the reliability coefficients for waist circumference ranged from 0.94 to 0.99 and for hip circumference from 0.94 to 0.98. Six-month follow-up measurements on the same subjects to describe biological variation showed a reliability coefficient of 0.90 for waist circumference and 0.89 for hip circumference. The quality of circumference measurements may be regarded as relatively good.

8.3.4. Sagittal measures

Sagittal height of the body, i.e. height of abdomen when lying in a supine position is one of the indicators used in assessing the amount of fatty tissue in the abdominal region (Kahn et al. 1996). Sagittal height was measured using an instrument developed by the NPHI technical unit, which uses the distance between two parallel wings to determine body dimensions. Instruments were specially commissioned for each field team in the Health 2000 Survey.

For the measurement of sagittal height the subject was asked to lie down on the examination table. One of the wings of the measuring device was placed under the subject's back, at the high point of the iliac crest. The subject was asked to relax and breathe regularly. The other wing of the measuring device was now lowered gently onto the subject's stomach. Measurements were performed after normal expiration. Sagittal height was recorded to an accuracy of 0.5 cm.

While the subject was lying supine on the examination table, the same instrument was used to measure body width. Measurements were taken at the high point of the iliac crest and recorded to an accuracy of 0.5 cm.

The quality of sagittal measurements showed somewhat more variation than circumference measurements. Nonetheless according to the results of continuous quality control the reliability coefficient for sagittal measurements was high at 0.88. The reliability coefficients for measurements taken by different people in the field groups and for biological variation were roughly of the same order.

8.3.5. Resting ECG

After the sagittal measurements, while the subject was still lying supine on the examination table, a routine 12-lead resting ECG was recorded for each subject. The procedures followed were the same as those recommended for clinical practice and population studies (Heikkilä 1982, Rose et al. 1982). Recordings were taken using Marquette Hellige MAC 500 electrocardiograms (Freiburg, Germany and Milwaukee, WI, USA), which record the electric impulses measured directly into the computer's memory and produce an ordinary hard-copy printout showing both pulse amplitude and distance and offering an automatic diagnostic assessment. Resting ECG was recorded at a paper speed of 50 mm/s.

While the subject was lying supine on the examination table, limb and chest electrodes were positioned in accordance with instructions for standard 12-lead ECG measurement. Heavy skin hair especially on the chest was shaved if necessary, but as a general rule no other procedures were applied to the skin. Disposable electrodes were used as far as possible throughout the examination, but on a few occasions these ran out and reusable electrodes had to be used. If the subject had a fully or partly amputated limb, the limb electrode was attached to the amputated stub.

Before ECG recordings were taken, the subject's ID code was entered into the ECG computer together with data on the subject's age, sex, blood pressure, height and weight. Blood pressure and height were entered according to the measurements just taken. Weight was entered on the basis of self-report, because the subject's weight had not yet been measured. No data were entered on medication.

Once this information had been fed into the computer and the subject was relaxed and breathing regularly, the recording button was pressed. If the technical quality of the ECG recording was not satisfactory, a new recording was taken, possibly after checking the attachment of the electrodes. As a general rule disturbance elimination was not used unless the subject had significant muscular tremors that interfered with the recording.

The recordings were simultaneously entered into the computer memory and printed out in 6-lead views. Two sets out printouts were taken for each subject, one of which was later given to the subject in connection with the clinical examination for delivery to their own GP, once the examining doctor had interpreted any abnormalities in the ECG strip. The other printout was attached to the subject's other results.

The electrical signals recorded in the electrocardiogram were transferred on 2HD diskettes in batches of around 100 recordings. These disks were sent for further

analysis to the Social Insurance Institution's research centre in Turku where the contents of the 2HD diskettes were stored to a local network serversystem (MUSE CV, Marquette Electronics Inc., Milwaukee, Wi, USA). Originally the plan was to transmit the ECG results over telephone lines, but this had to be abandoned because most of the rooms where the ECG recordings were taken did not have the necessary landlines.

Abnormalities identified visually in the ECG strips were coded in accordance with the Minnesota coding scheme (Prineas et al. 1982, Rose et al. 1982). The electric ECG signals were further processed using computer software developed for the evaluation of ECG changes. In the first phase, the identification numbers were checked and corrected if needed. The technically most valid recording of the often several recordings of the same subject was chosen and stored, and additional recordings were discarded. The checked and corrected ECG data were then transferred with 2HD diskettes to a separate PC where the electrical recordings were analysed with Magellan software programme (Marquette Electronics Inc, Milwaukee, Wi. USA). At this phase, the measurement points were checked and corrected if needed. At the end of this procedure, an Excile file was created of every single ECG containing several different measurements. These files included durations and amplitudes of P waves, T waves and QRS complexes, durations of conduction times (QT and PR intervals) and diagnostic suggestions of the 12SL software (Marquette Electronics Inc, Milwaukee, Wi, USA).

8.3.6. Comparability with the Mini-Finland Survey

The Mini-Finland Survey (Aromaa et al. 1985) and the Health 2000 Survey used similar mercury manometers and the same guidelines for the measurement of blood pressure. The results are directly comparable.

Both surveys followed the same procedures for recording the 12-lead rest ECG on paper printouts (Aromaa et al. 1985). The ECG units were not the same, but the plotted ECG curve is exactly identical because the units use the same technology. In addition to the standard 12 leads, the Mini-Finland Survey also took recordings from the three Frank orthogonal leads: the signal from the Frank lead system was later used in the automatic interpretation of the ECG results. The Health 2000 Survey, on the other hand, used the ECG unit's automatic reading system and later separate signal processing software to interpret the ECG findings. In both studies the same procedures were followed in coding the ECG findings in accordance with the Minnesota scheme, and therefore the results are fully comparable.

8.4. Measurement point 2

Markku Heliövaara

8.4.1. Spirometry

Spirometry is primarily used for diagnosing respiratory diseases, monitoring the effectiveness of treatment and assessing severity (Cotes 1975). In population studies spirometry has been used for purposes of monitoring respiratory inadequacies and measuring respiratory function (Aromaa et al. 1985). Spirometry tests are 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.

Measurements for the Health 2000 Survey were taken with a Vitalograph 2150 bellow spirometer, which is very similar to the device used in the Mini-Finland Survey. It has a long and proven track record and it has been widely used in clinical work.

Spirometry measures ventilation, the movement of air into and out of the lungs; the nature of any lung dysfunction (bronchial obstruction and restricted inspiration volume); and its severity. If obstruction was detected, a bronchodilator test was performed to determine reversibility. The key measurements taken were:

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

The spirometer was checked for proper function every day before the measurements were started. Every night the nurse removed the metal filter from the capsule in the spirometer tube in order to allow it to dry. The bellows were replaced according to the manufacturer's instructions after approximately 1,000 measurements: this was done during the Christmas break, and the old bellows were set aside as spares. The spirometer was calibrated daily by the nurse using a one-litre calibration pump and a separate calibration table. She then blew into the spirometer to produce a spirometer curve, recorded the date and her own code on the paper, measured FEV1 and VC from the curve, entered her results into the diary of measurement point 2 and filed the curve. All calibration data for the whole duration of the fieldwork were collected in this file.

During the test the subject stood erect with a straight back. The spirometer was on the table, usually placed an upturned removal box to hold the tube as straight

as possible. If the subject was unable to complete the test standing up, they were advised to sit down; in this case the spirometer was lowered to table level. A noseclip was not used. The disposable mouthpiece was placed in the subject's mouth between the teeth, lips firmly around the mouthpiece. The nurse instructed the subject to fill their lungs with air and then to exhale as forcefully as possible, urging them on towards the end of the exhalation.

When the subject had learned the technique and had had one or two dry runs, the aim in the test proper was to produce three as consistent curves as possible. The maximum permissible difference between the two highest FEV1 and VC values was 10%. This often required 4–6 efforts. The subject was allowed to rest and to remove the mouthpiece between measurements. The nurse scored each subject for cooperation on a scale: 0 =poor, 1 =good. If no measurements were obtained, the nurse recorded the reason in the subject's file.

From the three recordings, the nurse selected the curve with the highest FEV1 value, used a ruler to measure FEV1 and VC values, calculated FEV%, recorded the results on the curve and entered the results into the file. A broncholidator test was performed if the subject showed good motivation and if the test performance as such was satisfactory but FEV% was less than 70%. The nurse sprayed two 0.1 mg doses of salbutamol aerosol (Ventoline Evohaler) into a holding chamber (Volumatic) and made sure that the mouthpiece was firmly and tightly in the subject's mouth. The subject was 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 10 minutes after the administration of salbutamol, usually after the bioimpedance test and ultrasound examination.

The results for FEV% are highly comparable between the Mini-Finland Survey and the Health 2000 Survey, which indeed was the main objective in designing the measurements. The Mini-Finland Survey did not, however, include broncholidator tests. Furthermore, the spirometer used in the Mini-Finland Survey was calibrated with a calibration pump after 1,000 measurements, whereas in the Health 2000 Survey it was calibrated every morning. Atmospheric pressure was not measured in Health 2000 since these figures can be obtained if necessary from the Finnish Meteorological Institute. Differences in atmospheric pressure mainly have a bearing on the comparability of VC and FEV1 values.

The success of spirometry and the reliability of its results depend most crucially on full adherence to the procedural guidelines set out. By all accounts this has been achieved in both surveys (Aromaa et al. 1985). Test-retest reliability coefficients for VC and FEV1 values at the field stage were within the range of 0.96–0.99. However spirometric measurements of breathing function are dependent not only

on the health of the respiratory system (Cotes 1975), but also on such factors as the mechanical properties of the chest, muscle function and the condition of teeth or dental prostheses, which must be taken into account in interpreting the results.

8.4.2. Bioimpedance

Electrical bioimpedance analysis was developed in the mid-1980s for the determination of total body water content, fat free mass and fat mass. The method has since become well established in clinical practice. Bioimpedance makes use of the fact that different types of tissue vary in their electrical conductivity: muscle tissue consists of about 80% water and therefore shows much better electrical conductivity than fatty tissue. Electrical resistance is directly proportional to the length of the conductor (i.e. the extremities and body) and inversely proportional to its cross-section areas.

Early instruments had four electrodes, one for each limb. The results were affected by the subject's position, and it took around 10 minutes to complete the measurements because of the requirement of stabilising the distribution of body fluids. The attachment of the electrodes was often considered very awkward. For these reasons four-polar bioimpedance analyzers have not gained very much popularity in epidemiological research. The Health 2000 Survey opted to use a newly developed eight-polar tactile-electrode impedance meter (InBody 3.0, Biospace, Soul, South Korea) which weighs the subject, measures resistances separately for each segment at AC frequencies of 5, 50, 250 and 500 kHz and – working on the assumption that the body consists of five cylindrical segments (four limbs and the body) – uses the formulae fed into its computer to calculate estimates of the composition of the body and its five segments. This eliminates the need to stabilise body fluid distribution for the assessment of total body water content and the ratios fat free mass and fat mass. Contact electrodes are attached to the palms of the hand, the thumbs, the heels and the balls of the foot. The analyzer was chosen for use in the field experiment on account of its speed, comfort and ease of use.

Each morning before the measurements were started, the technician checked to make sure there was enough paper in the printer, switched on the analyzer and waited for about five minutes for it to upload, and then performed a self-measurement. Subjects were prepared for the measurement by wiping the soles of the feet and the palms with an electrolytic wipe. This served the dual purpose of cleaning and disinfecting as well as improving electrical conductivity.

The subject was instructed to step onto the analyzer, placing the heel of the foot on the round electrode and the ball of the foot on the oval electrode pad and to grip the hand electrodes so that the thumbs and the palms were firmly in contact with

the electrodes. The thumb electrode was to be gently pressed down. The subjects were asked to stand with their arms relaxed at their sides, slightly apart from the body. The subject's position was adjusted if necessary. The technician entered the subject's data, which showed up on the analyzer display. Age was entered to an accuracy of one year and height to an accuracy of one centimetre, gender was indicated with separate F and M keys. These data were not asked of the subject but obtained from the form completed at measurement point 1. After the Start button was pressed, the measurement lasted about two minutes.

If the measurement was interrupted and an error message appeared on the display indicating poor contact, the position of the subject's feet and their grip of the hand electrodes were checked. If necessary their feet, hands and electrodes were wiped again. The measurement was restarted and in most cases was successfully completed. Following the measurement two standard forms were printed out, one of which was given to the subject. The technician offered a brief interpretation of the findings. If necessary the subject discussed the meaning of the findings later on with the examining doctor.

Not all bioimpedance measurements were successful, although failures were very rare. During the first days of the field examination some of the field teams had to make do without the analyzer because of problems with delivery. Subjects with pacemakers were not tested (the word "pacemaker" was entered into the subject's file). Overall the number of subjects not tested was very small, but non-participation was selective by both region and health status.

InBody 3.0 yields highly accurate and repeatable measurements of total body water content (Bedogni et al. 2002), and therefore it is reasonable to assume that assessments of fat free mass and fatty tissue are also reliable. It is unlikely that the results are any more accurate than those produced by four-polar analyzers (Thomas et al. 2003), but the new instrument is certainly faster and easier to use. As yet there is only indirect evidence on the reliability of InBody 3.0 estimates of body and limb composition (Bedogni et al. 2002), and therefore validation by reference to DXA measurements was conducted by reinvestigating the sample from Kuopio and environs about 12 months after the base examinations. If the segmental estimates prove to be reliable, as initial findings seem to suggest (Pietrobelli et al. 2004, Salmi et al. 2004), the results from bioimpedance tests have great potential application.

Upon completion of the field examinations, the operation of all five analyzers was tested by bringing them into the same room and conducting measurements on the same person. The results were quite consistent: for weight within 0.4 kg, for muscle mass within 0.7 kg and for body fat percentage 1.5.

8.4.3. Ultrasound examination of calcaneal bone

Fractures associated with osteoporosis are a major and expensive problem that is set to increase with population ageing. There is a strong inverse correlation between bone density and the risk of fractures, with the risk decreasing by one half per one standard deviation of bone density (Marshall et al. 1996). The WHO definition of osteoporosis is based on a bone mineral density of more than 2.5 standard deviations below the mean value for young adults (WHO Study Group 1994). Originally the purpose of this definition was to standardise comparisons of female populations, but it has been increasingly applied for purposes of clinical diagnosis and assessing the need for medical therapy.

The most reliable method for the measurement of bone mineral density is DXA or dual energy x-ray absorptiometry of the hip and spine, but epidemiological studies have also experimented with less expensive and simpler methods in screening osteoporosis. Broadband ultrasound measurement of the properties of the calcaneus is the most promising of these alternative methods.

Ultrasound examination determines the speed of sound (SOS) and broadband ultrasound attenuation (BUA) through the calcaneal bone and on this basis provides an estimate of bone density. Both SOS and BUA are higher in healthy than in osteoporotic calcaneus bone. It has been shown that the results correlate strongly with DXA measurements of bone mineral density both in the calcaneus and in the neck of the femur and that they are predictive of hip fracture (Gonnelli et al. 1995, Hans et al. 1996, Cepollaro et al. 1997, Garnero et al. 1998, Plujim et al. 1999).

Ultrasound results for the calcaneus have also been found to predict hip fracture independently of DXA results. This is more pronounced for BUA than for SOS, which has been explained by the fact that ultrasound examination demonstrates changes not only in the density but also the structure of ageing bone (Hans et al. 1996, Masud and Francis 2000, Prins et al. 1998). The measurements for the Health 2000 survey were conducted with a Sahara Clinical Bone Sonometer (Hologic, Waltham, Massachusetts), which was chosen for accuracy and reliability of measurement, the absence of radiation exposure, ease and comfort of use as well as transferability.

The instructions say that the unit must be at room temperature when used. This presented no problems in connection with transits between examination locations. Upon arrival at a new location the nurse removed the protective rubber pads used for transportation and inserted the plug into the mains socket. If an error message appeared on the display, the manual advised to remove the plug from the mains and to reinsert. The same rule applied to any malfunctions. The nurse always took a quality check measurement before the arrival of the first subject of the day. First,

she conducted a phantom test using the standard object that was delivered with the unit and entered the results on a separate form. She then conducted a measurement on herself, recorded the results for BUA and SOS, the date and identity code on a new form and placed the form in the file for measurement point 2.

The nurse then set to work with the list of subjects for the day. Before every measurement she applied an approximately 1 cm layer of gel at the angled end of the rubber transducer pads and placed a protective sheet of paper on the unit. The subject was instructed to place their right foot on the Sahara unit so that their second toe was aligned with the line on the base of the unit and the heel was placed in the middle of the measurement slot. Using a separate velcro brace, the nurse then attached the subject's leg to the unit, making sure that the cannon bone was aligned with the brace both from the front and from the side.

Once the measurement had been taken the unit emitted a signal tone to indicate that the velcro brace could be removed and the bone density result recorded from the display. If the display indicated a result and an error message (an asterisk), the nurse repeated the measurement on the same foot. If the error message appeared again, the subject's left foot was tested. If the unit still displayed an error message, the nurse did not enter the results into the computer but recorded the figures from the last measurement on a paper form. Finally, the nurse recalled on the display the BUA and SOS results and recorded the figures both on a carbonless copy form (different forms for men and women) and into a computer file. She handed the top sheet to the subject and attached the copy to the subject's health examination file.

The Sahara unit proved a very reliable piece of equipment. The velcro braces had to be replaced from time to time, but each field team always carried a replacement. There were very few malfunctions that required service or repairs, and any problems that did occur were promptly addressed by the company that had supplied the units. Indeed when a result was not obtained that was mostly due to an unusual calcaneus structure, which caused the unit to report an unreliable measurement.

The utility of ultrasound examination of the calcaneus and the applicability of its results for purposes of epidemiological research are as yet uncertain. In the measurement of bone characteristics and the diagnosis of osteoporosis it is important to note that bone mineral density varies individually between different limbs and spinal vertebrae (Greenspan et al. 1996, Varney et al. 1999). It also remains disputed whether ultrasound examination of the calcaneus is suitable for mass screenings of osteoporosis, because it seems that the results are better in assessing fracture risk than in detecting DXA osteoporosis (Masud and Francis 2000).

Repeatability tests conducted during the field examinations yielded reliability coefficients of 0.91–0.96 for the BUA and SOS results.

8.5. Laboratory

Jaana Leiviskä, Irma Salminen and Jouko Sundvall

8.5.1. Sampling

Samples of whole blood, serum, lithium heparin plasma and a spot urine sample were collected from all subjects (Figure 8.5.1). Faeces and saliva samples were obtained from a subsample of subjects. The samples were divided into aliquot tubes as illustrated in Figure 8.5.2. Lipids and some other clinical parameters were measured immediately at the laboratory and the results were sent to the subjects (see Laboratory analyses 8.5.7). Sampling and sample processing are flowcharted in Figure 8.5.3.

Blood sample

Ten tubes of blood were drawn from each subject (10 ml plastic-walled, evacuated Terumo Venoject II tubes, gel tubes for serum). Sampling order was determined by the purpose of the samples. Three sampling trials were made, with an approval given by the subject, to get at least one serum tube for lipid determinations and one EDTA tube for DNA isolation.

Venous blood samples were collected from a vein in the arm, with the subject in sitting position. The tourniquet was released as soon as the blood began to flow to prevent hemolysis. 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 Litium Heparine tubes were carefully inverted six times against the plugs.

After sampling, the nurse affixed ID labels to the tubes. The serum was allowed to clot for 20 minutes after the final tube had been collected. The plasma samples were kept at room temperature for the same time as serum samples to get them centrifuged together.

Saliva sample

Saliva samples were collected from approximately 1,500 participants in the area covered by the Helsinki field team. The nurse provided a labelled cup to the subject, who was asked to chew a paraffin pellet for about two minutes and spit at least 3 ml of saliva in the measuring cup. A tongue sample was obtained by scraping the surface of the tongue with a wooden stick and by immersing the stick in a tube containing an alkaline solution.

Spot urine sample

A spot urine sample was requested from all subjects. This did not require aseptic collection. The urine samples were divided into labelled tubes either during or at the end of the day.

Faeces sample

Kits for collecting faecal samples were handed out at the final examination point to every sixth subject. The samples were taken at home and posted in prepaid packages. Once received at the National Public Health Institute (KTL), each sample was divided into two container tubes and frozen to –70°C.

Figure 8.5.1. Blood tube chart.

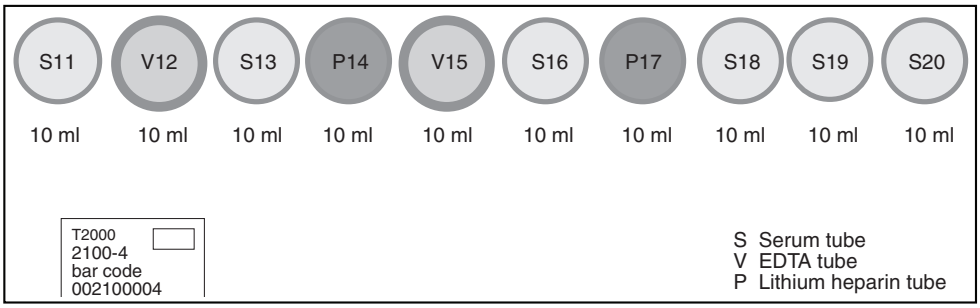
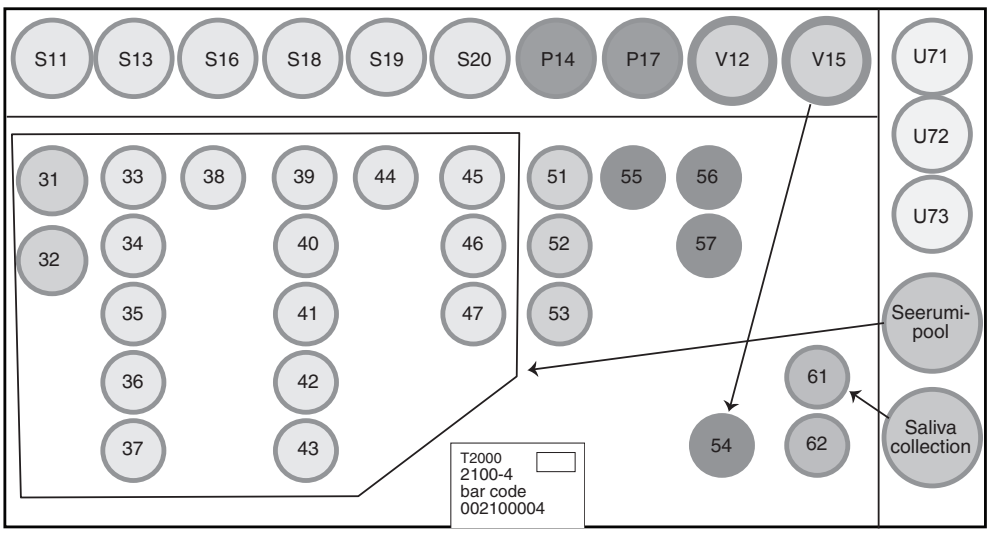
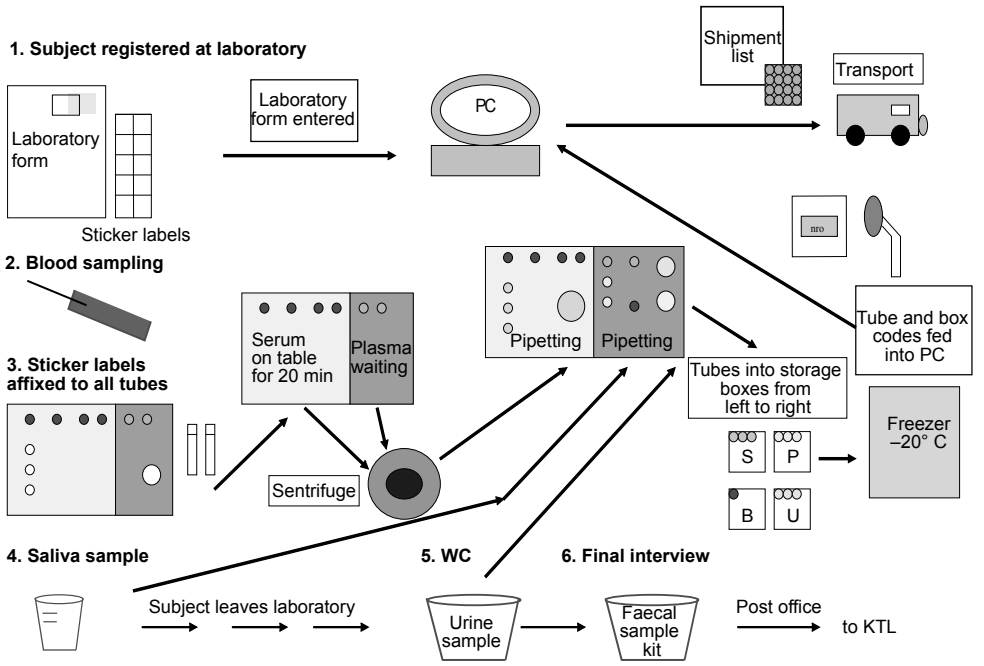


Figure 8.5.2. Aliquot tube chart.



S Serum-gel tube 10 ml
V EDTA tube for whole blood 10 ml
P Lithium heparin tube 10 ml
U Urine separation tube 10 ml
Aliquot tubes located under blood tubes are 1.5 ml Nalgene Cryo tubes, except for tubes no. 31 and 32, fitted for the autoanalyzer used

Figure 8.5.3. Flowchart of sampling and sample processing.



8.5.2. Sample processing

The first EDTA tube for DNA isolation was frozen at –20°C without opening the tube. A sample of 1 ml of whole blood was drawn from the second tube before freezing. The serum and plasma tubes were centrifuged at 1600–1800 G for 10 minutes. The sera from the six centrifuged gel serum tubes were collected into one large pooling tube. The pooled serum was mixed by carefully inverting the tube five times. Any hemolysed sera were not added to the pool but pipetted into separate aliquot tubes. Both serum and plasma were pipetted into 1.5 ml tubes in the order indicated by the aliquot chart (Figure 8.5.2).

The date code, time code, nurse ID code, the number and type of samples obtained, and the volume of the last serum and plasma aliquot sample were entered into the laboratory form. Any deviations in sampling and sample processing were also entered into the form using deviation codes.

The same rigorous procedures could not be followed with samples collected in connection with the home health examinations (see Chapter 9). Serum and plasma samples were separated with a field sentrifuge at a suitable place to work usually within 3–8 hours from venipuncture, and the samples were mailed to KTL unfrozen if facilities were not available for freezing.

8.5.3. Sample labels and sample management

The subjects gave the nurse their folder including a laboratory form and a sheet of sticker labels for labelling sample tubes, arranged in the order that the samples were drawn. Each label in the sheet carried the same recurring secondary key for that particular set of labels; an unequivocal primary key in bar code format; the survey code (T2000); and a code describing the type of sample. In the laboratory, the secondary key sticker was affixed to the top corner of the laboratory form; at the registration point, it was affixed to the subjects' health examination schedule for verification. The subject's examination code and the secondary key on the laboratory form were linked to each other (allowing later identification of the samples) when data from the forms were entered into the field laboratory computer during the same working day. The sample labels included no personal data.

A bar code sticker affixed on the cover and the side of a storage box, and the bar codes of the individual tubes in a box were read one by one with a bar code scanner into the field laboratory computer.

All data were transferred into the database at KTL headquarters. Sample ID codes, types of samples, quantities and comments were forwarded to the KTL sample management database. The location of the box in the freezer storage rack and the location of the rack in the freezer were later fed directly into the sample management database at KTL. Any individual sample tube can be located at the level of freezer/rack/box/individual tube.

8.5.4. Storage and shipment of samples

Serum, plasma, whole blood and urine samples were immediately frozen to –20°C on site, normally within 45–60 min but no later than 90 min from sampling. Underneath the freezer lid, a shelf was set up with 12–13 open, empty and labelled boxes ready to be filled with tubes according to a preplanned box chart, with 1–5 tubes from each set of samples placed into each box. To make this easier, the plasma aliquot tubes were colour-coded to distinguish them from serum tubes. The samples in the boxes were packed in dry ice and transferred from the field storage points to the their final storage location (–70°C), no later than 1–2 weeks after sampling, either by land (max. 700 km) or by air freight (over 700 km). The samples were followed by a shipment list and a mechanical thermometer, packed in the transport container.

Saliva samples were pipetted into sample tubes immediately upon receipt of the samples. The tubes were placed in dry ice where they were stored until transport. The samples were transported by land to the laboratory for storage at –70°C at least every other day.

8.5.5. Equipment and supplies

All the necessary supplies were delivered to the field laboratories in advance. Sampling supplies were subsequently ordered directly to the field examination points from suppliers with broad and reliable delivery networks. Equipment that moved with the laboratory included a field centrifuge and chest freezer. Electrical and manual pipettes were used in aliquoting the samples.

8.5.6. Field laboratory quality assurance

Procedures for sampling and sample processing were tested during two pilot phases. Staff were trained in advance, and the sampling stations were audited once or twice during the field examinations. Guidelines for field laboratory procedures were provided in writing. These included "Laboratory sampling process in the Health 2000 Survey", the KTL guidelines concerning the risk of infection, equipment manuals, service and maintenance contact information, and supplies lists complete with ordering information.

One of the laboratory's major planning considerations was to minimize the potential for error. Therefore the samples of only one subject were centrifuged at a time and then divided into storage tubes. In the event of problems, the field laboratory staff contacted the person in charge at KTL 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 staff responsible for sample collection, sample processing and the duties of examination nurse at measurement point 1, exchanging jobs at about one-week intervals.

8.5.7. Laboratory analyses

The samples collected in the Health 2000 Survey were analysed for cholesterol, HDL cholesterol, LDL cholesterol, triglycerides and glucose at the Social Insurance Institution's (SII, Kela) Research and Development Unit (Olympus, AU400, Germany) and for glutamyltransferase and uric acid at KTL's Analytical Biochemistry Laboratory (Optima 909, Thermo Electron, Vantaa, Finland). LDL cholesterol was also determined by calculation. The determinations were made from frozen samples within six months of sampling. Table 8.5.1 provides more detailed information concerning the methods used. The data produced in later analyses will be compiled into a cataloque of laboratory methods.

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). The laboratories took part in Labquality's External Quality Assessment (EQA) schemes. The accuracy of the methods (bias%) was calculated as the mean of the Short-term program organized by Labquality. Short-term quality assessment

samples are human serums of unknown concentrations analysed in the same way as other samples and reported monthly. The accuracy for lipid determinations was also calculated by Labquality's Lipids and lipoproteins program. 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 8.5.1.

Analysis	Method	Interassay coefficients of variation,(CV%)	Bias%
Cholesterol	Cholesterol, CHOD PAP, Olympus System Reagent, Germany	CV% = 2.1, mean = 5.4 mmol/l, N = 413 CV% = 2.2, mean = 7.1 mmol/l, N = 452	Labquality's Short-term and Lipid and lipoproteins programs 11/00 – 8/01 bias = 2.0% SD = 3.0
HDL cholesterol	HDL-C Plus, Roche Diagnostics GmbH, Germany	CV% = 4,8, mean = 1.30 mmol/l, N = 133 CV% = 5,3, mean = 1.37 mmol/l, N = 374	Labquality's Short-term and Lipid and lipoproteins programs 1/01 – 6/01 bias = –4.6% SD = 2.8
LDL cholesterol	LDL-C Plus, Roche Diagnostics, GmbH, Germany	CV% = 4,5, mean = 2.66 mmol/l, N = 347 CV% = 5,7, mean = 2.96 mmol/l, N = 356	Labquality's Short-term and Lipid and lipoproteins programs 2/01 – 5/01 bias = 0.4% SD = 3.5
Triglycerides	Triglycerides, GPO PAP, Olympus System Reagent, Germany	CV% = 2,1, mean = 1.39 mmol/l, N = 45 CV% = 3,2, mean = 1.47 mmol/l, N = 413	Labquality's Short-term and Lipid and lipoproteins programs 11/00 – 8/01 bias = 5.8% SD = 4.0
Glucose	Glucose, Hexokinase, Olympus System Reagent, Germany	CV% = 2.1, mean = 9.3 mmol/l, N = 411 CV% = 2,3, mean = 5.2 mmol/l, N = 432	Labquality's Short-term program 11/00 – 8/01 bias = 0.7% SD = 3.8
Glutamyltransferase	Gamma-GT, (IFCC/ ECCLS), Konelab, Thermo Electron Oy, Finland	CV% = 2.2, mean = 172 U/l, N = 396 CV% = 2.8, mean = 81 U/l, N = 349	Labquality's Short-term program 10/00 – 12/01 bias = –1.5% SD = 3.6
Uric acid	Uric Acid, Uricase PAP, Konelab, Thermo Electron Oy, Finland	CV%= 2.1, mean = 413 µmol/l, N = 478 CV%= 2.3, mean = 604 µmol/l, N = 399	Labquality's Short-term program 10/00 – 9/01 bias = 0.0 % SD = 2.2

Results for lipid values from SII (Kela) and KTL laboratory measurements were compared based on approximately 1,000 samples. The results are summarised in Table 8.5.2.

Table 8.5.2.

Method	SII (Kela), mean	KTL, mean	Difference	SD
Cholestrol	5.97 mmol/l	5.81 mmol/l	2.6 %	3.1
HDL cholesterol	1.33 mmol/l	1.45 mmol/l	–8.1 %	7.4
Triglycerides	1.62 mmol/l	1.48 mmol/l	9.5 %	5.5

8.6. Oral health examination

Miira Vehkalahti, Anne Nordblad, Liisa Suominen-Taipale, Sisko Arinen, Dorrit Hallikainen, Hannu Hausen, Matti Knuuttila, Anna-Lisa Söderholm, Sinikka Varsio, Kari Soikkonen, Mauno Könönen and Terttu Pietilä

8.6.1. Overview

The health examination included a clinical and radiographic examination of the mouth. Other aspects of oral examination were included in the home-visit interview and questionnaires. A short clinical oral examination, in which a nurse counted the number of teeth and recorded data on removable dentures, was included in an abbreviated health examination conducted at home among those who did not attend the health examination.

Clinical and radiographic examination of the mouth

The clinical and panoramic radiographic examination of the mouth was conducted by a dentist with the assistance of a dental nurse (or oral hygienist). Overall the examination lasted 15 minutes. The mouth was examined by the dentist and the dictated observations were entered into computer by the nurse, who also took the radiograph. In one case the nurse both took the radiographs and conducted part of the clinical examination because the dentist was prevented from being present. During software freeze-ups the clinical data were recorded on standard preprinted forms. These data were entered into the database as soon as possible after the problem was resolved.

At the end of the clinical examination the dentist reviewed the main findings of the radiographs on screen. The radiographs were later studied in closer detail by four specialist dentists.

Feedback returned to subjects

Following the clinical examination the dentist offered the subjects a summary of the clinical observations and where necessary urged them to seek dental care, or wrote an admission note either to the public dental care or to an outpatient department for oral diseases. All subjects received a hard copy of their radiographs as well as a written summary of the findings of the clinical and radiological examination. Following the more detailed examination of the radiographs, letters were sent out to 58 persons who had serious findings.

8.6.2. Health interview and questionnaire

In the health interview the subjects were asked whether they had teeth and/or removable dentures. They were also asked to assess their own oral health; to identify any symptoms caused by their teeth or dentures; and to say whether they had any difficulties in eating. Furthermore, the subjects were asked about their oral hygiene habits and the equipment they used as well as about their visits to dental care providers and the contents and costs of their dental care. The questions were largely the same as in earlier population studies (Vehkalahti et al. 1991, Arinen et al. 1998), but closer attention was paid in Health 2000 to hygiene habits, dental care received and attitudes to dental care. Some of the new questions were based on the UK population survey (Kelly et al. 2000), some were developed for this survey.

The response options to the question concerning teeth and/or dentures were “full dentures and no own teeth or tooth remnants”, “dentures and own teeth”, “no dentures but own teeth”, and “neither dentures nor own teeth”. On the basis of the responses the subjects were categorised as either dentulous or edentulous. The category of edentulous subjects comprised those who said they had no teeth of their own, the dentulous were those who according to self-report did have teeth of their own. The use of dentures was not considered in this classification. The distinction between dentulous and edentulous subjects was even used at the interview stage in that some of the questions were worded differently for the two groups. It was also used in grouping the results of the interview and questionnaire material, for instance regarding the use of dental care services.

Subjects were asked to rate the condition of their teeth and their oral health on a five-tiered scale (good–poor) and to say whether they had experienced toothache or other dental or denture problems during the past 12 months. Ability to chew food was assessed by asking whether the subjects could eat dry bread or biscuits without drinking anything (Fox et al. 1987) and whether they could chew hard or tough food.

Oral hygiene habits were inquired by asking the subjects how often ("once a day", "twice or more than twice a day", "less frequently than once a day", and "never"); what time of the day ("before a meal", "after a meal", "before I go out to the shops or to a party", and "in the evening before I go to bed"); with what equipment ("ordinary toothbrush", "electric toothbrush", "dental sticks" and "interdental brush") the subjects cleaned their teeth; and what agents ("fluoride toothpaste", "fluoride pills and solutions") they used for the home care of teeth. Furthermore, they were asked whether they were capable of cleaning their teeth and their mouth by themselves and how often they cleaned their dentures.

To assess the use of dental care services, the subjects were asked whether they were in the habit of going to see a dentist for check-ups and if so, how often ("once a year", "every other year", "less often") and whether they had their own dentist and if so, whether their dentist worked at the public dental care, at a private practice or somewhere else. Dental care during the past 12 months was inquired separately for visits to the public dental care, private and other dentists, dental technicians, and other dental care visits. Subjects who had been to the public dental care and a private dentist were asked how much they had paid for their treatment during the past 12 months. The subject responded by selecting one out of a set of cards displaying a certain range of sums. If the subject had not been to see a dentist during the period concerned, they were asked when they had last paid a visit to a dentist.

Subjects who had visited a dentist in past five years were further asked how their last appointment had been made; what treatment measures out of a list of 14 had been performed; and how satisfied or dissatisfied they had been with the treatment they had received. All subjects were asked whether they thought they currently were in need of dental care, whether they had ever received orthodontic treatment, and how afraid they were of visiting a dentist.

The questionnaire also included items on the use of sweet snacks as well as on the subjects' own perceptions of how their oral and dental health or any diseases or disorders impacted their quality of life.

Eight types of sweet snacks were listed: coffee or tea with sugar, other sweet beverages, toffee or liquorice or dried fruit, chocolate or filled biscuits and lozenges or chewing gum, both with and without xylitol. The subjects were asked how often they usually consumed these items, with the same preset response options

throughout: “3 or more times a day”, “once or twice a day”, “2–5 times a week”, “less often” and “never”.

Associations between oral health and quality of life were assessed with the 14-item Oral Health Impact Profile (OHIP-14; Slade and Spencer 1994, Slade 1997). Subjects were presented with a list of problems relating to the mouth, teeth or dentures and they were to say how often they had experienced those problems during the past month. The questions concerned seven dimensions i.e. functional limitation, physical pain, psychological discomfort, physical disability, psychological disability, social disability and handicap.

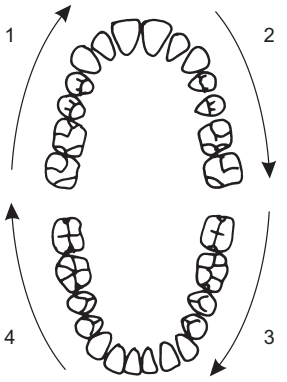
8.6.3. Clinical examination of the mouth

Clinical examinations were carried out with a portable dental treatment unit (Dentronic Mini-Dent®, Planmeca Oy) including a built-in compressor, saliva suction and a high-powered suction motor. In addition the team had the use of a portable patient chair, fibre optic light (Novar®), fibre optic head lamp (Tekmala Oy) and a letter scale. A file had been prepared for the dentist conducting the examination that contained detailed instructions on the various stages of the clinical examination, measurement determinations and for making computer entries. The subjects were first asked: “Do you have any health condition for which your doctor or dentist has said you require antibiotic protection in connection with dental care?” Periodontal pockets were not measured in subjects who answered in the affirmative.

The clinical examination always followed the same order. First, maximum mouth opening was measured and the subject’s jaw joints and masticatory muscles were palpitated. Information was recorded on the presence of dentures, which were checked for condition, fit and cleanliness. After the examination of the mucous membrane of the mouth and measurement of occlusion, the patient chair was adjusted to reclining position. Next, intraoral images were taken of any mucous membrane findings (still frame with video camera); dental plaque and any spaces in dental arches were measured; and the teeth, periodontal pockets and periodontal bleeding were checked.

The subjects’ 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 left quadrant (see Figure 8.6.1). The clinical examination was concluded with a panoramic radiographic. If the subject had removable dentures, they were finally asked how old the dentures were, how the subjects used them, and whether any repairs had been made or were needed.

Figure 8.6.1. Sequence of examining dentition.



Clinical measurements

The measurements were developed on the basis of methods used in earlier population studies as well as in clinical patient practice. For reasons of comparability with the Mini-Finland Survey (Vehkalahti et al. 1991), all methods used in Health 2000 were either identical or expanded versions of those earlier methods. Most of the new measurements are based on population studies in the UK (Todd and Lader 1991, Kelly et al. 2000) and the United States (Drury et al. 1996), others come from extensive surveys of oral health in adult populations and from textbooks.

Masticatory muscles and temporomandibular joints

Measurements of masticatory muscles and temporomandibular joints were based on the guidelines of Dworkin and LeResche (1992), but were confined to the most common aspects of masticatory function. Maximal interincisal distance was measured as the distance between the tips of the upper and lower front teeth and was recorded as being constricted if it was less than 40 mm. The temporomandibular joint was palpitated with the index finger, applying a pressure of 0.5 kg to a point about one finger's width in front of the ear while the subject opened their mouth two times consecutively. A clicking and a crepitation sound appearing in connection opening the mouth and their location on the left or right hand side were recorded. The subjects were asked whether they experienced pain during the palpitation ("yes", "no"), and their answer was recorded. The temporalis and masseter muscles were palpitated separately on both the left and right hand side. The temporalis was palpitated in the temple region, at about 2 cm from the corner of the eye, and the masseter above the jaw angle. The muscles were palpitated with one finger, applying a pressure of around 1 kg. In connection with each palpitation the subjects were asked whether they experienced pain ("yes", "no"), and the response was recorded.

Dentures

The presence and type of removable dentures was recorded separately for the upper and lower jaw: full dentures or partial dentures with a metal or acrylic base. If the subject had a partial denture, the status of the occlusion was measured by pulling occlusion paper from the side of clenched teeth. If the paper did not move, occlusion was defined as tight.

To determine the need for relining or repairing dentures and the need for cleaning, they were removed from the patient's mouth but not rinsed. Dentures were deemed to be clean if no plaque or calculus were present. Mucosal ulcers and gingival hyperplasia observed in connection with dentures were recorded separately for the upper and lower jaw. The presence of denture sore mouth was only studied in subjects with upper jaw dentures.

Oral mucosa

The procedure of the examinations and the definitions of findings were based upon WHO guidelines and an earlier major population survey (WHO 1980, Zain et al. 1995). Each dentist had a file with written descriptions and colour images of typical oral mucosal lesions (collections of A-L Söderholm) to assist with their determinations.

Oral mucosa were always examined in the same order, starting with lips and vermilion junction of the lips and then proceeding to the upper sulcus and cheeks, palate, lower sulcus and tongue and floor of the mouth, and finally to the soft palate and pharynx. Intraoral examination was started from the right upper tuber. The sulcuses, alveolar crests and cheeks were examined in the order shown in Figure 8.6.1 using a dental mirror. For examination of the sides and floor of the tongue as well as the floor of the mouth, the subject was asked to lift the tip of their tongue up to the palate. The dentist then tied a strip of gauze around the tip of the tongue and turned it right and left for better visibility.

The following oral mucosal lesions were recorded: angular cheilitis, pseudo-membranous fungal infection, rhomboid glossitis, fistula, white and red mucosal lesions, mucosal ulcer, swelling or tumor and gingival hyperplasia. Furthermore, it was recorded whether the finding was on the floor of the mouth and/or in the ventral surface of the tongue or elsewhere. Lesion size was classified into diameter categories of under 1 cm and 1 cm or over. Need for treatment was classified into one of three categories: "no", "yes" and "yes urgent". An intraoral camera (Intracam, Planmeca Oy) was used to take still video images of oral mucosal lesions (Intracam®, Planmeca Oy).

