An updated version of this publication is available at
http://urn.fi/URN:ISBN:978-952-302-700-8

EHES
Measuring the Health of Europeans

EHES Manual

PART A. PLANNING AND PREPARATION OF THE SURVEY

Hanna Tolonen (editor)

Directions 2013_001

Edited by

Hanna Tolonen

EHES MANUAL

PART A. PLANNING AND PREPARATION OF THE SURVEY

Helsinki 2013

The EHES Pilot Project received funding from the European Commission/DG Sanco. The views experessed here are those of the authors and they do not represent the Commission's official position.

Publisher National Institute for Health and Welfare

PO Box 30, FI-00271 Helsinki, FINLAND

http://www.thl.fi

ISBN (pdf) 978-952-245-842-1

ISSN (pdf) 2323-4172

http://urn.fi/URN:ISBN:978-952-245-842-1

Helsinki, Finland 2013

Reference:

Tolonen H (Ed.) EHES Manual. Part A. Planning and preparation of the survey. National Institute for Health and Welfare, 2013, Directions 2013_001. URN:ISBN:978-952-245-842-1, URL: http://urn.fi/URN:ISBN:978-952-245-842-1

Layout: Hanna Tolonen

Cover: Hanna Tolonen

Graphics:

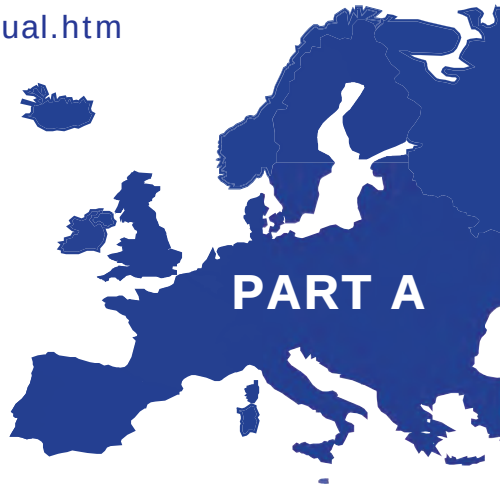
Chapter 3: Johan Heldal and Susie Jentoft

Chapter 10: Georg Alfthan

Chapter 12: Ari Haukijärvi

Other Chapters: Hanna Tolonen

Contents



Introduction

The European Health Examination Survey (EHES) Manual provides guidelines and specifies the requirements for the implementation of standardized national health examination surveys (HES) in the European countries. Recommendations based on past experiences from national and international surveys were prepared by the Feasibility of a European Health examination Survey (FEHES) Project (Tolonen 2008). The EHES manual builds on these recommendations and on further experience obtained during the EHES Pilot Project in 2009-2012. The EHES Manual has three parts:

A. Planning and preparation of the survey
B. Fieldwork procedures
C. European level coordination

The EHES Manual is maintained by the EHES Reference Centre, and the plan is to update it with further clarifications and additional relevant topics. The latest version of the EHES Manual is available in the Internet at www.ehes.info.

This is Part A of the EHES Manual. It provides guidelines for the planning and preparation of national health examination surveys.

As part of the planning of a national HES, each country has to prepare a national HES Manual. The procedures described in the national manual should follow the European standards specified in the EHES Manual. The national manual should be specific also in issues where the EHES manual can only give alternatives or general guidelines. The EHES manual is unspecific in situations where the national circumstances vary and there is no common procedure which could be reasonably followed in all countries. When the European recommendation differs from the procedure used in earlier national surveys, the procedure to be adapted in the new national HES needs to be considered carefully. Sometimes there may be need to compromise between European comparability and the possibility to follow national trends from the past. The countries should prepare the national manuals in collaboration with the EHES Reference Centre.

References

- Tolonen H, Koponen P, Aromaa A, et al. (Eds.) Recommendations for the Health Examination Surveys in Europe. B21/2008, Publications of the National Public Health Institute, Helsinki 2008. Also available from http://urn.fi/URN:ISBN:978-951-740-838-7

1. Survey planning and management

Päivikki Koponen[1], Hanna Tolonen[1], Arpo Aromaa[1], Kari Kuulasmaa[1]

[1] National Institute for Health and Welfare (THL), Helsinki, Finland

High quality planning and management are the keys to achieving the survey's objectives. The planning process ensures that the survey can be effectively implemented in the shortest reasonable time, within the budget and with the highest quality that is affordable and consistent with the aims and purposes (Franklin & Walker 2003). All survey plans need to be repeatedly overhauled depending on the progress of action. This requires efficient management. This chapter focuses on national activities in the planning and preparation of the national surveys, in particular survey management.

1.1