Occlusion and deviations in occlusion

For the examination of occlusion, entries were made on the number of opposing teeth, crossbite, scissors bite, open bite and the intercuspal relationship (Angle's classification). Neither missing teeth nor teeth in removable dentures were taken into account. Opposing teeth were determined in the same way as in the Mini-Finland survey (Vehkalahti et al. 1991), All other determinations were based on general orthodontic practice and earlier population surveys (Todd and Lader 1991, Drury et al. 1996).

Number of opposing teeth in contact (0–5) was calculated separately for the lateral teeth on the left and right side of mouth. The canine was accepted as one tooth of a most front pair of opposing teeth. Crossbite and scissors bite were both recorded in one of two categories, and wisdom teeth were not taken into account. Crossbite was determined for both frontal and lateral teeth and was recorded when a maxillary tooth came inside the opposing mandibular tooth, i.e. on the lingual side. Scissors bite was determined for side teeth and was recorded when a maxillary tooth overlapped the opposing mandibular tooth without any contact between the occlusal surfaces.

Open bite was measured from the central incisor in the upper right quadrant (d. 11) and in its absence from the corresponding tooth in the upper left quadrant (d. 21). Measurements were taken from teeth in occlusion using a WHO periodontal probe with a ball end (Plandent Oyj, no. 19577). Overjet was measured as the distance between the incisal tip of the upper incisor and the anterior surface of the lower incisor and recorded in one of four categories: less than 0 mm, 0–6 mm, 7–9 mm and more than 9 mm (Figure 8.6.2). Overbite was determined according to the position of the tip of the lower incisor in relation to the upper incisor when biting the teeth together. There were four classification options, normal and three malocclusion categories (Figure 8.6.3). The sagittal relation between the upper and lower jaw was determined using a modification of Angle's classification as normal; upper canine clearly distally from lower canine; upper canine clearly anteriorly from lower canine; and upper and lower canine in the cusp to cusp relationship.

Dental plaque

Dental plaque was measured from one surface of three different teeth using a scale modified from that developed by Silness and Löe (1964): the buccal surface of the last tooth in the upper right quadrant, the lingual surface of the last tooth in the lower left quadrant and the labial surface of the canine tooth. Teeth in removable dentures were not studied for plaque. Observations were recorded in three categories: "no plaque", "gingival plaque only" and "gingival and other plaque".

Figure 8.6.2. Classification of overjet.

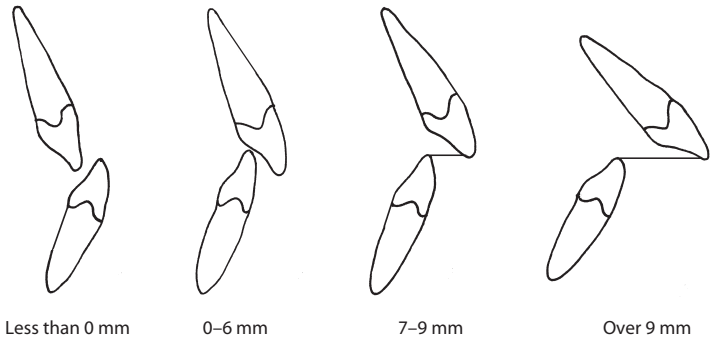
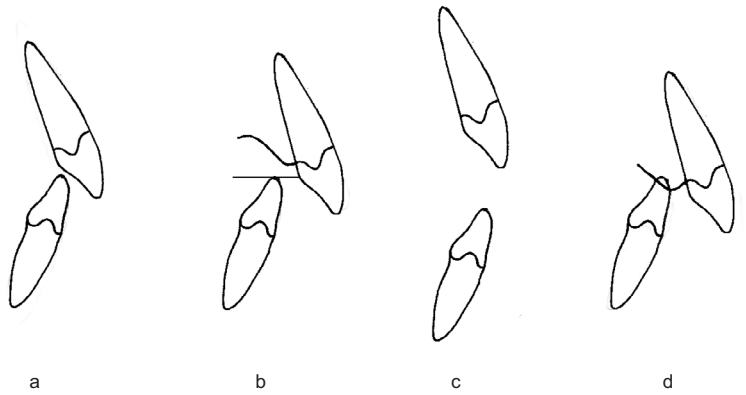


Figure 8.6.3. Classification of overbite: a) normal, b) extending to gingival third of upper incisor, c) open bite d) traumatic overbite.



Spaces in dental arches

Spaces in dental arches were measured using the same method as in the UK population survey (Kelly et al. 2000), although a much wider space was required in order to meet the criterion. Spaces were measured from the entire dentition with the exception of wisdom teeth. Spaces were determined according to the missing tooth by groups of teeth separately in the upper and lower jaw and classified according to possible dentures as follows: "no spaces", "space present and filled by false teeth", and "space present but not filled by false teeth". A space was recorded as being filled by false teeth when either a fixed bridge or a removable denture was observed. A single tooth replaced with a bridge was always entered as a missing tooth. A tooth remnant or a diastema were not classified as spaces.

A space was recorded in the molar region when both the first and second molars were missing. A space in the premolar region was recorded when at least one tooth was missing and a clearly observable space (6 mm) or a bridged tooth was detected in its place. A space in the frontal region was recorded when one or more teeth were missing and there were one or more clearly observable spaces (6 mm) or a bridged tooth.

Condition of the teeth

For the assessment of the condition of the teeth they were dried with an air spray, and a saliva ejector (Hygoformic®) and cotton rolls were used to make sure that they remained dry. The examination was performed using a mouth mirror, fibre optic light and a WHO periodontal probe with a ball end (Plandent Oyj, no. 19577). The sequence of the examination is shown in Figure 8.6.1. Tooth identification and the determination of tooth condition were based on the Mini-Finland survey (Vehkalahti et al. 1991) and WHO (1997) guidelines.

Teeth were identified on the basis of both their location and shape, and each tooth was identified as one of 32 permanent teeth or 20 primary teeth. Premolar teeth were identified as first premolars if they were located no more than 2 mm from a canine tooth and molars as first molars if they were located no more than 2 mm from the last premolar. When the subject had three remaining lower incisors in a good, perfect row, they were recorded as teeth 32, 31 and 41. If the subject had a primary tooth in the place of the corresponding permanent tooth, the data for the primary tooth were entered in lieu of the data for the permanent tooth, with the added note that this was a "primary tooth". If the subject had both the primary and the corresponding permanent tooth in their mouth, entries were only made for the permanent tooth, and a note was made regarding the presence of an extra primary tooth.

All surfaces of all teeth were examined, and observations on each tooth were classified as follows: "sound" (intact), "filling but no caries", "requires attention but no caries" (e.g. a fractured filling or tooth), "caries" (coronal and root caries separately) or "decayed to the root" (a distinction was made between tooth remnants with and without caries).

Caries was recorded as present when the lesion clearly extended into dentin and when the lesion was so extensive that proper treatment would have required at least a filling. Further criteria were that the lesion was cavitated; that it had penetrated the fissure and undermined the enamel; or the dentin walls of the lesion were to show clear signs of softening. Caries was not recorded if there was any uncertainty about the observation. Depending on the location of caries the tooth was recorded as having either coronal or root caries or both.

In addition to ordinary fillings, entries as fillings were made for prosthetic crowns, bridges and façades, but not for sealants or artificial bulges made to improve the stability of dentures. Teeth were recorded as being in need of repair when they showed no signs of caries but were fractured, or a filling in the tooth was fractured or had come loose or it had dissolved or was otherwise clearly incomplete, or if the tooth had a temporary filling. A fracture in need of repair was to extend clearly into the dentin.

A residual root was recorded when more than half of all vertical surfaces of the tooth had been damaged. Supporting teeth for overdentures, for example, were entered as non-caries roots.

Periodontal health

The periodontal health was determined by measuring the depth of periodontal pockets and the presence of gingival bleeding on probing (BOP) The depth of periodontal pockets was measured from all teeth except for wisdom teeth and tooth remnants. The sequence of the examination is shown in Figure 8.6.1. The measurements were made using a WHO periodontal probe with a ball end (Plandent Oyj, no. 19577) and markings at 3.5 and 5.5 mm. A force of 20 g was used in the measurements, and the instrument was calibrated each morning by the examining dentist using a letter scale.

The depth of periodontal pockets around each tooth was measured at four points in the following order: distal corner and midpoint on the buccal surface, midpoint and mesial corner on the lingual surface. The deepest measurement for each tooth was recorded in one of three categories: “no periodontal pocket”, “4–6 mm pocket ” and “6 mm pocket or deeper”.

Presence of bleeding on probing was recorded immediately after the measurement of pocket depth: “yes” or “no”. Observations of bleeding on probing in the upper jaw were recorded immediately after the measurement of periodontal pockets in the maxillary teeth. This was completed before proceeding to the measurement of pockets around the mandibular teeth and the assessment of their gingival bleeding on probing. Observations were recorded from three sextants in both jaws: frontal area, left side and right side.

Use of removable dentures

At the end of the clinical examination, the dentist asked all subjects who wore removable dentures how old these were (“less than 5 years” or “more”), whether the dentures had been repaired during the past five years, and whether the subjects themselves felt the dentures were currently in need of repair. Finally, the subjects were asked when they used their dentures: “rarely or never”, “usually while awake”

or “always”. All questions concerning the use of dentures were presented separately for maxillary and mandibular dentures.

8.6.4. Radiographic examination

The field teams had access to digital panorama imaging equipment (Planmeca Oy 2002 CC Proline®) as well as laptop computers with software (Dimaxis®, Plandent Oyj) for the preliminary examination of the radiographics taken and printers (HP DeskJet 930C®) for creating hard copies (HP two-sided Photo Paper Glossy, C 1847A®). At each examination site the equipment was installed and calibrated by a Planmeca expert. In addition, a representative of the Radiation and Nuclear Safety Authority visited each site to check the settings. The radiographic settings were adjusted according to the size of the subject within the range of 58–68 kV and 4–10 mA.

The panorama radiographic was taken by the nurse immediately after the clinical examination of the mouth. A radiographic was not taken if the subject was pregnant or if they had a postural anomaly in the cervical or thoracic spine, or if they refused. Prior to the radiograph the subjects were asked to remove any jewellery in the head and neck area as well as any dentures. The nurse made sure the subject was properly positioned with the aid of positioning lights. During the radiographic the subject was to place their tongue against the palate and breathe normally. The quality of the radiographic was immediately assessed by the dentist, and where necessary the subjects were asked permission to take another radiographic. A total of 77 retakes were done.

The field dentist conducted a preliminary examination of the radiographics. For all patients whose radiographics were taken, the following data were recorded in their files in one of two categories (“yes” or “no”): permanent implant, root canal treated tooth, vertical bone pocket, suspected cyst, suspected periapical finding, and structural change in jaw bone. The subjects were informed of any findings.

Stored in electronic format, the images were later examined and interpreted in closer detail on a 17-inch or larger computer screen. The quality of the images and the accuracy of imaging were assessed by four radiology dentists, who also interpreted the images. Their interpretation was based on a visual examination of a normal-sized, on-screen image. For the closer inspection of details, the image was magnified as necessary so that the interpretation could be confirmed. Determinations were based on standard dental practice (Langland and Langlais 1997); whenever there was any uncertainty about an observation, it was not recorded. A more cautious criterion than normal was applied in the assessment of

root fillings: only a root filling ending 3 mm or more from the apex was defined as underfilling (Kerekes and Tronstad 1979). Table 8.6.1 summarises the findings made from the radiographics, the objects of measurement and the classifications used.

Table 8.6.1. Targets of panorama radiographic interpretations and classifications used.

AREA/ PHENOMENON EXAMINED	Object of measurement	Determination / categories
DIAGNOSTIC QUALITY OF IMAGE	Image as a whole and each item of measurement	Good / Reasonable / Useless
CONDITION OF DENTITION	Identification and positioning of teeth: primary teeth and permanent teeth separately (total 20+32 cases)	Missing / Retinated / Root, partly embedded in bone / Root, wholly embedded in bone / Implant / Fixture / Caries / None of the above
ROOT TREATMENTS	Tooth: primary teeth and permanent teeth separately	No root treatment / Amputation / Faultless root treatment / Unsatisfactory root treatment
DEFECTS IN ROOT FILLING	Only unsatisfactory root treatments: over- and underfilling separately	No overfilling / Clear overfilling / No underfilling / Underfilling, > 3 mm from apex
PERIAPICAL CHANGES	Tooth: primary teeth and permanent teeth separately	No / Yes, diameter < 10 mm / Yes, diameter at least 10 mm
VERTICAL BONE POCKETS	Tooth: primary teeth and permanent teeth separately	No / Yes, depth > 3 mm and extends to middle third of the root / Yes, extends to apical third
FURCAL LESIONS	Tooth: only permanent molars	No / Yes
PERICORONITIS	Tooth: only wisdom teeth and permanent 2nd molars	No / Yes
HORIZONTAL BONE LOSS	Mouth quadrants, each assessed only if teeth present	No / Yes, extending to cervical third / Yes, extending to middle third / Yes, extending to apical third
JAW JOINTS	Articular ends, right and left separately	Normal / Arthrosis / Arthritis / Post-traumatic condition
SINUS MUCOUS MEMBRANES	Sinuses, right and left separately	Normal / Mucous membrane swollen / Mucous cyst / Other change
JAWBONE ATROPHY	Edentulous jaw, maxilla and mandible separately	No / Minor / Moderate / Intense
STRUCTURE OF MANDIBULAR BONE	Mandibular bone	Normal / Porous
OTHERS: CYST, SUSPECTED TUMOUR, FIBROTIC LESION, FOREIGN OBJECT	Whole image	No / Yes, what

8.6.5. Quality of clinical measurements

Quality assurance

Clinical measurements were based on validated methods used in previous dental population surveys. Measurement determinations were designed with a specific view to clarity and unambiguity. The piloting of the clinical investigation, the use of electronic forms, induction training to field examination staff and the detailed guidelines provided for the field teams all served as important steps in maintaining measurement quality assurance. The oral examination in the first pilot for the Health 2000 survey involved 34 persons, the second pilot 93 persons.

As part of the quality programme for the field stage of the Health 2000 survey, repeat measurements were conducted for 111 subjects and parallel measurements for 269 subjects in the oral examination. In addition, during Quality Day 3 all five field dentists and a reference dentist examined the same 42 non-sample adults. At the training stage all four specialist dentists who were recruited to interpret the radiographics examined 50 radiographics from subjects who took part in the second pilot test. In connection with the interpretations proper, some of the radiographics interpreted during previous days were submitted for re-interpretation at an interval of approximately 30 radiographics. In all, repeat assessments were made of 327 radiographics. A more detailed account of quality assurance procedures in the oral examination is provided elsewhere (Suominen-Taipale et al. 2004).

Assessing agreement of measurements

Agreement of measurements was described in terms of percentages of unanimous diagnoses, kappa values of measurements and the McNemar skewness test (Fleiss 1981). Comparisons of the parallel measurements conducted by the field dentists and the reference dentist on the group of 42 non-sample adults are shown in Table 8.6.2. Comparisons of the parallel interpretations made of the radiographs at the training stage are shown in Table 8.6.3.

On the basis of these preliminary data it seems that quality assurance of the clinical measurements has been highly successful. Overall the level of agreement between the measurements was very high, particularly so in the measurements of teeth, fillings and caries. Levels of agreement were somewhat lower in areas that are more difficult to measure, but that is consistent with earlier experiences from similar surveys.

Table 8.6.2. Quality of clinical oral examinations: comparison of parallel same-day measurements by field dentists and reference dentists on the same subjects (N = 42).

Measurement	Same[1] %	Kappa	95% CI[2]	Skewness[3]
Vertical mouth opening	93	0.31	0.19–0.44	–
Mucous membrane findings (no / yes)	88	0.34	0.32–0.62	ns
Cross and scissors bites (all categories)	92	0.78	0.65–0.91	ns
Angle's classification (all categories)	87	0.72	0.63–0.81	ns
Dental plaque, one area (upper molar)	57	0.22	0.12–0.32	– – –
Spaces in dental arch, one area (dd. 34–35)	99	0.86	0.76–0.97	ns
Condition of teeth (dd. 16, 21, 35, 46)	95	0.86	0.80–0.93	ns
Measurement of periodontal pockets (dd. 16, 21, 35, 46)	82	0.32	0.25–0.39	ns
Gingival bleeding on probing, two areas (dd. 24–27, 44–47)	57	0.17	0.08–0.25	– – –

[1] Percentage of unanimous diagnoses

[2] Kappa value 95% confidence interval

[3] McNemar skewness test: ns = no skewness, – sign: field dentist made less findings than reference dentist; number of signs indicates intensity of skewness (low, moderate or high)

Table 8.6.3. Interpretation of panorama radiographics (N = 50) at training stage: comparison of three field dentists' interpretations with reference dentist's interpretations.

Measurement	Same[1] %	Kappa	95% CI[2]	Skewness[3]
Readability of images	98	0.96	0.80–1.12	ns
Atrophy				
• upper jaw	90	0.41	0.30–0.52	ns
• lower jaw	91	0.38	0.27–0.49	ns
Periapical lesions_				
• yes / no	90	0.79	0.63–0.95	ns
• number	78	0.58	0.47–0.68	– – –
Horizontal bone loss	61	0.27	0.15–0.40	+++
Vertical bone pocket, extends to				
• middle third of root	77	–0.04	–0.12–0.05	+++
• apical region	94	0.72	0.56–0.88	ns
Changes in temporomandibular joints	82	0.30	0.23–0.37	+++
Root fillings				
• yes / no	98	0.96	0.80–1.12	ns
• incomplete	65	0.31	0.19–0.43	+++
• imperfect	69	0.16	0.04–0.28	+++
• overfilling	97	0.69	0.55–0.83	++

[1] Percentage of unanimous diagnoses

[2] Kappa value 95% confidence interval

[3] McNemar skewness test: ns = no skewness, + sign: field dentist made more findings than reference dentist, – sign: field dentist made less findings than reference dentist; number of signs indicates intensity of skewness (low, moderate or high)

8.7. Infection questionnaire (questionnaire 2) and snack

Petri Ruutu, Sami Heistaro and Arpo Aromaa

After the dental and oral examination the subjects were offered a snack, usually of coffee or tea, juice, rolls and fruit. While having this snack the subjects completed a questionnaire on infectious diseases (questionnaire 2), which was to be returned at the final checkpoint.

The questionnaire inquired about acute gastrointestinal symptoms, i.e. vomiting and diarrhoea; respiratory infections; and finally vaccinations. The latter specifically concerned vaccinations for adults: influenza vaccinations, pneumococcal vaccinations, hepatitis A, tetanus and polio vaccinations. The subjects did not normally have difficulty answering the questions.

Web link: Infection questionnaire (questionnaire 2)
http://www.terveys2000.fi/lomakkeet/en/t2004en.pdf

8.8. Functional capacity

Päivi Sainio, Seppo Koskinen, Sanna Natunen and functional capacity team

For the measurement of functional capacity, tests were carried out on eyesight and hearing, cognitive functioning, balance, perceptual-motor speed and hand grip strength. Furthermore, the endurance of back extensor muscles was tested in subjects aged 30–54. Subjects over 54 took a chair stand test and walking test, and the function of their lower and upper limb joints was assessed. All the methods used by the experts in these measurements were tools that are well-established and widely used in population surveys and clinical studies. The aim and purpose of the measurements was to obtain other than subjective, self-report information on functional capacity and to complement the picture created by the interview and questionnaire data.

Functional capacity tests were done at two parallel measurement points. They were scheduled to last 30 minutes. Each field team had a staff of four examiners, who alternated weekly between functional capacity measurements and mental health interviews. Backup was provided by the team's two field nurses who were doing the home health examinations; they had also received training to conduct the functional capacity measurements.

In connection with the home health examinations, measurements of functional capacity were done in the same way as in the health examination proper apart from a few minor exceptions (see Chapter 9).

The results were entered in the data collection program under the functional capacity component. If for some reason the examination or part of it could not be completed, the reason for missing data was entered under the notes section for the test concerned. Balance and reaction time results were recorded automatically in separate files and collated later on the basis of ID data into the rest of the subject's functional capacity results.

8.8.1. Study of visual acuity

Sirkka-Liisa Rudanko and Seppo Koskinen

Binocular visual acuity was measured using well illuminated (> 350 lux) distant and near vision charts. Visual acuity was measured with spectacles or contact lenses if normally worn by the subject. The adequacy of illumination was checked using EC1 digital luxmeters (Hagner, Sweden) every morning and whenever there were changes in the lighting conditions. Illumination was adjusted using the lights at each examination site, additional spotlights and blackout curtains.

For the examination of near vision, the subjects held the chart at a distance where they thought they could see it best. They 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 on a line of five, they were asked to move one line down towards smaller letters. The result entered for the subjects was the lowest line on which they correctly identified at least four letters. If the subject was unable to see even the biggest letters, the result was entered as 99, indicating a visual acuity lower than could be measured using the letter chart (e.g. blind).

For the examination of distant vision, the subjects stood four metres from 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.

For scotopic vision testing the level of lighting on the surface of the distant vision chart was reduced to 9–11 lux. The test was conducted within 30 seconds of turning down the lighting; otherwise the same procedure was followed as in the distant vision test.

The line numbers entered as the results of the visual acuity tests were later converted into corresponding visus values.

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

8.8.2. Hearing test

Timo Marttila and Seppo Koskinen

Air conduction hearing threshold was measured using a screening audiometer (Micromate 304, Madsen Electronics) in both ears at three frequencies (500, 1000 and 2000 Hz) in a silent room. Headphones with padded earpieces were used in order to minimise any environmental noise. The lowest stimulation level used was 5 dB. Subjects wearing hearing aids were asked to remove them for the duration of the test.

The test was started in what the subjects regarded as their better ear, or on the right hand side if there was no difference. Testing was started at a frequency of 1,000 Hz by first sounding the test tone (at least 25 dB, more for older people and those who seemed hard of hearing). The volume was lowered from its initial level at increments of 10 dB (e.g. 25dB – 15dB – 5dB) until the level was reached where the subject no longer could detect the audio signal. Then, the volume was increased 5 dB at a time in order to establish the hearing threshold, i.e. the lowest volume at which the subject heard the test sound. If the subject had not heard the initial test sound, the volume was increased at increments of 10 dB and the hearing threshold was determined in the manner described above.

The same procedure was then followed to determine the subject's hearing in the same ear first at the frequency of 2,000 Hz and then at the frequency of 500 Hz. Once these measurements were completed, the test proceeded to the subject's other ear. At each frequency the lowest volume audible to the subject was entered as the test result. If the subject did not hear the signal sounded at 90 dB, the result was entered at 99 dB.

If the average of the hearing thresholds in the better ear was higher than 35 dB, further questions were asked to establish whether and where the subject had previously had their hearing tested. If their hearing had never been tested before, the subjects were urged to contact a health centre GP for further tests.

8.8.3. Tests of cognitive functioning

Timo Suutama, Raimo Sulkava and Seppo Koskinen

Cognitive functioning was examined by selected tasks from the CERAD neuropsychological test battery, originally developed for assessing early phases of dementia and memory disturbances (Morris et al. 1989, Hänninen et al. 1999, Pulliainen et al. 1999). The cognitive functions assessed were speech production and encoding and retaining verbal material. In addition, during the home interview an abbreviated version of the Mini-Mental State Examination (MMSE) (Folstein et al. 1975) was administered to subjects aged 55 or over, providing a rough overall estimate of cognitive functioning (see Chapter 6).

In the test of verbal fluency, the subjects were to list as many animals as possible in one minute. The examining nurse measured the time with a stopwatch and used a tally chart to count the number of correctly and incorrectly cited animals, as well as any repeat references to the same animal. The number of correctly cited animals was recorded in the data collection program.

In the memory test, the subjects were shown 10 words one after another that they were to read aloud and commit to memory. Then, the subjects were asked to list the words they remembered; they were given 90 seconds to recall the words. If the subject failed to remember all 10 words the first time round, they were shown the words twice again, in a different order. The examining nurse recorded all correctly and incorrectly recalled words on a separate form. Words correctly recalled after each showing were recorded in the data collection program. If the subject was unable to see or read the words, the nurse read them out loud.

The delayed recall of the words was tested by asking the subjects to repeat the same list after about five minutes, after the reaction time and tapping test. The subjects were asked to list the words they remembered. The number of recollected words was recorded on a form and into the data collection program.

8.8.4. Perceptual-motor speed

Pertti Era and Päivi Sainio

The system used for measuring reaction and movement time (Good Response, Metitur Oy, Jyväskylä) (Era et al. 1986) consisted of a user panel, power source and computer program (Good Response). The panel had a waiting switch, four sets of lights and switches for turning off those lights. Once the subject was seated, the system was placed at a suitable distance and the height of their seat was adjusted.

Additional spot lights were switched off if they interfered with the subjects' sight of the lights on the test equipment. The subjects completed the tests with the index finger of their writing hand.

The test was started with three dry runs, or more if the subject still had difficulty understanding the test. The purpose was to react as quickly as possible to the lights that lit up on the panel by shifting one's index finger from the waiting switch to the switch that turned off the set of lights.

In the first dry run test a light lit up 12 times at random intervals at the same point in the panel (simple test). The subjects were instructed to turn off the light as quickly as possible by moving their index finger from the waiting switch to the switch that turned off the light. Next, the same 12-light test was repeated, but this time the lights lit up in different parts of the panel (multi choice test).

In both these tests, the subject's responses (switching off the light) were not accepted if they did not come within the programmed time limits or if they differed by more than two standard deviations from the subject's mean performance. If there were more than six failed performances or if the whole test was discontinued because of slow responses, it was usually repeated once after a few dry runs.

8.8.5. Measurement of hand grip strength

Pertti Era and Päivi Sainio

The equipment used in the measurement of hand grip strength consisted of a sensor, amplifier, a cable running between the sensor and amplifier, and a power source (Good Strength, IGS01, Metitur Oy, Jyväskylä). The handle-shaped sensor was placed on a table at a distance where when the subject's elbow was resting against the table, the grip handle fitted into the hand with the wrist in a neutral position (i.e. in slight dorsal flexion). Elbow joint angle was adjusted as closely as possible to 110 degrees by adjusting the height of the chair, and the forearm pointed forward in a 45 degree angle. The test was performed with the writing hand, with the opposite upper limb resting either on the table or in the lap. The method was developed from that used by Viitasalo et al. (1985).

The size of the grip handle was adjusted according to the size of the subject's hand so that when the handle was gripped, the second (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.

Once the zero level on the measuring device had been checked, the subjects were asked to grip the handle as hard as they could for 3–5 seconds; throughout this

time they were urged to do their best. Once the test was completed, the maximum grip strength was recorded in the data collection program (in Newtons) The second test was done 30 seconds later. If the difference between the two measurements was greater than 10%, a third test was done again 30 seconds later.

8.8.6. Measurement of balance

Pertti Era and Päivi Sainio

Balance was measured in the health examination proper using a computer-based measurement system (Good Balance, IGB01, Metitur Oy, Jyväskylä). Its main components are a triangular force platform and an electronics unit, including an amplifier, analogue/digital converter and a computer. Following the protocol introduced by Guralnik et al. (1994) four different measurements of balance were conducted:

1) Feet side by side, eyes open: Standing in a normal, natural posture, feet side by side, arms crossed in front of the body, eyes locked to a fixed point about two metres away, duration of measurement 30 seconds.

2) Feet side by side, eyes closed: Standing in the same position as above, eyes closed, duration of measurement 30 seconds.

3) Semi-tandem: Standing eyes open in a semi-tandem position with feet placed one after the other so that the proximal phalanx of the big toe of the trailing foot touched the inside of the heel of the leading foot, weight evenly on both feet. Hands rested freely and where necessary could be used to maintain balance. No separate reference was made to the subject about a visual fixed point. Duration of measurement 20 seconds.

4) Tandem: Standing eyes open in a tandem position, with the feet placed one after another along the same line so that the big toe of the trailing foot touched the heel of the leading foot, feet in a straight line, weight evenly on both feet. Hands rested freely and where necessary could be used to maintain balance. No separate reference was made to the subject about a visual fixed point. Duration of measurement 20 seconds. Tandem test was only performed if the subject had managed to stand at least 10 seconds in the semi-tandem position.

All measurements were made without shoes. The subjects were asked to stand in each position for as long as they could without swaying. While they were trying to find the right test posture, the subjects were allowed to steady themselves by hand against the wall. The measurement was started as soon as the subjects had found the right position by asking them to let go of the wall. The examiner used a

stopwatch to measure the time interval between the beginning and end of the test and recorded that time in the data collection program to an accuracy of one second. If the subjects swayed and lost their test posture before the measurement time expired, the time of their swaying was recorded. Measurements were not usually repeated unless the swaying was clearly due to some external disturbance factor.

At home health examinations and during equipment malfunctions, balance was measured using a simple field test (Guralnik et al. 1994), without the computerised system. For this test, the subjects were to remain standing in a semi-tandem position for 10 seconds. If successful, they were then asked to stand in a tandem position again for 10 seconds. If the subjects were unable to stand for 10 seconds in the semi-tandem position, they were asked to take an easier position, with their feet side by side, touching each other (10 seconds). In each position the results were recorded to an accuracy of one second.

In order to get as homogeneous a dataset on as large a sample of people as possible (time standing in different postures, maximum 10 seconds), field-test results were also produced for subjects studied on the force platform by converting their results for the semi-tandem, tandem or feet-together tests into corresponding field test values.

8.8.7. Endurance of back extensor muscles

Jaana Suni and Päivi Sainio

The endurance of back extensor muscles (Biering-Sörensen et al. 1984, Suni 2000) was only measured in subjects aged 30–54, and the test was not included in the home examination schedule. The test was not done if the subject had suffered acute myocardial or cerebral infarction during the past six months. If there were any other reasons preventing the subject from taking the test (e.g. intense acute back pain), these were considered separately and entered into the data collection program. The general principle was to try and complete the test with every subject.

The subjects lied down on a padded stepping board in prone position so that their lower body, from the iliac crest downward, was on the stepping board and their upper body rested down on the floor, some 20 cm lower. The examiner sat astride the subject's feet, on the lower portion of their calves. The subjects were then instructed to cross their hands behind the neck and to lift their upper body into horizontal position. Elbows were to remain in a horizontal line, neck straight, eyes fixed in an angle to the floor. The object was to remain in this position for as long as possible, but no more than four minutes. The subjects were told the maximum time in advance, and they were informed of the passage of time at about 30 second intervals. If their upper body dropped below horizontal level, the subjects were

instructed to adjust their position. The test was continued until the subjects were no longer able to lift their upper body back to horizontal. The result was recorded to an accuracy of one second.

8.8.8. Tests of physical capacity administered to subjects aged 55 or over

Päivi Sainio and Jarmo Malmberg

Subjects aged 55 or over took a ten-item joint function test, a chair stand test and a walking test over a distance of 6.1 metres.

Joint function test

The test to determine any functional limitations of the joints (Sievers et al. 1985) involved 10 separate movements, the first four of which were designed to test lower limb function and the remaining six upper limb function:

Walk on level surface
Walk on toes
Climb two stairs
Crouch
Raise upper arms
Extend elbow joints
Flex elbow joints
Backs of hands against each other (volar flexion of the wrist)
Clench fingers
Clench thumbs

Each of the component tests were performed in sequence following the examiner's demonstration. The subjects' performances were rated as normal (0), degraded (1) or failed (2) on the basis of detailed classification instructions.

Chair stand test

The chair stand test (Guralnik et al. 1994), which is widely applied in assessments of functional capacity in older people, used a standard chair with no arm rests and a seat height of 43 cm from the floor. The back of the chair was placed against the wall. The subjects were asked to sit down in the chair, with their hands across their chest and feet slightly apart. From this position, they were asked to stand up.

It was recorded in the data collection program whether the subjects managed to get up with or without using their hands or not at all. If the subjects did manage to get up without using their hands, they were asked to get up and sit down five times as quickly as possible, still without their hands. A stopwatch was used to time the

subject's performance, and the result was recorded in the data collection program. It was also recorded whether the subjects had managed to get up from the chair five times in a row with or without using their hands. The test was discontinued if not completed in 60 seconds, or if it posed any risk to the subject's safety.

Walking speed

Walking speed was measured over a distance of 6.1 metres (Fiatarone et al. 1994). Start and finish lines were marked out on the corridor floor using coloured tape. In home health examinations walking speed could be measured over shorter distances if the necessary space was not available; in this case the distance was recorded in the data collection program. Subjects were asked to walk the 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. They were allowed to use walking aids. The stopwatch was started at the beginning of the test and stopped when the subject crossed the end line. At the same time, the number of steps taken by the subject was calculated. The time and number of steps were entered in the data collection program. A separate entry was made if the subjects were unable to cover the distance within 60 seconds.

8.8.9. Quality assurance

Päivi Sainio, Seppo Koskinen, Sanna Natunen and functional capacity team

In order to complete the studies of functional capacity, the team of field examiners needed to have a solid knowledge and command of many different kinds of measurements, measuring devices and related software. Various quality control procedures were put in place to make sure that any external factors influencing the measurement results remained as constant as possible throughout the examinations in all field teams and at all points of measurement so that the variation seen in the results would reflect true variation among the subjects studied.

In the measurements of functional capacity, external factors with a potential impact on the results included the measurement instruments and their operation, the conditions under which measurements were taken and the actions of the field examiners, such as the instructions given to the subjects, the assessment of their performance as well as the recording of results.

Training, support and feedback for field examiners

The 20 field examiners involved in the measurement of functional capacity received advance training for their jobs during a dedicated training period. The 10 nurses

who were responsible for the home health examinations also took part. The training was organised in groups of 7–8, with supervision for practical training provided by six members of the functional capacity team. The measurement procedures and written instructions were prepared jointly with experts in each area. The training programme also included a video and expert lectures on the various areas of measurement.

A major point of emphasis in training as well as in the written guidelines was that all measurements must be rigorously standardised, including the verbal instructions given to subjects. Training for any one given test was provided by the same instructor to all field examiners and home health examination nurses. Upon completion of the training, skills tests were organised to assess the examiners' performance with regard to the instructions given to the subjects, the use of measurement tools, the assessment of the subjects' performance, interaction and the provision of feedback as well as the general fluency of executing the examination. Before the health examinations proper, the field teams conducted dry runs for one week at their respective locations, with the training supervisors monitoring the teams' performance with voluntary test subjects and where necessary providing guidance and instruction.

During the health examinations the training supervisors maintained regular contact with the field examiners both by phone and by visits to the examination sites. Any problems were recorded, and the teams also received instructions and further guidance via e-mail.

The performance of the nurses on the teams was monitored by means of video recordings. The aim was to make sure that all measurements were being correctly implemented and where necessary to root out any bad practices. Monthly videos were taken of each nurse's performance and sent back for evaluation to the training supervisors at survey headquarters. The supervisors made their assessments using a structured form and consequently provided feedback both to the individual nurses concerned and collectively to the whole team of nurses. In addition, a half-day refresher course was organised for each functional capacity team midway through the examinations. This provided an opportunity to review the execution of the examination by reference to materials compiled from the videos.

Calibration of equipment and quality assurance of operations

The equipment used in measuring functional capacity was checked and tested at regular intervals throughout the examination to ensure accurate operation. The hand grip strength meter and balance equipment were calibrated using standard weights whenever they were moved to a new site or at least once a week. The constancy of audiometry and reaction speed measurements was monitored by the members

of the functional capacity team, who conducted measurements on one another (using the same equipment on the same person) on a weekly basis and whenever the equipment was relocated. The results of these check-up measurements were recorded on separate forms for each examination site and mailed back monthly to survey headquarters.

If any equipment or computer software malfunctioned during a measurement and if the result consequently was unreliable or unavailable, this was recorded by the nurse in the data collection program and in the site diary. In the event of hardware malfunctions, the equipment set aside for home health examinations was taken into use and any scheduled home visits were postponed to a later date. In these situations the nurse on duty would directly contact the supplier of the equipment and also inform survey headquarters of the nature of the problem, where an entry was made in a separate malfunction diary. The supplier either rectified the problem as soon as possible, or supplied a new, replacement unit. All these steps served to ensure the quality of the measurement results in the event of malfunctions.

Once a week, the first subject of the day had repeat measurements of eyesight, hearing, reaction speed, hand grip strength, balance, joint function, chair stand test and walking speed, with changes made to the order of the people taking the measurements. These repeat measurements were taken about one hour after the measurements proper. Home visit nurses also took repeat measurements of functional capacity. Tables 8.8.1 and 8.8.2 show the results of these parallel measurements as well as the number of people who took part in them. For the purposes of analysis, the parallel measurements taken by the functional capacity examiners and the home-visit nurses have been combined. Repeatability between two measurements is described by reliability coefficients (for continuous variables) and by kappa coefficients (for categorical variables) (Winer 1971, Fleiss 1981).

Agreement was high (kappa over 0.6) in tests with an adequate number of cases (deviant findings), such as in certain components of the joint function test (walking on toes, crouching and climbing stairs) and in the tests of near and distant vision (Table 8.8.1). The number of cases in the walking and upper limb components of the joint function tests was very low (prevalence less than 5%). Likewise, the majority (97%) of subjects made no mistakes in the balance test. Reliable assessment of agreement is complicated by the small number of cases.

Repeatability for hearing and hand grip strength was excellent (repeatability coefficient $R > 0.90$) and for the chair stand test and the walking test good or moderate (Table 8.8.2). In several tests there were indications of the subjects' performance improving with repetition.

Table 8.8.1. Agreement of same-day measurements of functional capacity (categorical variables).

	N	Pr1	Δ	p-value	κ	95% CI
Eyesight tests						
Near vision (visus, 16 categories)	270			1.00	0.67	0.61-0.73
Distant vision (visus, 15 categories)	270			1.00	0.70	0.65-0.76
Balance field test (Guralnik)						
Balance variable, 4 categories[1)]	266			0.68	0.45	0.28-0.62
Joint function test[2)]						
Walk on level surface: degraded or failed	155	3.9	0.7	0.76	0.00	0.00-0.00
Walk on toes: degraded or failed	155	10.3	0.6	0.71	0.75	0.57-0.93
Climb two stairs: degraded or failed	143	13.3	1.4	0.48	0.75	0.58-0.91
Crouch: degraded or failed	155	24.5	–6.5	0.03	0.65	0.52-0.78
Raise upper arms: degraded or failed	154	7.8	0.7	0.71	0.67	0.44-0.90
Extend elbow joints: degraded or failed	154	2.0	1.3	0.32	0.00	0.00-0.00
Flex elbow joints: degraded or failed	154	1.3	–0.7	0.56	0.39	0.00-0.94
Backs of hands against each other: degraded or failed	153	4.6	–1.3	0.48	0.47	0.16-0.78
Clench fingers: degraded or failed	154	3.9	–1.9	0.32	0.37	0.05-0.69
Clench thumbs: degraded or failed	154	2.6	0.0	1.00	0.49	0.06-0.92

N = number of subjects tested

Pr1 = prevalence or proportion of degraded performances in the first measurement (in dichotomous variables)

Δ = difference between prevalences in first and second measurements

p-value = significance of difference between first and second measurements (McNemar test for dichotomous variables and modification of the test in variables with a larger number of categories)

κ = kappa

95 % CI = 95% confidence interval of kappa

[1] Performance in the balance test was classified into four categories as follows: 1 = feet side by side less than 10 sec, 2 = feet side by side 10 sec, semi-tandem less than 10 sec, 3 = semi-tandem 10 sec, tandem less than 10 sec, 4 = tandem 10 sec

[2] The variables of the joint function test were divided into two categories as follows: 0 = normal performance; 1 = degraded or failed performance

Table 8.8.2. Repeatability of same-day measurements of functional capacity (continuous variables).

	N	Mean1 (sd)	Δ	p-value	R
Hearing					
Hearing threshold in better ear at average frequencies of 500, 1000 and 2000 Hz (dB)	262	14.5 (12.8)	0.7	<0.001	0.97
Hand grip strength					
Maximal hand grip strength (N)	265	358 (133)	–11	<0.001	0.95
Chair stand test					
Time for 5 stand-ups (s)	151	13.5 (3.7)	0.9	<0.001	0.75
Walking test					
Walking speed over a distance of 6.1 metres (m/s)	153	1.59 (0.41)	–0.002	0.94	0.77
Force platform balance test					
Feet side by side, eyes open					
– X speed (mm/s)	106	3.81 (1.98)	0.12	0.46	0.51
– Y speed (mm/s)	106	5.72 (2.21)	–0.02	0.88	0.73
Feet side by side, eyes closed					
– X speed (mm/s)	106	4.96 (3.10)	–0.08	0.71	0.74
– Y speed (mm/s)	106	10.42 (6.69)	1.06	0.002	0.83
Semi-tandem test					
– X speed (mm/s)	105	13.39 (5.02)	0.27	0.44	0.70
– Y speed (mm/s)	105	11.47 (5.55)	0.68	0.07	0.74
Tandem test					
– X speed (mm/s)	89	18.80 (6.19)	1.09	0.04	0.64
– Y speed (mm/s)	89	15.67 (7.26)	1.72	0.006	0.59

N = number of subjects
Mean1 (sd) = mean and its standard deviation (first measurement)
Δ = difference between means (first and second measurement)
p-value = significance of differences between means (t-test)
R = repeatability coefficient

8.8.10. Checking and correcting data from the functional capacity measurements

Päivi Sainio, Päivi Haavisto, Noora Kuosmanen and Sanna Natunen

The results of the functional capacity tests were checked and corrected in accordance with the general principles of data correction (see Chapter 16). The software used for data collection allowed for the entry of free form comments on

each measurement. If no result was fed into the program, the reason for missing data was to be entered in a text field. As the field procedures could obviously not provide an answer to every eventuality, the field examiners were instructed to make a note of any situations where they were not sure how they should record the results. This proved particularly useful in the home health examinations and in the measurements taken at institutions, where the functional capacity of some subjects was very poor.

When the comments and observations entered by the field examiners for each test were reviewed one by one during data checking, some differences were seen in the ways they had entered the results. Especially in the case of subjects who were bedridden and incapable of hardly any physical performance, reporting differed quite considerably. Some field examiners had made no entry at all in the results section, others had noted that the subject was unable to complete the test. At this stage all entries for the different tests were harmonised in keeping with the written guidelines.

8.9. Clinical medical examination

Antti Reunanen and Markku Heliövaara

All subjects who attended the health examination were also to undergo a clinical medical examination. However not all people who turned up at the health examination reached this stage because they dropped out for one reason or another. In some instances non-participation in the clinical medical examination was indeed due to the fact that it was scheduled towards the end of the programme.

8.9.1. General overview

The purpose of the clinical examination was to obtain a medical assessment of the subject's main chronic diseases, to determine the need for treatment for those diseases and to assess the subject's level of functional capacity. All this had to be done on average within just 30 minutes. The aim was to give as standardised an examination as possible, a sound GP's examination. Training for the field physicians who conducted the clinical examinations focused particularly on standardisation of the clinical status examination and on the application of a consistent set of diagnostic criteria. The general guidelines applied in assessments of functional capacity were also the same across the board, but in this particular area there was greater variation among the examiners than in the results for clinical status and

diagnostic assessments, because only limited background data were available and no follow-ups were possible.

The main focus of the clinical medical examination was on circulatory diseases and musculoskeletal disorders. Within the given time and information constraints, the examining physician also tried to identify any other somatic and mental diseases. Particularly in the case of circulatory diseases and musculoskeletal disorders a conscious effort was made to ensure diagnostic comparability with the corresponding assessments in the Mini-Finland Survey.

During the clinical examination all results were entered into a Blaise computer program. An ordinary paper form was used as standby for the rare instances when the program crashed. However these data would be fed into the program as soon as it was up and running again.

On completion of the clinical examination the field physician offered a summary to the subjects of the main results from their previous examinations, such as rest ECG, spirometry tests, bioimpedance analysis and ultrasound examination of calcaneal bone. If necessary, the physician advised the subject to have further tests done or to seek treatment for a disease or abnormality detected.

8.9.2. Known diseases

At the start of the clinical examination, the physician asked the subject, firstly, what symptoms they had that hampered their daily activities; and secondly, what diseases they remembered having been diagnosed with.

At the time that the survey was designed, the assumption was that the field physicians would have access to information collected at earlier home-visit interviews on reported diseases and medications and possibly on the respondents' entitlements to special reimbursements for medical costs. Unfortunately, though, these data were not available by the time that the survey got underway. The situation improved later on, but for reasons of consistency and comparability it was decided that these background data should not be used in any part of the examination. Thus this part of the clinical examination was solely based on information given by the participants, without data from official health registries or from the home-visit interviews. Since questions concerning disease history were only briefly asked in the early part of the clinical examination, it is possible that the subjects did not recollect and mention all their diagnosed diseases at that time.

8.9.3. Clinical examination

The clinical examination proper was designed with a view to maximum economy of time use, above all so that subjects would not need to move back and forward between different positions for different examinations. The physicians used regular medical equipment in their examinations.

When indications of circulatory or respiratory diseases were detected, the physician proceeded with the examination in the conventional order from auscultation of the heart and lungs to an examination of peripheral arteries. In addition to any abnormalities in cardiac rhythm detected in the examination, the physician asked whether the subject had previously experienced any rapid atrial arrhythmias.

Examination of the musculoskeletal system was started with a nerve-pinch test in the lower limbs, proceeding then to the knee and hip joints and finally to the lumbar spine. The lumbar spine examination made use of a rather unconventional piece of equipment, namely an electric toothbrush. Using a specially commissioned rubber cap, the vibrating head of the toothbrush was gently applied to the processus spinosus of the vertebrae in the lumbar spine to see whether there were any obvious signs of pain in the nerve root. This simple test has only recently been adopted in clinical practice, but the evidence suggests it is quite reliable in detecting nerve root symptoms in the lumbar spine region (Yrjämä and Vanharanta 1994, Yrjämä et al. 1997).

The clinical examination then proceeded via an assessment of shoulder and neck movement to elbow and wrist joints. Wrist examinations focused especially on tests to determine constriction of the carpal tunnel (e.g. Tinell and Tetro). As for the examination of joints, the main concern was with the active and passive range of movement. Angle measurements were used to make these assessments if and as necessary.

With respect to other findings indicating diseases, special attention was paid to skin symptoms and neurological findings, particularly those suggestive of parkinsonism. Otherwise the clinical examination followed the principles of good clinical practice.

8.9.4. Diagnostic assessments

A diagnostic assessment was subsequently made of all diseases and disorders detected in the clinical examination. The main points considered were the certainty of the diagnosis (certain/suspected), the year of diagnosis, the need for treatment

(no need/receiving treatment/inadequate treatment/no treatment in spite of need) and the provision of treatment.

The occurrence of major diseases was inquired using a structured set of questions. For other diseases, the examining physician searched the appropriate ICD 10 diagnosis and recorded any necessary additional information. The data collection program had separate fields for the entry of comments and additional information.

The circulatory diseases inquired in the structured questions were angina pectoris, myocardial infarction, heart failure, hypertension, arrhythmia, valvular disease, peripheral arterial occlusive disease of the lower limbs and cerebrovascular disorders. For subjects with coronary heart disease, it was recorded whether they had had an invasive coronary procedure (bypass or angioplasty). For arrhythmias, valvular disorders and cerebrovascular disorders, the specific type of condition was entered where possible. Finally, the physician recorded their assessment as to whether there were clear changes in the recorded ECG.

The respiratory diseases listed were asthma, chronic obstructive pulmonary disease and allergic rhinoconjunctivitis.

In the category of musculoskeletal disorders, the structured items concerned the presence of inflammatory arthritis, knee and hip arthrosis, chronic neck syndrome, chronic lower back syndrome, chronic shoulder impingement syndrome, chronic epicondylitis, chronic carpal tunnel syndrome as well as sequelae of amputations and injuries to knee and ankle ligaments. In the case of arthritic conditions, neck syndromes, lower back syndromes and shoulder syndromes, the specific type of disease was recorded.

Other somatic diseases inquired using structured questions were diabetes, hyperlipidaemia, hypothyroidism, Parkinson's disease, cataract, glaucoma, chronic eczema and chronic hand eczema. For diabetes and chronic eczema, the specific type of disease was recorded. Structured items for mental disorders concerned psychotic disorders, depression and dementia.

8.9.5. Functional capacity and need for rehabilitation

Following the diagnostic assessments, the examining physician proceeded to consider the subject's functional capacity. Mobility was assessed for all subjects, and the status of those with heart conditions was additionally evaluated using the New York Heart Association (NYHA) functional classification. Level of functional capacity was assessed for all subjects and level of work ability for subjects of working age. The main underlying reasons for all limitations of functional capacity and work ability were identified as far as possible. Both permanent and temporary

functional limitations were recorded. The latter were particularly relevant in cases where the temporary limitation hampered the assessment of permanent functional capacity.

In addition to the assessment proper of functional capacity, the examining physician rated each subject's potential work ability in four hypothetical occupations: teacher, builder, school caretaker and systems analyst.

The examination of functional capacity concluded with the physician's assessment of whether the subject was in need of any form of rehabilitation. Finally, the medical examination ended with an overall assessment of the subject's general health status.

8.9.6. Quality control

Measures of quality assurance for the clinical examination included repeat examinations conducted on the same subject by another physician on the team as well as separate quality control tests in which the examination was conducted by a physician from another field team on a subject specifically recruited from outside the study sample. However these tests involved only relatively small numbers of people (93 in continuous quality control and 173 in a separate study), and therefore only very general conclusions can be drawn about differences observed.

The κ-values for the repeatability of diagnostic assessments fell within the range of 0.31–1.00, with the highest figures recorded for asthma, arthritis and myocardial infarction diagnoses. The diagnosis of chronic obstructive pulmonary disease seemed to show particularly poor repeatability, although its overall prevalence was also very low.

In a comparison of the diagnoses by physicians from different field teams, the corresponding κ-coefficients fell in the range of 0.34–0.89. Here, the best repeatability was observed for diagnoses of arthrosis, asthma and diabetes. The poorest repeatability was again seen for chronic obstructive pulmonary disease (COPD), which may have to do with the low prevalence figures.

Unfortunately, quality control in the clinical medical examination was confined to such a limited material that the number of diagnoses for many major diseases remained too small. Nevertheless it seems that the standard of diagnostics for serious conditions and diseases that impact on functional capacity was high or at least satisfactory. Having said that, the less common the disease and the less accurate the diagnosis, the greater also the diagnostic variation.

Diagnostic assessment of the clinical examination was in some cases not accurate enough because of the limited time available for the examination and lack of

background information. Therefore, several of the clinical diagnoses were later checked and corrected (if needed) using additional information from the laboratory results analysed after the field examination and from all available register information.

8.9.7. Comparability with the Mini-Finland Health Survey

The clinical medical examinations were conducted along the same general lines in both the Mini-Finland Health Survey carried out in 1978-80 (Aromaa et al. 1985, Sievers et al. 1985, Aromaa et al. 1989b) and the Health 2000 Survey. The diagnoses and assessments of functional capacity, therefore, are largely comparable. However there were some differences in the design and contents of the health examinations, which inevitably are reflected in comparability.

In both the Mini-Finland Survey and the Health 2000 Survey an extensive health interview was conducted with the subjects at their homes. All interviewees were then invited to a two stage health examination. The first stage involved an investigation of a wide range of risk factors, symptoms and findings. On the basis of the results of this first stage, all subjects with some identified risk factor or finding suggesting a disease were invited to attend the second stage of the health examination, which included a clinical medical examination.

Although almost two-thirds of the subjects attending the first stage of the health examination also took the medical examination, a physician did not see all of the participants in the Mini-Finland health examination. In Health 2000, all of those who attended the health examination also took part in the medical examination. According to the methodological assessment of the medical examination in the Mini-Finland Survey, this selection did not significantly affect the sensitivity or accuracy of the diagnoses of major diseases (Heliövaara et al. 1993).

Another important difference between the research designs in the Mini-Finland Survey and Health 2000 derived from the organisation of the health examination in two stages. In the Mini-Finland Survey the examining physicians had access to more extensive and detailed information about the subjects than was the case in Health 2000. This was because the physician was aware at the time of the examination of the findings made in the first stage of the Mini-Finland Survey, which led to the clinical examination in the second stage. In the Mini-Finland Survey the second stage of the health examination was conducted several months after the first stage. Therefore the physicians conducting the clinical examination had access to information on reported diseases, medications, symptoms, ECG findings and laboratory results, whereas the physicians in Health 2000 had far less information at their disposal.

The clinical examinations in the two surveys also differed in some of their details, but it is unlikely that these had any impact on the diagnostic assessment of major diseases. In Health 2000 the decision was made to refrain from any attempt to diagnose new diseases, firstly because the chances of doing this were remote and secondly because it was thought the information would not have had very much value.

The major diseases identified in the physician's diagnostic assessments were the same in both surveys. In the case of less common diseases, there were some differences between the structured diagnostic items. The new structured diagnoses included in Health 2000 were diseases that have gained increasing significance over the past two decades, or that nowadays can be more effectively diagnosed and treated. Assessments of functional capacity and work ability were made in largely the same way in both surveys, although the Mini-Finland Survey did not include an assessment of work ability in four hypothetical occupations.

8.10. Mental health interview

Sami Pirkola, Jouko Lönnqvist and mental health team

There is a relative scarcity of reliable and up-to-date epidemiological data on the prevalence of several mental health disorders in Finland, as it is more than 20 years now since the previous major health survey in the country. The reliable diagnosis of mental health disorders is a time-consuming business and requires the use of structured or standardised interview methods.

In the Health 2000 Survey, the interview on mental health disorders or symptoms focused on four principal categories: mood disorders, substance addiction, psychoses and anxiety disorders. Questionnaires and other studies were administered to obtain additional information from the subjects.

8.10.1. Main areas of interest

The mental health team collected information on factors related to mental health and well-being in three different ways:

1. Assessments based primarily on the mental health interview (mood disorders, alcohol use disorders, psychotic symptoms, and anxiety disorders);
2. Assessments based primarily on questionnaires (Beck Depression Inventory (BDI), General Health Questionnaire (GHQ-12), Toronto Alexithymia Scale

(TAS), factors predisposing to alcoholism, hypochondria and somatisation disorder, and eating disorders in subjects aged 18–29);

3. Assessments based on the home-visit interview (perceived need for treatment, and quality of care received).

Less extensive data were collected for younger subjects aged 18–29, although the questionnaire for this age group included a set of items concerning eating disorders, which in turn were not included in the survey for people aged 30 or over.

8.10.2. Methods

The contents of the home-visit interview (see Chapter 6) and questionnaires (see Chapters 7 and 8.12) are described in more detail elsewhere in this report. As part of the clinical medical examination (see 8.9 above), the doctor was to answer questions about the subject's mental health: did they feel that the subject suffered from any mental health disorder and if so, what. The following provides a more detailed account of the mental health interview and the laboratory samples used in the study of mental health.

Mental health interview

The interview method chosen for the study of mental health disorders in the Health 2000 Survey was the Composite International Diagnostic Interview or CIDI. CIDI is a structured interview developed by the World Health Organization (WHO) for purposes of epidemiological research (WHO 1990). 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. It has been tested in studies conducted in several countries (Wittchen et al. 1998a, Kessler 1999).

Work is currently underway to set up a joint international consortium on psychiatric epidemiology that is built around the CIDI. As the interview is used all around the world, it provides an invaluable tool for country comparisons of the prevalence of different disorders. By 1997, the CIDI had been used in eight major epidemiological surveys in North and South America as well as in Europe. All told, almost 34,000 people had been interviewed in these surveys. At the moment there are seven ongoing epidemiological studies with a combined total of almost 32,000 interviewees. Furthermore, the World Mental Health 2000 project involved six major epidemiological studies around the world, interviewing a total of 48,000 people (Kessler 1999).

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. 1998b). Its main drawback is that the interview lasts on average 75 minutes, which is why it could not be implemented in full in the Health 2000 Survey.

For the purposes of the Health 2000 Survey, the mental health team opted to go with the most recent CIDI version available, i.e. CIDI 2.1. At least two public domain versions are available of the program. The team looked at the CIDI Auto version used by Australian researchers and the German M-CIDI, and eventually decided to use the latter (Wittchen et al. 1998a,b).

A group of researchers from the National Public Health Institute (KTL) received training on the use of the interview method and its computer version from Professor Hans-Ulrich Wittchen from the University of Munich, who was involved in developing the M-CIDI program as well as DSM diagnostics. Some of these researchers conducted pilot interviews for the Health 2000 Survey, others provided the training for its field interviewers.

The M-CIDI program was translated into Finnish in collaboration with a Munich-based computer operator, who produced Finnish-language M-CIDI versions of different interview components for preliminary testing. Rough translations were tested both at the pilot stage of the survey and at other stages with volunteer subjects. The interview items were translated from English into Finnish by health care professionals with a knowledge of DSM diagnostics. These translations were then edited on the basis of the feedback from the pilot stage as well as from lay respondents in test interviews.

Once testing with volunteers was completed and the team was happy with the interview package, the translation was checked for accuracy by an authorised translator. Some further revisions were made to the questions on the basis of this feedback. The final translation and the computer version produced of that translation were completed shortly before the field stage survey got underway in August 2000.

Based on questions about symptoms experienced during the past 12 months, the CIDI interview used in the Health 2000 Survey inquired about the occurrence of depression, chronic depression, generalised anxiety, social phobia, panic disorder, alcohol use disorder and possible psychotic disorder during that one year. As for alcohol or other addictions, subjects were also asked whether they had experienced symptoms earlier on in their life.

For purposes of quality assurance of the mental health interview in the Health 2000 Survey, inter-interview reliability (κ-coefficient for test-retest reliability)

was determined for diagnoses of depression. To this end 20 interviewers who had been involved in conducting the mental health interviews paired up in teams of two and interviewed 49 randomly selected clients of a private occupational health service. The interviewers were blind to each others' diagnoses, and the order of the interviews between interviewers was varied.

The quality control interviews were based on the depression component of the M-CIDI mental health interview, the purpose of which is to establish whether the interviewee meets the diagnostic criteria of depression or chronic depression (dysthymia) during the past 12 months. The κ-value obtained for depression was 0.88 (95% confidence interval 0.64–1.00, percentage of agreement 94%) and for dysthymia 0.88 (95% confidence interval 0.64–1.00, percentage of agreement 98%). Based on these results, the inter-interview reliability of the M-CIDI depression components was excellent (Pirkola et al. 2005).

Laboratory samples

Recent research has shown a growing interest in the role of cortisol and HPA axis in stress-related conditions and psychiatric disorders (McBurnett et al. 2000, Bandelow et al. 2000). Among these conditions and disorders, key areas of interest for the mental health team in the Health 2000 Survey were mood disorders, anxiety disorders, substance abuse disorders as well as perceived and measured stress. Measurement of the role of HPA axis in different conditions yields crucial information on the biological dimensions of these phenomena. One particular concern is with whether cortisol concentrations differ in mood, anxiety and substance abuse disorders when they are or are not associated with acute or long-term stress.

The mental health team has planned to use the laboratory samples collected in the Health 2000 Survey to determine plasma cortisol concentrations. Records show the exact time at which the laboratory samples were collected, and since the subjects were separately asked at what time they got up on the day of the examination, it is possible to determine cortisol concentration levels in relation to the expected concentrations based on the amount of time the subjects had been awake.

The laboratory samples can also be used to determine plasma concentrations of medication, which provide useful information for monitoring adherence to psychiatric medication regimens. The DNA collected allows for studying different genetic polymorphisms. The purpose is to test various research hypotheses primarily based on associations.

8.11. Final interview

Pirkko Alha

The last stage of the health examination in the Health 2000 Survey was the final interview, which summed up the whole examination. Checks were made to ensure that the subject had taken all assigned tests and examinations and completed all forms and questionnaires. The subjects were also informed about the forms they would be given to take home, and about any other examinations that they possibly would still have.

Once the subject arrived at the final interview point, the nurse examined the file they had brought along. Having checked that the subject had had all their assigned tests, the nurse entered this information into computer. The reasons for any missing tests were also entered.

Next, the nurse proceeded to check that the data had been correctly entered in all the subject's forms. The subjects were asked whether they had given a urine sample. Questionnaires 1 and 2 were checked for any missing items and completed if necessary. The subjects were then handed instructions for questionnaire 3 and the dietary questionnaire. Advice was given on how to complete the dietary questionnaire, and in two regions (Turku and Oulu) the subjects were informed that they would be receiving their food diaries by mail. Return envelopes were handed out for questionnaire 3 and the dietary questionnaire.

Some of the test results that the subjects carried in their file were handed over to the subjects themselves. Copies were taken for archiving. Subjects also received a feedback form (T2022) which detailed their tests results, an ECG printout, feedback from the dental examination, ortopantomography x-ray, bioimpedance printout, and calcaneal bone ultrasound printout. The nurse also recorded on computer which forms had been given to the subject. If the subject had received a referral from a doctor for further examinations (T2027), a copy was taken and archived.

The subjects were informed that the laboratory results would be mailed to them at home. A sample of the subjects also received instructions on how to collect and return a faecal sample. Subjects were also given the opportunity to provide their own anonymous feedback on the survey.

At the final interview point in the Tampere region, the subjects received a written invitation to attend an examination organised by the UKK Institute for Health Promotion Research. Invitations were made to all those who met the admission criteria, i.e. who reported that they could walk 2 km without aids or who could complete the chair stand test.

Finally, the nurse thanked the subject for their participation and asked whether they had any questions. This was the subject's final point of contact with the survey and therefore very important to what impressions they took away. The nurses were also instructed to make a note any unusual events or situations in the computer program's 'other comments' section.

All forms collected at the final point of the health examination were archived and organised and mailed back to headquarters once a week. Since the first subjects did not begin to arrive at the final point until 3–4 hours after the examination started, the nurse spent this time trying to contact people who had not turned up, trying to make new appointments before the team moved on to the next location. Separate forms were completed for people who did not attend the health examination (T2044).

The nurse in charge of the final examination point was usually the head nurse and also had a host of other duties. These included organising the necessary transport for moves from one location to another, making other arrangements for the removal, scouting the new examination facilities, staff accommodation, among other things.

8.12. Complementary questionnaire (questionnaire 3)

Arpo Aromaa

On leaving the health examination, the subjects received two further questionnaires that they were asked to complete and return by mail within the next few days. The first, complementary questionnaire (questionnaire 3) was intended to collect further background data, the second was a dietary questionnaire on the use of food.

8.12.1. Sleep

The first part of the questionnaire dealt with sleep, difficulties falling asleep, waking early, the tendency to fall asleep during the daytime and snoring and possible sleep apnea.

8.12.2. Disadvantages in housing conditions

The questionnaire included 12 items concerning possible problems with housing.

8.12.3. Attitudes regarding health

Arpo Aromaa and Sami Pirkola

Items measuring attitudes to health concerned the tendency to be overly concerned about various symptoms and the possibility of illness. Similar questions are commonly used to help determine the presence of hypochondriasis.

The Whiteley index is an originally 14-item self-rating scale developed for the assessment of hypochondriacal symptoms (Pilowsky 1967). Hypochondriasis is characterised by the fear or conviction of ill health, in spite of clinical evidence to the contrary. The Health 2000 Survey used the 7-item version of the Whiteley index, which has proved to be equally reliable as the original instrument (Fink et al. 1999). The degree of hypochondriasis is measured as a sum score. The Mini-Finland Survey used the original Whiteley index. The subjects were handed the questionnaire in connection with the home interview and asked to bring it along to the health examination.

8.12.4. Oral health and quality of life

It has been reported that oral health has a very significant impact on quality of life. The oral health impact profile (OHIP) is a set of 14 questions designed to measure those impacts (Slade et al. 1994). It was translated into Finnish for use in this survey. Questions concern various symptoms related to oral and dental health and their impacts on everyday life and functional capacity.

8.12.5. Experience of everyday life

Experience of everyday life and sense of life control was measured using Antonovsky's Sense of Coherence scale (Antonovsky 1993).

8.12.6. Seasonal variations

Arpo Aromaa and Sami Pirkola

The change of seasons affects not only people's activities, but also their mood. Questions were asked about the perceived existence of seasonal changes and about the problems they may cause. One of the areas that has received increasing attention in recent years is seasonal affective disorder (SAD) and its treatment (Partonen and Lönnqvist 1998). In the Health 2000 Survey, seasonal effects on well-being were inquired with two items (14 and 15) under the questionnaire heading Seasonal changes.. The questions were drawn from the Seasonal Pattern Assessment Questionnaire (SPAQ; Rosenthal et al. 1984), which has been reported to have good validity (Magnusson 1996). The sensitivity of the SPAQ scale is 94% and accuracy 73%, but its repeatability is at best satisfactory (Lund and Hansen 2001).

SAD can be screened by linking the data from the Seasonal changes items in the questionnaire and the depression items in the CIDI interview. When data on the prevalence of mood disorders, the time of year and earlier results on the prevalence of SAD are all taken into account (Blazer et al. 1998, Magnusson 2000), it is estimated that around 100 out of the 6,900 adults interviewed suffer from SAD. These data will be used to determine the prevalence of SAD and to study its regional variation and underlying factors.

Because of shortcomings in the screening method, the purpose is later on to conduct a separate interview (SCID) to ascertain the validity of the SAD diagnosis and at the same time to validate a new screening tool developed for the detection of SAD (Seasonal Health Questionnaire, SHQ; Thompson and Cowan 2001) in a representative population sample.

8.12.7. Health-related quality of life

The next battery of questions (15D) concerned functional capacity, activities of daily living and problems in these activities. The items are defined as a scale measuring health-related quality of life (Sintonen 1981, Sintonen 2001).

8.12.8. Experiences of the influence of alcohol

Items under this heading measured the impacts of alcohol and the development of alcohol tolerance.

8.12.9. Emotions and feelings

Sami Pirkola

Alexithymia refers to a style of personality that is characterised by lack of awareness about one's own emotions and failure to differentiate them from somatic symptoms; an inability to talk about emotions; lack of imagination; and an oversimplified, concrete thinking style (Taylor et al. 1997). The alexithymic individual may, for example, give an endless description of external details and completely disregard emotions. Because of their complete lack of emotional expression, sufferers are often considered dull and uninteresting.

Alexithymia is associated with many psychiatric and somatic diseases as well as with health behaviour. Originally the condition was associated with so-called psychosomatic diseases (Sifneos 1973), and it has even been suggested that it could emerge as a new paradigm of psychosomatic diseases (Taylor et al. 1991). More recently, however, it has been shown that alexithymia is also connected with alcoholism (Ziolkowski et al. 1995), panic disorder (Joukamaa and Lepola 1994), eating disorders (Cochrane et al. 1993), somatisation and increased reporting of somatic symptoms (Kauhanen et al. 1994, Lumley et al. 1997) and depression (Saarijärvi et al. 1993). It predicts male mortality independently of other risk factors (Kauhanen et al. 1996). Recent studies have shown that, unlike mental strain symptoms and/or repressed hostility, it is an independent risk factor for elevated blood pressure (Jula et al. 1999). Alexithymia is the opposite of psychological creativity and insight (Conte et al. 1990).

All in all, alexithymia is a relatively novel concept that is characterised by general difficulties in identifying, verbally expressing and controlling emotions and that would appear to be an independent risk factor for several psychiatric and somatic diseases. Research will help us better understand the role of emotional factors in illness and health as well as in health behaviour. Some epidemiological research has been done into the associations of alexithymia with sociodemographic factors (Lindholm et al. 1990, Salminen et al. 1999), but there are no population-level studies on its associations with other morbidity.

Description of method

It was not until the 1980s that a scientifically based instrument was developed for the assessment of alexithymia: the Toronto Alexithymia Scale (TAS; Taylor et al. 1985). The psychometric properties of the instrument has been found to be sound (Bagby et al. 1990, Taylor et al. 1988, 1990a,b). A new, 20-item version (TAS-20) has since been developed which is considered even better than the original tool (Bagby et al. 1994a,b). It consists of 20 statements that are rated on a 5-point

Likert scale from strongly disagree to strongly agree. Scores for items 4, 5, 10, 18 and 19 are reversed. Alexithymia is assessed on the basis of the sum score which ranges from 20 to 100: the higher the score, the more severely alexithymic the individual. For purposes of assessing the prevalence of alexithymia, the sum score can be dichotomised using the cut-off point of 60/61. Three component factors can be identified in TAS-20: 1) difficulty describing emotions and differentiating them from physical sensations, 2) difficulty describing emotions to others, and 3) externally oriented thinking style.

TAS-20 has shown better internal consistency, test-retest reliability and validity than the original instrument (Parker et al. 1993, Bagby et al. 1994a,b). The validity of the Finnish translation of TAS-20 has been rated as good. In a study among the northern Finland 1966 birth cohort, the index figures used in the measurement of validity were even better than those in the original Canadian validation study (Joukamaa et al. 2001). The instrument is simple and easy to complete, taking no more than 10 minutes or so. TAS-20 was included as part of questionnaire 3, which was given to the subjects for completion at home.

8.12.10. Questions for women and men

These questions concerned infections and diseases in the genital area as well as the examination of breasts and testicles.

8.12.11. Driving

Finally, a battery of questions was asked to elicit information on exposures possibly linked to the amount of time spent driving, particularly on stress on the back.

8.12.12. Experiences

Overall the questionnaire was clear and the questionnaires returned were completed with very little missing information. There were some problems, as expected, with questions concerning the influence of alcohol, and unexpectedly with questions concerning driving. Since the questionnaire was to be completed at home and returned by mail, the response rate of just 75% came as no surprise.

Web link: Complementary questionnaire (questionnaire 3)
http://www.terveys2000.fi/lomakkeet/en/t2005en.pdf

8.13. Dietary questionnaire

Satu Männistö, Jukka Montonen, Pirjo Pietinen and Paul Knekt

Food frequency questionnaires (FFQs) have become a primary method in epidemiological studies concerned with the association of diet and the risk of diseases (Willett 1998, Pietinen 1999). These questionnaires provide information on the subject's diet during a specified period prior to the study. Food consumption over a longer period sheds more meaningful light than the current moment on the development of chronic diseases, which can take years and even decades.

Often these questionnaires are designed to measure the subject's diet as a whole. The main aim is to rank-order the subjects according to their food or nutrient intakes, not so much to measure the absolute intakes. The processing of FFQs is readily computerized and inexpensive, and they are easy for subjects to complete. The development of the questionnaire itself, on the other hand, is a time-consuming business, and it is necessary always 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). The participants were asked to describe their ordinary diet over the past 12 months. The questionnaire listed 125 food items, mixed dishes and alcoholic beverages commonly used in Finland, grouped in the following categories: dairy products, cereal products, fat spreads, vegetables, potatoes, pasta and rice, meat, fish, poultry and eggs, fruit and berries, desserts, confectionery and other snacks, and beverages.

The subjects were asked to estimate how often they had used the foods listed on a nine-point scale: never or rarely, 1–3 times a month, once a week, 2–4 times a week, 5–6 times a week, once a day, 2–3 times a day, 4–5 times a day and 6 times or day or more. Participants could compare the size of their portions with predefined sizes printed on the questionnaire. There was also an empty space for entering any foods that the participants used frequently but that were not listed in the questionnaire. The questionnaire also included separate items on special diets and the use of dietary supplements.

The FFQ was handed to the participants in connection with the health examination or home 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. It was requested that the participants complete and return the questionnaire to the National Public Health Institute within two weeks.

A first reminder was sent out to subjects who had not returned the questionnaire within 1–2 months of the health examination. If necessary a second reminder was sent out 2–4 months after the health examination. All in all, after two reminders, 6,373 questionnaires were received.

Once the forms had been returned to the National Public Health Institute they were given a preliminary check by a nutrition researcher, with special reference to their general credibility, the food amounts indicated and the number of empty lines. If there was a whole page that had not been filled, a copy of that page was sent back to the participant with the request that he or she fill it out. Based on this preliminary check, the returned questionnaires were divided into three groups: questionnaires acceptable for analysis, incomplete or unfeasible questionnaires, and blank questionnaires. Incomplete or unfeasible questionnaires were reconsidered.

Once the acceptable forms had been entered into the database, the maximum and minimum frequency values for each food item in the questionnaire as well as the logical credibility of each item and response were checked. At this stage, 368 blank or otherwise incorrectly completed questionnaires were excluded. Intake of food and nutrients was eventually calculated for 5,998 participants (75% of the Health 2000 sample).

Consumption of different foods and daily intake of nutrients were calculated by using the Finnish National Food Composition Database (FINELI®) as well as FINESSI software of the National Public Health Institute. Checks were performed on range of nutrient intakes, especially extremely high or low intakes. The final dietary dataset comprises around 100 food groups, some 100 ingredients and 100 nutrients that can be used for research purposes.

Short-term repeatability was assessed by repeating the questionnaire within eight months of the first measurement. In the reproducibility study, the questionnaire was sent out to 209 randomly selected participants. The second FFQ was obtained from 180 participants. The intraclass correlation coefficient for foods ranged between 0.16 and 0.82, and for nutrient intakes the figures were between 0.22 and 0.72.

For the validation study, a random sample of 470 participants from the Health 2000 population was collected from eight different locations. In that study, the participants were asked to keep a three-day food record of all foods and beverages they consumed during those consecutive days, including the time and place of each meal. To help them assess the size of their portions, the subjects were given a picture book with 153 colour photographs of different portion sizes of commonly used foods (Haapa et al. 1985, additional photographs 2002). The food record and the picture book were mailed to the subjects once they had returned the FFQ to the National Public Health Institute.

The three-day food record was returned by 334 (87%) participants. The same principles were applied to checking the food records as to the FFQs. Forty food records were excluded, leaving 294 (76%) records for the final validation data from participants who had provided acceptable food consumption data both in the FFQ and in the food record. The correlations between the FFQ and the food records ranged from 0.14 (retinol) to 0.66 (fiber and alcohol) in men, and from 0.20 (n-3 fatty acids) to 0.70 (alcohol) in women (Paalanen et al. 2006).

8.14. The interaction with clients

Päivikki Koponen

Every effort was made in the health examination to create an atmosphere that would inspire a positive attitude and motivation on the part of the subjects to participate in the survey, and ensure that the situation in the examination was ethical. It was thought that a relaxed atmosphere would help to minimise any undue interference caused by the client's nervousness in the tests and measurement results. Furthermore, in an atmosphere of confidence it would be easier for subjects to speak openly even about sensitive problems. Even though the aim otherwise was to create as standardised a situation as possible, each client was treated as individually as possible. It has been reported earlier that these aspects of interaction between health examination staff and subjects have a direct bearing on levels of participation (Groves and Coupier 1998, Koponen and Aromaa 2004).

Feedback forms were handed out to every other subject to obtain their assessments of how the field teams had performed in their job. Specifically, they were asked to identify aspects of the examination that made it a positive experience and/or what problems they had encountered or what they thought could have been done better. The feedback forms were returned at regular intervals back to headquarters for summary. In addition the nurse in charge of each field team monitored the client feedback received, and where necessary discussions were held with staff in the field or during visits by management from survey headquarters.

The majority of the clients who returned the forms took the view that participation in the examination had been a positive experience by virtue of the polite and friendly attitude of staff members. As one client pointed out: "The staff were exceptionally friendly, it is very rare for an old person to enjoy that kind of attention nowadays" (a 81-year-old woman). Around half of the subjects felt that the positive experience was in large part due to the interest shown in and the personal attention given to them. Other positive feedback included comments that the examinations

had been so thorough; that people had received new information about their health; and that the results were explained clearly and in great detail.

Less than one-third of those who returned the feedback form had criticisms or suggestions for improvements. The main problems reported were the difficulty of answering a large number of questionnaires and the occasional queues to different examinations. There were only a few isolated negative comments on the examinations or interaction with clients. Some of the participants would have wanted to receive more information and guidance, for instance: "I personally would have wanted as much new information on health issues as possible … I was rather left with the feeling that you got all this information out of me, but what do I do now with all my complaints." (a 44-year-old woman)

8.15. Feedback to subjects

Sami Heistaro

The people who took part in the survey received quite a lot of information about their health in connection with the examinations, and more was provided in a letter that was sent home to them later on. The subjects considered the information they received both interesting and valuable, which helped to encourage those recruited to take part in the survey.

At the health examination the subjects were given a form detailing their blood pressure values, height, weight, hip and waist circumference and spirometry results. In addition, the subjects received a preliminary assessment of their distant and near vision as well as their hearing. A doctor interpreted their ECG printouts as well as the results of their calcaneal bone ultrasound measurements, which were also given to the subjects to take home. A dentist provided an evaluation of the orthopantomographic x-ray, a copy of which was prepared for the subject to take along.

The subjects received a letter on other feedbacks a few months after the health examination. These included the results of the subjects' laboratory tests (total and HDL cholesterol, triglycerides, gamma-GT, urate, blood sugar), an explanation of the meaning of these results and what to do if they were higher or lower than the recommended values. If the laboratory results differed significantly from reference values and the result could not be explained by earlier known diseases, the subject was personally contacted by phone or (if this was not possible for one reason or another) by mail. Unfortunately in some cases the sending of feedback letters was considerably delayed.

Furthermore, the feedback letter provided more detailed information on the results for vision and hearing as well as on functional capacity. The letter was clear because it prompted only very few inquiries. However, many subjects had expected to receive their letter more quickly than they eventually did, which caused a number of phone calls to the NPHI.

If either the medical or dental examinations disclosed findings that required urgent attention, the subjects concerned were referred to further examinations or to receive treatment at their local health centre or other appropriate health care unit.

9. HOME HEALTH EXAMINATION

Seppo Koskinen, Päivi Sainio, Irma Salminen and Antti Reunanen

Home health examinations were conducted on all consenting subjects who did not attend the health examination proper. The home health examination included many of the same measurements as the examination proper and a short version of the interview for subjects who had not taken the home-visit interview proper.

9.1. Duties of home-visit nurses

Each of the five field teams had a health examination staff of 14–15 appointed to different posts and measurement points. In addition, each team had two home-visit nurses whose duties included the following:

- standing in for other health examination staff in the event of illness
- providing door-to-door transportation for subjects who were unable to travel to the health examination site either of their own accord or even by taxi
- trying to contact subjects who did not show up at the health examination (regardless of whether they had taken the home-visit interview)

when subjects who had not attended the health examination were contacted,

- making an appointment with the subject for a health examination proper

if the subject refused,

- making an appointment for a home health examination at the subject's home, in an institution or elsewhere

if the subject refused to have a home health examination,

- conducting a telephone interview (see Chapter 10)

in the case of subjects who agreed to a home health examination,

- conducting the home health examination either at the subject's home, in an institution or elsewhere
- supporting the rest of the field team, among other things, by
  - assisting elderly or disabled subjects in the health examination
  - assisting to clear backlogged examination points

- standing in for the examination nurse in charge to allow her to attend to other duties
- making necessary acquisitions
- providing tea, coffee and snacks
- conducting other tasks to facilitate the health examination and to promote the quality of the data collected

As long as the health examinations proper were continuing, the main priority for the home-visit nurses was to assist in keeping that process running smoothly. Next on the list of priorities was trying to contact dropouts and people who had not been reached and making appointments for health examinations. During the course of the health examinations proper, the home-visit nurses would continue their preparations for and conduct home visits to the extent that their other duties permitted.

9.2. Training for home-visit nurses

During the three-week training period (21 August – 8 September 2000) ahead of the health examinations, the home-visit nurses received detailed instruction and guidance on how to conduct the functional capacity tests and the mental health interview. In addition, they were given training on the laboratory skills they would need in their home visits, on taking blood pressure and body measures, on performing brief oral examinations and on IT issues relevant to the home health examinations. The training session was concluded with the home-visit nurses conducting one dry-run home health examination in teams of two.

During the early stages of the fieldwork, the home-visit nurses concentrated on assisting the field teams. When the field teams had been at work for a month and the health examinations began to run more smoothly without the need for continuous assistance from the home-visit nurses, a two-day refresher training course was arranged for them in Helsinki on 9–10 October 2000. A further training day was held on 29 November 2000, which included a review of the experiences from the first 20 home health examinations.

As well as dealing with aspects of the home health examinations themselves, the training courses gave special attention to the motivation of subjects. The nurses were explained in detail the impacts of non-response to the validity of the results and shown what kinds of arguments they could use to try to persuade the subjects to take part in the examination. Furthermore, experiences from earlier population surveys were discussed so as to give them a sense of what kind of people they could expect to show the most resistance (marginalised groups, people with illnesses and

disabilities, people covered by good occupational health services and people who generally were critical towards interview and other similar surveys) and how best to motivate them.

The home-visit nurses were encouraged to stress the point that the health examination offered was quite unique in its diversity and comprehensive and that it was crucially important to collect the data from every person recruited into the sample so as to gain a representative picture of the population's health and related factors, for this information would provide a valuable foundation for the development of health and other services for Finnish people.

9.3. Content of home health examination

The home health examination was designed with a view to including as many elements of the health examination proper as possible. However, none of the elements of the clinical medical examination nor any measurements requiring heavy equipment could be included. It was also agreed that the maximum duration of the home health examination should not exceed 2–3 hours. The content of the home health examination and the estimated duration of its different stages are shown in Table 9.1.

The home health examination followed essentially the same basic guidelines as were issued for the health examination proper and for the Statistics Finland home-visit interviews. However the following additions were made to those guidelines:

- if the subject is unable to provide the information required, as many of those items as possible shall be obtained from a family member, nurse, etc.
- if the subject is unable to complete questionnaires 1 and 2, the home-visit nurse shall compile this information by interviewing the subject
- separate guidelines issued for home-visit measurements shall be followed in the measurement of blood pressure, heart rate, height, weight, body circumference and peak expiratory flow; the main difference compared to the guidelines for the health examination proper were 1) measurement of blood pressure using an Omron M4 monitor, 2) measurement of weight using portable scales (EKS), 3) measurement of height using a standard metallic measuring tape and an erasable mark made on the wall with the subject standing back against the wall, 4) PEF reading instead of spirometry measurement
- separate instructions were followed for the extraction and processing of blood samples; the main differences compared to the health examinations proper

were that the amount of blood collected was smaller (30 ml), the serum and plasma were centrifuged after storage at room temperature for 3–8 hours, and the centrifuged samples were stored in a refrigerator, frozen within 1–3 days of sample collection, unfrozen samples were posted by mail if necessary to the National Public Health Institute (KTL) and storage of whole blood DNA sample at room temperature the whole time before freezing

- the oral examination was conducted following the procedure described in the home health examination data collection form (T2074)
- For the stair-climb and chair-stand tests in the home health examinations, the stairs and chairs available in the subject's home were used. The height of the stairs and the seat of the chair (ideally at 43 cm) were entered in the data collection program. The walking test was taken over a shorter distance than the regulated 6.1 metres if that was not possible; in this case the distance was recorded. A field version of the balance test was conducted following the protocol described by Guralnik et al. (1994) (see Chapter 8.8). Scotopic vision was not tested at home.

9.4. Implementation of home health examination

As soon as the field health examinations were properly up and running and the continuous assistance of home-visit nurses was no longer required, the home health examinations were started – in some field teams they got underway as early as October 2000, in others at the turn of 2000–2001. Where possible, the home health examinations were arranged while the field teams were working in the subject's home municipality, but especially towards the end of the health examinations the home-visit nurses often had to travel quite far from their field teams. However as long as the field teams continued with their work, the home-visit nurses based themselves at the field team's current location. For instance, they would routinely deliver their blood samples to the field team for freezing usually on the same day.

By the time that the health examinations proper ended in early March 2001, 165 out of a total schedule of 416 home health examinations had been completed. The home health examinations continued until the middle of June 2001. When the field teams were finished with their work, new offices were arranged for the home-visit nurses in each of the country's five cities with university hospitals, providing them with the facilities they needed to store their equipment, freeze the blood samples and try to contact their subjects.

Ordinarily the home health examinations were conducted by one home-visit nurse, where necessary by a team of two. The nurses had the use of a car with which

they carried the 20 kg worth of equipment they needed (see Table 9.2). During the course of one day, the home-visit nurses managed to do 1–4 visits. Much of their time was taken up by trying to contact the subjects by phone, letter and personal visits if all other methods of contact failed: the nurses made up to several dozens of attempts to contact people who did not show up at the health examination. Travel to and from the subjects taking the home health examination also took up a great deal of time.

It can safely be estimated that the costs of one home health visit were somewhat higher than the costs of one health examination proper, even though it included a smaller range of tests and measurements. On the other hand, the home health examinations did manage to reach a very large number of people who on average were in very poor health and who otherwise would not have taken part in the survey. It is by virtue of the home health examinations that the data compiled in the Health 2000 Survey are highly representative of the population: they reached a substantial proportion of people who declined or who were unable to take part in the health examination proper for reasons of restricted mobility (Sainio et al. 2006) and other health reasons. The home health examinations were also successful in the sense that they involved no hazardous situations or any other problems.

Table 9.1. Stages of data collection in the home health examination (see also health examination proper, Chapter 8).

Duration (min)	
10	Information to subject (oral, written, informed consent form) and installation of IT equipment
10	Registration point basic data
10	Measurement of blood pressure
7	Laboratory: extraction and processing of blood sample (30 ml)
3	Oral examination: edentulousness, dentures, number of teeth
20	Snack + home-visit version of questionnaire 2 + where necessary short questionnaire 1 (where necessary completion of questionnaires or missing questionnaire items by interview) and installation of equipment for functional capacity measurements
either	
5–15	Check questionnaires 1 and 2
or	
40–60	Interview (short home-visit interview and interviews based on home-visit versions of questionnaires 1 and 2, unless home-visit interview conducted and questionnaires 1 and 2 not completed)
10	Home-visit version of symptoms interview
10	Measurements: height, weight, waist and hip circumference, PEF
40	Measurements of functional capacity: vision (no scotopic vision), hearing, cognitive capacity, reaction and movement time, hand grip strength, balance (without force platform according to Guralnik); in addition, a joint function test, chair stand test and walking speed test for subjects aged 55 years or over
10	Final information: handing out of feedback form, questionnaire 3 and dietary questionnaire, review of any further examinations necessary
10	Re-assembly of equipment

150 minutes (190 min, if including short home-visit interview and short questionnaire 1)

Table 9.2. Equipment used in the home health examination.

Mobile phone, contact information (for field team, Statistics Finland interviewers etc.)	
Equipment suitcase, including	
– laptop computer, mouse and mouse pad, battery charger, power cable, standby battery	
Office supplies etc.:	
– ballpoint pens (including Health 2000 pen for subject), pencil, eraser	
– scissors, blue tac, coloured masking tape, string	
– 2 extension cords (3 m and 5 m)	
– disinfectant, first aid kit	
Forms and sticker labels: personalised file for subject including	
– home visit programme	T2078
– home health examination form	T2074
– information note for participants	T2058
– informed consent forms (2)	T2050
– written informed consent given on behalf of subject (2)	T2053
– short home interview (in Finnish and Swedish)	T2075
– MMSE drawing task related to the above	T2007
– short questionnaire 1 (in Finnish and Swedish)	T2070
– CERAD form	T2026
– home-visit version of questionnaire 2 (in Finnish and Swedish)	T2072
– home-visit version of symptoms interview (in Finnish and Swedish)	T2071
– home-visit feedback form to subject (2)	T2076
– home-visit laboratory sampling form	T2073
– questionnaire 3 + dietary questionnaire + SAE envelope	T2005, T2006
– sticker labels with subject's name and examination code	
– subject's laboratory stickers	
Home interview file, including	
– display cards (in Finnish and Swedish)	
– interview instructions for Statistics Finland interviewers	
Blood pressure and anthropometry instruments	
– blood pressure gauge + stopwatch	
– measuring tape, metallic measuring tape (8–10 m), and 4 m and 6.1 m measuring strings	
– scales	
– PEF measuring device + equipment	
Blood sampling equipment	
Equipment for oral examination	
– torch, spatulas	
Snack	
– filled roll + fruit + paper plate	
– juice + disposable cup	
Equipment for functional capacity measurements	
– near vision chart + laminated blow-up	
– distant vision chart + pointer	
– spot light + spare bulb	
– audiometer + headset + press button + transformer + batteries	
– book of CERAD tasks in Finnish and Swedish	
– reaction time gauge and transformer (10 V)	
– laminated instruction board for balance test and focusing point	

10. PARTICIPATION, COMPLEMENTARY DATA COLLECTION AND OTHER MEASURES TO INCREASE PARTICIPATION RATE

Seppo Koskinen, Johanna Laiho, Sirkka Rinne, Noora Kuosmanen, Pirkko Alha, Harri Rissanen and Tuija Martelin

Participation in epidemiological health surveys varies considerably. Sometimes participation is confined to just a small minority of the sample population, sometimes data are obtained from virtually everyone in the sample. A high proportion of non-participants will undermine the reliability of the results, chiefly because the characteristics of this group are likely to differ from those of the participants. This means that the results on the prevalence or level of the phenomena studied do not necessarily provide a true picture of the situation in the study population and the results cannot be generalised to that population. Furthermore, non- participation may distort the results on how the phenomena concerned vary from one population group to the next, if non-participation and/or its associations with the phenomena studied varies between population groups.

The size of the bias caused by non-response or non-participation varies widely depending on the phenomenon concerned and on the methods of data collection. For instance, a survey on the prevalence of homelessness cannot produce reliable results by sending a postal questionnaire to the most recent known address of the people selected into the sample. People's willingness to participate may also be influenced by other living conditions factors Evidence on the associations of non-response with lifestyle factors is strongest for smoking: smoking is consistently more common among non-participants than among participants (e.g. Barchielli and Balzi 2002, Brøgger et al. 2003). Based on earlier surveys it seems that another important factor in this respect is mental health. For example, in a Norwegian survey on the population of Tromsø, register sources indicated that severe psychiatric disorders were 2.5 times more prevalent among non-participants than among participants (Hansen et al. 2001). Many other health problems also tend to be more common among non-participants. Mobility problems inhibit participation especially in health surveys outside the home, but if the data are collected in the subjects' homes, mobility does not seem to have a major impact on participation (e.g. Aromaa et al. 1989a, Sainio et al. 2006). Mortality data confirm that participation in health surveys is associated with health problems. Follow-ups in different datasets have shown that mortality among non-participants is about twice as high as among participants, but the excess mortality of non-participants varies according to cause of death (e.g. Hara et al. 2002, Jousilahti et al. 2005).

One of the main aims of Health 2000 was to produce as reliable assessments as possible of the prevalence of various health problems and associated factors in the adult population in Finland. Therefore, every possible effort was made to collect data from as many people in the survey sample as possible. Another motivation for minimising non-participation was that response rates in the Mini-Finland Survey 20 years earlier were extremely high (96% in the interview and 90% in the health examination, see Aromaa et al. 1989a), and our purpose was to compare the two sets of results.

10.1. Strategies to maximise participation

A major concern at all stages of planning and implementation of the Health 2000 Survey was to obtain as comprehensive and reliable data on as many people in the sample as possible. Crucial to achieving this goal were the information efforts both ahead of and during the field stages of the project; the training provided to field examination staff on how to minimise non-participation; and the steps taken to facilitate participation in the home-visit interview and the subsequent health examination. If in spite of all the motivation and persuasion the subjects still did not want to or were unable to take part in the health examination proper, they were offered the option of an abridged interview and/or health examination at home or at the institution where they lived. If this was not possible, the team tried to collect key interview and basic questionnaire (questionnaire 1) data by telephone. If the subjects could not be contacted at all or if for some other reason they did not take part in any of the above data collections, they were sent a brief postal questionnaire covering the same items as the telephone interview and a return envelope. These processes and the respective forms are described in more detail below.

10.1.1. Communications

Health 2000 received extensive exposure in both the national and local media shortly before the launch of the project and its field examinations. A major point of emphasis in the stories and interviews published was that the data collected in the survey would be hugely important to gaining a reliable and up-to-date picture of the nation's health and related factors. It was also stressed that the data collected would be used in developing health care services and preventive measures so that they better meet people's needs. Each field team assumed responsibility for the provision of information locally and issued press releases ahead of the start-up of the field examination. In many areas local journalists came to see how the health examinations were organised and reported the findings to their audience. A more detailed account of project communications is provided in Chapter 20.

10.1.2. Training

It was made clear in the training and written guidelines for Statistics Finland interviewers and field staff that all the people involved in the project were expected to contribute to motivating and persuading all people included in the sample to take part in the study. The guidelines prepared for the interviewers offered the following arguments to help them in these efforts:

1) A more detailed health examination could not be obtained anywhere else: participants would receive, for instance, an ECG and a doctor's assessment of its results, a complete dental examination including x-rays, the results of respiratory function tests, estimates of body fat percentage, eyesight and hearing tests, the results of a wide range of physical functional capacity tests, a detailed medical examination and an appointment with a doctor to talk through the results of their tests.

2) The data collected in the home-visit interviews are an important part of the survey and provide the basis for the health examination, and particularly for the medical examination.

3) The data obtained from every individual in the sample are important so that services and benefits can be properly allocated according to people's true needs. The survey data will also help to inform decisions on the future development of health care and other services in Finland. In order that these decisions can be based on accurate information, it is important that these data are obtained from all the people who have been recruited into the sample.

4) The survey is primarily funded from the public purse, and to make sure that it has the greatest possible benefit to both individual subjects and to the whole population it is important that everyone selected into the sample can be examined.

5) If the interviewee is unable to travel to the health examination because of restricted mobility or some other handicap, KTL will reimburse their taxi fare to and from the clinic. In special cases a home-visit nurse may even be able to pick up the subjects from their home.

Training also covered different aspects related to making appointments and conducting the interviews and health examinations with a view to facilitating the decision to participate.

10.1.3. Implementation of interviews and health examinations

To make it easier for employed persons to participate, the project team persuaded employer associations to recommend to their members that people in their employ be allowed to participate in the field examination during working hours. In addition, wherever possible, employed persons were given the opportunity to take the interview and health examination after office hours. Every effort was made to find a suitable time slot for each subject. The same flexibility was applied to the choice of interview location: where possible the interviews were arranged at the subject's home (or for institutionalised persons at the institution where they lived), but they could also be done at the health centre, library, or other place that was most convenient to the interviewee. If necessary the interviewers would try to contact the subjects several times both by letter and by phone in order to make an appointment for the interview. If they failed to contact the interviewee at home, they tried other places. If the subject refused to participate, another interviewer would take over the case and try again. Chapter 6 provides a more detailed description of the implementation of interviews.

Upon completing the interview, the Statistics Finland interviewer made an appointment for the interviewee to take the health examination. If none of the time-slots allocated to the interviewer suited the subject, the interviewer phoned the appointments centre to find a slot that did. The subjects could also change their appointments themselves by phoning the appointments centre; the number was given to each interviewee. Where necessary the subjects could also take the health examination outside their own place of residence, if the same or some other field team visited a location that was more easily accessible to the subject. Persons who did not take the interview or who for some other reason did not have an appointment for the health examination received an invitation by mail and/or were contacted from the appointments centre. In other words, every person in the sample was invited to the health examination. If the subject was unable otherwise to get to the clinic where the health examination was held, their taxi fare was reimbursed or they were picked up by a member of the health examination team, usually a home-visit nurse.

All persons who did not turn up at the health examination on time were contacted by the field team, either on the day of the appointment or as soon as possible. At the same time form T2044 was completed. The data recorded on this non-participation form included the subject's contact information, date and time of health examination appointment and home-visit interview (if taken) as well as other data relevant to participation, such as the new appointment made for the health examination proper or home health examination, the completion of a telephone interview or information on the subject's death. For subjects who refused to attend

the health examination, the reason for their refusal was recorded. For those who were not contacted, a plan of action was entered in the form on how efforts to reach the person would be continued. The form for non-participants was updated as efforts to contact them were continued; spaces were provided for data on seven attempts to reach them. If necessary a second form was used. When the field teams closed down, the forms for non-participants were handed over to the examination nurses who continued with the home visits in that area, unless the subject had by that time taken the examination or refused to take part. The health examination appointments procedure is described in more detail in Chapter 5.4.

If the subjects did not want to or were unable to travel to the health examination, they were offered an alternative, slightly abridged health examination at home or at the institution where they lived. Each field team included two nurses who had received special training to conduct home health examinations. If the person concerned had not taken the interview, a home interview was taken that was slightly shorter than the home interview proper (form T2075). The home health examinations were started when the field teams had got to work and they were continued after the completion of the health examinations through to the end of June 2001. Where possible the subjects were contacted in advance to make an appointment for home health examinations, but if the nurse failed to contact the subject, they would nevertheless visit them at the time indicated in the letter sent out in advance – and where necessary and possible at a later time as well if the subject was not at home during the first visit. The content and implementation of the home health examinations is described in detail in Chapter 9.

As full data as possible were collected on the subjects who took part in the examination. The interview and large parts of the health examinations were computer-aided, and responses were obtained on as many items as possible. The basic questionnaire handed out in connection with the interview and the infection questionnaire given at the beginning of the health examination were checked at the health examination and any missing items were completed. If the subject was unable to answer, factual data were obtained from a family member or nurse, but questions involving the subject's opinion or interpretation were not asked. Similarly, in the health examination measurements or tests were not taken if it was thought that they might be dangerous to the subject or if it was clear they would not be able to complete them. In these cases, the reason for not taking the test was recorded.

10.1.4. Telephone interview

If the subjects refused to take the interview and the health examination, either at the clinic or at home, they were contacted by phone to try and get answers to

key interview and basic questionnaire items (telephone interview form T2077). In the telephone interview, the subjects were given a further opportunity to take the health examination proper or the home health examination. Only the telephone interview items drawn from the basic questionnaire were presented to those telephone interviewees who had participated in the home interview but who had not attended the health examination or returned the basic questionnaire. The telephone interviews were taken both by home-visit nurses and by staff working at the appointments centre.

10.1.5. Questionnaire for non-participants

In August 2001, a post-questionnaire (containing the same items as the telephone interview) was mailed to all those subjects who had not been contacted and who refused to take part in all other stages of the examination. This form (T2095) contained items from the home interview and basic questionnaire that were considered the most important. A brief covering letter was attached in which the recipients were encouraged to complete the form; a stamped and addressed return envelope was also attached. The covering letter said among other things that by completing the questionnaire the respondents can contribute to the development of the national health care system, that the questions are easy to answer and will only take a short time, and that all the information provided would be treated in strictest confidence.

10.2. How did efforts to maximise participation succeed?

10.2.1. Whole sample

The survey overall was highly successful. Almost 89% of the people in the final sample took part in the interview (short or long version of the home-visit interview), almost 85% attended the health examination (at a clinic or at home) and at least some information was obtained from 93% of the people recruited in the sample (Table 10.1). The figures were even higher in the Mini-Finland Survey 20 years ago (Aromaa et al. 1989a), but in other Finnish and in all foreign health surveys participation rates have been much lower (Aromaa et al. 2003a,b). The participation rates were high at all stages of the survey (Table 10.1) and in all age groups (Table 10.2).

The success of the interviews was mainly attributable to Statistics Finland's well-honed interview organisation. Most interviewers were highly experienced and had

taken part in the data collection for the earlier Finnish health care survey (Arinen et al. 1998), which included many of the same elements as Health 2000. Success was achieved even though the interview was longer than normal and even though the training provided to interviewers in Health 2000 was limited to just one day and 31 pages of written guidelines. Special attention was given in the survey to supporting the efforts of interviewers out in the field. In addition, the interviewers had the opportunity to take the health examination themselves, which may have served to strengthen their commitment to the data collection process. Furthermore, this first-hand experience of the examinations will have given them a firmer background in motivating people to take part. Another factor contributing to the overall success of the project was the interesting subject-matter and the opportunity to have a full health examination. Furthermore, the good reception and positive coverage of the project in the media and the Health 2000 brochure that was mailed to all people in the sample probably had the effect of encouraging participation.

In connection with the home interviews conducted by Statistics Finland interviewers, the subjects were informed about the health examination and appointments were made; if necessary those appointments could later be easily changed. This probably helped to increase participation rates. Home health examinations were crucial to increasing participation among older people in particular: visits were made to homes and to institutions to attend to those people who were unable to get to the health examination proper. Even in the age group 85 or over, around 80% of the sample took part in the interview and around 70% attended the health examination. It was largely by virtue of the home health examinations that participation rates in older age groups were higher than was the case earlier in the Mini-Finland Survey. There were some difficulties with the interviews and examinations conducted at institutions, which were partly due to staff at these institutions wanting to be rather cautious.

Table 10.1. Original sample, final sample, participation in different stages of data collection and non-participation.

	Number		%	
Sample	8,028			
deceased before fieldwork		49		
Final sample	7,979		100.0	
Participants in home-visit interview	7,087		88.8	
long interview		6,986		87.6
short interview		101		1.3
Participants in health examination	6,354		79.6	
symptoms interview		6,238		78.2
measurements: measurement point 1		6,351		79.6
measurement point 2		6,339		79.4
laboratory		6,354		79.6
oral examination		6,335		79.4
functional capacity measurements		6,329		79.3
clinical examination		6,326		79.3
mental health interview		6,005		75.3
Participants in home-visit health examination instead of health examination proper [1]	417		5.2	
Questionnaire respondents [2]				
basic questionnaire (questionnaire 1) [2]		6,736		84.4
infection questionnaire (questionnaire 2) [2]		6,734		84.4
complementary questionnaire (questionnaire 3)		6,269		78.6
dietary questionnaire [3]		5,998		75.2
Participants in telephone interview [4] or post-questionnaire	306		3.8	
telephone interview[4]		243		3.0
post-questionnaire		63		0.8
Participants in any stage of data collection[5]	7,415		92.9	
Non-participation	564		7.1	
refused		451		5.4
abroad		30		0.4
not reached		68		1.1
other reasons		15		0.2

[1] A total of 417 home visits were made. Home visit measurements consisted mainly of those taken at measurement point 1 and functional capacity measurements. Blood samples were also taken. The number of symptoms interviews completed (short version) was 393.

[2] Includes abridged (short) versions of questionnaires.

[3] Population weight calculated for 6,005 persons.

[4] Of the 892 persons in the final sample who did not take the home interview, 243 took the telephone interview. In addition, 211 persons who took part in the home-visit interview but who did not attend the health examination, provided responses to four key items inquired in the health examination through the short telephone interview.

[5] Participants in the home-visit interview (7,087) or in the telephone interview and post-questionnaire (306) and 22 persons who only took the health examination or who returned some questionnaire.

Table 10.2. Participation in different stages of data collection by sex and age.

	Final sample	Interview (short or long)		Health examination (at clinic or at home)		Telephone interview or post-questionnaire	
	number	number	%	number	%	number	%
Men							
30–44	1,276	1,075	84.2	1,018	79.8	66	5.2
45–54	973	848	87.2	826	84.9	40	4.1
55–64	618	555	89.8	527	85.3	26	4.2
65–74	432	398	92.1	379	87.7	16	3.7
75–84	236	214	90.7	203	86.0	6	2.5
85–	79	72	91.1	58	73.4	2	2.5
Total	3,614	3,162	87.5	3,011	83.3	156	4.3
Women							
30–44	1 322	1,185	89.6	1,148	86.8	37	2.8
45–54	943	863	91.5	843	89.4	27	2.9
55–64	703	645	91.7	634	90.2	20	2.8
65–74	551	498	90.4	478	86.8	18	3.3
75–84	557	485	87.1	448	80.4	31	5.6
85–	289	249	86.2	209	72.3	17	5.9
Total	4,365	3,925	89.9	3,760	86.1	150	3.4
Both sexes							
30–44	2,598	2,260	87.0	2,166	83.4	103	4.0
45–54	1,916	1,711	89.3	1,669	87.1	67	3.5
55–64	1,321	1,200	90.8	1,161	87.9	46	3.5
65–74	983	896	91.1	857	87.2	34	3.5
75–84	793	699	88.1	651	82.1	37	4.7
85–	368	321	87.2	267	72.6	19	5.2
Total	7,979	7,087	88.8	6,771	84.9	306	3.8

10.2.2. Participation in different population groups

A separate report by Statistics Finland on the statistical quality of the interview material collected for the Health 2000 Survey provides a detailed description of the variation in non-response to the home interview survey by sociodemographic factors (Laiho 2004). In the present report, this analysis is extended to cover non-participation in the health examination (Table 10.3). The sociodemographic data for this analysis are based on Statistics Finland register sources. The focus here is on unweighted non-response rates, which provides a more accurate account of the success of the fieldwork, but the results cannot be directly generalised to the target population and the representativity of the sample material.

A slightly higher proportion of men (8%) than women (6%) did not take part in any stages of the survey. The proportion of non-participants in the health interview was also slightly higher among men (14%) than women (12%). Participation in the

health examination proper did not vary by sex, but women took part in the home health examination somewhat more often than men.

Participation by age varied among women and men. Among men aged 50 or over, only 5–7% did not take part in any stage of the survey, while among younger men nearly one in ten did not participate. Among women, those in the age bracket 40–69 were the most active participants: in this age group only 5–6% did not take part in any stage of the survey, whereas in younger and older age groups the proportion of non-participants was 7–8%. The proportion of non-respondents in the health interview was also highest among men under 50 (16%) and among women over 70 (16–17%). Among women the most active participants in the health examination were found in the age group 40–69, among men in the age group 50–79. Among women aged 80 or over, no more than around half and among men almost two-thirds took part in the health examination proper, but even in the oldest age group almost 80% took part either in the shorter home-visit examination or in the health examination proper.

Finnish-speakers were the most active participants. Among Swedish-speakers the participation rate was somewhat lower, and it was lowest of all among the women and particularly among the men whose mother tongue was other than Finnish, Swedish or Sami.

By place of residence, non-response varied in the same way among both men and women. The participation rate was highest among people living in the catchment area of Oulu university hospital, and non-response was highest in the Helsinki and Turku areas. People living in rural areas took part in the survey somewhat more actively than those living in urban centres.

Variation by socioeconomic status was otherwise rather insignificant, but among men non-participation was higher in students and in the categories of "others" and "unknown". Farmers and upper-level employees were the most active participants.

Unemployment did not significantly impact the participation of women, but among men a much higher proportion of those who had been out of work for most of the year failed to participate as compared to those who were employed or who had been unemployed for shorter spells. A comparatively large proportion of the small number of male students also failed to participate. The same was true of both men and women for whom no data were available on their main activity. Among retired women, a very large proportion did not take part in the health examination proper, but almost half of them were reached through the home health examination.

Only minor differences were seen by level of education. Failing to participate in any stage of the examination was most common in the highest and lowest education groups, as was non-participation in the health (home-visit) interview. Non-

participation in the health examination was highest in the group of respondents who had the least schooling. The participation rate was lowest in the lowest income category. Among men in particular, those with low incomes were less keen to participate.

Married people were the most active participants, but even in the group of those living in common-law marriage non-participation was at a much lower level than among those who had no family, among children living with their parents and among those whose family status was unknown. By size of household-dwelling unit, non-response was highest among people who lived alone. Non-response declined with increasing size of household-dwelling unit among men to five and among women to six persons.

Apart from sociodemographic characteristics, many other factors also impact survey participation, above all health and functional capacity. In particular, participation in health surveys is markedly lower among people with severe mental health problems. Mobility restrictions (Sainio et al. 2006) and reduced eyesight (Laitinen et al. 2005) reduce participation especially in health examinations.

Table 10.3. Percentage of non-participants among persons aged 30 or over by population group.[1]

	Non-participant in any stage		Non-participant in long interview		Non-participant in health examination proper or home health examination		Non-participant in health examination proper		Number in sample	
	Women	Men	Women	Men	Women	Men	Women	Men	Women	Men
Number	279	285	505	488	605	603	887	738	4,365	3,614
Percentage	6.4	7.9	11.6	13.5	13.9	16.7	20.3	20.4		
Age										
30–39	7.7	9.8	12.0	15.9	14.5	19.8	16.8	21.8	911	864
40–49	5.7	9.6	9.7	15.8	11.1	18.1	13.6	20.7	937	914
50–59	5.8	7.2	8.8	12.6	9.8	15.0	11.9	17.7	807	816
60–69	4.9	4.6	8.6	11.0	9.6	13.3	12.1	16.2	594	519
70–79	7.9	4.9	16.7	9.5	18.8	11.8	27.6	20.3	522	305
80–	6.6	7.1	16.2	9.2	22.7	20.4	49.7	36.2	594	196
Mother tongue										
Finnish or Sami	6.1	7.4	11.1	13.0	13.5	16.1	19.7	19.7	4,027	3,373
Swedish	7.8	9.0	15.0	15.3	16.0	20.6	26.3	25.4	281	189
Other	21.1	34.6	26.3	40.4	29.8	46.2	33.3	51.9	57	52
University hospital district										
Helsinki	8.6	10.7	14.4	17.1	17.0	19.9	22.2	23.8	1,419	1,197
Tampere	3.6	8.0	8.2	12.9	12.3	15.4	16.6	17.6	609	488
Turku	6.6	7.9	13.6	14.6	13.8	17.9	22.2	21.7	1,006	814
Kuopio	5.9	6.2	9.1	11.9	12.4	15.3	20.6	18.7	759	615
Oulu	4.2	3.2	7.9	5.6	9.8	10.0	16.1	15.0	572	500
Municipal classification (Stakes)										
Cities/Urban centres	7.8	9.4	13.3	16.2	16.0	19.1	21.9	22.9	2,383	1,837
Urban municipality	8.1	6.1	11.5	12.8	14.4	15.9	19.5	19.5	174	164
Urban-adjacent rural area	5.1	6.9	10.5	11.6	12.2	15.3	19.0	19.1	949	810
Core rural area	4.1	6.4	9.0	10.9	9.6	13.4	19.9	16.8	543	469
Sparsely populated rural area	3.1	4.8	5.9	7.5	9.5	12.0	14.2	15.6	325	334
Socio-economic status										
Farmers	6.6	5.2	8.2	7.8	8.2	10.3	11.5	11.2	61	116
Other entrepreneurs	7.3	4.3	9.3	10.6	12.0	15.4	13.3	18.3	150	208

	Non-participant in any stage		Non-participant in long interview		Non-participant in health examination proper or home health examination		Non-participant in health examination proper		Number in sample	
	Women	Men	Women	Men	Women	Men	Women	Men	Women	Men
Upper-level employees	4.3	7.7	8.9	12.5	9.1	12.9	10.1	14.3	438	441
Lower-level employees	5.6	6.4	8.6	10.7	10.6	12.9	11.8	14.3	979	419
Manual workers	6.6	9.0	10.5	15.2	11.1	18.1	14.3	19.7	504	869
Students	3.8	22.9	7.6	25.7	15.1	34.3	20.8	37.1	53	35
Pensioners	6.6	5.7	14.1	11.0	16.8	13.8	29.9	21.4	1,777	1,151
Others	7.7	13.4	11.9	21.2	14.1	29.3	17.0	33.2	312	283
Unknown	15.4	21.7	19.8	31.5	35.2	40.2	36.3	42.4	91	92
Main activity										
Employed	5.5	7.2	8.8	12.7	10.1	15.0	11.6	16.6	2,078	2,076
Unemployed	7.5	13.5	11.5	21.3	13.8	29.4	16.4	33.3	305	282
Students	3.8	18.2	7.6	21.2	15.1	30.3	20.8	33.3	53	33
Pensioners	6.5	5.7	14.3	11.2	17.0	13.9	30.7	22.0	1,704	1,090
Others outside the labour force	3.2	1.8	4.8	5.4	6.4	7.1	7.9	7.1	63	56
Unknown	16.7	36.4	22.2	41.6	32.1	55.8	35.2	57.1	162	77
No. of weeks unemployed										
0	6.2	7.2	11.5	12.7	13.8	15.4	20.7	19.2	3,811	3,141
1–26	7.7	10.0	12.2	15.2	13.9	19.0	17.1	20.9	287	211
27–52	7.5	14.1	11.6	22.1	15.4	29.8	18.0	34.4	267	262
Education										
Basic level or no education	7.6	9.2	13.8	14.8	17.1	19.2	27.9	25.0	1,905	1,451
Intermediate level	5.2	6.6	9.3	13.3	11.5	16.2	14.9	18.7	1,305	1,245
Lowest-level tertiary education	5.5	6.6	10.1	11.0	11.0	13.6	13.8	15.4	711	455
Lower-degree level tertiary education	4.5	6.8	10.8	10.4	10.1	11.5	12.9	12.5	178	192
Higher-degree level tertiary education	7.5	9.6	11.7	14.0	12.4	14.0	14.7	17.7	266	271
Disposable income*										
I (lowest income) decile	9.7	13.3	16.0	20.0	19.5	24.0	28.9	31.1	318	225
II	5.3	8.9	12.4	14.5	14.6	20.6	26.8	24.9	507	325
III	6.2	10.1	11.8	16.4	14.1	21.4	20.8	26.1	467	318
IV	5.9	6.9	11.5	14.1	13.9	15.6	22.2	20.7	460	334
V	6.7	5.9	12.8	10.9	14.3	14.9	21.0	16.5	405	357
VI	6.2	4.5	9.3	11.2	10.3	14.9	14.1	17.7	389	357
VII	3.7	9.3	9.1	13.7	10.5	15.7	14.4	17.5	438	343
VIII	6.8	5.8	12.2	13.2	15.4	14.9	17.6	18.2	410	395
IX	5.2	6.2	9.0	10.4	11.1	12.3	12.3	15.6	424	422
X (highest income) decile	7.1	7.3	10.0	11.6	10.8	12.4	13.5	14.4	452	467
Unknown	16.8	26.7	23.1	29.6	37.9	46.5	75.8	67.6	95	71
Family status										
No family	8.1	10.0	15.2	17.5	17.7	23.1	28.6	28.3	1,377	802
Married	4.9	5.8	8.8	10.3	10.2	12.2	13.2	15.1	2,402	2,191
Child	12.0	10.7	21.3	21.4	22.2	27.5	29.6	30.5	108	131
Cohabiting	5.6	9.4	10.3	15.6	14.9	18.4	18.0	20.4	377	403
Unknown	12.8	28.7	21.8	35.6	35.6	47.1	74.3	67.8	101	87
Size of household										
1 person	8.1	10.5	15.1	18.3	17.6	23.6	28.7	29.1	1,383	818
2	5.4	6.7	9.7	11.8	11.8	13.7	15.5	17.7	1,418	1,311
3	6.1	7.6	10.7	13.2	12.5	16.6	15.4	18.9	606	589
4	3.7	6.5	8.6	10.4	9.7	12.3	12.0	13.1	567	527
5	6.1	4.7	8.0	8.0	9.4	9.4	12.3	10.9	212	212
6	3.2	3.3	3.2	13.1	6.5	18.0	8.1	21.3	62	61
7 or more persons	4.6	4.0	13.6	16.0	18.2	16.0	27.3	16.0	22	25

[1] Figures are unweighted percentages describing the success of fieldwork, but unweighted figures cannot be generalised to the whole population.

* The deciles are based on household disposable income derived from register sources, which have been scaled with the number of consumption units in a household (OECD definition of consumption unit concept). The decile boundaries have been defined based on the target population's income data.

10.3. Study of young adults

The field survey of persons aged 18–29 was conducted after the interviews with the population aged 30 or over in spring and summer 2001. The survey included a home-visit interview, a basic questionnaire that was handed to the respondents in connection with the interview and a dietary questionnaire, which was mailed to the respondents after they had returned the basic questionnaire. The study of young adults is described in closer detail in Chapter 12. As in the case of the main survey with people over 30, special attention was given to maximising participation, which was supported and encouraged using the same methods as applicable. Almost all of the interviewers had taken part in data collection in the main survey and in the training provided ahead of that survey, and therefore they were well prepared to interview the young adults as well – even though the training remained rather short and cursory. The project received extensive coverage in both the national and local media before and during the field stages. Since the study of young adults did not include a health examination, that could not be used as an argument to motivate participation. Participation was encouraged by sending home interviewees a ticket for a draw which they were to return together with a completed questionnaire: prizes for the draw donated by sponsors of the project included a mobile phone, bicycle, vouchers and various other smaller prizes.

In autumn 2001, a post-questionnaire was mailed to all those who refused to take the interview and who were not contacted. This questionnaire included key items from the health interview and the basic questionnaire. Two cinema tickets were sent to all those who returned the post-questionnaire. The same incentive was used to encourage the completion of basic questionnaires by those who had taken part in the interview but who had not yet completed and returned the questionnaire. Reminders were sent out once. The basic questionnaires were returned by the end of 2001.

The study of young adults was another success. Health interviews were obtained from almost 80% and basic questionnaires from almost 70% of those in the sample. If the core data received through post-questionnaires are taken into account, the participation rate was 90% (Table 10.4). Participation was clearly higher than in any other recent Finnish health survey among young adults, let alone those conducted elsewhere in Europe (Aromaa et al. 2003a,b). It is particularly worth noting that just over half (205) of the 391 persons who did not take the interview completed and returned the 16-page post-questionnaire, which was based on core items from the interview and basic questionnaire.

Table 10.4. Sample of young adults aged 18–29, participation in different stages of data collection and non-participation.

	Number	%
Sample	1,894	
deceased before field survey	0	
Final sample	1,894	100.0
Participants in		
interview	1,503	79.4
basic questionnaire	1,282	67.7
dietary questionnaire	789	41.7
post-questionnaire	205	10.8
at least one of the above	1,710	90.3
Non-participation	184	9.7
refused	114	6.2
abroad	12	0.6
not contacted	55	2.9
other reasons	3	0.2

Participation in the survey was much higher among young women as compared to young men: 7% of women as opposed to 12% of men did not take part in any stage of the survey (Table 10.5). Among women differences between age groups were marginal, whereas among men non-participation was highest in the age group 25–29. Participation rates were roughly the same among Finnish and Swedish-speakers, but more than one-quarter of those with some other mother tongue did not take part in any stage of the survey.

Participation was highest in the catchment area of Kuopio university hospital, lowest in Helsinki and Tampere. Among men the participation rate was also higher than average in the Oulu area and lower in the Turku area. Non-participation was slightly higher in urban than in rural areas.

By socio-economic status, the participation rate was lowest in the small group of pensioners and in the categories of “others” and “unknown”. Participation among upper-level employees and students was slightly above average. Among unemployed men non-participation was clearly higher than among those who were employed, regardless of the duration of unemployment. Among women short-term unemployment did not increase non-participation, but among women who had been out of work for more than 6 of the 12 months preceding the survey, non-participation was high.

Among both women and men non-participation was highest among those with the lowest level of education, and lowest in the group with an academic degree. The participation rate was slightly lower than average in the lowest income brackets, but otherwise the differences between income groups were quite marginal.

Among women, differences by marital status were minor, whereas among men those who were married or co-habiting were the most active participants. Men who lived alone did not take part in the survey quite as actively as those who lived in larger household-dwelling units, but otherwise the size of household-dwelling unit did not show a strong correlation with participation rate.

Table 10.5. Percentage of non-participants among young adults aged 18–29 by population group.[1]

	Non-participant in any stage		Non-participant in health interview		Non-participant in basic questionnaire		Number in sample	
	Women	Men	Women	Men	Women	Men	Women	Men
Number	65	119	175	216	233	379	913	981
percentage	7.1	12.1	19.2	22.0	25.5	38.6		
Age								
18 – 19	5.2	11.4	20.7	22.7	24.1	36.9	174	176
20 – 24	6.4	9.4	19.5	18.2	23.7	37.3	375	413
25 – 29	8.8	15.3	18.1	25.8	28.0	40.8	364	392
Mother tongue								
Finnish or Sami	6.7	12.0	18.4	22.1	24.0	37.9	821	911
Swedish	3.1	8.5	20.3	17.0	28.1	34.0	64	47
Other	28.6	26.1	39.3	30.4	64.3	78.3	28	23
University hospital district								
Helsinki	9.3	14.8	24.1	28.0	30.7	46.7	332	332
Tampere	10.3	18.3	22.2	24.6	29.1	34.9	117	126
Turku	4.6	12.4	17.3	24.3	22.8	41.0	197	210
Kuopio	4.1	5.7	9.0	12.7	14.5	29.9	145	157
Oulu	5.7	7.7	18.0	13.5	25.4	30.1	122	156
Municipal classification (Stakes)								
Cities/Urban centres	8.3	12.3	22.1	24.4	27.8	41.5	596	578
Urban municipality	5.9	18.8	20.6	21.9	29.4	46.9	34	32
Urban-adjacent rural area	4.6	10.4	13.2	20.7	21.1	35.2	152	193
Core rural area	3.3	17.8	11.0	21.1	18.7	34.4	91	90
Sparsely populated rural area	9.3	6.8	16.3	10.2	20.9	28.4	43	88
Socio-economic status								
Farmers	0.0	0.0	0.0	0.0	0.0	5.3	1	19
Other entrepreneurs	6.3	18.5	25.0	22.2	31.3	59.3	16	27
Upper-level employees	6.5	7.3	12.9	11.6	19.4	27.5	62	69
Lower-level employees	4.8	8.2	15.4	21.1	21.9	38.8	228	147
Manual workers	5.5	10.9	20.4	23.4	28.4	40.3	201	320
Students	5.1	9.1	15.2	16.7	21.1	31.8	256	198
Pensioners	37.5	35.7	56.3	35.7	56.3	35.7	16	14
Others	9.9	19.2	26.4	30.1	34.1	46.6	91	146
Unknown	23.8	26.8	35.7	34.2	35.7	51.2	42	41
Main activity								
Employed	5.2	9.6	17.8	19.5	24.7	36.8	482	584
Unemployed	8.4	22.7	26.5	35.1	34.9	51.6	83	97
Students	4.4	9.1	14.4	16.8	20.0	31.5	250	197
Pensioners	50.0	30.8	70.0	30.8	70.0	30.8	10	13
Conscripts, civilian servicemen	.	4.7	.	14.0	.	30.2	0	43
Unknown	19.3	36.2	27.3	53.2	31.8	74.5	88	47
No. of weeks unemployed								
0	7.3	10.4	19.1	19.9	25.2	35.4	655	723
1-26	4.7	16.3	16.6	27.2	23.2	45.7	211	184
27-52	14.9	18.9	31.9	29.7	40.4	52.7	47	74

	Non-participant in any stage		Non-participant in health interview		Non-participant in basic questionnaire		Number in sample	
	Women	Men	Women	Men	Women	Men	Women	Men
Education								
Basic level or no education	9.4	17.6	25.1	29.5	32.3	47.0	235	302
Intermediate level	7.6	10.0	18.8	20.0	24.7	36.4	499	569
Lowest-level tertiary education	1.9	10.7	13.5	14.3	24.0	28.6	104	56
Lower-degree level tertiary education	2.4	6.7	9.5	13.3	9.5	33.3	42	30
Upper-degree level tertiary education	6.1	4.2	12.1	4.2	15.2	16.7	33	24
Disposable income*								
I (lowest income) decile	10.3	17.2	20.0	25.5	27.1	38.9	155	157
II	8.7	12.8	20.2	25.6	29.8	43.0	104	86
III	7.5	4.8	21.3	12.1	23.8	27.7	80	83
IV	2.3	12.9	18.6	20.0	23.3	35.3	86	85
V	7.8	11.2	22.2	20.4	27.8	39.8	90	98
VI	4.2	9.0	12.6	20.2	22.1	37.1	95	89
VII	6.7	18.6	18.0	29.4	19.1	51.0	89	102
VIII	3.9	6.8	13.0	18.8	19.5	35.0	77	117
IX	8.5	13.0	16.9	22.0	23.9	37.0	71	100
X (highest income) decile	1.8	6.7	26.3	20.0	35.1	36.7	57	60
Unknown	55.6	75.0	55.6	75.0	66.7	100.0	9	4
Family status								
No family	6.4	16.5	18.9	27.0	24.5	41.6	233	269
Married	6.3	9.9	18.1	20.7	24.9	38.6	365	435
Child	0.0	11.5	0.0	23.1	0.0	30.8	2	26
Cohabiting	6.6	9.1	18.9	16.5	25.3	33.7	301	243
Unknown	58.3	70.0	66.7	80.0	75.0	100.0	12	10
Size of household								
1 person	7.2	17.6	19.9	28.2	25.4	42.9	236	273
2	7.0	8.5	20.7	17.3	25.9	36.5	328	260
3	8.1	10.3	20.5	20.6	26.7	37.4	161	214
4	3.6	8.2	14.4	19.0	24.3	36.7	111	158
5	0.0	14.6	4.1	25.0	14.3	39.6	49	48
6	33.3	20.0	33.3	30.0	33.3	30.0	9	10
7 or more persons	0.0	14.3	10.0	14.3	20.0	21.4	10	14

[1] Figures are unweighted percentages describing the success of fieldwork, but unweighted figures cannot be generalised to the whole population.

* The deciles area based on household disposable income derived from register sources, which have been scaled with the number of consumption units (OECD definition of consumption unit concept). The decile boundaries have been defined based on the target population's income data.

10.4. Dataset for the Mini-Finland resurvey

The participants of the Mini-Finland Health Survey in 1978–1980 who in November 2000 were alive and lived in Helsinki, Turku, Salo, Lahti, Tampere, Kuopio or Oulu, received an invitation to a health examination in spring 2001. They were also sent a questionnaire that they were asked to complete and bring along to the health examination. The examination included almost all the same items as the full Health 2000 programme. In connection with the examination a short home interview was conducted (form T2075). The content and implementation of the resurvey is described in more detail in Chapter 11.

As in other components of the Health 2000 Survey, the aim in the resurvey was to maximise participation. The invitation letter encouraged the subjects to take part. A few days after the invitation letter had been sent out, the subjects were phoned to ask whether they could make the appointment. If the time was not suitable, a new appointment was arranged. Free taxi vouchers were offered to those who otherwise would have difficulty travelling to the health examination. If the subject was unable or unwilling to take part in the health examination, a telephone interview was conducted if possible (form T2077). If the subject refused altogether, the reason for non-participation was recorded. As in the Health 2000 Survey, a post-questionnaire (form T2095) containing the same items as the telephone interview was sent to all those who had not been contacted and to all others who had refused to take part in other stages of the examination (see Chapter 10.1.5 above).

The Mini-Finland resurvey was also highly successful. Interview and questionnaire data were obtained from almost 80% of the subjects. Similarly, 80% participated in the health examination proper or in the home health examination. In addition, almost 9% took part in the telephone interview, or returned the post-questionnaire. Just 11% of the sample gave no information at all (Table 10.6).

Table 10.6. Sample invited to take part in the Mini-Finland resurvey, participation in different stages of data collection and non-participation.

	Number	%
Sample	1,278	
deceased before field survey	8	
Final sample	1,270	100.0
Participants in		
interview	986	77.6
basic questionnaire	994	78.3
health examination proper	923	72.7
home health examination	95	7.5
telephone interview	86	6.8
post-questionnaire	25	2.0
at least one of the above	1,130	89.0
Non-participation	140	11.0
not invited to the survey	18	1.4
refused	80	6.3
abroad	0	-
not contacted	42	3.3

10.5. Conclusions

One of the key considerations in the planning and implementation of the Health 2000 Survey was to maximise participation. This did have the effect of driving up the survey costs somewhat, but on the other hand it also helped to achieve a uniquely high participation rate. As a result, the results provide an exceptionally reliable picture of the health of the population and related factors. Nevertheless it is important to bear in mind that even at low levels, non-response may significantly distort the results when the phenomenon studied is closely associated with participation. For this reason, it is important to use register data to study to what extent the health of the subjects differs from the health of non-participants both at the time of the field examination and after that examination.

11. RESURVEY OF PERSONS WHO PARTICIPATED IN THE MINI-FINLAND SURVEY

Sami Heistaro and Seppo Koskinen

One part of the Health 2000 project was a resurvey of persons who had taken part in the Mini-Finland health examination 20 years earlier. A total of 1,278 participants from that earlier examination were recruited to take part. These persons took a health examination (Appendix 5) that was very similar to the examinations in the Health 2000 Survey, but somewhat shorter. The Statistics Finland interviewers were not involved in this part of the project, but only a brief health interview was carried out in connection with the health examination.

Table 11.1. Sample of Mini-Finland resurvey and participations by region.

Region	Time (2001)	Sample
Helsinki	27.02.–30.03.	657
Kuopio	08.01.–11.01.	57
Lahti	01.02.–07.02.	89
Oulu	05.02.–19.02.	82
Salo	03.01.–05.01.	68
Tampere	13.02.–23.02.	173
Turku	25.01.–07.02.	160
Total		1,278

12. STUDY OF YOUNG ADULTS

Seppo Koskinen, Tuija Martelin, Laura Kestilä and Arpo Aromaa

12.1. Introduction

Young adults – those who have come of age but who are still living out their youth – remain consistently out focus in national health surveys of the adult population. This may be due to the relative absence of chronic diseases at this stage of life and the consequent lack of interest in the health of young adults; it is just not considered equally important as the health of older people. The scarcity of research is also explained by the inherent difficulty of studying the associations between different background factors and health in young adults. In many areas this stage of life is still characterised by constant change, and most young adults will not yet have reached their final educational level, socio-economic status or family structure, for instance.

Yet health and health problems are just as important at this stage of life as they are in others. Furthermore, health in later adulthood is largely determined by factors from earlier stages of life. Health differentials between population groups can also be traced back to early behavioural and environmental factors: it is during young adulthood that many important environmental and behavioural traits are established that are crucial to future health. Indeed information about health, its determinants and factors shaping these determinants at this age are crucially important to the promotion of health in the whole adult population.

There is only very limited up-to-date research knowledge on health and related factors in young adults, particularly in the age group under 25. In Finland the Social Insurance Institution (SII, Kela) has conducted representative interview studies on health security at irregular intervals since 1964 (Arinen et al. 1998). The National Public Health Institute (KTL) has compiled information on the health and health behaviour of the population of working age in annual postal questionnaires among people aged 15–64 (see e.g. Helakorpi et al. 2003) as well as in its five-yearly FINRISK surveys in the population aged 25–64 (Vartiainen et al. 2003). Statistics Finland's living conditions surveys (Huuhka et al. 1996) also include data on the health of Finnish people aged 15 or over. Furthermore, interview studies conducted once every three years on employment and health (Piirainen et al. 2003, Kauppinen et al. 2004) have compiled information on health-related factors in the population over 25.

A major difficulty with health surveys in young adults is the high level of non-response. In KTL's annual postal questionnaires, for instance, the sample is

representative of the population, but the picture of the health of young males aged 15–24 in particular may well be distorted since no more than 60% take part. Similarly, the health questionnaire among university students in 2000 achieved a response rate among men of just over 50% (Kunttu and Huttunen 2001). More information is available on people over 25 and in the age group 11–18 (Rimpelä et al. 2003, Currie et al. 2004), and response rates are reasonably high, suggesting that the results are probably quite reliable as well.

The separate study conducted in the Health 2000 Survey on the health of young adults aged 18–29 included an extensive interview, and the participants furthermore completed various questionnaires. The home visit interviews were taken by Statistics Finland in April-July 2001, and data collection on those who did not take part in the interviews was completed in December 2001.

This chapter describes the data collection for the study of young adults and the contents and methods of that study. Many of the areas and items covered are identical to those surveyed in the population over 30, and full descriptions of the methods used can be found in earlier chapters of this report. Further information is also provided in a separate report in Finnish (Koskinen et al. 2005).

12.2. Sampling and data collection

The sample for the study of young adults consists of persons belonging to the sample of the Health 2000 Survey and aged 18–29 at the time when the sample was drawn; they numbered 1,894. The sampling design and selection process for Health 2000 is described in Chapter 3 above (see also Laiho and Djerf 2004). Sampling was done on the basis of information as at 1 July 2000, i.e. the age and place of residence of the young adults within the sample was determined according to the situation at the beginning of July 2000.

The field examination of young adults included an interview conducted by 127 Statistics Finland interviewers; the basic questionnaire administered at the time of the interview; and the dietary questionnaire that was mailed to all respondents who returned the basic questionnaire. The interview data were recorded on laptops and the results were sent back in encrypted form to Statistics Finland at regular intervals.

Various steps were taken to achieve as high a response rate as possible. Before and during the field examination the survey was widely publicised in national and local media. Those who took the home interview could enter draws for prizes donated by sponsors of the survey. Those who did not return the basic questionnaire by the deadline were sent a reminder and a new form; if they still did not answer, a

short post-questionnaire was sent later on containing core questions from the basic questionnaire. Two cinema tickets were offered to those who acceptably completed the basic questionnaire after the reminder or the post-questionnaire.

Interviews were conducted with 1,503 persons (79%). The basic questionnaire was returned by 1,282 persons (68%), and the post-questionnaire was returned by 205 persons (11%). Overall then, about 90% of the young adults in the sample answered the most important interview questions. The most common reason for non-response was refusal (Table 10.4). The sample and the numbers participating in different stages of the survey are shown in Table 12.1; Table 12.2 provides the corresponding data by university hospital district. Data on the proportion of non-respondents at different stages of the survey are given in Table 10.5.

Table 12.1. Sample and participants at different stages of the survey by gender and age, young adults aged 18–29.

	Sample	Participants					
		Interview		Basic questionnaire		Post-questionnaire	
		number	% of sample	number	% of sample	number	% of sample
Men							
18–24	**589**	474	80,5	370	62,8	56	9,5
25–29	**392**	291	74,2	232	59,2	41	10,5
Total	**981**	765	78,0	602	61,4	97	9,9
Women							
18–24	**549**	440	80,1	418	76,1	75	13,7
25–29	**364**	298	81,9	262	72,0	33	9,1
Total	**913**	738	80,8	680	74,5	108	11,8
Total	**1894**	1,503	79,4	1282	67,7	205	10,8

Table 12.2. Sample and participants at different stages of the survey by gender and place of residence, young adults aged 18–29.

	Sample	Participants					
		Interview		Basic questionnaire		Post-questionnaire	
		number	% of sample	number	% of sample	number	% of sample
Men							
Helsinki	**332**	239	72.0	177	53.3	44	13.3
Turku	**126**	95	75.4	82	65.1	8	6.3
Tampere	**210**	159	75.7	124	59.0	25	11.9
Kuopio	**157**	137	87.3	110	70.1	11	7.0
Oulu	**156**	135	86.5	109	69.9	9	5.8
Total	**981**	765	78.0	602	61.4	97	9.9
Women							
Helsinki	**332**	252	75.9	230	69.3	49	14.8
Turku	**117**	91	77.8	83	70.9	14	12.0
Tampere	**197**	163	82.7	152	77.2	24	12.2
Kuopio	**145**	132	91.0	124	85.5	7	4.8
Oulu	**122**	100	82.0	91	74.6	14	11.5
Total	**913**	738	80.8	680	74.5	108	11.8

12.3. Methods of data collection

Although the dataset collected for the study of young adults was less comprehensive than for the survey of the population aged 30 or over, the aim was to collect key data on all participants aged 18 or over. Therefore the health interview of young adults included some of the items that in the main survey were covered in the health examination, and accordingly the basic questionnaire included some items from the complementary questionnaire for the population aged 30 or over. Table 12.3 provides a side by side comparison of the data collection for the main survey and the study of young adults.

Table 12.3. Comparison of the data collection for the main survey and the study of young adults in Health 2000.

Persons over 30	Young adults aged 18–29
Health interview[1]	Health interview[1]
Basic questionnaire (questionnaire 1)[1]	Basic questionnaire[1]
Health examination:	–
– symptom interview	– parts of health interview (section BE)
– measurements	– height and weight: health interview (section BG)
– laboratory	–
– oral examination	–
– infection questionnaire (questionnaire 2)	– infections and vaccinations: health interview (sections BF and DG)
– functional capacity measurements	– linguistic fluency: health interview (section HF)
– clinical examination	–
– mental health interview	– parts of health interview (section BH)
Complementary questionnaire (questionnaire 3)	– parts of basic questionnaire
Dietary questionnaire	Dietary questionnaire

[1] The health interview and basic questionnaire administered to the population aged 30 or over and young adults aged 18–29 are largely identical.

Health interview

The Statistics Finland interviewers checked the accuracy of each participant's home address and contacted them by letter to make an interview appointment. Before starting with the young adults, the interviewers had already interviewed people aged 30 or over, and a few training sessions were arranged focusing on the content of the interviews. The interviews were computer-assisted. In connection with the interview the respondents received an information letter and a form of consent to sign. During the interview some of the responses were coded using computerised classification systems, which was supported by a comprehensive list of local municipalities, occupations, illnesses, medications and surgical operations. The respondents were handed the basic questionnaire and a return envelope and

asked to mail the completed form back to KTL. The interview lasted on average 75 minutes; printed out on paper, it ran up to over 100 pages. Before data collection proper the interview schedule for young adults was tested in a small-scale pilot survey with a dozen or so young adults who did not belong to the survey sample.

The interview schedule was designed to collect key background data, information on health status and diseases, medication use, use of health services, living habits, current and childhood living environment, functional capacity, employment and work ability and rehabilitation. The schedule for young adults aged 18–29 was designed on the basis of the interview for the population aged 30 or over. Items that were irrelevant to the health of young people were omitted, and new ones were added that had not been asked of the population aged 30 or over (such as more detailed items on studying and difficulties at school, or on the use of snuff). Furthermore, some items that were included in the health examination of people aged over 30 were adopted as part of the interview schedule for young adults (see Table 12.3). Musculoskeletal disorders and symptoms were inquired with questions that in the population aged 30 or over were included in the symptom interview. Infections and vaccinations were inquired using the items of the infection questionnaire that participants aged 30 or over completed in connection with the health examination. Mental health items included some of the main questions included in the mental health interview that was incorporated in the health examination of the population aged 30 or over. Furthermore, the cognitive capacity of the participants was assessed with a word-listing task, which was part of the functional capacity measurements of the population aged 30 or over. The health interview consisted of the following main components: background information, health and diseases, questions concerning parents and siblings, health services, oral health, living habits, living environment, functional capacity, work and work ability, rehabilitation and interviewer's assessments.

Basic questionnaire

The basic questionnaire for young adults aged 18–29 was designed on the basis of the corresponding instrument for the population aged 30 or over, making relevant adjustments. Using largely the same questions as for the population aged 30 or over, the 31-page basic questionnaire included items on participation in health-promoting activities, search for information about health and diseases, quality of life, psychological experiences, experiences of work and studying, symptoms, physical exercise, eating sweet snacks, alcohol use, time use and leisure activities, computer use, childhood living conditions and working conditions. Furthermore, items concerning sleep and genitourinary infections were included from questionnaire 3 for the population aged 30 or over (see Table 12.3). The basic questionnaire additionally included items on weight control, attempted suicide, drug use and sex

life, which were not asked of respondents aged 30 or over. The content of the basic questionnaire is shown in Table 12.4, and the form is posted on the KTL website at www.terveys2000.fi. A reminder and a new questionnaire were mailed to those who did not return the basic questionnaire within about four weeks; two cinema tickets were offered as an incentive to those who returned an acceptably completed questionnaire.

Table 12.4. Content of basic questionnaire for young adults (form T2140).

Health promotion	Eating and drinking sweets
Search for information about health and diseases	Alcohol use
Quality of life	Treatment of drink problems
Psychological experiences	Drugs
Experiences of work and studying	Time use and leisure activities
Symptoms and infections	Computer use
Sleep and sleeping	Safety of immediate environment
Physical exercise	Childhood
Anabolic hormones	Human relations and sex life
Weight control and slimming	Working conditions

Dietary questionnaire

Diet was measured using the same frequency-type food use questionnaire as in the survey of the population aged 30 or over (see Chapter 8). The dietary questionnaire was mailed to all those people who returned the basic questionnaire they received in connection with the home-visit interview. A total of 797 forms were returned to KTL. The forms were checked and eight incomplete questionnaires were rejected. In the end data on food use were obtained from 789 young persons aged 18–29.

Post-questionnaire

A separate post-questionnaire was mailed to persons who did not participate in the interview. Two cinema tickets were offered as an incentive to those who returned an acceptably completed form. The 16-page questionnaire contained the most important items from the interview and basic questionnaire.

Finalising the data

A large part of the data collected was directly recorded by computer, which meant that these data were practically error-free. Preliminary checks were carried out on the questionnaires at KTL and the forms were then sent to an outside company for recording. All data were checked for formal correctness, and logical checks and corrections between different data items were carried out. The staff who were assigned to work on these checks and corrections were highly experienced; much of this work was done by people working on their postgraduate studies following completion of a degree programme in health care sciences.

12.4. Assessing the success of the field examination of young adults

The field examination of young adults aged 18–29 was highly successful, despite some difficulties in the stages of survey planning and field staff training that were caused by time pressures. Health interviews were completed with almost 80% and basic questionnaires with almost 70% of those in the sample. If the core data obtained by post-questionnaires are taken into account, the participation rate was 90%. Participation was clearly higher than in any other recent Finnish health survey among young adults, let alone in those conducted elsewhere in Europe (Aromaa et al. 2003a, b). A high participation rate is extremely important in that it reduces the errors caused by selective non-response.

The success of the interviews was chiefly attributable to Statistics Finland's experienced interview organisation. The project was a success despite the fact that the interview was longer than normal and despite the fact that there was not enough time to provide sufficient training for the interviewers. The interesting subject matter was obviously a contributing factor as well. Furthermore, it is noteworthy that over half of those who did not take the interview returned the 16-page post-questionnaire, which included the core items of the interview and basic questionnaire.

The methods used in the survey represented well-established methods in population surveys, and various steps were taken to ensure their high quality. Some of the methods were specially developed for this project, and they too were used in as standard format as possible. Nevertheless it is possible that there have remained inter-interviewer differences that may particularly influence regional comparisons. The interviewers would certainly have benefited from more comprehensive training. It would also have been useful to conduct a more extensive pilot study covering the whole survey schedule before the actual field examinations, to fine-tune the methods and only then start the examination proper. However for reasons of funding and scheduling the Health 2000 survey was forced into compromises that are not unfamiliar from other survey settings. The data collection process is assessed in closer detail elsewhere (Nieminen and Kuusela 2004).

13. DATA PROTECTION AND ETHICAL ISSUES

Sami Heistaro

A major consideration at all stages of the Health 2000 Survey was with the provision of data protection and with the appropriate handling and storage of all the data and materials collected. Every possible precaution was taken to prevent unauthorised access. In the examination files, personal data were replaced by examination codes; even the researchers analysing the final data no longer have access to any personal data. However since these data will be needed for follow-up purposes as well as for linking with other data, they are accessible to a small number of authorised personnel for these specified purposes.

The plans and protocols for the Health 2000 Survey were submitted for approval to the relevant ethical committees. The application was first reviewed by the National Public Health Institute's Ethical Committee in September 1999. Following changes in legislation, the more detailed project plan was submitted to the Ethical Committee for Research in Epidemiology and Public Health at the Hospital District of Helsinki and Uusimaa (HUS) in May 2000. At both these stages, the plans received favourable opinions.

Separate permission was sought from and granted by the Advisory Committee on Radiation and Nuclear Safety for panorama x-rays of the mouth and upper and lower jaws (OPTG), and the Radiation and Nuclear Safety Authority (STUK) granted a safety licence for this examination.

An information letter was handed out to the subjects in connection with the home interviews by Statistics Finland staff and later in connection with the health examination, and in both situations trained staff were available to answer any questions. Once the subjects had read the information letter, they were asked to sign the informed consent form. Identical consent forms were signed separately for the home interview and the health examination because some of the subjects only took part in one or the other stage of the survey.

14. QUALITY MAINTENANCE AND QUALITY CONTROL

Paul Knekt, Pirkko Alha, Sirkka Rinne and Harri Rissanen

Survey data collection always gives rise to different kinds of variation, which complicates the interpretation of the results. This variation may be systematic or random, and it may be due to ”biological” variation occurring over time in the individual concerned, or to technical variation in the measurement. “Biological variation” may be due to the time of day or year, the conditions in which the measurements are taken or the state of the object of measurement. Changes happening over time in the subject’s health behaviour or health status are included under the definition of “biological variation”. Common sources of variation related to measurement include the failure to calibrate measuring instruments, or the failure to provide adequate instructions or training to the staff. Poor staff motivation may also give rise to additional variation. Uncontrolled variation may cause substantial loss of information. Systematic variation causes bias in the indicators derived from the data and may have a significant impact on comparisons between different subpopulations and results obtained at different points of time.

In Health 2000, key quality objectives for the data collection included the production of as valid and reliable measurements as possible, the maintenance of as stable data levels as possible throughout the duration of measurements, and maximum comparability of data measured by different staff and by different measurement units. For this reason the following factors were particularly relevant to quality considerations in this survey: most of the staff taking measurements were involved in this kind of population survey for the first time; several different people took these measurements at different examination sites; and the survey took place over a relatively long period of time. Furthermore, some of the instruments were new and no information was available on their validity and/or reliability.

The achievement of the quality objectives was supported by training and education, standardisation of measurements (quality management), continuous quality assurance in connection with data collection and the use of specific measurement designs during specific quality days (Table 14.1). In addition, separate quality assurance measures were followed both during the survey proper and more in-depth examinations. The following designs were used for purposes of quality assessment:

- repeat measurements (same person took two measurements of the same item)
- parallel measurements (two or more persons took measurements of the same item)
- reference measurements (measurement results were compared with a valid reference measurements)
- standard measurements (measurements were taken of standard items)

14.1. Quality management

The aim of quality management was to maintain rigorous and standardised measurement procedures throughout the field examination. Written instructions were prepared for all measurements and training was provided both ahead of and during the examinations. Video recordings were taken during the early stages of the examination to ensure that the staff in the field teams were working consistently according to the instructions given.

Furthermore, all measurement conditions were standardised and certain measuring instruments were calibrated or their operation was checked. For example, the five instruments used in the ultrasound examination of calcaneal bone were calibrated at the end of the examination by examining the same five persons on each instrument (Table 14.1, quality measure 1). The bioimpedance devices were checked so that upon completion of the examination, the same operators took measurements of themselves on all five devices used. Both hand grip strength meters (N = 325) and the equipment used in the balance test (N = 245) were calibrated at regular intervals during data collection by using standard weights.

14.2. Continuous quality control in the field team

All field teams followed daily a rigorous quality control regime to ensure the constancy and quality of their measurements throughout data collection. The data obtained were used to assess the repeatability of the results obtained by a single operator or measuring device and to assess the agreement of the results between different operators and instruments. Repeat and parallel measurements were taken both from field staff and the subjects.

At the beginning of each day the operators took breathing function, calcaneal bone ultrasound and bioimpedance measurements of themselves (Table 14.1, quality measure 2.1). All in all, depending on the method, a total of 346–484 measurements

were taken. Furthermore, staff working on functional capacity assessments took measurements of one another to measure hearing (N = 299) and reaction time (N = 245) when moving to a new examination site or at least once a week.

Parallel measurements by two field staff members and repeat measurements by one field staff member were taken daily in accordance with the weekly programme for each examination point on the first or first two subjects of the day (Table 14.1, quality measure 2.2). Both the field teams' measurement operators and home-visit nurses took part in conducting the parallel measurements. The aim was to have the same number of different combinations of staff during the course of the examination.

Almost 100 subjects took part in parallel clinical examinations and symptoms interviews as well as in measurements of blood pressure, waist circumference, sagittal measure, respiratory function and calcaneal bone density. Furthermore, parallel functional capacity measurements were taken of over 100 subjects (eyesight, hearing, reaction time, hand grip strength and balance). Parallel joint function tests and chair stand tests and walking tests were taken on 86 subjects aged 55 or over. Repeat clinical examinations of the mouth were conducted on 111 subjects.

14.3. Estimating total measurement variation

For the estimation of total measurement variation (i.e. simultaneous analytical and biological variation), some 40 persons were invited from each group for repeat measurements 6–8 months after the field examination (Table 14.1, quality measure 3). The total material thus comprised some 200 subjects.

This examination included questionnaire 1 (basic questionnaire), the short home-visit interview and the short symptom interview. In addition, measurements were taken of blood pressure, waist circumference and various aspects of functional capacity (eyesight, hearing, reaction time, hand grip strength, balance (Guralnik) and cognitive capacity for all subjects and joint function, chair rise and walking speed tests for those aged 55 or over). Furthermore, blood samples were taken. The total variation of the Mini Mental State Examination was estimated in a sample of 105 subjects, and total variation in the diet questionnaire in a sample of 180 subjects in repeat questionnaires 5–9 months later.

14.4. Quality days

In addition to the quality data collected during the field examinations proper, complementary quality data were also compiled during three different kinds of quality days. These days completed quality data collected at the field study. Quality days I and II were held during the field stage and quality day III at the end of the examination; this was arranged jointly for all five groups.

14.4.1. Quality day I

Each field team's programme included six type I quality days, with around 10 subjects from the sample invited to take part each day. In all, the quality material comprised some 300 subjects. During these quality days results from different measuring devices were compared, parallel measurements were conducted between home-visit nurses and measurement operators, and reference measurements were taken.

For the comparison of different measuring devices, the same operator used two different devices to take measurements of blood pressure, respiratory function, weight and hand grip strength (Table 14.1, quality measure 4.1.1). Blood pressure measurements were taken with a mercury sphygmomanometer and an automatic device (N = 265), respiratory function measurements using spirometry and PEF measurements (N = 267), weight was measured on the bioimpedance device's scale and the home-visit nurses' digital scale (N = 276) and hand grip strength was measured with Mini-Finland and Health 2000 instruments (N = 46).

Parallel measurements were taken of waist circumference and certain aspects of functional capacity (eyesight, hearing, hand grip strength and Guralnik balance) by home-visit nurses and measurement operators in a subsample of some 180 subjects (Table 14.1, quality measure 4.1.2). In addition, joint function test, chair rise and walking speed tests were taken in 78 persons aged 55 or over.

Parallel and reference measurements for the clinical examination of mouth (Table 14.1, quality measure 4.1.3) compared the findings of the examining dentist and a reference dentist in a sample of 269 subjects. Reference measurements of joint function tests were based on videotapes of examinations with 18 subjects, which were assessed by a reference group of 27 nurses and 2 researchers.

14.4.2. Quality day II

Each field team's programme included four type II quality days, for which 10 voluntary participants aged 45–82 were recruited from outside the survey sample (Table 14.1, quality measure 4.2). The total material thus comprised 200 subjects. Repeat measurements were conducted on this material so that measurements were taken both by the operators themselves and by measurement operators from another field team. The subjects visited each examination point twice. Parallel measurements were taken of blood pressure, respiratory function, calcaneal bone ultrasound, functional capacity (eyesight, hearing, reaction time, hand grip strength, balance, joint function, chair rise and walking speed test) and clinical examination.

14.4.3. Quality day III

Quality day III was arranged during one calendar week in Helsinki and Vantaa (Table 14.1, quality measure 4.3). Persons from outside the survey sample were invited to take part, inter-measurer consistency for operators in different field teams was assessed, and the consistency of their results with those obtained by national reference measurers was studied.

Reference and parallel measurements were taken for the symptom interview, blood pressure, waist circumference, sagittal measure and functional capacity (hearing, joint function, chair rise and cognitive capacity) in a sample of 138 persons aged 25–83 so that three sets of measurements were taken on each person. Measurements for all subjects were taken by the operator for the reference area. In other areas measurement operators took parallel measurements, the aim being to have the same number of subjects in each of the six combinations of measurement operators in the four areas. This yielded a repeat dataset of 138 subjects and a parallel dataset of 2 x 138 subjects.

The consistency of the depression component of the M-CIDI mental health interview was assessed by conducting pairwise parallel measurements on 49 subjects (see Chapter 8.10).

A total of 42 persons aged 27–66 took part in the clinical oral examinations at the University of Helsinki Department of Dentistry. Parallel measurements were taken by six persons: the dentist from each field team and one reference measurer.

14.5. Separate quality monitoring

In addition to the quality measures listed above, a number of separate quality assurance observations were conducted. All laboratory determinations for the survey were carried out at laboratories that had standard quality control systems in place to ascertain the level and repeatability of measurements. The validity of the dietary questionnaire was assessed using a 3-day diary kept by 294 participants in the survey proper. Repeatability of x-rays of the mouth was estimated by coding 327 x-rays and the repeatability of the Minnesota Code by repeat analysis of 200 ECGs.

A total of 1,526 persons aged 45–74 took part in the complementary investigation of the circulatory system (SVT+D, see Chapter 21) 6–29 months after the survey proper. Together with the main dataset, this data subset allows for the assessment of the total variation. In addition, the SVT+D in-depth examination included repeat ultrasound measurements and comparisons of the equipment used. Repeat ultrasound measurements were taken in a sample of 44 persons. Two programs (Prosound and Provin) were used in the measurement of carotid artery intima media thickness (IMT) from ultrasound images. The results were compared in a sample of 200 persons. The material for repeat studies a few days later consisted of 18 subjects. On both occasions the measurements from these people were taken by three field examiners, and the dataset of six images was analysed by the same reader.

To estimate the total variation of measurements, the back symptoms interview was repeated with 89 subjects, and the endurance test of back extensor muscles with 660 subjects.

Table 14.1. Summary of the size of the quality data collected in the field examinations proper by measurement and quality measure.

Measurement		Quality measure								
		1	2.1	2.2	3	4.1.1	4.1.2	4.1.3	4.2	4.3
Home:	Questionnaire 1				208					
	Home-visit interview				208					
	Mini Mental State Examination				105					
Point 1:	Symptom interview			96	208					138
Point 2:	Blood pressure			92	207	265			198	138
	Waist circumference			92	204		175			138
	Sagittal measure			92						138
Point 3:	Respiratory function		482	93	267				193	
	Calcaneal bone ultrasound	25	484	93					193	
	Bioimpedance	5	346							
	Weight				276					
Point 4:	Laboratory				201					
Point 5:	Clinical examination of mouth			111				269		42
Point 6:	Functional capacity									
	– eyesight			111	208		177		200	
	– hearing		299	111	207		169		200	138
	– reaction time		282	106	198				190	
	– hand grip strength	325		113	205	46	169		190	
	– balance	245		106	206		173		189	
	– joint function			86	112		78	18	182	64
	– chair stand test			86	112		78		181	138
	– walking speed test			86	112		78		181	
	– cognitive capacity				208					137
Point 7:	Clinical examination			99					173	
Point 8:	Mental health interview									49
Point 9:	Dietary questionnaire				180					

1. Quality management

2.1 Repeat and parallel measurements on field team staff

2.2 Repeat and parallel measurements on subjects in same examination group

3. Total variation at about 6–8 months

4.1 Quality day I (4.1.1 Comparison of measuring instruments, 4.1.2 Parallel measurement, 4.1.3 Reference measurement)

4.2 Quality day II (Parallel measurements between groups)

4.3 Quality day III (Reference measurements and parallel measurement between groups)

15. IMAGE DOCUMENTATION AND QUALITY ASSURANCE

Jussi Valkeajoki, Päivi Sainio, Antti Reunanen, Markku Heliövaara and Sami Heistaro

15.1. Video recording

Video recording in the Health 2000 Survey was initially motivated by needs of quality assurance: project headquarters in Helsinki wanted to have image documentation of how the field examinations in the five teams around the country were progressing. Video recordings were made 1) independently by the field teams themselves and 2) by a video photographer.

A Canon V40 Hi8 video camera was purchased for each of the five teams. This is a simple, compact and easily transportable automatic camera that provides adequate image quality for purposes of monitoring and studying the field procedures. The camera has a small colour LCD display to facilitate shooting. The subjects were asked separately to give their written consent to be video recorded.

15.1.1. Field teams' video recordings of functional capacity measurements

In the functional capacity tests the health examination staff were to take a wide range of different measurements using different types of equipment. Video recordings were shot by field team members with a view to ensuring the consistency of these measurements. Initially the plan was to take weekly recordings of the measurements conducted by each staff member, but this was subsequently scaled down to one recording a month. On each team one of the members was appointed to take charge of the video recordings. Training was provided to these nurses while the health examinations were already underway using written, illustrated instructions.

Nurses conducting the measurements mounted the video camera on a tripod and set it to record for the duration of the examination. However, to keep the objects within the frame and to make sure their backs were not turned to the camera, the nurses had to go back to check their angles and move the camera between different stages of the examination. Therefore, whenever timetables permitted, the video recordings were done by another nurse.

The nurses in charge of the video recordings mailed the Hi8 camera cassettes back to project headquarters, where they were transferred to VHS video cassettes for viewing and archiving. The cassettes were then sent back for reuse. The videos were reviewed once a month at headquarters by the officer responsible for staff training in functional capacity measurements. The performance of the nurses was assessed using a structured schedule: the main focus was on their adherence to the procedural guidelines specified for each test, the overall fluency of their performance, their interaction with the subject, their use of IT equipment and measuring devices and their recording of the results.

Any problems detected from the videos were recorded in a feedback form, which was e-mailed to all nurses working on the functional capacity measurements. The videos were also used in the field team training session organised midway through the health examinations, where they were jointly analysed by the functional capacity nurses and training personnel.

15.1.2. Video monitoring by the Finnish Institute of Occupational Health Musculoskeletal Group

Sirpa Rauas

Videos shot of the clinical examinations performed by the field physicians in the Health 2000 Survey were used for purposes of quality assurance of the musculoskeletal examinations. The recordings were taken by a visiting video photographer. In October 2000 the first video was viewed by the three members of the Finnish Institute of Occupational Health's Musculoskeletal Group. The experiences from this review provided important direction for the organisation of quality assurance later on in this project. Based on the observations made from the video, comments on the examination practices were sent out in the early stages of the survey to all field teams.

In October the first videos were also obtained on the performance of all field physicians. The videos were viewed by an occupational physiotherapist, who provided written comments for each individual physician on practices where they departed from the guidelines and reported on the findings to the chief physician. The chief physician prepared a summary of these comments, which was e-mailed to all field physicians and members of the health examination teams. There was also some open e-mail communication to provide more accurate guidelines for the clinical examination of the musculoskeletal system.

In January 2001 the occupational physiotherapist viewed a second set of tapes that were recorded around the turn of the year. The purpose was to assess the consistency of the examiners' procedures and the quality of the examination as routines began to set in. Each field physician's current performance was compared with their first examinations of the patient's musculoskeletal system. Summaries were produced of these comparisons as well.

Based on these experiences it was indeed necessary to have quality assurance processes in place in order to maintain consistent musculoskeletal examinations. Individual routines and customs were clearly reflected in the way that the field physicians went about the examinations. The biggest differences were seen in those tests with which the field physicians had the least prior experience. On the other hand when they were conducting more familiar examinations they tended not perhaps to be quite as rigorous in their insistence on initial postures, for example.

15.1.3. Video recording of other examinations

Apart from project staff who had overall responsibility for the functional capacity measurements and the clinical examinations, those responsible for other examination points (such as the oral examination and measurement point 1) were also keen to use video materials to monitor the progress of the examination. As in the case of the medical examination, the decision was made not to shoot the examinations weekly but only on a few occasions.

15.1.4. Experiences of the use of video recordings for quality assurance

Measurements of functional capacity

It was unfortunate that there was not enough time to provide hands-on training to the examination nurses on the techniques of video recording, but this training had to be done simply by sending out written instructions. The nurses had some difficulty adopting video recordings as part of the examination protocol at such a late stage of the process, in some cases when the field examinations were already underway. For some nurses it seemed rather difficult to come to terms with the idea of being in the lens of the camera. On the other hand this was quite a simple and straightforward quality assurance measure that yielded very cost-effective information on how the examinations were progressing.

In spite of the human problems involved the quality of the video material was reasonably high and it offered a reasonably clear picture of what was happening

out in the field. The field teams themselves came upon the idea of using another nurse to take the video recordings, which made it much easier to follow the events unfolding in the tape and gave nurses a valuable opportunity to see how their colleagues were working. This contributed to an improved consistency in measurement procedures. However nurses were not always available to take on this job, and in this case the examining nurse had to make arrangements for the video recording in accordance with the original plan. This option was considered more demanding and laborious.

With time, the field teams gained more confidence with the job of video recording, but never to the extent that it became a fluent routine until after the field examinations had ended. During and after the field examinations, video recordings were taken of some of the home-visit examinations.

Back at project headquarters, the videos were viewed and assessed by two staff involved in training for the functional capacity measurements. It took them about three hours to view the videos sent back by one team and to compile feedback. The original plans of viewing one weekly video of one nurse's performance proved too ambitious and time-consuming in view of the reviewers' other duties, and therefore these plans were revised downwards to once a month. This still yielded several videos of each nurse's performance and provided more than adequate material for assessing how the guidelines and feedback provided were being followed.

Other examination points

At the other examinations points, the image documentation was collected by an outside photographer. At each location the photographer used the team's own equipment to shoot the necessary examinations during the course of one day. In most cases the examination nurses were able to assist the photographer in recording the measurements at the functional capacity point.

During the autumn of 2000 the project photographer, Jussi Valkeajoki, made five trips to visit each of the field teams. During these visits he took shots not only of the functional capacity measurements, but also of the physicians' medical examinations, as well as of oral examinations. In spring 2001, the photographer made a further five field trips to take another set of shots of the functional capacity measurements, medical examinations and measurement point 1.

The subjects for the video recordings were chosen at random by shooting the person who happened to be next in line for the examination when the photographer was ready to start. If the subject refused, the next person was chosen. Most subjects agreed to the video recordings, even though they sometimes had to take almost all their clothes off in front of the camera. The attitude of the examination nurses was clearly a key factor in whether or not the subjects gave their consent.

Project training staff used the videos recorded at the different examination points for purposes of training and monitoring field examination staff. A major focus here was on individual differences in how the examinations were conducted as well as on how the instructions given were followed.

15.1.5. Conclusions

Video recordings proved to be a useful method for purposes of quality maintenance in the health examinations. They provided information on procedures that deviated from the guidelines and instructions issued and allowed for immediate intervention. It is important that enough time is made available for reviewing and assessing video materials and that this process is done in a structured and systematic way so that all the components of the performance concerned are properly taken into account. It is also worthwhile to plan in advance in what form and how often feedback is provided to the people conducting the experiments. By and large the video recordings were up to or even exceeded expectations. There were some problems with the video recordings performed by the teams themselves, particularly when the examinations were to be videod by the same nurse. Sometimes the camera was less than steady even when the shooting was done by another nurse, and the frames were not always focused on relevant activities. However these were quite rare instances and dependent on who was operating the camcorder.

To improve the quality of video recordings by the field teams themselves in future surveys, it would be necessary for the nurses to receive a more thorough training and for them to have the opportunity to practise in advance. Even so this method can never be expected to produce other than material of an average quality. It is obviously always the best option to use a professional photographer wherever this is possible.

15.2. Photography

Photography was used in the Health 2000 Survey for illustrating the video recording instructions, for general documentation and for producing presentation materials. Among the subjects shot were: 1) selected stages of the examination in October 2000, 2) all stages of the examination in March 2001, 3) the disassembly of examination points in March 2001 and 4) the removal of the examination equipment in March 2001.

15.2.1. Shooting in the field

Images for the illustration of the video recording instructions were shot in October 2000 on black and white film using a 35–135 mm lens and flash. Because of the small size of the examination rooms, most images were shot at 35–50 mm.

The documentation and presentation images in March 2001 were shot using an ever wider lens (24mm) on colour film without flash. This meant that in the small rooms it was possible to fit in even more information in one shot about the environment, the equipment used and the performance of the examination.

For the presentation materials, the photographer went through all the examination points with a model subject, from registration through to the final interview. At each examination point even the simplest routines were shot so as to get as detailed a picture as possible of the various stages of the examinations.

During the last days of the examination the photographer took a few rolls of shots of the examination nurses and the general milieu of the examination points in Helsinki. These shots provide a glimpse into how the examination points were assembled and later disassembled and how the equipment was moved from one examination site to the next.

15.2.2. Processing of presentation images

The images were converted into electronic format directly from the negatives using a film scanner and a file size of around 10 MB per image. Paint Shop Pro was used to adjust colour balance and other image features so as to make them as sharp and clear as possible. The green yellow light cast onto the film by the fluorescent lamps in the examination rooms were adjusted towards blue and magenta. Natural light manipulated by the image processing software was more pleasant to the eye than the harsh and technical flash light from the camera, which would additionally have created dark shadows behind each object. The use of natural light (and by the same token a large aperture) reduced the depth of field and softened the contours. This slight softness may be regarded as a drawback of the use of natural light, but on the other hand it made for a more pleasant viewing experience.

The colour, brightness and contrast of the images were adjusted and the necessary captions as well as the Health 2000 logo were added to them. The images were then produced into presentation materials: slide shows, transparencies and PowerPoint presentations.

16. CHECKING AND RECORDING DATA

Pirkko Alha and Sirkka Rinne

16.1. General

The data for the Health 2000 Survey were collected in many different ways. Some of the data were collected in interviews, some were obtained from paper questionnaires completed by the subjects. In addition, various tests and samples were taken.

Most of the data were collected in electronic format using a Blaise program. The data collected using paper forms were recorded by an outside company. Various procedures were applied to check and correct both types of data. In the case of paper forms, checks could be carried out on the original data, whereas other methods were needed to check the electronic data. As a rule these checks were carried out only on data from the same examination point or form, i.e. no checks were carried out between different forms.

The data were checked and corrected in two stages. In the first stage, the checks focused on the variables included in the basic results report. For this purpose, only one round of checks was performed. Once this report was out, a chain of checks was started that involved several different stages. The first step was to set out the principles and procedures to be followed in this work. The actual job of checking and correcting the data was coordinated by two project planning staff with extensive experience of similar earlier surveys.

Not surprisingly, the process of data checking was a time-consuming operation – after all there was a huge amount of data to get through. Although the plan had been to incorporate permissible limits and logical conditions in the Blaise program, this was not fully successful. Operators could bypass conditions specified in the program by entering comments and notes, for instance. The checks were carried out on one variable at a time, and the whole process involved several rounds of checks. The definition of logical conditions was particularly time-consuming. This was partly because the people doing this job had no prior experience and needed training.

During the course of these checks it was realized that the conditions needed updating. As most of the series of questions did not include the options “don’t know” or “not applicable”, the amount is missing data was very high. Likewise, the

nurses sometimes found that they were unable to slot their interviewees' responses in any of the given categories. Programming errors also added to the amount of time taken up by the checks: the elimination of these errors would have required more preparation time.

All the documents, files and SAS program files used in the checks are saved in such a way that they show the name and date of the data.

16.2. Preliminary processing of data

Data were sent back from the field teams to project headquarters once a week. The Blaise files were received electronically, and the paper forms were either collected or mailed. The number of paper forms soon began to accumulate, and they were initially filed by type of form. Receipt of each form was recorded using a separate program. When the code for the type of form and the subject's examination code were entered, the program displayed the name of the person in question. This was compared with the personal data shown in the forms. Any errors were addressed, and finally the form was stamped as "registered". The same procedure was followed for forms completed by the subjects themselves and those completed at the health examination site.

16.2.1. Preliminary checks and recording

Once registered, the paper forms were ordered according to the examination number and archived. A preliminary check was carried out on each form before being passed on for data recording. Unclear and illogical responses were corrected as far as possible so as to facilitate the job of staff doing the actual recording. Instructions on conducting the preliminary checks were compiled in writing for all the forms. These provided guidance for instance for situations where the subject had chosen two response options instead of the permitted one, and what action was to be taken in the event of missing data.

Because there was so much data to process, these preliminary checks and data recording took up a great deal of time. The data returning from recording staff were checked at the level of examination number before approval of the recording. Part of the data were recorded twice to identify recording errors. Overall there were very few such errors.

16.2.2. Documentation

The data were divided into different categories: total material (01), Health 2000 sample proper (11), Mini-Finland resurvey (21), and quality assurance and additional subjects (31).

The SAS program files, data files, conditions and correction rules created in connection with the data checks were saved in accordance with the planned directory structure. The following subdirectories were created for each examination point: SAS program files, data, description, lists, instructions and check conditions. The programming statements which revealed any erroneous variables or with which the corrections were entered into files, were stored in SAS program files. Data comprise the examination data proper, correction files and the corresponding SAS correction statements and assignments. The Lists directory includes a list that defines the variables that are entered into the correction file because of an error. Data record descriptions and comparative documents created of each form were saved in the Description directory. Check conditions were saved in their own directory, as were the correction rules (Instructions directory). The files were tagged in accordance with the standards set for each examination point and stage of check procedure. The last characters of the file names indicate the date.

The various stages of the checking and correction process were recorded in the Checks files document. Data were entered in initial checks, checks of upper and lower limits, corrections required by notes and comments, logical checks as well as other corrections and additions. The document includes a record of all SAS program files used, correction files created and of the version of the data created at each stage.

16.2.3. Initial checks

Registration proved extremely useful from the earliest stages of the checking process. The registered file was used to extract preliminary data on participation. When the first versions of the data collected out in the field were received from the IT team, they were compared with the registration data. These checks revealed that large numbers of data records were missing. The IT team looked into this and discovered that there had been problems with the transfer of files from the field teams to project headquarters. Consequently it was decided that checks should be run on the computers of all field teams. Hundreds of data records turned up in this search, and in the end the number of lost data records was quite small.

The actual check process was started with comparisons of the examination numbers. If it was possible to change an erroneous number and replace it with the correct one, this was done. The most important item to be checked was the examination schedule, which indicated the examination points that each subject had visited. The material included some duplicate data records. Part of the reason for this was that the program had been opened on several occasions for the same subject. On the other hand, data records were also found for subjects who had not been to the examination points. After these initial checks a locked file version was created that shall be retained as the permanent baseline file.

16.2.4. Existence

Registration data were used to form a follow-up file with a classification for data on participation in the home-visit interviews, health examination or home visit and on attempts to contact non-participants, telephone interviews and non-response questionnaire. This file was updated at different stages of the check process and used in all checks related to existence. It also includes those persons who did not eventually take part in the health examination, for instance, but who only turned up to return the blood pressure gauge or for other such reasons.

During both the home-visit interview and the health examination or home visits, the subjects were asked to sign an informed consent form. All these forms were checked, which yielded a classification of the existence of consent.

Existence criteria were specified for each examination point or form. As a rule at least two variables were chosen for which acceptable data were to be available. The data records that did not satisfy these criteria were examined more closely to determine whether the person in question could be classified as having participated in the survey. Registration files, consent data and follow-up files were also consulted for this check. If the relevant data record for the subject had gone missing, but the feedback form included data from the examination point in question, these data were entered as information for that examination point. Data records indicating zero for participation were removed from the material.

On the basis of the existence checks participation data were formed for each person and for each examination point or form. Finally, data were entered for each subject's examination dates and reasons for non-participation.

16.2.5. Description

Descriptions of each examination point or form were entered in an Excel file on the basis of the Blaise program or recording instructions. The variables from different forms have been compared with each other, and variables based on identical questions were labelled according to the variable tag used in the principal form. A prefix was attached to all variables other than those from the home-visit interview so that they can be associated with the form in question. For instance, the variables from questionnaire 1 and questionnaire 2 are distinguished by the prefixes Kys1_ and Kys2. Versions have been produced of the different forms that include the relevant variable tags, allowing data order forms to be posted on the Internet.

16.2.6. Remarks

Staff conducting the health examinations and interviews out in the field could enter any relevant notes and comments in the program's OPEN and REMARK fields. Indeed, this option was frequently used. Initially the purpose was that any notes and comments were to be related to a specific question, but as it turned out they often contained a lot of other information as well.

In the early stages of the check process there was some ambiguity as to how the data in these fields could be used. Eventually the decision was made that all the remarks made in connection with the health examination proper or home-visit interviews would be classified. This was done without any set variable values, i.e. the classifications were made purely on the basis of the remarks. For each examination point there was one or more classifications, e.g. for blood pressure and other measurements separately. As similar, general remarks were made across all examinations points, the classifications had certain common codes. In addition, there were ten or so subject codes that operators could use.

Partial classifications were made of the home-visit interviews and the interviews proper for the purposes of the basic results report; these may later be updated. For research use, the remarks will later be formed into a separate file containing all the classified cases.

16.3. Corrections

The first step was to create a check conditions document defining the permissible values or logical conditions for each variable and the variables to be entered in the event of an error. These documents were then used to form SAS program files

which carried the check conditions. Since the purpose was to create a corrections file including data on corrections and any additional comments, conditions were programmed into a SAS program file that allowed for the inclusion of the desired variables and the classification of remarks.

In each case all remarks were first checked. In some cases this was sufficient to obtain the correct value. In some instances data had to be rejected because they were not considered reliable. Various letter codes were used to indicate missing data. Most corrections were entered in a correction file, with the correct value entered in the correction column. A brief comment was added in the remark section on the reasoning behind the decision. In some cases the corrections were made directly in the form of SAS scripts.

The corrections were entered into the file in a separate run, which might include several different corrections concerning the examination point in question. The correction run was followed by a proc compare run, the results of which were used to check the accuracy of the corrections entered into the file. In addition, printouts were always taken to check the contents run. If the corrections had been successful, a new file version was created for that date.

16.3.1. Upper and lower limit checks

The check process started with upper and lower limit checks. Even though permissible values had been entered for some materials in the Blaise program, they were sometimes too broad. For instance, the lower limit for height was 100 cm and the upper limit 230 cm. Errors consisted not only of values under and over those values, but also of missing data. In this instance it was possible to consult the entries made in the paper forms and to make the corrections accordingly.

Upper and lower limits checks were repeated until no new cases appeared in the correction file.

16.3.2. Corrections due to remarks

The next step was to address the corrections necessitated by comments made in the remarks section. This concerned the examination points or forms for which the remarks had been classified. At each point codes were identified that were expected to cause changes to variables. If values were found that needed correcting, the changes were entered into the file.

16.3.3. Logical checks

The same procedure was followed for logical checks as for upper and lower limits. The main focus was on the logical consistency of the responses to the main question and follow-up questions. Precedence was given to the follow-up question. For instance, if the respondent mentioned the number of cigarettes smoked, it was assumed that the response to the main question (Do you smoke?) had to be 'yes'. Missing data to main and follow-up questions again caused a lot of work.

16.3.4. Blank lines and classifications

Many series of questions ended with the response option, "other please specify?" If this had been answered, the first step was to check whether any of the preset response options would have been a more appropriate answer, and the corresponding changes were made to the variables. Where necessary these responses were separately classified into new variables. Checks on the rehabilitation and aid variables required a particularly large number of corrections.

The interviews included large numbers of items (such as those on medications, illnesses and surgical operations) with preset classifications. These classifications had been incorporated in the Blaise program to facilitate the interviewer's job. The interviewer entered in a text variable the medicine, illness or operation reported by the interviewee, and was then to choose the corresponding code from a menu that appeared on the screen. For example: the interviewee reported using a medicine called Burana. The interviewer was then to select the corresponding product from a list (in this case of three medicines) and approve the choice, whereby the code was saved into the computer.

In many cases the interviewer had not written down the whole name of the medicine in question, but only the first part of the name (Burana); however the code revealed that the medicine in question was in fact Burana-Caps. In these cases the decision was made to err on the side of the interviewer's choice. In cases where products have the same ATC code, this obviously makes no difference. Medicines were coded using the SII system and translated to correspond with the ATC coding system.

Natural remedies were corrected so that at least one main category and one effective substance was mentioned for each product. Two variables were set aside for main categories and three variables for effective substance. The classification of natural remedies was produced by Anna-Liisa Enkovaara from the National Agency for Medicines.

The self-reported illnesses in the health interviews are based on a slightly modified classification previously used by the SII. ICD-9 and ICD-10 disease codes were added to this classification under the headings of different disease categories. The classification of surgical procedures was compiled on the basis of the procedures component of the Hilmo file by Ilmo Keskimäki from the National Research and Development Centre for Welfare and Health (Stakes). Diseases and surgical operations were checked and corrected in accordance with the classifications mentioned above.

16.3.5. Reclassifications

During the course of the checks process some of the spirometry and calcaneal bone results from measurement point 2 had to be reassessed. Furthermore, a separate BMI file was created to compile height and weight data from all possible sources. This was done because the bioimpedance device was not available during the earliest stages of the examination, and later on it suffered from occasional malfunctions. Data on height were also missing from some of the results from measurement point 1.

16.3.6. Final checks

In addition to the data checks proper, intervariable frequency runs were done using the LIST MISSING command, which yielded tables from all possible combinations between the variables concerned. All erroneous combinations were listed and corrected.

17. DATA OBTAINED FROM REGISTERS AND REGISTER FOLLOW-UP

Seppo Koskinen, Paul Knekt, Pirkko Alha, Sirkka Rinne, Esa Virtala, Harri Rissanen and Arpo Aromaa

Data were extracted from various register sources to complement the main body of data collected in the field examinations. This was done for three purposes: Firstly, register data were used to achieve a more accurate and complete picture of the subjects' health and related factors before and at the time of the field examination. Secondly, a comparison of the participants' and non-participants' health in the light of register data provides a good idea of the accuracy of the picture of public health drawn on the basis of the field examination data. Thirdly, the purpose of using register data is to find out how accurately the data collected in the field survey predict the development of the subjects' health: this is done by linking the cross-sectional data collected in the field examination with follow-up data on the subjects' causes of death and illnesses, service use and the payment of benefits granted on grounds of illness.

At all stages of the register data process, special attention was given to data protection. This was ensured by close adherence to the relevant legislation, the rules of the National Public Health Institute and the bodies maintaining the registers, and the guidelines of good research practice. The subjects were informed of the use of register data in writing, and they signed an informed consent form (see Chapter 13).

The linking of register data was designed and carried out in close cooperation between the project organisation and the bodies maintaining the registers concerned. The most important data items linked to the field materials concern causes of death, hospital treatments, entitlements to special medication reimbursements and certain other illness-related benefits, purchases of prescribed medicines, cancers, work disability and employment as well as housing.

18. RELEASE OF FILES FOR RESEARCH PURPOSES

Pirkko Alha and Sirkka Rinne

In order to obtain access to Health 2000 files, researchers are required first of all to submit their research plans to the team responsible for the subject area concerned. The plan will then be forwarded for review by the project group or its working committee. A code number shall be given to each approved research plan; that code number is needed to place a data order.

Data can be ordered via the Health 2000 website, which has all the forms used in data collection and the corresponding variables; selections are made simply by the click of the mouse. By clicking on the "Send" button, a copy is automatically sent to the person placing the order and the request is e-mailed to the persons controlling the release of the data. This eliminates the need to key in long variable titles, and the researcher can be sure they will get exactly the data they require.

Once the data order has been received, an agreement of cooperation is drafted. The draft is checked by the researcher, who shall return it to the National Public Health Institute (KTL). The agreement is printed in two copies. At KTL, on behalf of the Health 2000 Survey, it is signed by the head of department. The paper copies are then mailed to the researcher for their signature. One copy of the agreement remains with the researcher, the other is returned to KTL for archiving.

Data requests received via e-mail facilitate the task of assembling the required variables because the names can be copied directly from the e-mail. The dataset is formed by starting out from the total sample and its background data (e.g. persons aged 30 or over N = 8,028, young adults N = 1,894 and Mini-Finland follow-up study N = 1,278). All available requested variables are then attached to this dataset, unless the researcher has specified restrictions.

The data collected are usually sent in zipped SAS or STATA form as an e-mail attachment. This is followed by a background data document indicating the content of the background variables. In addition, various classifications used in the variables concerned are attached if and as necessary.

19. WEIGHTING AND STATISTICAL ANALYSIS

Kari Djerf, Johanna Laiho, Risto Lehtonen, Tommi Härkänen and Paul Knekt

19.1. Sampling weights and their use

The purpose of sampling weights is to return the observed data to correspond to the distribution of the target population. Sampling weights can thus be used to adjust for sampling variability and errors caused by non-response. Weights should be used in all statistical analyses and estimations based on survey data. If sampling weights are not used, the results cannot be generalised to the target population, and therefore they will not be comparable with other similar surveys.

The derivation of sampling weights is started by calculating the original inclusion probabilities for the individuals in the sample from which overcoverage cases were excluded. The individuals' inclusion probability is dependent on the regional distribution of the population and the age of the target person as follows:

One-stage sampling – Area clusters

– 15 largest health centre districts were sampled with a probability of 1
– 65 health centre districts were sampled with unequal selection probabilities

Two-stage sampling – Stratification of sampled clusters

– By age
– A finer sampling interval for people aged 80 or over

The sampling interval for people aged 80 or over was one half smaller than for others, which means that their inclusion probability was twice as high as for people in younger age groups living in the same area.

The basic form of sampling weights is thus a design weight, which in the case of most sampling designs is the inverse of the inclusion probability of an ultimate sampling unit. The probability of each sampled target person being selected is therefore derived on the basis of one- and two-stage sampling rules and population distributions.

For most practical situations, however, the design weight is not sufficient because after drawing the sample and collecting the data it may be discovered that the data are skewed due to errors in the sampling frame, sampling, non-response or measurement. Design weights have to be adjusted on the basis of various model assumptions. The weighting thus derived is called re-weighting. This is

a commonly used method by which the original design weights are adjusted by means of auxiliary information either about the population, the sample or both (Oh and Scheuren 1983, Särndal and Lundström 2005). The simplest re-weighting methods are called post-stratification and raking ratio adjustments. In post-stratification the sample is weighted according to known distribution information for the population. In individual-based surveys, weighting is done according to demographic information, such as age and gender groups and area of residence (Särndal et al. 1992, Djerf 2000).

The calibration of original design weights thus has two effects that adjust sampling weights:

– Adjusting for the effect of non-response on the final attained sample
– Generalising the final data to be representative for the target population of the survey.

19.1.1. Calibration of design weights

The original design weights were calibrated using the SAS-macro CALMAR (Sautory 1993). Due to the multi-stage nature of the survey data, sampling weights were produced for target persons on four different levels as follows:

Main survey (persons aged 30 or over)
Survey of young adults (aged 18–29)

– All respondents: participated in any section of the survey or separate non-response interviews or inquiries,
  Main survey: n = 7,415; Survey of young adults n = 1,710
– Respondents' union: participated in any section of the survey,
  Main survey n = 7,112; Survey of young adults n = 1,505
– Nutrition: participated in several sections and especially in the nutrition inquiry, Main survey n = 6,005; Survey of young adults n = 789
– Intersection: participated in interviews or corresponding inquiries and in most clinical sections and inquiries, Main survey n = 5,482; Survey of young adults n = 1,292

It should be noted, however, that in addition to groups above it is possible for different kinds of respondent groups to develop in various survey variables or even in entire sections. The clearest example is the inquiry on mental health, which was only available in Finnish. Therefore only those people were selected as respondents who understood the questions well enough, regardless of their native language. As

far as is known, item non-response in other sections occurred mainly for random reasons (e.g. due to temporary faults in the measuring instruments).

The design weights were calculated separately for these different groups. These weights were based on stratum and cluster-specific inclusion probabilities calculated for the largest respondent group above, i.e. some accepted response to any phase of the study. First, the number of clusters $n = 80$ was divided into university hospital strata using equal allocation: $n_h = 80/5 = 16$. However, the stratum-specific sample sizes were defined with proportional allocation calculated from the population size: $n_h = n\ (N_h / N)$. The sample size was further divided into two types of samples: the 15 largest towns with proportional allocation and two-stage cluster design. The inclusion probabilities for the 15 largest towns are those using simple random sampling in each stratum:

$$\pi_{h1} = \frac{n_{h1}}{N_h}\frac{N_h}{N_1}$$

and the design weight is obtained from the inverse: $w_{h1} = \frac{N_1}{n_{h1}}$

The inclusion probabilities for two-stage sampling design are product of two inclusion probabilities. Let f_1 denote the primary sampling unit (cluster) selection carried out with probabilities proportional to size (PPS) sampling, and respectively f_2 the inclusion probability of the elements from the selected cluster:

$$\pi_{h2} = f_1 f_2 = \frac{N_c n_{h2}}{N_{h2}} \times \frac{m_c}{N_c} = \frac{n_{h2} m_c}{N_{h2}}$$

where N_h is to the population size in stratum h, N_c is the population size in cluster c, n the number of clusters to be selected and m_c cluster-specific sample size. Sub-index 2 refers to second-stage sampling where the population counts were adjusted not to include those of the 15 largest towns. The sampling weight is obtained by

$$w_{h2} = \frac{N_{h2}}{n_{h2} m_c}$$

At the weighting stage the sample weights were calculated for the four weighting groups listed above so that the sample size was replaced by the number of respondents by stratum. In the case of the main survey it was performed separately for respondents under age 80 and for those aged 80 or over to adjust for the oversampling of the latter group. The purpose of this weighting strategy was to keep the additional variation of initial weights caused by non-response as harmless

as possible. The weight was constant inside each weighting group, and the primary variation of design weights was caused by stratum-specific differences between individuals aged under 80 and those over 80.

The following variables and demographic information about sample persons derived from the sampling frame were used for calibrating the weights of the health interview data:

- Design weight based on adjusted inclusion probability
- Health centre district indicator
- University hospital district indicator
- Age (10 year categories in main study, 3 categories for study of young adults)
- Gender
- Native language (2 categories)

Table 19.1. Weight average, variation coefficient and correlation between weights of the same group in the Health 2000 Survey.

Main study group (n)		average	coefficient of variation (%)	correlation
All	– design weight	438.9	15.9	
(7,415)	– calibrated weight	438.9	16.8	0.947
Union	– design weight	457.6	15.7	
(7,112)	– calibrated weight	457.6	16.7	0.939
Nutrition	– design weight	542.0	9.5	
(6,005)	– calibrated weight	542.0	13.8	0.692
Intersection	– design weight	593.6	9.6	
(5,482)	– calibrated weight	593.7	17.6	0.325
Survey of young adults group (n)		**average**	**coefficient of variation (%)**	**correlation**
All	– design weight	444.6	12.6	
(1,710)	– calibrated weight	444.6	18.1	0.692
Union	– design weight	505.1	17.4	
(1,505)	– calibrated weight	505.1	22.2	0.782
Nutrition	– design weight	963.5	24.0	
(789)	– calibrated weight	963.5	39.6	0.604
Intersection	– design weight	588.4	19.7	
(1,292)	– calibrated weight	588.4	25.7	0.765

Calibration brought additional information into the survey design and consequently increased the variation of weights. In the first two groups the distributions for the respondents and the sample are relatively close to each other, and therefore the weights did not change very much. However, for those taking part in the nutrition survey and those in the intersection group, the additional information had a fairly large effect. In these groups the proportion of elderly respondents fell sharply, which is visible in the small coefficient of variation of the design weights. Since the calibration technique returns the relations between groups correctly relative to the different background variables, the variation of weights can significantly increase.

The variables used in calibration may influence the final estimates obtained through weighting. The selection of auxiliary variables for weighting in the Health 2000 Survey required critical consideration as the data will be used for various kinds of surveys and diverse and complex analyses. The decision was made not to use derived variables such as socio-economic group, but only basic demographic variables in calibration: this was in order to retain the usability of the data over time. The use of too many weighting cells should also be avoided in the calibration of sampling weights; therefore language groups were combined and the size of age groups was kept sufficiently large.

19.1.2. Recommendations on the use of weights

The four expansion weights described above inflate the sample observations to the level of the population. In addition, we also need to have weights whose statistical features remain unchanged, but whose sum is the number of respondents and whose average is 1. A weight called analysis weight is derived for each weight coefficient. The use of analysis weights is recommended for most examinations containing averages and based on modelling and similarly, expansion weights for estimating population and subgroup totals.

It is difficult to recommend any one of the derived weight coefficients for overall use because data from the Health 2000 Survey are used for a wide variety of purposes. A rule of thumb could be that in order to achieve optimal weighting, the features of the studied group should be as close as possible to one of the above-mentioned groups. Therefore, studies concerning the nutrition questionnaire, for example, should always use the weight calculated for that group. Similarly, when examining a combination of several variables, item non-response in different variables cumulates, whereby the group studied should instead be a group involved in the calculation of the intersection weights. The “all” weight describing maximum participation should probably only be used in analyses derived from registers and

concerning mainly background information. The most useful weight for research purposes will therefore be either a union weight or an intersection weight.

19.2. Sampling variance and standard error estimation

The sampling and estimation design in the Health 2000 Survey may be described as complex in that most of the elements were selected using two-stage stratified cluster sampling. In the case of complex sampling designs the derivation of the formulas for the sample variance of point estimates requires proper attention to the sample design. The sampling design of the Health 2000 Survey was described under Chapter 3.2. To derive the appropriate variance estimators, the data were adjusted so that the method of one-stage cluster sampling that was applied to the 15 largest towns is separated from two-stage cluster sampling. This is described in more detail below.

The sampling design for the 15 largest towns was one-stage stratified element-level sampling. Therefore a technical adjustment was made to the survey data by which each of these towns forms a stratum and each survey respondent is treated as a cluster. For example, the number of people in these towns who participated in the survey was 2,695, whereby the number of degrees of freedom in the variance estimator is 2,680 in the whole country (2,695 clusters/elements−15 strata).

In the actual two-stage part, each stratum has its own university hospital district code. The first sampling stage involved sampling the health centre districts (clusters) and the second stage the sampling of persons from the selected clusters. The clusters were consecutively numbered to make processing as simple as possible. In the whole country the number of these clusters selected by PPS-type sampling was 65, so the number of degrees of freedom in the variance estimator formula is: number of clusters – number of strata, i.e. 65−5=60.

Design-based estimation is described in modern sampling theory textbooks, such as Chambers and Skinner (2003), Lehtonen and Pahkinen (2004), and Lohr (1999), as well as in the manual of SUDAAN, one the most widely used analysis programs (Research Triangle Institute 2004).

Analytical derivation of sample variance estimators is not always possible for complex sampling designs. The parameters to be estimated or their functions, for example on the level of sub-populations, are often non-linear. In the case of multi-stage cluster sampling it is necessary to use approximation, which may be based either on Taylor series expansion or sample reuse techniques. These methods are mentioned in all the volumes listed above.

Next, we present one of the most commonly used approximations for sampling variance in the case of a complex design. We assume that the sample was selected with replacement, but with unequal inclusion probabilities. The approximation takes the between-cluster variation into account, but neglects the intra-cluster variation. We must have at least two clusters per stratum, and we anticipate the sampling fraction of clusters to be small. For estimation we need the following information on the design: stratum and cluster identifier and sampling weight (final weight).

The estimator for the population mean of some study variable is

$$\hat{\bar{Y}} = \sum_{h=1}^{H} \sum_{i=1}^{n_h} \sum_{j=1}^{m_c} w_{hij} y_{hij} \Big/ \sum_{h=1}^{H} \sum_{i=1}^{n_h} \sum_{j=1}^{m_c} w_{hij}$$

where indices *h* defines strata, *i* clusters and *j* individuals.

The variance estimator for the mean using the Taylor series approximation can be presented as:

$$\hat{V}(\hat{\bar{Y}}) = \sum_{h=1}^{H} \sum_{i=1}^{n_h} \frac{n_h}{n_h - 1} \left(Z_{hi} - \bar{Z}_h \right)^2$$

where $Z_{hi} = \sum_{j=1}^{m_c} w_{hij} \left(y_{hij} - \hat{\bar{Y}} \right) \Big/ \sum_{h=1}^{H} \sum_{i=1}^{n_h} \sum_{j=1}^{m_c} w_{hij}$,

and $\bar{Z}_h = \sum_{i=1}^{n_h} Z_{hi} / n_h$.

The standard error estimates are obtained by:

$$StdErr(\hat{\bar{Y}}) = \sqrt{\hat{V}(\hat{\bar{Y}})} \ .$$

In sample surveys the efficiency of the complex design can be compared with that of simple random sampling. As a measure we use the design effect:

$$deff\left(\hat{\theta}\right) = \frac{\hat{V}(\hat{\theta})_{p(s)}}{\hat{V}(\hat{\theta})_{srswr}}$$

The cluster sampling design typically suffers from homogeneity of responses within the clusters, called intra-class correlation, i.e. variation is smaller than when simple random sampling is used. Positive intra-class correlation yields design effect estimates that are typically larger than 1. However, analysis according to various

sub-groups which do not coincide with the sampling design result in the reduced design effects, even close to one.

The table below shows some basic health variables from the Health 2000 Survey tabulated according to the basic design feature: simple random sampling in the largest towns vs. two-stage cluster design. As we can assume, the SRS design yields design effects equal to one, whereas the effects of the complex design differ considerably. The analyst should not assume that the study variables are free from intra-class correlation, but the data should be analysed with programs that can take those design complexities into account. For data analyses, therefore, it is safest to use software that is capable of taking into consideration the survey sample design, such as SUDAAN, certain SAS or SPSS procedures or Stata, or to apply methods of analysis that take account of the hierarchic nature of the data when constructing the model (for more details see Lehtonen et al. 2003).

Table 19.2. Some study variables of the Health 2000 Survey (main study) by sampling design. Union weigths were applied.

Study variable Design	Number of observations	Mean	Standard error for the mean	Design effect
Diastolic blood pressure (mmHg)				
– all	6,334	81.52	0.29	4.41
– 1-stage	2,468	81.23	0.21	0.96
– 2-stage	3,866	81.70	0.46	6.39
Systolic blood pressure (mmHg)				
– all	6,336	133.49	0.43	2.72
– 1-stage	2,468	131.15	0.40	0.99
– 2-stage	3,868	134.99	0.65	3.74
Visits to doctor for disease				
– all	6,957	0.74	0.01	1.27
– 1-stage	2,685	0.76	0.01	1.03
– 2-stage	4,272	0.72	0.01	1.42
Health status: good				
– all	6,986	0.61	0.01	1.18
– 1-stage	2,695	0.65	0.01	1.01
– 2-stage	4,291	0.59	0.01	1.28
Chronic disease: exists				
– all	6,981	0.53	0.01	1.68
– 1-stage	2,693	0.48	0.01	1.00
– 2-stage	4,288	0.56	0.01	2.12
Body mass index				
– all	5,979	26.59	0.06	1.17
– 1-stage	2,445	26.14	0.09	1.02
– 2-stage	3,534	26.90	0.08	1.25
Waist circumference				
– all	6,289	92.92	0.18	1.23
– 1-stage	2,449	91.48	0.27	1.02
– 2-stage	3,840	93.83	0.25	1.38

19.3. Statistical analysis

Since the Health 2000 Survey employed a complex sampling design, it is obvious that statistical analyses should be carried out in such a way that its effects can reliably be taken into account. Certain statistical programs can account for design complexities and, furthermore, one can introduce those effects directly in statistical analysis. Some alternatives are outlined below.

19.3.1. Study designs and statistical methods

Data from the Health 2000 Survey will probably be used primarily in cross-sectional studies, cohort studies and embedded case control studies. For cross-sectional analysis, the most often used methods include linear and generalised linear models, which are discussed under 19.3.4 below. It is important to select a method that takes into account the design complexities (stratification and clustering) as well as sampling weights.

Results from Health 2000 and its predecessor, the Mini-Finland Survey, are compared in the same way as in a cross-sectional survey. When the comparison concerns basic statistics, e.g. descriptive statistics, or those applying model-standardisation, we encourage the use of expansion weights which adjust for the effects due to unequal inclusion probabilities and both unit and item non-response. Both surveys contain proper sets of weights. For cohort studies one of the basic methods is probably the Cox proportional hazards model. Embedded case control studies often apply conditional logistic models. These are not dealt with in this chapter, but the reader is advised to consult the extensive literature.

19.3.2. Sub-population analysis

The analysis of sub-populations may sometimes cause problems, particularly when analysing a very small sub-population or rare event. The combination of one- and two-stage cluster sampling in the Health 2000 Survey may also lead to similar problems when comparing data from the 15 largest towns with other geographic domains selected with a two-stage design. Some specific studies in the Health 2000 Survey were based on screening and thus resulted in a smaller subset of respondents. If the analysis is confined to those observations, the risk of unreliable results will be increased. So far there have been no indications of any dramatic problems in sub-population analyses, but it is still better to be safe. Fatal errors may result from improper analysis (see e.g. Korn and Graubard 1999, pages 207-

211). Such problems can be avoided, first of all, by always keeping all observations in the dataset and selecting the observation for analysis using the proper statement in the analysis program. For example, the SUDAAN program has the option ”SUBPOPN”, SAS: “Domain” statement, and Stata likewise ”SUBPOP()”,avoiding the use of options like ”IF…”, or ”IN..”.

If the analysis is restricted to a very small subset, the results cannot be generalised to the appropriate population. For example, the analysis may yield unreliable results due to such factors as native language, small geographic region or rare event. The response rate for Finnish or Swedish-speaking citizens was much higher than for people with some other native language. Similarly, the number of respondents from smaller health care centre areas may be too small or contain selection bias for any comparative analysis (see Chapter 19.1.1 above).

19.3.3. Model adjustment

Model adjustment by using predicted margins is a method for assessing differences in means or prevalences caused by some factor while controlling for the effects of confounding factors by using a linear or logistic regression model (Lee 1981). The model-adjusted mean (predictive margin) is calculated as an average of individual predicted means

$$\mathrm{PM}_{\mathrm{linear}} = \frac{1}{n}\sum_{i=1}^{n} X_i\beta$$

and the model-adjusted prevalence as an average of the individual predicted probabilities

$$\mathrm{PM}_{\mathrm{logistic}} = \frac{1}{n}\sum_{i=1}^{n} \frac{\exp\{X_i\beta\}}{1+\exp\{X_i\beta\}} \quad ,$$

where i indexes the n individuals, β denotes the column vector of the regression coefficients and X_i denotes the row vector of the corresponding covariate values of individual i.

In calculating predictive margins, the values of some covariates are modified by setting them to the fixed values chosen by the researcher. Confounders, which are included in the model as the other covariates, would have no effect on the difference of these predictions, because the distribution of the confounders would be the same for each prediction.

For example, let the binary outcome variable be high social participation, and assume that education, age and gender are included in the model as covariates. By setting education = high over all individuals, the resulting model-adjusted prevalence has the interpretation "if everyone had high education, then the prevalence of high social participation would be $\mathrm{PM}_{\mathrm{logistic}}^{\mathrm{education=high}}$". Similarly, by setting education = basic, we would get the predicted prevalence of the opposite scenario, "if everyone had basic education, then the prevalence of high social participation would be $\mathrm{PM}_{\mathrm{logistic}}^{\mathrm{education=basic}}$". A comparison of the predictions $\mathrm{PM}_{\mathrm{logistic}}^{\mathrm{education=high}} - \mathrm{PM}_{\mathrm{logistic}}^{\mathrm{education=basic}}$ would then illustrate the effect of education on the prevalence of high social participation. The potential confounding effect of age and gender would be adjusted, because the age and gender distribution would be exactly the same for both predicted prevalences $\mathrm{PM}_{\mathrm{logistic}}^{\mathrm{education=high}}$ and $\mathrm{PM}_{\mathrm{logistic}}^{\mathrm{education=basic}}$. The variance estimators in the case of complex sampling designs are based on Taylor series linearization and presented in Graubard and Korn (1999).

19.3.4. Multivariate analysis and software

To demonstrate multivariate analysis for the Health 2000 Survey data, we used a dataset with minimum editing and imputation and other confounding factors. The dataset consists of 5,954 observations. Our setting included a total of 2,495 primary sampling units (PSU), of which 65 are in the five two-stage strata, and the remaining 2,430 in the 15 certainty strata (where each individual constituted a "cluster"). A total of 2,475 degrees of freedom were thus available for design-based variance estimation in most cases considered. For the construction of the weights, we also made a simplified assumption of ignorable unit non-response. The analysis weights were re-scaled weights such that the mean was equal to one. Note that the results presented may differ from other publications of the Health 2000 Survey. The material for this section is based on Lehtonen et al. (2003).

The variable systolic blood pressure was selected for analysis. The variable indicated a relatively strong clustering effect. We fitted different models involving main effects and interactions constructed using some of the available potential predictors. Our substance matter interests were focused on predictors relating to physical conditions and socio-economic characteristics. Because we deal with a health-related variable, the effects of demographic variables (sex and age) were adjusted for. Our initial set of predictor variables was the following. Demographic variables consist of age (in years) and gender (1 male, 2 female). Education measured in years spent for completed education was used as a socio-economic variable. A variable related to physical condition is waist circumference (in centimetres). For illustrative purposes we also constructed new variables from the

education and waist circumference variables by grouping the observations into three nearly equal-sized groups (three-parts) according to the values of a given predictor.

We empirically compared multivariate survey analysis methods that allow for the inclusion of the complex sampling design effects in estimation and testing procedures. We were especially interested in methods that account for clustering effects. In generalised linear modelling under the so-called nuisance approach (Lehtonen and Pahkinen 2004), it is possible to attain design-consistent estimation and asymptotically valid testing. This approach uses "pseudo" maximum likelihood (PML) and related estimation techniques, "sandwich" type empirical standard error estimators and design-based Wald test statistics. Techniques based on generalised estimating equations (GEE) with multivariate quasi-likelihood (QML) estimation are used as a more advanced alternative (Liang and Zeger 1986, Diggle et al. 1994). In GEE, an assumption of an exchangeable correlation structure of observations in a cluster is applied. In these methods, the main inferential interests are in model coefficients; intra-cluster correlations as such are not necessarily of scientific interest. Multilevel or mixed models (Goldstein 2003, McCullogh and Searle 2001) involving fixed and random effects were used as another alternative. Cluster-specific random effects were incorporated in a model in order to account for the clustering effects. In this approach, intra-cluster correlations and cluster-level inferences are also often of scientific interest. We compared the methods empirically, by calculating point estimates and their estimated standard errors and design effects, as well as t-test and similar statistics. More detailed results are presented in Lehtonen et al. (2003).

Analysis options

To manage a comparison of the selected methods, we formulated a set of analysis options to be referred to in our multivariate analysis exercise (Table 19.3). With Options 1 and 2, based on a fixed-effects model, we used the generalised estimating equations (GEE) approach. The GEE method with an independent correlation structure (Option 1) relates to the standard PML method where observations are assumed to be independent within clusters for the estimation of the regression coefficients, but are allowed to be intra-cluster correlated in the estimation of the covariance matrix of the estimated regression coefficients (using a "sandwich", or robust, or empirical, variance estimator; e.g. Lehtonen and Pahkinen 2004, p. 285). In the GEE method assuming an exchangeable correlation structure (Option 2), observations are allowed to be intra-cluster correlated in the estimation of both the regression coefficients and the covariance matrix of the estimated regression coefficients, and QML estimation is used. With GEE estimation, we also wanted to examine whether we will end up with closely comparable estimation and testing

results with software sometimes labelled as "design-based" (such as SUDAAN), and software with a more "model-based" orientation (such as SAS). We assumed that this comparison might have some practical relevance for users of the Health 2000 database. The GEE methods are available in all SUDAAN modelling procedures and in the SAS procedure GENMOD. SUDAAN and SAS also allow for the incorporation of element weights in the estimation procedures. This method is sometimes called the weighted GEE method (e.g. Preisser et al. 2002).

Option 3 uses a mixed model formulation involving both fixed and random effects. For simplicity, we adopted a variance components model where cluster-specific random intercepts are included in the model in addition to the fixed effects. REML (residual, or restricted, ML) estimation was used for the estimation of the model parameters. A "sandwich" type robust standard error estimator was again used for the estimated fixed effects. We also incorporated element weights in our modelling in Option 3. These analyses can be carried out with standard statistical software products. For Option 3 we used the SAS procedure MIXED. We postulated a linear model for the continuous response variable and a logistic model for the binary response variable.

In the Reference Option, which serves as a baseline method, a standard fixed-effects model was postulated and simple random sampling with replacement was assumed. This option ignores all the sampling complexities, corresponding to an analysis with a conventional OLS or ML estimation of the regression coefficients and model-based standard error estimators. By definition, design effect statistics for this option are equal to one. The analysis can be carried out by any standard statistical software package (such as SAS or SPSS).

Table 19.3. ***Analysis options used in multivariate modelling exercise.***

Option	**Model formulation and estimation method**	**Aiming to account for…**		
		Weighting	Stratification	Clustering
Option 0	Reference option Fixed-effects model, ML, model-based SE method	No	No	No
Option 1	Fixed-effects model, PML, robust SE method	Yes	Yes	Yes
Option 2 a) and b)	Fixed-effects model, QML, robust SE method a) SUDAAN application, b) SAS application	Yes	a) Yes b) No	Yes
Option 3	Mixed model, REML, robust SE method	Yes	No	Yes

SE: Standard error
ML: Maximum likelihood estimation
PML: Pseudo maximum likelihood estimation
QML: Quasi maximum likelihood estimation
REML: Residual maximum likelihood estimation

To weight or not to weight in a complex analytical survey? There seems to be no unique solution to this important theoretical and practical problem (see e.g. Pfeffermann et al. 1998 with discussion). For demonstration purposes, however, we took a design-based position and incorporated the weights in all options except the Reference Option. We also motivated this choice by reasons of design consistency and because it provides some protection against possible model failure. The cost we had to pay was in a somewhat reduced efficiency. However, that effect was small.

Table 19.4 summarises some examples of software products and their analysis procedures that are available for options 1 to 3. Fitting regression models by GEE methods is discussed among others by Horton and Lipsitz (1999), who give a review of software for GEE estimation, covering options offered by SAS, Stata, SUDAAN and S-Plus software products; and by Ziegler et al. (1998), who present an extensive literature review on GEE methodology. Multilevel modelling using SAS procedure MIXED is discussed e.g. in Singer (1998).

Table 19.4. ***Software available for different types of variables under analysis options 1 to 3.***

Software product	Coverage of options	Type of response variable	Analysis procedure
SUDAAN (version 8.0.1)	**Option 1** GEE for fixed-effects models with independent correlation structure	Continuous	REGRESS
		Binary	RLOGIST
		Polytomous	MULTILOG
		Count	LOGLINK
	Option 2a GEE for fixed-effects models with exchangeable correlation structure	Continuous	REGRESS
		Binary	RLOGIST
		Polytomous	MULTILOG
		Count	LOGLINK
SAS (version 9.1)	**Option 1** GEE for fixed-effects models with independent correlation structure	Continuous	GENMOD
		Binary	SURVEYREG
			SURVEYLOGISTIC
		Polytomous	GENMOD
		Count	GENMOD
	Option 2b GEE for fixed-effects models with exchangeable correlation structure	Continuous	GENMOD
		Binary	GENMOD
		Polytomous	(not available)
		Count	GENMOD
	Option 3 REML for mixed models	Continuous	MIXED
		Binary	GLIMMIX
		Polytomous	GLIMMIX
		Count	GLIMMIX

Models

A linear mixed model specification can be written compactly as

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{v} + \boldsymbol{\varepsilon} \tag{1}$$

where $\mathbf{y}$ denotes the $n \times 1$ vector of survey variable measurements, $\mathbf{X}$ is the $n \times p$ design matrix for the fixed part of the model, $\boldsymbol{\beta}$ is the corresponding $p \times 1$ fixed parameters vector, $\mathbf{Z}$ is the $n \times q$ design matrix for the random part of the model, $\mathbf{v}$ is the corresponding $q \times 1$ vector of random effects, and $\boldsymbol{\varepsilon}$ is the $n \times 1$ residual vector. We make the standard assumptions that $\mathbf{v}$ and $\boldsymbol{\varepsilon}$ follow normal distributions with

$$\mathrm{E}\begin{bmatrix}\mathbf{v}\\ \boldsymbol{\varepsilon}\end{bmatrix} = \begin{bmatrix}0\\ 0\end{bmatrix} \quad \text{and} \quad \mathrm{Var}\begin{bmatrix}\mathbf{v}\\ \boldsymbol{\varepsilon}\end{bmatrix} = \begin{bmatrix}\mathbf{G} & 0\\ 0 & \mathbf{R}\end{bmatrix}$$

where $\mathbf{G}$ is the $q \times q$ variance-covariance matrix of $\mathbf{v}$ and $\mathbf{R}$ is that for $\boldsymbol{\varepsilon}$. Thus, the variance of y is $\mathrm{Var}(\mathbf{y}) = \mathbf{Z}\mathbf{G}\mathbf{Z}^T + \mathbf{R}$. Further, $\mathrm{Var}(\mathbf{y})$ can be modelled by setting up the random-effects design matrix $\mathbf{Z}$ and by specifying covariance structures for $\mathbf{G}$ and $\mathbf{R}$. A simple random effects model is a special case of the general specification with $\mathbf{Z}$ containing dummy variables, $\mathbf{G}$ containing variance components in a diagonal structure, and $\mathbf{R} = \sigma^2\mathbf{I}_n$, where $\mathbf{I}_n$ denotes the $n \times n$

identity matrix. The general linear model $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ is a further special case with $\mathbf{Z} = 0$ and $\mathbf{R} = \sigma^2 \mathbf{I}_n$. More details can be found for example in McCullogh and Searle (2001).

In model (1), we are especially interested in the beta parameters of the fixed part of the model. We therefore insert our subject matter predictors in the fixed part design matrix of model (1). The design matrix $\mathbf{Z}$ in the random part of (1) allows for the modelling of the clustering effects. Random intercepts will constitute the random part parameters in our simple modelling exercise. Thus, model (1) simplifies to

$$y_{hik} = \mathbf{x}_{hik}^T \boldsymbol{\beta} + v_{hi} + \varepsilon_{hik}, \tag{2}$$

where $\mathbf{x}_{hik}$ are the x-vectors and *h* refers to strata, *i* refers to clusters within strata, and *k* refers to elements within clusters. In our setting, the cluster-specific random effects v_{hi} and residuals ε_{hik} are assumed to be mutually independent and normally distributed with zero means and variances of σ_v^2 and σ_e^2, respectively. For systolic blood pressure (continuous response) we use linear fixed-effects models and linear mixed models. For chronic illness (binary response) we use fixed effects binomial logistic models.

For GEE estimation, consider a generalised linear model of the form

$$E_m(F(\mathbf{y})) = \mathbf{X}\boldsymbol{\beta} \tag{3}$$

where *F* refers to the link function and *m* to an expectation under the model. We specify an identity link function for our continuous response variables and a logistic link function for the binary response variables. In the GEE approach, we model the covariance structure of observations within clusters by $\mathbf{V}_{hi} = \phi \mathbf{A}_{hi}^{1/2} \mathbf{R}(\alpha) \mathbf{A}_{hi}^{1/2}$ where $\mathbf{A}_{hi}$ is a diagonal matrix of variance functions, α refers to the correlation of pairs of observations in a cluster, and ϕ is the scale parameter. For the independent correlation structure, α is set to zero, and the "working" correlation matrix $\mathbf{R}(\alpha)$ reduces to an identity matrix. For exchangeable correlation structure, α denotes the off-diagonal elements of $\mathbf{R}(\alpha)$. More details can be found in Liang and Zeger (1986) and Diggle et al. (1994).

Results

The main substance matter predictor for systolic blood pressure was the waist circumference variable. For illustrative purposes we used it as a three-category variable (RCIRCUM). In the analysis we adjusted for sex and age effects. It is known that mean systolic blood pressure tends to increase with age, and for a given age group, means increase with increasing waist circumference level.

Our aim is to examine the structure of the variation in the mean levels of systolic blood pressure according to the values of the predictors. A linear ANCOVA model was fitted under the five analysis options, including the main effects of all predictors as well as their pair-wise interaction terms. Ignoring the design complexities, the analysis under the reference option suggests that a reasonable model includes all the main effects and pair-wise interaction effects of waist circumference with age and sex. A similar model was also obtained under the options that account for the design complexities. However, under these options, the significance of the interaction effect of sex with waist circumference was weaker than for the reference option (Table 19.5).

Table 19.5. Testing the interaction effect of sex with waist circumference in modelling systolic blood pressure under selected analysis options.

	DF	F value	Prob.
Reference Option	2	5.51	0.0041
Option 1	2	5.24	0.0054
Option 2 a	2	4.69	0.0093
Option 2 b	2	4.65	0.0096
Option 3	2	4.72	0.0090

Let us examine in more detail the interaction of sex with waist circumference (Table 19.6). The most liberal results were given by the Reference Option. Options 1 to 3 give closely agreeing point estimates and estimated standard errors. However, for Options 2a, 2b and 3, the t-test statistics were slightly smaller than for Option 1. In Option 1, estimation is based on the PML method where clustering effects are accounted for in the estimation of the standard errors (but not in the estimation of the fixed effects parameters). In Options 2a, 2b and 3, clustering effects are accounted for both in the estimation of the fixed effect parameters and in standard error estimation, either by the GEE method with an exchangeable correlation structure or by fitting a linear mixed model with cluster-specific random intercepts.

Table 19.6. Testing the interaction effect of sex (males) with waist circumference (middle class) in modelling systolic blood pressure under the analysis options.

	Estimate	Standard error	Design effect	t-test (Studentized value)	Prob.
Reference Option					
sex(1).rcircum(2)	0.019	0.0088	1.00	2.14	0.0324
Option 1 (SUDAAN/REGRESS)					
sex(1).rcircum(2)	0.018	0.0091	1.06	2.02	0.0430
Option 2a (SUDAAN/REGRESS)					
sex(1).rcircum(2)	0.017	0.0091	1.05	1.87	0.0614
Option 2b (SAS/GENMOD)					
sex(1).rcircum(2)	0.017	0.0089	1.02	1.90	0.0573
Option 3 (SAS/MIXED)					
sex(1).rcircum(2)	0.017	0.0089	1.02	1.93	0.0532

Discussion

Health surveys often have complex sampling designs, involving stratification, clustering and unequal inclusion probabilities, as well as imputation and re-weighting schemes in the estimation design. These surveys require statistically sound and practically manageable strategies of analysis. Our aim here was to compare the methodologies available and to apply the corresponding computational tools in a fairly simple analysis setting for a recent dataset collected in Finland in the Health 2000 Survey. It appeared that to a reasonable extent, the major design complexities can be accounted for by using appropriate computational tools that are readily available in commonly used statistical software products.

We concentrated on methods that permit the inclusion of the most important survey design complexities in estimation and testing procedures. We were especially interested in clustering effects. We therefore constructed a setting where complicated re-weighting and imputation schemes were excluded. The weight variable was a design weight with only minor manipulation. There was only little variation in the weights, and therefore weighting had only minor effect on the results of the analysis. The selected survey variable indicated relatively strong positive intra-cluster correlation.

The analysis options covered methodologies commonly used in complex surveys. The main differences in the methodologies were in the inferential framework, model formulation, estimation strategy and software application. We applied linear fixed-

effects type modelling and modelling with mixed models. Estimation techniques were used that are constructed to be flexible enough for modelling situations of varying complexity. In addition, our choice of computational tools was such that the methods are readily available in standard statistical software products. We used software products under the headings of "software for design-based analysis" and "software for model-based analysis".

Consistent numerical results, and similar inferential conclusions, were obtained under the different analysis options. This suggests that to a reasonable extent, the clustering effects can be accounted for by any of the methods examined (except the reference method which ignores all the sampling design complexities). The methods using pseudo-likelihood and generalised estimating equations under the design-based approach cover many typical analysis situations for generalised linear modelling. The use of mixed or multilevel models under the model-based approach would allow for more complex modelling of covariance structures.

20. COMMUNICATIONS

Sanna Natunen and Arpo Aromaa

The aim of communications in the Health 2000 Survey was to achieve maximum visibility and positive exposure in the early stages of the field examinations and to sustain this image throughout the project. This was intended to increase awareness of the survey among the general public and in this way to encourage more active participation.

An important part of communications to the general public and to the participants in the survey consisted of information to national and local health authorities as well as to employers. This was successful in that employer associations recommended that staff be allowed to attend the health examination during working hours.

20.1. Organisation

Communications for the Health 2000 Survey was centrally coordinated by the Department of Health and Functional Capacity (TTO) at the National Public Health Institute (KTL). Its tasks involved communications to the national media during the early stages of the field examinations, coordination of internal communications to the field teams, including weekly news and new instructions, and information to the subjects.

Once the examinations got underway, the field teams themselves assumed greater responsibility for communications. Local communications on arrival to a new area was the responsibility of each field team's head nurse. They contacted local radio stations and newspapers in advance to gain maximum exposure and coverage.

20.2. Channels

National communications was done via both nationwide and regional newspapers and television and radio news. A press release had been prepared for the field teams that they could use for local communications, which was primarily through local newspapers and radio stations. Photo and interview requests from the media were accepted. Communications between personnel out in the field and project headquarters was by e-mail or phone.

The subjects were approached by letter and project brochures. The brochures prepared for those over 30 and for young adults (18–29 years) were designed to

appeal to these particular age groups, and therefore they had some differences. The services of an outside agency (Pentagon Design Oy) were secured to assist with the design process. The same agency also designed the project logo (see the cover of this report), which was intended to give added credibility to the project and create a coherent visual image that was used in all communications. The logo was used in all printed matter, communications materials as well as on other incidental material such as pens and plastic bags.

The Health 2000 website continues to serve as an important channel of communication (http://www.terveys2000.fi). The website carries a wide range of information about the survey, including the forms used in data collection and a series of images from different stages of the field examinations. It also contains an account of all research plans and research results. The website is continuously updated, and it also provides a search facility for survey publications.

20.3. Partners and sponsors

The interests of the survey's various partners were also taken into account in communications. The survey was carried out jointly with a number of national research institutes (see Chapter 2), and it was funded by several social insurance institutions and other organisations. All of these partners are listed in the project's publications and on its website. The survey for young adults was sponsored by a number of business companies that donated prizes for draws among the participants. The logos of all partners were printed on the brochures and other materials handed out to the subjects.

20.4. Exposure and publicity achieved

Media monitoring services for the project were provided by Oy Observer Finland Ab Media Intelligence, who tracked all media hits, collected relevant newspaper clippings and supplied monthly reports to headquarters on the media exposure received by the project.

In August 2000, before the actual field examinations got underway, project headquarters issued a few advance media releases on the survey. The launch of the field examinations in October-November received good coverage in the electronic media as well as in the nationwide and local press. According to statistics compiled by Observer Finland, the volume and reach of publicity received by the Health 2000 Survey were at their highest immediately after the field stage got underway in autumn 2000 and at the beginning of 2001. A second peak in project publicity

was achieved in spring 2001 when information was released on the focus turning to young adults and finally on the conclusion of the survey.

The first media releases on the project's results were issued and the first information briefing was held in connection with the publication of the basic results report at the beginning of June 2002. The information briefing was a great success: the survey featured on the main news broadcasts of all the national television channels and on several radio channels. The next day, national and local newspapers carried prominent stories of the survey. Helsingin Sanomat, the country's leading daily newspaper, carried a leader on the project. All news items and other articles were objective and positive.

Journalists have shown a keen interest in the results and observations of the survey. The results published by the survey have frequently been in the news, and the survey itself has been mentioned in several news items and articles. Monitoring and tracking results show that newspapers and magazines continue to be interested even after seven years from the field stage.

Overall, project communications has been highly successful. It gained extensive visibility and exposure in both the national and local media. The main channel of communications was via local newspapers, but the project also achieved considerable publicity in national and regional papers. Trade journals also took note of the launch and progress of the survey.

21. IN-DEPTH EXAMINATIONS

Antti Reunanen

Despite the vast scale of Health 2000, the survey did still not cover everything that the project group would have wanted to cover. Partly this had to do with the tight schedule of the main survey, which could not be extended without putting undue burden on the participants; partly it was because some examinations require special methods that simply would not have been feasible within the survey context. Planning for some of the in-depth examinations only got under way after the master plan for the main survey had been completed.

All in-depth examinations were carried out on people who had participated in the main survey. They consisted either exclusively of questionnaires or interviews, or the participants were invited to take further examinations after the basic health examination.

The most extensive in-depth examinations concerned cardiovascular diseases, diabetes, mental health and oral health. The key aspects of the in-depth examinations that have been started up to date are shown in Table 21.1.

In the in-depth cardiovascular examinations, the subjects underwent an ultrasound examination of carotid arteries to detect any early changes of atherosclerosis in arterial walls, and an oral glucose tolerance test. Initially the plan was to incorporate both of these tests in the health examination, but for practical reasons they had to be omitted. The sample size was determined in such a way that if the projected re-survey is carried out in a few years' time, the expected changes in atherosclerotic indicators can be statistically significant.

In-depth mental health examinations consisted of an extensive survey of people who had experienced psychosis (PIF) and a mental health survey of young adults (NAPS). For the former, all those subjects were invited to a neuropsychological examination who had shown indications of an earlier psychosis. The latter examination focused on young adults who had previously completed only the health interview and questionnaire. A post questionnaire on mental health was sent to all participants, and on the basis of the responses some people were invited to a psychiatric interview.

To assess short-term changes in oral health, a sample of participants who had attended the initial health examination were invited to a repeat study on average four years later. The repeat examination was structured in the same way as the baseline examination. In addition to tracing changes in oral health, another aim of the repeat survey was to assess the operation of the dental health care system: a

major overhaul of that system had taken place in-between the two surveys, and the purpose was to see how effective that overhaul had been.

Table 21.1. In-depth examinations in Health 2000 Survey.

Examination	Sample	Main areas of interest
Complementary investigation of circulatory system (SVT+D)	Persons aged 45-74 living near central university hospital N: 1,526	Ultrasound examination of carotid arteries Oral glucose tolerance test Long-term ECG recording, endothelial function and hemodynamics
Psychosis in Finland (PIF)	Persons who had suffered psychosis N: 897, of whom 174 controls	Neuropsychological examinations
Mental health of young adults (NAPS)	Total sample of young adults N: 1,710	Mental health questionnaire, psychiatric interview and post questionnaire
Monitoring of oral health	Sample of 1,500 drawn from participants in baseline survey	Examination of oral health similar to baseline examination and inquiry of the use of dental services four years after baseline
Arthrosis and osteoporosis validation	N: 130 persons from the surroundings of Kuopio	DXA examination to validate calcaneal bone ultrasound results
Prevalence of celiac disease	All participants in baseline survey N: 6,403	Serum antibody examination and small intestine biopsy for subjects with positive findings
Health fitness test	Persons participating in baseline survey in Tampere region N: 660	Establish reference values for fitness tests
Eyesight in Finland	Everyone with eyesight problems in baseline survey N: 677	Questionnaire, inspection of medical records
Bioimpedance validation	Sample of people examined in Kuopio region	Repeat bioimpedance measurements and DXA measurement one year after baseline survey

22. DISCUSSION AND CONCLUSIONS

Arpo Aromaa

The main object of the Health 2000 Survey was to collect information for health policy and social security purposes as well as for research on public health and its promotion. The survey covered a broad range of issues, from living conditions and living habits, risk and protective factors, symptoms, findings indicative of disease, the treatment of illnesses, factors affecting work ability and functional capacity, to the use of aids and the need for help. Indeed in terms of content and coverage this was a more comprehensive and up-to-date survey than any previous corresponding study anywhere in Europe or in any OECD country outside of Europe. It also addressed current health policy issues and needs for scientific information more directly and more effectively than national health surveys do on average. The biggest planning challenge was collecting the large and diverse dataset within rather strict time constraints.

Overall, the Health 2000 Survey was quite an exceptional undertaking, even by international standards. The previous corresponding survey in Finland, the Mini-Finland Health Survey, was conducted by the Social Insurance Institution (SII, Kela) 20 years ago. The National Public Health Institute (KTL) had been setting aside funds for the purposes of the Health 2000 Survey for several years, although the bulk of the necessary funds were provided by the various partners involved. Without the contribution of these partners, the survey would not have been possible.

Earlier similar surveys in the United States (NHANES) and Finland (Mini-Finland Health Survey, Aromaa et al. 1989b) have had the use of specially commissioned mobile clinics with which the necessary equipment has been carried from one examination site to the next and where examinations have been conducted by a permanent trained staff. For the Health 2000 Survey, the staff were specially recruited and trained and the necessary equipment was selected and acquired for this specific purpose. The health interviews at people's homes and at institutions were conducted by Statistics Finland's professional interview staff.

Five field teams were set up to conduct the health examinations. The equipment they needed was transported by lorries from one examination site to the next. Many of the instruments, particularly those used in functional capacity measurements, were quite subtle and susceptible to breakage during transport in cold winter conditions. This caused some loss of measurements, even though every necessary preparation was made for immediate repairs and replacements.

The field staff for the health examinations were recruited by advertising the vacancies in newspapers in June 2000. Their three-week training started in August. As such the amount of time available for training was certainly adequate, but it would have been important for each field team to have the chance subsequently to run their own pilot survey followed by a one-week training and feedback period before launching the examinations proper. Indeed, this is what would have been done had it not been for the scarcity of resources, which forced the decision to abandon these plans. In any event additional training would definitely have been beneficial.

The health interviews started in August, which is also when the training for interviewers took place. All the Statistics Finland interviewers were competent and experienced people, yet many of them had only limited experience of health surveys. In this light, the one to two days of training that they received was not enough. In future surveys it is necessary to allow for several days of training at the outset and at least one or two additional days of refresher training during the survey.

At each of the localities visited by the health examination teams of 16–17, it was necessary to have the appropriate facilities, complete with waiting rooms and refreshment facilities. In most cases these facilities were secured at local health centres or other municipal buildings. In some municipalities the local authorities arranged for alternative rental facilities, and in some instances appropriate facilities had to be rented on the open market. For this reason the rental costs varied widely from case to case. An important search criterion was always that the facilities should always have disabled access, which indeed they usually did. However in a few cases the facilities were on the first floor of a building with no lift.

One of the most critical success factors for the survey was communications, which played an important part in promoting a favourable attitude among both the participants and their employers as well as in obtaining funding and the support of various organisations. Initially the main emphasis was on direct oral communications, later the accent shifted to written communications to the media and to the participants. Responsibility for communications was delegated early on from the survey management team to the head nurses in the field team, which proved to be a good decision. Overall the project was highly successful with its communications, as is evidenced by the fact that in 2005, the mass media were still familiar with the Health 2000 Survey as a concept.

The subjects were treated individually and with every courtesy and respect. It was believed that this would have the added benefit of encouraging other people to take part in the examination. The feedback collected suggests that the participants' experiences overall were very positive.

22.1. Sample

For practical reasons of organising the health examinations, they had to be carried out in selected localities so that there was a sufficient number of subjects at each site. Therefore the sampling design was a two-stage cluster sample comprising 80 examination sites and in smaller localities 50–100 subjects, which corresponded to 3–5 days' work. The sample was compiled by Statistics Finland to be representative of continental Finland. The probability of persons aged 80 or over being included in the sample was twice as high as in other age groups; this was to ensure that a sufficient number of older people were included. The sample consisted of two age groups: those aged 18–29 and those 30 or over. The participants were systematically sampled from the population register by the SII (Kela). In the interpretation of the results weights are used that take account of the sample structure and aim to adjust for the effect of non-response. The results, therefore, describe the situation in the whole Finnish population in 2000. However the complex sample structure did cause some difficulty to the researchers.

22.2. Planning stage of the survey

It is impossible to overestimate the importance of the planning stage to the outcome of the survey. Planning got underway in late 1998 and started officially with a joint meeting of the various planning teams in March 1999. In February and April 2000, two pilot surveys were carried out to test the overall design and the content areas covered. At this stage the content of the survey was still incomplete; it was only completed just in time before the launch of the field examinations. Ideally, the project should have had much more time to plan and finalise the contents and the procedures for the survey. There are several important considerations: 1) while the involvement in the project of a large number of partners meant it had access to a wide range of different skills and competencies, this also slowed down the overall process; 2) the planning team included a number of people with experience of major population surveys, particularly the Mini-Finland Survey, which undoubtedly helped to speed up completion; and 3) some of the time problems experienced with programming could have been resolved by hiring more people, but that was no longer feasible in the latter stages of the project. In spring 2000 it was no longer possible to extend the programming team's deadlines, both for financial reasons and because the staff, equipment and facilities were all booked and set to go.

In any event the complex machinery was successfully launched on schedule, even though both programming and training did find themselves working under considerable time pressure.

22.3. Preparations for fieldwork

The most important preparations ahead of the fieldwork were to contact the chief physicians at the local health centres and to arrange for the necessary facilities. Both of these were the responsibility of the project field manager and were successfully completed: attitudes to the survey were extremely positive throughout, and in most cases the necessary facilities were made available either at the local health centre or in some other municipal building. A major factor in the implementation of the health examinations themselves was that the forms and examination files had be compiled in advance at project headquarters.

22.4. Stages of the examination

The health examinations were conducted in the population aged over 30 primarily in two stages. First, home-visit interviews were conducted, and in this connection appointments were made for a health examination a few weeks later. At the health examination, the subjects were to return the questionnaires they had been given in connection with the interviews. An infection questionnaire was completed during the health examination. A third, complementary questionnaire was taken in connection with the health examinations, and a diet questionnaire was given for the subjects to complete at home and to return to KTL. To minimise non-response, short health examination visits were made to those people who had not attended the health examination proper. Furthermore, telephone interviews were conducted with non-participants, who also received brief questionnaires to complete.

More than 89% of the sample took part in the health interviews, which is a very high figure. Part of this can be attributed to the professionalism of the interview staff, part of it perhaps to good communications, but there is no doubt that the forthcoming detailed health examination was a very major incentive.. Appointments for the health examination were made in connection with the interviews. At this point it was also made clear that if the subjects needed a taxi ride or some other form of transportation, the costs would be covered by KTL. All in all, 79% of the sample took part in the health examinations; when the home visit examinations are included, the figure rises further to 84%. Because of mobility problems the rate of participation in the age group 85 or over was somewhat lower, but even in the oldest age group the home visits helped to push up the participation rate to almost 70%.

In the light of these statistics, the average rates of around 70% for domestic questionnaire studies or just 50–60% in many European national surveys probably mean that their results are distorted by selection bias. Because of the high

participation rates in the Health 2000 Survey, the chances of accurate and reliable results are much higher than normally.

22.5. IT environment

Data collection for the survey took place both at the subjects' homes and at the health examination sites. The aim was to collect most of the data in electronic format and in such a way that the need for checks and corrections after the survey would be minimised and that the data would be available for research uses as quickly as possible. Most of the data were entered directly on laptops, mainly by means of electronic forms created using the Blaise system, a software package developed by Statistics Netherlands.

The interviewers sent back the data they had collected in encrypted form via modem to Statistics Finland, where they were compiled into a database. At each measurement and recording point in the health examinations, the field teams had laptops that were connected to each other via a wireless local area network. The teams had access to an ISDN connection that they used to download updates and other materials and to send out encrypted data to project headquarters, where the material was compiled into the central database. On each team, one of the members was appointed to take charge of IT operations.

The field teams' computers were backed up, and all the data collected during the day was supposed to be automatically transferred in the evening onto the main server. A small amount of data was lost due to hardware and software malfunctions; some of these problems could have been avoided by better training. Some problems were also caused by the system accepting erroneous identifications and dual entries, which had to be addressed after the completion of data collection. During the checks process it also transpired that some of the data collected had never been sent back to KTL at all. However, these data were recovered from the servers and were therefore not lost. Further precautions are needed to prevent the recurrence of these kinds of problems in the future.

The main difficulties for the field teams were caused by the lack of training and the late arrival of the software. Even though the bugs in the programs were resolved during the first few days and weeks of the fieldwork, they did cause some loss of data. Delays in telecom service provision also caused some difficulties and disruption to the project: every time they arrived at a new location, the field teams had to contact the service provider several times to make sure that the connections had indeed been installed as agreed.

Many checks and corrections still had to be made after data collection. The only way to effectively reduce this extra workload is to make sure that the forms are ready and available well ahead of time so that they can be tested and so that any problems and errors can be addressed at a much earlier stage than was possible in the present case.

22.6. Home-visit interview

The interview schedule was extremely comprehensive. This was because data collection for the survey was designed to serve so many different purposes, research projects and groups. There was some concern that, given its duration of at least 90 minutes, the interview would be too long to be feasible, but the interviewers showed greater professionalism to make them work. The only situation in which one might imagine a much shorter schedule is in a project that unlike the present, one-off survey would be repeated at regular intervals, and that furthermore could spend much more time on planning. Given the magnitude of the survey, the content and instruments of Health 2000 were designed within a fairly short space of time, i.e. in practice during 1999. The work done in spring 2000 was mainly a matter of adding the finishing touches and fine tuning.

The items included in the home-visit interview were selected with a view to the most important themes in the present survey and the Mini-Finland Survey, and they were worded as far as possible in the same way as in earlier studies, particularly the Mini-Finland project. Nevertheless quite a large number of questions were specially developed for this study. The items that could not be covered in the interview were spread out across four different questionnaires. This obviously meant that responses were obtained from a variable number of participants.

Key areas covered in the interview concerned health and illnesses, health services, medicines, living habits and living conditions, including childhood living conditions, education, occupation and employment, functional capacity and need for help, work ability, aids and rehabilitation. Most of the questions were well designed and worded and were based on previous surveys, or they concerned easily ascertainable facts.

Reasonably reliable data were obtained on the presence of different illnesses as well as on the use of services and medicines. As in previous surveys, assessments of work ability and functional capacity were based on self-report, and they were equally reliable as earlier measurements using this same method. The battery of questions on rehabilitation was the most difficult for the interviewers and particularly for the respondents, although this had nothing to do with how the

questions were worded but rather with conceptual ambiguity about rehabilitation itself. There was some concern that the respondents might consider what we called questions for men and questions for women (on infertility, pregnancies, etc.) rather sensitive, but even so the interviewers received a good response to these items.

The interviewers had the unusual task of measuring the respondents' cognitive capacity by using two simple tests. The verbal task was integrated into the interview schedule and generally went very well. The administration of the drawing test also caused no problems, but it was clearly too difficult for the interviewers to assess the respondents' performance. Therefore the tests had to be later reassessed. There is every reason to suggest that these difficulties could have been avoided had it been possible to provide much more advance training to the interviewers.

At the time of writing, analysis of the data collected in the Health 2000 Survey has been going on for several years. On the basis of the experiences so far, it seems fair to conclude that the results are reasonably reliable and that they are well suited for purposes of health research.

22.7. Measurement of blood pressure at home

Another special assignment for the interviewers was to hand out blood pressure gauges to the respondents and to provide instruction on how to use them. Again, the training for interviewers was not entirely adequate. However the biggest problems were caused by the logistics of distributing and reclaiming the gauges throughout the duration of the survey. Indeed, some results were not obtained because the project temporarily ran out of these gauges. In any event this effort did yield a valuable dataset on home blood pressure readings for more than 2,000 subjects, which usefully complements the measurements collected in the health examinations.

22.8. Basic questionnaire

The basic questionnaire that was left with the respondents after the home-visit interviews included the most important items that had to be omitted from the of the interview schedule. Examples include a number of items on psychological symptoms, depression, mood and job fatigue, which have originally been compiled as separate questionnaires. The other items in the basic questionnaire were such that they were expected to generate reliable responses: they concerned functional capacity and quality of life, time use and leisure activities, alcohol use, the eating of sweet snacks, health promotion and the living environment.

Responses were obtained from almost all people who attended the health examinations, but not from all interviewees. Nonetheless the decision to place these items in a separate questionnaire was definitely the right one.

22.9. Health examination

At the field health examination each subject visited ten different examination points which were scheduled to last either 15 or 30 minutes. On average, the whole health examination lasted 4.5 hours. A snack was offered midway through the examination. The examination was sequenced in such a way that the tests and measurements did not interfere with each other.

The purpose of the health examination and its measurements was to complement and specify the data obtained in the interviews and questionnaires. The clinical examinations were designed to provide as reliable information as possible on the presence of illnesses and the need for treatment in the disease categories that were the main focus of attention. The symptom interviews, anthropometric and clinical and physiological measurements as well as blood samples provided useful additional information on illnesses, findings indicative of diseases as well as on risk and protective factors. In principle similar studies have been conducted in other contexts as well, but very rarely with nationally representative samples. Completely novel measurements in this regard were the bioimpedance test and the calcaneal bone ultrasound examination.

Another novel feature was the oral and dental examination, complete with digital x-ray imaging. High-quality results were obtained within just 15 minutes, and the results so far suggest that the x-rays may reveal significant unmet needs for treatment. The x-rays required professional installation and separate inspections at each site by the radiation and nuclear authorities, which did increase costs, but overall everything went very smoothly.

The functional capacity measurements provided a significant addition to the range of data ordinarily collected in interviews and questionnaires. There is good reason to assume that measurements of physical performance at different points in time and in different circumstances offer more stable data on the determining factors for functional capacity than self-report responses. Therefore they may also help us to understand changes in functional capacity and the differences seen between different demographic groups.

In this case the examinations included tests of visual acuity, hearing, cognitive capacity, perceptual-motor speed, hand grip strength and postural balance, plus a number of more simple tests such as walking speed. All these measurements

focused on specific and limited aspects of functional capacity, but the results will be used to single out the most useful and recommendable tests. At the same time, it will be necessary to work on developing observational methods that simulate whole clusters of everyday activities.

The main focus of the clinical examinations carried out by doctors was on circulatory, respiratory and musculoskeletal disorders. The first step was to check anamnestic information concerning illnesses, to assess the patient's status, to record a large number of findings from examinations and to make diagnostic assessments. The examination was highly standardised and structured around an electronic schedule. In many cases the clinical assessments thus obtained draw a highly accurate picture of the subject's medical conditions and illnesses.

Initially the plan was that the field doctors would have access to the main preliminary data from the home interviews. This was a simple and straightforward matter of transferring the relevant data, yet it failed on account of inadequate programming capacity (or shortage of time). In principle this was a very significant departure from the original plans, and it may have undermined the validity of the clinical examination.

The diagnostic CIDI interview that was done by specially trained nurses proved to be a very useful tool. In spite of the sensitivity of the questions and the duration of the interview, participation rates were very high and the interview yielded a valid picture of the most important mental health problems. Not unexpectedly, the results of the CIDI interview became the cornerstone of the mental health inquiry.

22.10. Complementary questionnaire (questionnaire 3)

At the end of the health examination the subjects were given a questionnaire that they were asked to fill in at home and return to KTL. The items in this questionnaire were of a nature that lend themselves most readily to this method of data collection: they included questions on sleep and sleeping, disadvantages in housing conditions, oral health and quality of life, health-related quality of life, seasonal variations and emotions and feelings. Some of the questions were series that are widely used in international surveys, others were specifically developed for the purposes of this investigation.

Around 78% of the forms were returned, less than the figures for other survey stages. However, this was the only way that these data could have been collected in the first place.

22.11. Dietary questionnaire

For reasons of time restraints it was not possible in the health examination to collect data on eating and diet, and therefore the decision was taken to make use of a questionnaire that has been previously been used in a few other contexts. The method itself worked very well, and after two rounds the response rate climbed to a respectable 75%.

22.12. Feedback to subjects

The purpose was to provide feedback to the subjects on key results soon after the health examination. However this was delayed, again largely due to time constraints during the planning stages of the survey and in programming. The only way to avoid these kinds of delays would be to allocate more precious time resources to advance preparation. In this case the problems with feedback were certainly alleviated by the fact that subjects received some information on their results during the field examinations directly from the physicians and nurses, and both the ECG and oral x-ray were given to the subjects when they left the examination.

22.13. Health examination at home

Home health examinations were conducted on as many subjects as possible who did not attend the health examinations proper, either at their home or in an institution. These brief examinations were conducted by trained nurses from the field teams. Since the examinations were done by just one nurse, the range of tests carried out and equipment used was much more restricted.

A home health examination was included in the survey schedule because this was the only practical way to lower non-participation rates in the oldest age groups. If the home visits had been made by two instead of just one nurse, the examinations could have been more extensive. Unfortunately, for reasons of resource limitations, this was not possible.

22.14. Reasons for non-participation

Questionnaires and/or telephone interviews were conducted where possible to find out the reasons for non-participation in subjects who declined to participate even in the home health examination. Some of these were centrally administered

from KTL, some from the field teams. Persistent efforts to contact non-participants eventually paid off as many of those who initially dropped out and even declined to participate agreed to supply the requested information.

22.15. Other subjects

Some people who had earlier taken part in the Mini-Finland health survey were also invited to attend the health examination. The purpose here was to obtain a dataset for studying the appearance of key diseases and associated factors. This project was highly successful.

22.16. Examination of young adults

Both for technical reasons and for reasons of scarce resources, the full interview and health examination programme had to be confined to persons aged 30 or over. However there is an acute need in Finland for more information on the health of young adults. Once the plans for the principal survey were completed it was decided that much of the key information concerning young people could be obtained in interviews that would be conducted after the interviews for the principal survey.

The health interview for persons aged 18–29 was conducted in the same way as in the lead survey, but its content was modified to suit this age group. Participation was high for the age group, and was further improved by the use of prize draws and cinema tickets as incentives. Although there is a body of opinion that is opposed to the use of prizes for reasons of principle, the experiences from this survey suggest that they may become an increasingly important way in the future to ensure adequate response rates. The data collected among these younger people are a useful complement to the dataset on the population aged 30 or over, and they will provide a sound basis for many pioneering studies in Finland on the health of young adults.

22.17. Quality assurance

Several steps were taken to ensure the reliability of the measurements, including the calibration of equipment, the issuing of guidelines and training. More training would have been required both before and during the survey. Even though the performance of both interviewers and staff conducting the measurements were regularly monitored, further training and monitoring would certainly have been beneficial.

Staff performance was also monitored by means of video recordings. The experiences from this were very good, although it is clear that the results would have been even better had it been possible to provide staff with the necessary training in recording techniques.

Integrated and separate quality control tests were carried out to assess the repeatability and validity of the most important measurements as well as inter-rater differences. For reasons of time restraints most of these results were only available after the field examinations. It would certainly have been beneficial if these measurements and the monitoring of staff performances would have yielded systematic feedback during the field stages.

22.18. Data protection, ethical issues and data security

The people invited to take part in the survey and the participants received an information letter ahead of the project and two informed consent forms, one for the health interview and the other for the health examinations. The examination was safe and it involved no ethical problems. The information letter explained how the data collected from the subjects would be used and how they would be linked with certain data drawn from registers.

All the data were handled and processed in such a way that there was no possibility of unauthorised access. They were transferred in encrypted form, and data on paper forms (e.g. questionnaires) were held and transported in locked cabinets and boxes.

22.19. Checking and correcting data

The existence of persons, forms and data were checked early on so that the various subsets could be defined. For electronic forms, checks were also carried out to ascertain logical coherence. In this case it was possible to enter remarks that explained the reasons for missing data and any inconsistencies. It was also at this stage that classifications were done that could not be completed during the fieldwork. Preliminary checks were carried out on the questionnaires before they were saved. Data checks were then carried out as described earlier.

A whole range of measurement results were electronically fed directly into the material. Unfortunately it turned out that erroneous identifications in individual cases caused a great deal of extra work.

Overall the finalisation of the data collection was a huge undertaking that took almost two years after the completion of the field survey. It was, however, spread

out over several stages so that data processing could be started early on; indeed the first basic results were published within 12 months of the completion of the field survey. One of the major reasons why the finalisation took so long was the huge volume and diversity of the data collected, but lingering software problems at the outset of the field examinations was another contributing factor. It is important to explore new ways in which to speed up this process of finalisation. As was pointed out earlier, one of these ways is to incorporate further checks at the stage of data collection, which will require that more time is made available for planning and testing electronic forms.

22.20. Data obtained from registers

The plan for the project was to obtain complementary and cross-sectional data from certain national registers. Later on, the development of the subjects' health was to be monitored by using register data. This latter task concerned data on health and illness as well as the use of medicines, on which the researchers had long-standing experience. However data were also needed on other aspects at the time of the survey. These included data on education and occupation from Statistics Finland, data on income from the Tax Administration, data on institutional care from a combination of different sources as well as data on disability retirement by combining data from the national pensions and employment pension systems.

The linking of register data is nowadays considered a simple and straightforward process, indeed to such an extent that it is feared it may jeopardise data security. Once again experience showed that it is an extremely demanding and time-consuming job to get all the necessary permits and authorities and to check the validity and applicability of the data. For research purposes it is important to have a clear understanding of the applicability and reliability of register data for each specific task.

The seemingly simple task of ascertaining who amongst the subjects were in institutional care at the time of the survey was in fact extremely laborious, and there was still some uncertainty in the final results. The use of data from the Hilmo database on hospital admissions for purposes of studying the incidence of diseases still requires that checks are carried out for each individual case. It is also harder to ascertain who is on disability pension and to trace the number of new pensions than it was ten years ago when all pensions were recorded in one set of SII registers. Both for reasons of compiling accurate statistics and for research purposes it would be important that the holders of registers could themselves ascertain the high quality of the data and on the other hand regularly link the data from such registers that need to be used in unison in order to gain an overall view.

22.21. Release of files for research purposes

The Health 2000 project organisation includes several research institutes and research staff from those institutes, and many other bodies and researchers have also joined up. For purposes of coordinating analyses and reporting it is required that all studies based on materials collected in the survey submit a research plan to the research team in question. The plan is ultimately submitted for approval to the project group or executive committee; in urgent cases approval may be sought from the survey director or the director's deputy.

All applications have been handled smoothly and speedily in 2003–2004, and only a relatively small proportion of plans are returned for clarification. Once the research plan has been approved, the researcher or research team may request the file they require via internet or e-mail. Initially there were some delays here as not all of the data were ready yet. However since 2003 all orders and deliveries have gone smoothly, and by the end of 2004 hundreds of files had been released for research purposes.

22.22. Statistical analysis

Because of the two-stage cluster sampling process it is necessary to use methods of statistical analysis and calculation that take account of the sample structure and non-response. Since the patterns of non-response were different at different stages of the survey, it was necessary for practical reasons to single out the most important stages of the survey, for which Statistics Finland has compiled weighting coefficients. These coefficients are supplied to the researchers with each file.

The most suitable calculation methods are SAS/SUDAAN, which runs both on mainframes and PCs as well as STATA version 8, recently released for PC use. These tools will provide statistically accurate results that can be generalised to the whole population.

As these methods were not yet widely available during the field examinations, the researchers needed much support and guidance. Despite this it made more economic sense to provide that guidance rather than to conduct centralised statistical analyses, primarily because the research organisations would not have had the personnel resources to undertake such an effort. This route was, however, chosen for producing the results for the main project reports: indeed this proved a useful for prompt and uniform reporting and for efficient time use by researchers.

22.23. In-depth examinations

No field project can ever hope to carry out all the tests and measurements it would like to do in one fell swoop. It is useful, therefore, to set aside enough time to take any additional tests at another stage; that will greatly enhance the value of the survey. It is also necessary to look at the validity of some measurements in certain segments of the study population. Furthermore, it is known, or there is at least reason to suspect that participation is selective in relation to certain health problems (e.g. mental health problems).

For these various reasons the decision was made to carry out further tests and examinations that would shed additional light on cardiovascular diseases, diabetes, certain musculoskeletal disorders, psychoses and other mental health disorders and some order diseases and injuries. Furthermore, the incidence of oral diseases and related factors was now studied for the first time.

It is already known that the prevalence of psychoses is twice as high as was originally estimated on the basis of population studies. This survey has yielded invaluable information from calcaneal bone ultrasound examinations, which can be compared with results based on standard methods of studying osteoporosis. Ultrasound examination of carotid arteries provided a clear picture of the advance of atherosclerotic vascular disease, and the clinical assessment of diabetes was clarified by a two-hour glucose tolerance test and by the determination of plasma insulin. All in all the in-depth examinations provide a valuable addition to the results from the field examinations proper. However, there are still many other diseases and conditions that would warrant more in-depth investigation.

22.24. Conclusions

The Health 2000 Survey holds an exceptionally important place among the world's population surveys. It was successfully designed and completed within the target timetable and budget set. Finland now has a clearer picture than any other country in the world of the nation's major health problems and functional disabilities and, together with the Mini-Finland Survey, of how these have developed over time.

One of the biggest assets of this project was that so many of the people involved had prior experience of Finnish population surveys and particularly of the Mini-Finland Health Survey. Another key asset was the commitment of so many different partners, which gave it great breadth of expertise and competence. Furthermore, it was absolutely crucial that funding as well as services were received free of charge from various quarters. The experience factor was vitally important in situations

where quick decisions were needed on the most appropriate methods, where guidelines had to be drafted and preparations made for the necessary training.

It was essential that all the field staff as well as the people working at project headquarters in Helsinki were fully committed to their jobs. Working hours were often long and jobs had to be done under considerable time pressure, yet still with great accuracy.

There were, inevitably, problems and difficulties along the way. At the design and planning stage the involvement of so many different partners was not only an asset, but it also slowed down the process. Planning and preparations for the fieldwork were to be completed within 12–18 months. The finalisation of the plans was a slow process, and consequently work to compile the electronic forms and IT systems was delayed. Unfortunately, there was no way in which the planning of the survey contents could have been speeded up. The timetables for the field examinations were also fixed. One way in which the time pressures might have been relieved somewhat would have been to recruit more programming staff by the beginning of 2000.

Because of the lack of time and money at the planning stage and during the fieldwork, the project was unable to provide sufficient refresher training to field staff or to implement adequate quality assurances. Furthermore, the field physicians did not get the preliminary information they were supposed to receive from the health interviews, and some of the other problems could not be addressed until during the first month of the fieldwork. The same factors were partly responsible for the large number of checks and corrections that had to be made after data collection.

This and other reports published on the Health 2000 project and its successes and difficulties provide valuable lessons for future planning of similar surveys. The most crucial factors are the balance of research and other staff, timetables and funding. The Health 2000 Survey could have afforded a slightly more relaxed timetable if the budget at its disposal had been about one-fifth higher.

23. REFERENCES

Antonowsky A. The structure and properties of the sense of coherence scale. Soc Sci Med 1993;36:725-733.

Arinen S, Häkkinen U, Klaukka T, Klavus J, Lehtonen R, Aro S. Suomalaisten terveys ja terveyspalvelujen käyttö. Terveydenhuollon väestötutkimuksen 1995/96 päätulokset ja muutokset vuodesta 1987. SVT Terveys 5, Stakes ja Kansaneläkelaitos, Helsinki, 1998. (in Finnish)

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. Terveys, toimintakyky ja hoidontarve Suomessa. Mini-Suomi-terveystutkimuksen perustulokset. Kansaneläkelaitoksen julkaisuja AL:32, Helsinki ja Turku, 1989a. (in Finnish)

Aromaa A, Heliövaara M, Impivaara O, Knekt P, Maatela J. Mini-Suomi–terveystutkimuksen toteutus. Osa 1. Tavoitteet, menetelmät ja aineisto. Helsinki ja Turku: Kansaneläkelaitoksen julkaisuja ML:88, 1989b. (in Finnish)

Aromaa A, Koponen P, Tafforeau J, Vermeire C (eds.) Status and future of health surveys in the European Union. Final report of Phase 2 of the project Health Surveys in the EU: HIS and HIS/HES Evaluations and Models. Publications of the National Public Health Institute KTL/B15, Helsinki, 2003a (HIS/HES database: http://www.iph.fgov.be/hishes/).

Aromaa A, Koponen P, Tafforeau J, Vermeire C; HIS/HES Core Group. Evaluation of Health Interview Surveys and Health Examination Surveys in the European Union. Eur J Public Health. 2003;13(3 Suppl):67–72. (2003b)

Aromaa A, Reunanen A, Impivaara O, Heliövaara M, Knekt P, Maatela J. Mini-Suomi-tutkimuksen toteutus. Osa 2. Verenkiertoelinten ja hengityselinten sairauksien tutkimusmenetelmät. Helsinki ja Turku: Kansaneläkelaitoksen julkaisuja ML:49, 1985. (in Finnish)

Bagby RM, Parker JDA, Taylor GJ. The twenty-item Toronto Alexithymia Scale-I. Item selection and cross-validation of the factor structure. Journal of Psychosomatic Research;1994;38:23-32. (1994a)

Bagby RM, Taylor GJ, Parker JDA. The twenty-item Toronto Alexithymia Scale-II. Convergent, discriminant and concurrent validity. Journal of Psychosomatic Research J Psychosom Res 1994;38:33–40. (1994b)

Bagby RM, Taylor GJ, Parker JDA, Loiselle C. Cross-validation of the factor structure of the Toronto Alexithymia Scale. Journal of Psychosomatic Research 1990;34:47-51.

Bandelow B, Wedekind D, Pauls J, Broocks A, Hajak G, Rüther E. Salivary Cortisol in Panic Attacks (Brief report). American Journal of Psychiatry 2000;157:454-456.

Barchielli A, Balzi D. Nine-year follow-up of a survey on smoking habits in Florence (Italy): higher mortality among non-responders. Int J Epid 2002;31:1038-1042.

Beck AT, Ward CH, Mendelson M, et al. An inventory for measuring depression. Arch Gen Psychiatry 1961;4:561-571.

Bedogni G, Malavolti M, Severi S, Poli M, Mussi C, Fantuzzi AL, Battistini N. Accuracy of an eight-point tactile-electrode impedance method in the assessment of total body water. Eur J Clin Nutr 2002;56:1143-1148.

Biering-Sörensen F. Physical measurements as risk indicators for low-back trouble over a one-year period. Spine 1984;9:106–119.

Blazer DG, Kessler RC, Swartz MS. Epidemiology of recurrent major and minor depression with a seasonal pattern: the National Comorbidity Survey. Br J Psychiatry 1998;172:164–167.

Brøgger J, Bakke P, Eide GE, Gulsvik A. Contribution of follow-up of nonresponders to prevalence and risk estimates: a Norwegian respiratory health survey. Am J Epid 2003;157:558–566.

de Bruin AZ, Picavet HSJ, Nossikov A. WHO. Health Interview Surveys. Towards international harmonization of methods and instruments. WHO Regional Publications, European Series, no 58. WHO, Geneva, 1996.

Cepollaro C, Gonnelli S, Pondrelli C, ym. The combined use of ultrasound and densitometry in the prediction of vertebral fracture. Br J Radiol 1997;70:691-696.

Chambers R, Skinner C. (eds.). Analysis of Survey Data. Chichester: John Wiley & Sons, 2003.

Cochrane CE, Brewerton TD, Wilson DB, Hodges EL. Alexithymia in eating disorders. Int J Eating Disorders 1993;14:219-222.

Conte HR, Plutchik R, Jung BB, Picard S, Karasu TB, Lotterman A. Psychological mindedness as a predictor of psychothrapy outcome: a preliminary report. Compr Psychiatry 1990;31:426-431.

Cotes IE. Lung function. Assessment and application in medicine. 3. painos. Lontoo: Blackwell Scientific, 1975.

Craig CL, Marshall AL, Sjöström M, Bauman AE, Booth ML, Ainsworth BE, Pratt M, Ekelund U, Yngve A, Sallis JF, Oja P. International physical activity questionnaire: 12-contity reliability and validity. Med Sci Sports Exerc 2003;35:1381-1395.

Currie C, Roberts C, Morgan A, Smith R, Setterbulte W, Samdal O, Barnekow Rasmussen V. Young people's health in context. Health Behaviour in School-aged Children (HBSC) study: international report from the 2001/2002 survey. Health Policy for Children and Adolescents, No 4, WHO Europe, 2004.

Derogatis LR, Lipman RS, Covi L. SCL-90: an outpatient psychiatric rating scale -preliminary report. Psychopharmacol Bull 1973;9:13-27.

Diggle PJ, Liang K-Y, Zeger SL. Analysis of Longitudinal Data. Oxford: Oxford University Press, 1994.

Djerf K. Properties of some estimators under unit nonresponse. Statistics Finland. Research Reports 231. Helsinki: Statistics Finland, 2000.

Drury TE, Winn DM, Snowden CB, Kingman A, Kleinman DV, Lewis B. An overview of the Oral Health Component of the 1988-1991 National Health and Nutrition Examination Survey (NHANES III-Phase 1). J Dent Res 1996;75(Spec Iss):620-623.

Dworkin SF, LeResche L. Research diagnostic criteria for temporomandibular disorders: review, criteria, examinations and specifications, critique. J Craniomandib Disord. 1992;6:301-355.

Era P, Jokela J, Heikkinen E. Reaction and movement times in men of different ages: a population study. Percept Mot Skills 1986;63:111–130.

EuroQol Group. EuroQol: a new facility for the measurement of health-related quality of life. Health Policy 1990;16:199-208.

Fiatarone M, O'Neill E, Ryan N, Cleents K, Solares G, Nelson M, Roberts S, Kehayias J, Lipsitz L, Evans W. Exercise training and nutritional supplementation for physical frailty in very elderly people. N Engl J Med 1994;330:1769–1775.

Fink P, Ewald H, Jensen J, Sørensen L, Engberg M, Holm M, Munk-Jørgensen P. Screening for somatization and hypchondriasis in primary care and neurological in-patients: a seven item scale for hypchondriasis and somatization. J Psychosom Res 1999;46:262-273.

Finnish Hypertension Society Working Group. Suomen Verenpaineyhdistys ry:n asettama työryhmä. Kohonneen verenpaineen hoito. Duodecim 2002;118:110-126. (in Finnish)

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

Fletcher CM, Elmes PC, Fairbairn AS, Wood CH. The significance of respiratory symptoms and the diagnosis of chronic bronchitis in a working population. BMJ 1959;2:257-266.

Folstein M, Folstein S, McHugh P. 'Mini-mental State': a practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 1975;12:189–198.

Fox PC, Busch KA, Baum BJ. Subjective reports of xerostomia and objective measures of salivary gland performance. J Am Dent Assoc 1987;115:581-584.

Garnero P, Dargent-Molina P, Hans D, et al. Do markers of bone resorption add to bone mineral density and ultrasonographic heel measurement for the prediction of hip fracture in elderly women? The EPIDOS prospective study. Osteoporosis Int 1998;8:563-569.

Goldberg DP. The detection of psychiatric illness by questionnaire. Maudsley Monograph No 22. Oxford University Press, Oxford, 1972.

Goldstein H. Multilevel Statistical Models, 3nd Edition. London: Arnold, 2003.

Gonnelli S, Cepollaro C, Agnusdei D, ym. Diagnostic value of ultrasound analysis and bone densitometry as predictors of vertebral deformity in postmenopausal women. Osteoporosis Int 1995;5:413-418.

Graubard BI, Korn EL. Predictive Margins with Survey Data. Biometrics 1999;55:652–659.

Greenspan SL, Maitland-Ramsay I, Myers E. Classification of osteoporosis in the elderly is dependent on site-specific analysis. Calcif Tissue Int 1996;58:409-414.

Groves RM, Coupier MP. Nonresponse in household interview surveys. Wiley series in probability and statistics, John Wiley & Sons, Inc. New York, USA, 1998.

Gumm HP. A new class of check digit methods for arbitrary number systems. IEEE Transactions on information theory 1985;31:102–105.

Guralnik JM, Simonsick EM, Ferrucci L, Glynn RJ, Berkman LF, Blazer DG, Scherr PA, Wallace RB. A short physical performance battery assessing lower extremity function: association with self-reported disability and prediction of mortality and nursing home admission. J Gerontol 1994;49: M85–94.

Haapa E, Toponen T, Pietinen P, Räsänen L. Annoskuvakirja. Helsinki: Kansanterveyslaitos ja Helsingin yliopisto, 1985. (in Finnish)

Hans D, Dargent-Molina P, Schott AM, ym. Ultrasonographic heel measurements to predict hip fracture in elderly women: the EPIDOS prospective study. Lancet 1996;348:511-514.

Hansen V, Jacobsen BK, Arnesen E. Prevalence of serious psychiatric morbidity in attenders and nonattenders to a health survey of a general population. Am J Epid 2001;154:891–894.

Hara M, Sasaki S, Sobue T, Yamamoto S, Tsugane S, for the JPHC Study Group. Comparison of cause-specific mortality between respondents and nonrespondents in a population-based prospective study: ten-year follow-up of JPHC Study cohort I. J Clin Epidemiol 2002;55:150–156.

Heikkilä J (ed.). EKG, perusteet ja tulkinta. Espoo: Lääketehdas Orion, 1982. (in Finnish)

Heinonen OP. Autoklinikka. Duodecim 1966;82:1161–1164. (in Finnish)

Heiskanen M, Aromaa K, Niemi H, Sirén R. Tapaturmat, väkivalta, rikollisuuden pelko. Väestöhaastattelujen tuloksia vuosilta 1980–1997. Oikeuspoliittisen tutkimuslaitoksen julkaisuja 171, Tilastokeskus, Helsinki, 2000. (in Finnish)

Helakorpi S, Patja K, Prättälä R, Aro AR, Uutela A. Suomalaisen aikuisväestön terveyskäyttäytyminen ja terveys, kevät 2003. Publications of the National Public Health Institute KTL/B17, Helsinki, 2003. (in Finnish) (http://www.ktl.fi/publications/2003/b17.pdf).

Helakorpi S, Uutela A, Prättälä R, Puska P. Suomalaisen aikuisväestön terveyskäyttäytyminen ja terveys, kevät 2000. Kansanterveyslaitoksen julkaisuja KTL/B8, Helsinki, 2000. (in Finnish)

Heliövaara M, Mäkelä M, Sievers Kai, Melkas T, Aromaa A, Knekt P, Impivaara O, Aho K, Isomäki H. Tuki- ja liikuntaelinten sairaudet Suomessa. Kansaneläkelaitoksen julkaisuja AL:35, Helsinki 1993. (in Finnish)

Horton NJ, Lipsitz SR. Review of software to fit generalized estimating equation regression models. The American Statistician 1999;53:160–169.

Huuhka M, Lahelma E, Manderbacka K, Mattila V, Karisto A, Rahkonen O. Terveydentila ja sosiaalinen murros. Vuosien 1986 ja 1994 elinolotutkimukset. Tilastokeskus, SVT Elinolot 1996:2. Helsinki, 1996. (abstract: http://www.stat.fi/tk/tp_tied/tiedotteet/v96/leh78.html). (in Finnish)

Hänninen T, Pulliainen V, Salo J, Hokkanen L, Erkinjuntti T, Koivisto K, Viramo P, Soininen H, ja Suomen muistitutkimusyksiköiden asiantuntijaryhmä. Kognitiiviset testit muistihäiriöiden ja alkavan dementian varhaisdiagnostiikassa: CERAD-tehtäväsarja. Suom Lääkäril 1999;15:1967–1975. (in Finnish)

Joukamaa M, Lepola U. Alexithymic features in patients with panic disorder. Nordic Journal of Psychiatry 1994;48:33-36.

Joukamaa M, Miettunen J, Kokkonen P, Koskinen M, Julkunen J, Kauhanen J, Jokelainen J, Veijola J, Läksy K, Järvelin M-R. Psychometric properties of the Finnish 20-item Toronto Alexithymia Scale. Nord Psychiatry 2001;55:123-127.

Jousilahti P, Salomaa V, Kuulasmaa K, Niemelä M, Vartiainen E. Total and cause specific mortality among participants and non-participants of population based health surveys: a comprehensive follow up of 54 372 Finnish men and women. J Epidemiol Community Health 2005;59:310–315.

Jula AM, Salminen JK, Saarijärvi S. Alexithymia, a facet of essential hypertension. Hypertension 1999;33:1057-1061.

Kahn HS, Austin H, Williamson DF, Arensberg D. Simple anthropometric indices associated with ischemic heart disease. J Clin Epidemiol 1996;49:1017-1024.

Kalimo R, Toppinen S: Työuupumus Suomen työikäisellä väestöllä. Työterveyslaitos, Helsinki, 1997. (in Finnish)

Katz S, Downs TD, Cash HR, Grotz RC. Progress in development of the index of ADL. Gerontologist 1970;10:20-30.

Katz S, Ford AB, Moskowitz RW, Jackson BA, Jaffe MW. Studies of illness in the aged. The index of ADL: a standardized measure of biologocal and psychosocial function. JAMA 1963;185:914-919.

Kauhanen J, Kaplan G A, Cohen R D, Julkunen J, Salonen J T. Alexithymia and risk of death in middle-aged men. J Psychosom Res 1996;41:541-549.

Kauhanen J, Kaplan G A, Cohen R D, Salonen R, Salonen J T. Alexithymia may influence the diagnosis of coronary heart disease. Psychosom Med 1994;56:237-244.

Kauppinen T, Hanhela R, Heikkilä P, Lehtinen S, Lindström K, Toikkanen J, Tossavainen A. Työ ja terveys Suomessa 2003. Työterveyslaitos, Helsinki, 2004. (in Finnish)

Kelly M, Steele J, Nuttall N, Bradnock G, Morris J, Nunn J, Pine C, Pitts N, Treasure E, White D. Adult Dental Health Survey: Oral Health in the United Kingdom 1998. Her Majesty's Stationery Office. London, 2000.

Kerekes K, Tronstad L. Long-term results on endodontics treatment performed with a standardized technique. J Endod 1979;5:83-90.

Kessler RC. The World Health Organization International Consortium in Psychiatric Epidemiology (ICPE): initial work and future directions – the NAPE Lecture 1998. Acta Psychiatrica Scandinavica 1999;99:2-9.

Koponen P, Aromaa A. Survey Design and Methodology in National Health Interview and Health Examination Surveys. Review of literature, European Survey experiences and recommendations. In: Aromaa A, Koponen P, Tafforeau J, Vermeire C, Primatesta P, Marmot M, Kurth B, Lehto-Järnstedt U-S, Räty S. Health surveys: evaluation and recommendations. Subproject reports of phase 2 of the project Health Surveys in the EU: HIS and HIS/HES Evaluations and Models. Publications of the National Public Health Institute KTL/B16, Helsinki, 2004:100-167.

Korn EL, Graubard BI. Analysis of Health Surveys. New York: John Wiley & Sons, 1999.

Koskinen S, Kestilä L, Martelin T, Aromaa A (eds.). Nuorten aikuisten terveys. Terveys 2000 -tutkimuksen perustuloksia 18–29-vuotiaiden terveydestä ja siihen liittyvistä tekijöistä. Publications of the National Public Health Institute KTL/B7, Helsinki, 2005. (in Finnish)

Kunttu K, Huttunen T. Korkeakouluopiskelijoiden terveystutkimus 2000. Kela, Sosiaali- ja terveysturvan katsauksia 45, Helsinki 2001. (in Finnish)

Kuusela V, Parviainen A. An object-based case management system for CAPI surveys. In: Boyeldieu I. (ed.). Des Utilisateurs de BLAISE. INSEE: Paris, 1997. (http://www. blaiseusers.org/Ibucpdfs/1995-1998/kuusel97.pdf)

Kuusela V. A Windows based user interface and CAPI information system. 7th International Blaise Users´ Conference. Washington DC, USA, 2001. (http://www.blaiseusers.org/ibucpdfs/2001/Kuusela-IBUC_paper.pdf)

Kuusela, V, Tanskanen, V, Virtala, E. (2000): Data Collection in Two Phase and Multi Centre Health Survey. Paper presented at 6th International Blaise User's Conference, May 2000, Kinsale, Ireland http://www.blaiseusers.org/2000/papers/vesa_multi_centre_health_survey.pdf

Laiho J, Djerf K. Otanta-asetelma. In: Laiho J, Nieminen T (eds.). Terveys 2000 -tutkimus. Aikuisväestön haastatteluaineiston tilastollinen laatu. Otanta-asetelma, tiedonkeruu, vastauskato

ja estimointi- ja analysointiasetelma. Tilastokeskus, Tutkimuksia 239, Helsinki, 2004:21–27. (in Finnish)

Laiho J. 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, 2004:28-47. (in Finnish)

Laitinen A, Koskinen S, Härkänen T, Reunanen A, Laatikainen L, Aromaa A. A nationwide population-based survey on visual acuity, near vision and self-reported visual function in the adult population in Finland. Ophthalmology 2005;112:2227-2237.

Langland OE, Langlais RP. Principles of dental imaging. Williams & Wilkins, Baltimore, 1997.

Lawton MP, Brody EM. Assessment of older people: self-maintaining and instrumental activities of daily living. Gerontologist 1969;9:179-186.

Lee J. Covariance adjustment of rates based on the multiple logistic regression model. Journal of Chronic Diseases 1981;34:415–426.

Lehtinen V, Joukamaa M, Jyrkinen T, Lahtela K, Raitasalo R, Maatela J, Aromaa A. Suomalaisten aikuisten mielenterveys ja mielenterveyden häiriöt. Kansaneläkelaitoksen julkaisuja AL:33, Turku and Helsinki, 1991. (in Finnish)

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

Lehtonen R, Djerf K, Härkänen T, Laiho J. Modelling complex health survey data: a case study. In: Höglund R, Jäntti M, Rosenqvist G (eds.). Statistics, econometrics and society: essays in honour of Leif Nordberg. Helsinki: Statistics Finland, Research Reports 238, 2003:91-114.

Liao D, Arnett DK, Tyroler HA, Riley WA, Chambless LE, Szklo M, Heiss G. Arterial stiffness and the development of hypertension. The ARIC Study. Hypertension 1999;34:201-6.

Liang K-Y, Zeger SL. Longitudinal data analysis using generalized linear models. Biometrika 1986;73:13–22.

Lindholm T, Lehtinen V, Hyyppä MT, Puukka P. Alexithymic features in relation to the dexamethasone suppression test in a Finnish population sample. Am J Psychiatry 1990;147:1216-1219.

Lohr SL. Sampling: design and analysis. Pacific Grove: Duxbury Press, 1999.

Lumley M A, Tomakowsky J, Torosian T. The relationship of alexithymia to subjective and biomedical measures of disease. Psychosomatics 1997;38:497-502.

Lund E, Hansen V. Responses to the Seasonal Pattern Assessment Questionnaire in different seasons. Am J Psychiatry 2001;158:316–318.

Magnusson A. Validation of the Seasonal Pattern Assessment Questionnaire (SPAQ). J Affect Disord 1996;40:121–129.

Magnusson A. An overview of epidemiological studies on seasonal affective disorder. Acta Psychiatr Scand 2000;101:176–184.

Männistö S, Virtanen M, Mikkonen T, Pietinen P. Reproducibility and validity of a food frequency questionnaire in a case-control study on breast cancer. J Clin Epidemiol. 1996;49:401-409.

Marshall D, Johnell O, Wedel H. Meta-analysis of how well measures of bone mineral density predict occurrence of osteoporotic fractures. BMJ 1996;312:1254-1259.

Maslach C, Jackson SE, Leiter MP. Maslach Burn Out Inventory Manual (3rd ed.), Consulting Psychologist Press, Palo Alto, 1996.

Maslach C, Jackson SE. The measurement of experienced burn-out. J Occup Behavior 1981;2:99-113.

Masud T, Francis RM. The increasing use of peripheral bone densitometry. Better as assessing fracture risk than diagnosis of osteoporosis. BMJ 2000;321:396-8.

McBurnett K, Lahey B, Rathouz P, Loeber R. Low salivary cortisol and persistent aggression in boys referred for disruptive behavior. Archives of General Psychiatry 2000;57:38-43.

McCullogh CE, Searle SR. Generalized, linear, and mixed models. New York: John Wiley & Sons, 2001.

McWhinnie JR. Disability assessment in population surveys: results of the OP.E.C.D. Common Development Effort. Rev Epidemiol Santé Publique 1981;29:413-419.

Morris JC, Heyman A, Mohs RC, Hughes JP, van Belle G, Fillenbaum G, Mellits ED, Clark C. The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part I. Clinical and neuropsychological assessment of Alzheimer's Disease. Neurology 1989;39:1159–1165.

Nieminen T. Haastattelijapalautetta Terveys 2000 -tutkimuksesta. Hyvinvointikatsaus 2003;2:42-45. (in Finnish)

Nieminen T, Kuusela V. Terveyshaastattelu. In: Laiho J, Nieminen T (eds.). Terveys 2000 -tutkimus. Aikuisväestön haastatteluaineiston tilastollinen laatu. Otanta-asetelma, tiedonkeruu, vastauskato ja estimointi- ja analysointiasetelma. Tilastokeskus, Tutkimuksia 239, Helsinki, 2004:10–20. (in Finnish)

Nosikov A, Gudex G (eds). Eurohis. Developing common instruments for health surveys. WHO Regional Office for Europe, IOS Press, Amsterdam, 2003.

O´Brien E, Asmar R, Beilin L, et al. European Society of Hypertension recommendation for conventional, ambulatory and home blood pressure measurement. J Hypertens 2003;21:821-848.

Oh JL, Scheuren F. Weighting adjustment for unit response. In: Madow WG, Olkin I, Rubin DB (eds.). Incomplete data in sample surveys. Vol. 2. New York: Academic Press, 1983:143-184.

Paalanen L, Männistö S, Virtanen MJ, Knekt P, Räsänen L, Montonen J, Pietinen P. Validity of a food frequency questionnaire varied by age and body mass index. J Clin Epidemiol. 2006 Sep;59:994-1001.

Parker JDA, Bagby RM, Taylor GJ, Endler NS, Schmitz P. Factorial validity of the 20-item Toronto Alexithymia Scale. European Journal of Personality 1993;7:221-232.

Partonen T, Lönnqvist J. Seasonal affective disorder. Lancet 1998;352:1369–1374.

Pevalin DJ. Multiple applications of the GHQ-12 in a general population sample: an investigation of long-term retest effects. Social Psychiatry and Psychiatric Epidemiology 2000;35:508-512.

Pfeffermann D, Skinner CJ, Goldstein H, Holmes DJ, Rasbash J. (1998). Weighting for unequal selection probabilities in multilevel models (with discussion). Journal of the Royal Statistical Society 1998(Series B);60:23–40.

Pharmaca Fennica 1999. Lääketietokeskus, Helsinki, 1999. (in Finnish)

Pharmaca Fennica 2000. Lääketietokeskus, Helsinki, 2000. (in Finnish)

Pharmaca Fennica 2001. Lääketietokeskus, Helsinki, 2001. (in Finnish)

Pietinen P. Ravinnonkäytön tutkimusmenetelmät. In: Aro A, Mutanen M, Uusitupa M (eds.). Ravitsemustiede. Hämeenlinna: Karisto Oy, 1999. (in Finnish)

Pietrobelli A, Rubiano F, St-Onge M-P, Heymsfield SB. New bioimpedance analysis system: improved phenotyping with whole-body analysis. Eur J Clin Nutr 2004;58:1479-1484.

Piirainen H, Elo A-L, Hirvonen M, Kauppinen K, Ketola R, Laitinen H, Lindström K, Reijula K, Riala R, Viluksela M, Virtanen S. Työ ja terveys -haastattelututkimus v. 2000. Taulukkoraportti. Työterveyslaitos, Helsinki, 2000. (in Finnish)

Piirainen T, Hirvonen M, Elo A-L, Huuhtanen P, Kandolin I, Kauppinen K, Ketola R, Lindström K, Salminen S, Reijula K, Riala R, Toivanen M, Viluksela M, Virtanen S. Työ ja terveys haastattelututkimus 2003. Taulukkoraportti. Työterveyslaitos, Helsinki, 2003. (in Finnish)

Pilowsky I. Dimensions of hypochondriasis. British Journal of Psychiatry 1967;113:89-93.

Pirkola S, Isometsä E, Suvisaari J, Aro H, Joukamaa M, Poikolainen K, Koskinen S, Aromaa A, Lönnqvist J. DSM-IV mood-, anxiety- and alcohol use disorders and their comorbidity in the Finnish general population. Results from the Health 2000 study. Social Psychiatry and Psychiatric Epidemiology 2005;40:1-10.

Plujim SMF, Graafmans WC, Bouter LM, Lips P. Ultrasound measurements for the prediction of osteoporotic fractures in elderly people. Osteoporosis Int 1999;9:550-556.

Preisser JS, Lohman KK, Rathouz PJ. Performance of weighted estimating equations for longitudinal binary data with drop-outs missing at random. Statistics in Medicine 2002;21:3035-3054.

Prineas RJ, Crow RS, Blackburn H. The Minnesota Code. Manual of electrocardiographic Findings. Standards and procedures for measurements and classification. Boston: John Wright, 1982.

Prins SH, Jörgensen HL, Jörgensen LV, ym. The role of quantitative ultrasound in the assessment of bone: a review. Clin Physiol 1998;18:3-17.

Pulliainen V, Hokkanen L, Salo J, Hänninen T (eds.). CERAD, kognitiivinen tehtäväsarja, käsikirja. Kuopio, 1999. (in Finnish)

Purola T, Kalimo E, Sievers K, Nyman K. Sairastavuus ja lääkintäpalvelusten käyttö Suomessa ennen sairausvakuutusta. Kansaneläkelaitoksen julkaisuja A:1, Helsinki, 1967. (in Finnish)

Raitasalo R. Depressio ja sen yhteys psykoterapian tarpeeseen. Kansaneläkelaitoksen julkaisuja A:13, Helsinki, 1977. (in Finnish)

Reims H, Fossum E, Kjeldsen SE, Julius S. Home blood pressure monitoring. Current knowledge and directions for future research. Blood Press 2001;10:271-287.

Research Triangle Institute. SUDAAN Language Manual, Release 9.0. Research Triangle Park, NC: Research Triangle Institute, 2004.

Reunanen A, Aromaa A, Pyörälä K, Punsar S, Maatela J, Knekt P. The Social Insurance Institution´s Coronary Heart Disease Study. Acta Medica Scand 1983:Suppl 673.

Rimpelä A, Lintonen T, Pere L, Rainio S, Rimpelä M. Nuorten terveystapatutkimus 2003: Tupakkatuoteiden ja päihteiden käytön muutokset 1977–2003. Stakes, Aiheita 13/2003, Helsinki, 2003. (in Finnish)

Rose GA, Blackburn H. Cardiovascular survey methods. Geneva: World Health Organization, 1968.

Rose GA, Blackburn H, Gillum RF, Prineas RJ. Cardiovascular survey methods. Second edition. Geneva: World Health Organization, 1982.

Rosenthal NE, Bradt GH, Wehr TA. Seasonal Pattern Assessment Questionnaire. Bethesda, MD: National Institute of Mental Health, 1984.

Saarijärvi S, Salminen JK, Tamminen T, Äärelä E. Alexithymia in consultation-liaison patients. Gen Hosp Psychiatry 1993;15:330-333.

Sainio P, Koskinen S, Heliövaara M, Martelin T, Härkänen T, Hurri H, Miilunpalo S, Aromaa A. Self-reported and test-based mobility limitations in a representative sample of Finns aged 30+. Scandinavian Journal of Public Health 2006;34:378–386.

Salmi J, Pekkarinen H, Kröger H, Heliövaara M. Segmental assessments of fat free mass by bioimpedance in relation to DXA. Int J Obes 2004;28:S127.

Salminen JK, Saarijärvi S, Äärelä E, Toikka T, Kauhanen J. Prevalence of alexithymia and its association with sociodemographic factors in the general population of Finland. J Psychosom Res 1999;46:75-82.

Sautory O. La Macro CALMAR. Readressement d'un échantillion par calage sur marges. I.N.S.E.E. Série des documents de travail n:o F 9310. Paris: I.N.S.E.E, 1993.

Seidell JC, Kahn HS, Williamson DF, Lissner L, Valdez R. Report from a Centers for Disease Control and Prevention workshop on use of adult anthropometry for public health and primary health care. Am J Clin Nutr 2001;73:123-126.

Sievers K, Melkas T, Heliövaara M. Mini-Suomi -terveystutkimuksen toteutus. Osa 3. Tuki- ja liikuntaelinten sairauksien tutkimusmenetelmät. Helsinki ja Turku: Kansaneläkelaitoksen julkaisuja ML:50, 1985. (in Finnish)

Sifneos PE. The prevalence of 'alexithymic' characteristics in psychosomatic patients. Psychother Psychosom 1973;22:255-262.

Silness J, Löe H. Periodontal disease in pregnancy. II. Correlation between oral hygiene and periodontal condition. Acta Odontol Scand 1964;22:121-135.

Simpura J. Development of a common instrument for alcohol consumption. In: Nasikov A, Guder C (eds.). Eurohis. IOS Press, WHO Regional Office for Europe, Copenhagen, 2003.

Singer JD. Using SAS PROC MIXED to fit multilevel models, hierarchical models, and individual growth models. Journal of Educational and Behavioral Statistics 1998;24:323–355.

Sintonen H. An approval to measuring and valuering health states. Soc Sci Med [Med Econ] 1981;15:55-65.

Sintonen H. The 15D instrument of health-related quality of life: properties and applications. Ann Med 2001;33:328-36 (Review).

Slade GD. Derivation and validation of a short-form oral health impact profile. Community Dent Oral Epidemiol 1997;25:284-290.

Slade GD, Spencer AJ. Development and evaluation of the Oral Health Impact Profile. Community Dental Health 1994;11:3-11.

Suni J. Health-related fitness test battery for middle-aged adults with emphasis on musculoskeletal and motor tests [thesis]. Jyväskylä: University of Jyväskylä, 2000.

Suominen-Taipale L, Nordblad A, Vehkalahti M, Aromaa A (eds.). Suomalaisten aikuisten suunterveys. Terveys 2000 -tutkimus. Publications of the National Public Health Institute KTL/B16, 2004 (http://www.ktl.fi/attachments/suomi/julkaisut/julkaisusarja_b/2004b16.pdf).

Susitaival P, Husman T (eds.). Tuohilampikysymyssarjat. Kuopio: Tuohilampi-ryhmä, 1996. (in Finnish)

Särndal C-E, Swensson B, Wretman J. Model assisted survey sampling. New York: Springer Verlag, 1992.

Särndal C-E, Lundström S. Estimation in surveys with nonresponse. Chichester: John Wiley & Sons, 2005.

Taylor GJ, Bagby RM, Parker JDA. The alexithymia construct. A potential paradigm for psychosomatic medicine. Psychosomatics 1991;32:153-161.

Taylor GJ, Bagby RM, Parker JDA. Disorders of affect regulation. Alexithymia in medical and psychiatric illness. Cambridge University Press, Cambridge, 1997:29-30.

Taylor GJ, Babgy RM, Ryan DP, Parker JDA, Doody KF, Keefe P. Criterion validity of the Toronto Alexithymia Scale. Psychosomatic Medicine 1988:50:500-509.

Taylor GJ, Babgy RM, Ryan DP, Parker JDA. Validation of the alexithymia construct: A measurement-based approach. Canadian Journal of Psychiatry 1990;35:290-297. (1990a)

Taylor GJ, Parker JDA, Babgy RM. A preliminary investigation of alexithymia in men with psychoactive substance dependence). American Journal of Psychiatry 1990;147:1228-1230. (1990b)

Taylor GJ, Ryan D, Bagby RM. Toward the development of a new self-reprot alexithymia scale. Psychoterr Psychosom 1985;44:191-199.

Thomas BJ, Cornish BH, Pattemore MJ, Jacobs M, Ward LC. A comparison of the whole-body and segmental methodologies of bioimpedance analysis. Acta Diabetol 2003;40:S236-237.

Thompson C, Cowan A. The Seasonal Health Questionnaire: a preliminary validation of a new instrument to screen for Seasonal Affective Disorder. J Affect Disord 2001;64:89–98.

Todd JE, Lader D. Adult dental health 1988 United Kingdom. Her Majesty's Stationery Office. London, 1991.

Tuomi K, Ilmarinen J, Jahkola A, Katajarinne L, Tulkki A. Työkykyindeksi. Työterveyslaitos, Helsinki, 1992. (in Finnish)

Varney LF, Parker RA, Vincelette A, Greenspan SL. Classification of osteoporosis and osteopenia in postmenopausal women is dependent on site-specific analysis. J Clin Densitom 1999;2:275-283.

Vartiainen E, Jousilahti P, Juolevi A, Sundvall J, Alfthan G, Salminen I, Puska P. FINRISKI 1997. Tutkimus kroonisten kansantautien riskitekijöistä, niihin liittyvistä elintavoista, oireista ja terveyspalvelujen käytöstä. Tutkimuksen toteutus ja perustaulukot. Publications of the National Public Health Institute KTL/B1, Helsinki, 1998. (in Finnish)

Vartiainen E, Laatikainen T, Tapanainen H, Salomaa V, Jousilahti P, Sundvall J, Salminen M, Männistö S, Valsta L. Suomalaisten sydän- ja verisuonisairauksien riskitekijät FINRISKI -tutkimuksessa 1982–2002. Suomen Lääkärilehti 2003;58:4099–4106. (in Finnish)

Vehkalahti M, Paunio I, Nyyssönen V, Aromaa A (eds.). Suomalaisten aikuisten suunterveys ja siihen vaikuttavat tekijät. Kansaneläkelaitoksen julkaisuja AL:34, Helsinki ja Turku, 1991. (in Finnish)

WHO Consultation. Definition, diagnosis and classification of diabetes mellitus and its complications. Part 1: Diagnosis and classification of diabetes mellitus. Geneva: World Health Organization, 1999.

WHO. Guide to epidemiology and diagnosis of oral mucosal diseases and conditions. Community Dent Oral Epidemiol 1980;8:1-26.

WHO. Oral health surveys. Basic methods (4. painos) WHO, Geneva, 1997.

Viitasalo JT, Era P, Leskinen AL, Heikkinen E. Muscular strength profiles and anthropometry in random samples of men aged 31-35, 51-55 and 71-75 years. Ergonomics 1985;28:1563–1574.

Wilhelmsen L, Tibblin G, Werkö L. A primary preventive study in Gothenburg, Sweden. Prev Med 1972;1:153-160.

Willett W. Nutritional epidemiology. Oxford: Oxford University Press, 1998.

Williams HC, Burney PGJ, Strachan D, Hay RJ. The U.K. Working Party´s Diagnostic Criteria for Atopic Dermatitis. II. Observer variation of clinical diagnosis and signs of atopic dermatitis. Br J Dermatol 1994;131:397-405. (1994a)

Williams HC, Burney PGJ, Pembroke AC, Hay RJ. The U.K. Working Party´s Diagnostic Criteria for Atopic Dermatitis. III. Independent hospital validation. Br J Dermatol 1994;131:406-416. (1994b)

Winer BJ. Statistical principles in experimental design. Zud ed. Tokio: McGraw-Hill, 1971.

Wittchen HU, Reed, V, Kessler R. The relationship of agoraphobia and panic in a community sample of adolescents and young adults. Archives of General Psychiatry 1998;55:1017-1024. (1998a)

Wittchen HU, Lachner G, Wunderlich U, Pfister H. Test-retest reliability of the computerized DSM-IV version of the Munich-Composite International Diagnostic Interview (M-CIDI). Social Psychiatry & Psychiatric Epidemiology 1998;33:568-578. (1998b)

Yrjämä M, Tervonen O, Kurunlahti M, Vanharanta H. Bony vibration stimulation test combined with magnetic resonance imaging. Can discography be replaced? Spine 1997;22:808-813.

Yrjämä M, Vanharanta H. Bony vibration stimulation: a new, non-invasive method for examining intradiscal pain. Eur Spine J. 1994;3:233-235.

Zain RBM, Ikeda N, Yaacob MBH. Oral mucosal lesions survey of adults in Malaysia. University of Malaysia, Inoffice Equipment SDN. BHD, Malaysia, 1995.

Ziegler A, Kastner C, Blettner M. The generalized estimating equations: an annotated bibliography. Biometrical Journal 1998;40:115–139.

Ziolkowski M, Gruss T, Rybakowski JK. Does alexithymia in male alcoholics constitute a negative factor for maintaining abstinence? Psychother Psychosom 1995;63:169-173.

APPENDICES

Appendix 1. Health 2000 field personnel in September 2000.

Helsinki

Physicians	Hanna Laurén Helsinki	Ulla Litmanen Espoo	
Dentist	Minna Luoto- Heinonen Ryttylä		
Head nurse	Liisa Uusitalo Helsinki		
Dental nurse	Marja Raivio Helsinki		
Field nurses	Liisa Kärki Espoo Physiotherapist	Hanna Lehtonen Helsinki Nurse	Raija Nevala Vantaa Nurse
	Anna-Kaisa Nupponen Tuusula Nurse	Jaana Peltola Helsinki Nurse	Riitta Rasmussen Helsinki Nurse
	Pekka Suominen Hyvinkää Nurse	Tiina Uusitalo Helsinki Nurse	Maarit Williams Helsinki Nurse
Home-visit nurses	Salme Ahlers Hamari Laboratory technician	Maija-Liisa Lahti Helsinki Nurse	

Appendix 1 continued

Turku			
Physicians	Katriina Lagerroos Masku	Päivi Hartman Masku	
Dentist	Tarja Rutkiewicz Porvoo		
Head nurse	Riitta Sipilä Turku		
Dental nurse	Hanna Kylmämetsä Loimaa		
Field nurses	Anne Heino Espoo Nurse	Kirsi Karlsson Kuusisto Physiotherapist	Jaana Laaksonen Raisio Physiotherapist and M.Sc.
	Marjatta Lehtisaari Forssa Nurse	Leena Liljeroos Pori Laboratory technician	Sari Saaranto Raisio Nurse
	Eija Sunell Marttila Nurse	Virpi Teikari Turku Nurse	Eeva-Sirkku Tuuha Turku Nurse
Home-visit nurses	Päivi Aunio Märynummi Nurse	Maarit Saari Vaasa Nurse	

Appendix 1 continued

Tampere			
Physicians	Pia Väkevä Tampere	Arja Laitinen Tampere	
Dentist	Pirkko Koskela Jyväskylä		
Head nurse	Päivi Sirén Siivikkala		
Dental nurse	Riitta-Liisa Piikkilä Kangasala		
Field nurses	Kirsi Henttinen Hämeenlinna Nurse	Johanna Kormano Toijala Nurse	Eija-Liisa Kytölä Kangasala Laboratory technician
	Ella Lehto Tampere Nurse	Marja-Terttu Oksanen Tampere Nurse	Maire Selin Tampere Laboratory technician
	Marjo Virkki Lempäälä Nurse	Leena Yrjänheikki Tampere Nurse	Hanna Äikäs Lahti Nurse
Home-visit nurses	Hilkka Isotalo Tampere Nurse	Anne Laakso Tampere Nurse	

Appendix 1 continued

Kuopio

Physicians	Tuula Jokiniemi Kuopio	Marja-Leena Peltola Helsinki	
Dentist	Liisa Suominen-Taipale Kuopio		
Head nurse	Kirsi Tiihonen Kuopio		
Dental nurse	Ulla Tyyni Kuopio		
Field nures	Outi Himanen Mikkeli Physiotherapist	Kristiina Holopainen Kuopio Nurse	Mari Hänninen Kuopio Nurse
	Leila Kaatrasalo Lapinlahti Laboratory technician	Rauno Kallo Mikkeli Physiotherapist	Tuula Laaksovirta Leppäkaarre Nurse
	Kirsi-Marja Manninen Kuopio Nurse	Kirsti Pyy Valkeala Nurse	Sari Suutarinen Kuopio Nurse
Home-visit nurses	Carita Röpelinen Savonlinna Nurse	Tuomas Onnukka Kuopio Laboratory technician	

Appendix 1 continued

Oulu			
Physicians	Tuula Pulska Oulu	Aimo Korpilähde Posio	
Dentist	Mirka Niskanen Ylivieska		
Head nurse	Kristiina Väisänen Jääli		
Dental nurse	Mervi Konttila Oulu		
Field nures	Liisa Annila Kemijärvi Nurse	Maija Harvio Oulu Nurse	Anne Jaakkola Oulu Laboratory technician
	Marja-Liisa Kuvaja Jormua Nurse	Raija Salo Kokkola Nurse	Marianne Stelander Luumäki Nurse
	Tiina Säätelä Kokkola Physiotherapist	Marjatta Takala Oulu Nurse	Sarianna Vaara Oulu Nurse
Home-visit nurses	Terttu Törmänen Kuusamo Dental nurse, Nurse	Aira Pohjonen Jääli Nurse	

Standby			
Dentist	Anneli Ahovuo-Saloranta Tampere		
Dental nurse	Raila Forss Kuopio	Arja Salovaara Ryttylä	Mari Ylöstalo Lahti

Appendix 2. Phases of data collection and field personnel in the Health 2000 Survey.

AT HOME:

90 minutes **INTERVIEW** (by Statistics Finland's interview organisation)

30 minutes **FILLING IN QUESTIONNAIRE 1**

AT HEALTH CENTRE ETC.:

15 minutes **1 REGISTRATION** (observer 1)
- information, informed consent, Symptoms Interview
- handing Questionnaire 2 and the urine sample container

15 minutes **2 MEASUREMENTS: height, body circumference, ecg, blood pressure** (observer 2)

15 minutes **3 MEASUREMENTS: spirometry, bioimpedance, heel bone density** (observer 3)

15 minutes **4 LABORATORY** (observers 4 and 5)
- drawing blood samples (100 ml), sample processing

15 minutes **5 ORAL EXAMINATION** (observers 6 and 7)
- clinical oral examination, orthopantomography

15 minutes **SNACK, FILLING IN QUESTIONAIRE 2**

30 minutes **6a FUNCTIONAL CAPACITY TESTS** (observer 8) **6b FUNCTIONAL CAPACITY TESTS** (observer 9)
- physical and cognitive capacity, vision and hearing

30 minutes **7a CLINICAL EXAMINATION** (observer 10) **7b CLINICAL EXAMINATION** (observer 11)

30 minutes **8a MENTAL HEALTH INTERVIEW** (observer 12) **8b MENTAL HEALTH INTERVIEW** (observer 13)

15 minutes **9 FINAL INTERVIEW** (observer 14)
- checking that all examinations and questionnaires have been completed
- handing Questionnaire 3 and Dietary Questionnaire
- information about the previous and possible further examinations

altogether about 3 hours and 15 minutes

AT HOME:

(100 minutes) **(HEALTH EXAMINATION FOR THOSE NOT ATTENDING THE HEALTH EXAMINATION PROPER AT THE HEALTH CENTRE ETC.**) (observers 15 and 16)

40 minutes **FILLING IN QUESTIONNAIRE 3 AND DIETARY QUESTIONNAIRE**

AT UNIVERSITY HOSPITALS AND RESEARCH INSTITUTES:

FURTHER EXAMINATIONS FOR SUBSAMPLES

FROM REGISTERS:

REGISTER DATA

Appendix 3. Health 2000 organisation during the planning and data collection phases.

Steering group

Markku Lehto (chair), Seppo Koskinen (secretary), Arpo Aromaa, Jorma Back, Mikael Forss, Aino-Inkeri Hansson, Kaija Hasunen, Jussi Huttunen, Heli Jeskanen-Sundström, Satu Lahti, Simo Lämsä, Jorma Rantanen, Vappu Taipale, Matti Uimonen and Jukka Wuolijoki (at the beginning Timo Relander represented Statistics Finland and Esko Kalimo represented the Social Insurance Institution)

Managing group

Jussi Huttunen (chair), Seppo Koskinen (secretary), Arpo Aromaa, Sami Heistaro and Antti Reunanen

Project group

Arpo Aromaa (chair), Terhi Saarinen (secretary), Sami Heistaro, Markku Heliövaara, Unto Häkkinen, Olli Impivaara, Pekka Jousilahti, Paul Knekt, Seppo Koskinen, Jouko Lönnqvist, Tuija Martelin, Anne Nordblad, Veijo Notkola, Antti Reunanen, Hilkka Riihimäki, Petri Ruutu, Jouko Sundvall, Antti Uutela and Erkki Vartiainen

Executive committee: Arpo Aromaa (chair), Terhi Saarinen (secretary), Sami Heistaro, Seppo Koskinen and Vesa Tanskanen

Working group for living conditions

Tuija Martelin (chair), Tarja Nieminen (secretary), Matti Heikkilä, Sakari Kainulainen, Timo Kauppinen, Eero Lahelma, Simo Mannila and Veijo Notkola

Working group for health behaviour and psychosocial factors

Antti Uutela (chair), Ritva Prättälä (secretary), Anna-Mari Aalto, Hannu Alho, Arja R. Aro, Markku Heliövaara, Sari Isotupa, Paul Knekt, Päivikki Koponen, Marjaana Lahti-Koski, Esko Mälkiä, Satu Männistö, Pekka Oja, Pirjo Pietinen, Raimo Raitasalo, Antti Reunanen, Sakari Suominen, Jussi Vahtera, Liisa Valsta, Miira Vehkalahti and Eira Viikari-Juntura

Working group for cardiovascular diseases

Antti Reunanen (chair), Anna Kattainen (secretary), Matti Jauhiainen, Antti Jula, Risto Kaaja, Antero Kesäniemi, Katriina Kukkonen-Harjula, Mika Kähönen, Markku Laakso, Riitta Luoto, Silja Majahalme, Leena Mykkänen, Markku S. Nieminen, Janne Rapola, Veikko Salomaa, Marja-Riitta Taskinen, Jaakko Tuomilehto, Marjut Varpula and Erkki Vartiainen

Working group for musculoskeletal diseases

Hilkka Riihimäki (chair), Markku Heliövaara, Sami Heistaro (secretary), Olli Impivaara, Tuula Jokiniemi, Satu Luoto, Pirjo Manninen, Matti Mäkelä, Simo Taimela, Esa-Pekka Takala and Eira Viikari-Juntura

Working group for respiratory and skin diseases

Pekka Jousilahti (chair), Tari Haahtela, Sami Heistaro (secretary), Markku Heliövaara, Jussi Karjalainen, Kaj Koskela, Henrik Nordman, Timo Palosuo, Juha Pekkanen, Tuula Petäys, Kari Reijula and Päivikki Susitaival

Working group for mental health

Jouko Lönnqvist (chair), Sami Pirkola (secretary), Kirsi Ahola, Martti Heikkinen, Teija Honkonen, Erkki Isometsä, Matti Joukamaa, Raija Kalimo, Olli Kiviruusu, Teemu Kärnä, Eero Lahtinen, Ville Lehtinen, Kari Poikolainen, Raimo Raitasalo, Jouko Salminen and Jaana Suvisaari

Working group for oral health

Anne Nordblad (chair), Sinikka Varsio (secretary), Sirkkasisko Arinen, Dorrit Hallikainen, Hannu Hausen, Matti Knuuttila, Liisa Suominen-Taipale, Anna-Lisa Söderholm and Miira Vehkalahti

Working group for communicable diseases

Petri Ruutu (chair), Markku Kuusi (secretary), Juhani Eskola, Pentti Huovinen, Hannele Jousimies-Somer, Ilkka Julkunen, Eija Könönen, Pauli Leinikki, Tuija Leino, Anja Siitonen and Martti Vaara

Working group for cancer

Paul Knekt (chair), Lyly Teppo, Matti Rautalahti, Risto Sankila and Jarmo Virtamo

Working group for functional capacity

Seppo Koskinen (chair), Päivi Sainio (secretary), Arpo Aromaa, Pertti Era, Pauli Forma, Raija Gould, Päivi Haavisto, Jukka-Pekka Halonen, Kaj Husman, Juhani Ilmarinen, Jorma Järvisalo, Sirkka-Liisa Karppi, Jarmo Malmberg, Simo Mannila, Timo Marttila, Seppo Miilunpalo, Matti Ojamo, Sirkka-Liisa Rudanko, Sanna Räty, Raimo Sulkava, Timo Suutama, Reijo Tilvis and Mariitta Vaara

Working group for health services use

Unto Häkkinen (chair), Pirkko Alha (secretary), Ilmo Keskimäki, Timo Klaukka, Päivikki Koponen and Kimmo Räsänen

Working group for sampling, survey execution, data processing and analysis

Risto Lehtonen (chair), Kari Djerf, Pirjo Hyytiäinen, Tommi Härkänen, Paul Knekt, Kari Kuulasmaa, Vesa Kuusela, Johanna Laiho, Marjo Laine, Paula Lamberg, Jouni Maatela, Tuija Martelin, Erkki Nenonen, Mikko Nenonen, Tarja Nieminen, Mikko Nissinen, Timo Peltomaa, Harri Rissanen, Pentti Salmela, Matti Sarjakoski, Eero Tanskanen, Vesa Tanskanen, Tuula Tiainen, Kai Vikki and Esa Virtala

Working group on laboratory tests

Jouko Sundvall (chair), Marja Leena Kantanen (secretary), Georg Alfthan, Matti Jauhiainen, Kimmo Kuoppasalmi, Eija Könönen, Jaana Leiviskä, Jukka Marniemi, Kimmo Peltonen, Markus Perola, Irma Salminen, Vesa Tanskanen and Ismo Ulmanen

Working group on clinical medical examination

Antti Reunanen (chair), Markku Heliövaara, Anna Kattainen, Satu Luoto, Tuula Petäys and Esa-Pekka Takala

Working group on the preparation of the field work

Sami Heistaro (chair), Terhi Saarinen (secretary), Pirkko Alha, Arpo Aromaa, Jaason Haapakoski, Olli Impivaara, Pekka Jousilahti, Seppo Koskinen, Jaana Leiviskä, Jouni Maatela, Tarja Nieminen, Veijo Notkola, Panu Oksa, Marjatta Riisiö, Sirkka Rinne, Matti Sarjakoski, Jouko Sundvall, Vesa Tanskanen and Kai Vikki

People contributing to the project in the participating organisations

The sample was defined and formed at Statistics Finland under the supervision of Risto Lehtonen. The sampling process itself was conducted at the Social Insurance Institution's Information Systems Department under Erkki Nenonen. The weighting coefficients were constructed at Statistics Finland.

The interviews were carried out by Statistics Finland's interview staff of around 160.

The health examinations were conducted by KTL field units. Each of the five units had a staff of 16–17, working under the following nominated nurses-in-charge: Liisa Uusitalo (Helsinki), Päivi Sirén (Tampere), Riitta Sipilä (Turku), Kirsi Tiihonen (Kuopio), and Kristiina Väisänen (Oulu). Sami Heistaro had overall responsibility for the organisation and coordination of field operations.

Laboratories: The SII Biochemistry Laboratory under the supervision of Jukka Marniemi and the KTL Laboratory of Analytical Biochemistry under Jouko

Sundvall were responsible for most of the determinations. DNA isolations were performed at the KTL Department of Molecular Medicine under Ismo Ulmanen. Irma Salminen was in charge of the laboratory operations out in the field, while Jaana Leiviskä assumed responsibility for the transfer and storage of samples.

Project information management and data communications were planned and implemented by project staff as well as by the KTL's ADP unit under the supervision of Mikko Nissinen and the information system unit under Jaason Haapakoski.

During the fieldwork Virpi Killström at project headquarters was responsible for various aspects of the operation, including personnel resources, employment contracts and travel arrangements.

A large number of people at project headquarters contributed to final data editing and to the various tasks involved in the preparation of this report. Apart from the members of the working groups proper, these people included Heidi Alha, Pirkko Alha, Päivi Haavisto, Katri Hakulinen, Hannele Ikkala, Tarja Kiesi, Onni Koskinen, Noora Kuosmanen, Tomi Mäkinen, Marjatta Riisiö, Saila Rinne, Sirkka Rinne, Harri Rissanen, Sanna Räty, Terhi Saarinen and Ulla Tyyni.

At KTL Tommi Härkänen, Paul Knekt and Esa Virtala were responsible for the planning of statistical analyses and the selection and development of the methods employed. Esa Virtala was also in charge of creating the files as well as programming and designing the tabular results. In addition to the names mentioned in the previous paragraph, Mikko Pekkarinen worked on taking printouts.

Appendix 4. Research localities, periods and sample sizes (planned) [actual] by field group.

Research locality (sample size)	Municipalities included in the sample	Period
Field group 1		
Helsinki (900)	Helsinki	11.09.2000–22.11.2000
Espoo (309)	Espoo	27.11.2000–15.12.2000
Hyvinkää (88)	Hyvinkää	02.01.2001–05.01.2000
Tuusula (88) [85]	Tuusula	08.01.2001–12.01.2001
Loviisa (88)	Lapinjärvi, Liljendal, Loviisa, Pernaja, Ruotsinpyhtää	15.01.2001–19.01.2001
Vantaa (260) [259]	Vantaa	22.01.2001–16.02.2001
Porvoo (88)	Porvoo	19.02.2001–23.02.2000
Field group 2		
Tampere (310) [311]	Tampere	11.09.2000–13.10.2000
Hämeenlinna (101)	Hattula, Hauho, Hämeenlinna, Kalvola, Renko	16.10.2000–20.10.2000
Heinola (100)	Heinola	23.10.2000–27.10.2000
Nokia (100)	Nokia	30.10.2000–03.11.2000
Keuruu (80) [82]	Keuruu, Multia	06.10.2000–10.10.2000
Orimattila (100) [101]	Artjärvi, Myrskylä, Orimattila, Pukkila	13.11.2000–17.11.2000
Forssa (100)	Forssa, Humppila, Jokioinen, Tammela, Ypäjä	20.11.2000–24.11.2000
Somero (50)	Somero	27.11.2000–01.12.2000
Loimaa (50)	Alastaro, Loimaa, Loimaan kunta, Mellilä, Oripää	04.12.2000–08.12.2000
Valkeakoski (100)	Valkeakoski	11.12.2000–15.12.2000
Jämsä (80) [82]	Jämsä, Jämsänkoski, Kuhmoinen	02.01.2001–05.01.2001
Riihimäki (100) [102]	Hausjärvi, Loppi, Riihimäki	08.01.2001–12.01.2001
Karkkila (88)	Karkkila	15.01.2001–19.01.2001
Lahti (150) [151]	Lahti	19.02.2001–25.02.2001
Muurame (80) [82]	Korpilahti, Muurame	26.02.2001–02.03.2001
Field group 3		
Kuopio (140) [130]	Kuopio	11.09.2000–29.09.2000
Lappeenranta (90)	Lappeenranta	02.10.2000–06.10.2000
Joensuu (80)	Joensuu	09.10.2000–12.10.2000
Juuka (80)	Juuka	13.10.2000–17.10.2000
Ruokolahti (88)	Ruokolahti	18.10.2000–23.10.2000
Luumäki (88)	Luumäki	24.10.2000–27.10.2000
Kotka (90)	Kotka	30.10.2000–03.11.2000
Siilinjärvi (80) [83]	Maaninka, Siilinjärvi	06.11.2000–10.11.2000
Lieksa (80)	Lieksa	13.11.2000–16.11.2000
Pyhäselkä (80)	Kiihtelysvaara, Pyhäselkä	20.11.2000–24.11.2000
Imatra (88)	Imatra	27.11.2000–01.12.2000
Kerimäki (80)	Enonkoski, Kerimäki, Savonranta	04.12.2000–08.12.2000
Mikkeli (80) [83]	Anttola, Hirvensalmi, Mikkeli, Mikkelin mlk, Ristiina	02.01.2001–05.01.2001
Polvijärvi (80)	Polvijärvi	13.01.2001–16.01.2001
Hamina (88)	Hamina, Vehkalahti, Virolahti	22.01.2001–26.01.2001

Jyväskylä (120) [113]	Jyväskylä	10.02.2001–18.02. 2001
Savonlinna (80) [81]	Punkaharju, Rantasalmi, Savonlinna	19.02.2001–23.02.2001
Kouvola (88)	Kouvola, Valkeala	26.02.2001–02.03.2001

Field group 4

Oulu (180) [182]	Oulu	11.09.2000–06.10.2000
Ilmajoki (100)	Ilmajoki	09.10.2000–13.10.2000
Kajaani (60)	Kajaani	16.10.2000–20.10.2000
Haukipudas (60) [59]	Haukipudas	23.10.2000–26.10.2000
Utsjoki (60) [61]	Utsjoki	27.10.2000–31.10.2000
Kemi (60 [62]	Kemi	01.11.2000–03.11.2000
Nivala (60)	Nivala	06.11.2000–08.11.2000
Ylivieska (60) [62]	Alavieska, Sievi, Ylivieska	09.11.2000–13.11.2000
Pyhäjärvi (60)	Pyhäjärvi	14.11.2000–17.11.2000
Seinäjoki (100) [101]	Nurmo, Peräseinäjoki, Seinäjoki, Ylistaro	20.11.2000–27.11.2000
Sodankylä (60)	Sodankylä	28.11.2000–01.12.2000
Kiiminki (60)	Kiiminki	04.12.2000–08.12.2000
Lapua (100)	Lapua	11.12.2000–15.12.2000
Lapinlahti (80) [83]	Lapinlahti	02.01.2001–05.01.2001
Iisalmi (80)	Iisalmi, Vieremä	08.01.2001–12.01.2001
Raahe (60) [55]	Pattijoki, Pyhäjoki, Raahe, Ruukki, Siikajoki	15.01.2001–18.01.2001
Simo (60)	Kuivaniemi, Simo	19.01.2001–25.01.2001
Kuusamo (60)	Kuusamo	26.01.2001–28.01.2001
Rovaniemi (60) [57]	Rovaniemi	29.01.2001–01.02.2001
Taivalkoski (60)	Taivalkoski	21.02.2001–26.02.2001
Kokkola (60) [63]	Kokkola, Kälviä, Lohtaja, Ullava	27.02.2001–02.03.2001

Field group 5

Turku (330) [278]	Piikkiö, Turku	11.09.2000–11.10.2000
Kaarina (50) [51]	Kaarina	12.10.2000–13.10.2000
Pietarsaari (100) [99]	Luoto, Pedersöre, Pietarsaari	16.10.2000–20.10.2000
Ulvila (50)	Kullaa, Ulvila	23.10.2000–25.10.2000
Harjavalta (50)	Harjavalta, Kiukainen, Nakkila	26.10.2000–27.10.2000
Masku (50) [49]	Askainen, Lemu, Masku, Nousiainen, Vahto	30.10.2000–01.11.2000
Parainen (50)	Parainen	02.11.2000–03.11.2000
Rauma (50)	Eurajoki, Kodisjoki, Lappi, Rauma	06.11.2000–08.11.2000
Kristiinankaupunki (100)	Isojoki, Karijoki, Kristiinankaupunki	09.11.2000–17.11.2000
Vehmaa (50)	Kustavi, Taivassalo, Vehmaa	20.11.2000–22.11.2000
Salo (50)	Salo	23.11.2000–27.11.2000
Lohja (88)	Karjalohja, Lohja, Nummi-Pusula, Sammatti	28.11.2000–03.12.2000
Perniö (50)	Perniö, Särkisalo	07.12.2000–08.12.2000
Kokemäki (50)	Kokemäki	08.01.2001–10.01.2001
Uusikaupunki (50)	Uusikaupunki	11.01.2001–12.01.2001
Pori (120)	Pori	15.01.2001–24.01.2001
Uusikaarlepyy (100)	Uusikaarlepyy	13.02.2001–20.02.2001
Naantali (50) [52]	Merimasku, Naantali, Rymättylä	21.02.2001–23.02.2001
Vaasa (90) [89]	Vaasa	26.02.2001–02.03.2001

Appendix 5. Phases of data collection and field personnel in the resurvey of persons who participated in the Mini-Finland Health Survey in 1978-1980.

AT HOME:

30 minutes **FILLING IN QUESTIONNAIRE**

AT HEALTH CENTRE ETC.:

15 minutes **1 REGISTRATION** (observer 1)
– information, Symptoms Interview

15 minutes **2 MEASUREMENTS** (observer 2)
– height, ecg, blood pressure

15 minutes **3 MEASUREMENTS** (observer 3)
– spirometry, bioimpedance, heel bone density

15 minutes **4 LABORATORY** (observer 4)
– drawing blood samples (50 ml), sample processing

15 minutes **5 ORAL EXAMINATION** (observers 6 and 7)
– clinical oral examination, orthopantomography

15 minutes SNACK

45 minutes **6 INTERVIEW** (observers 5, 12 and 13)

30 minutes **7 FUNCTIONAL CAPACITY TESTS** (observers 8 and 9)
– physical and cognitive capacity, vision and hearing

30 minutes **8 CLINICAL EXAMINATION** (observers 10 and 11)
– physician's clinical examination

15 minutes **9 FINAL INTERVIEW** (observer 14)
– checking that all examinations and the questionnaire have been completed

altogether about 3.5–4 hours

(**AT HOME:** whenever possible for those not attending the health examination proper at health centre etc.)

(60 minutes) (interview and questionnaire (observers 15 and 16))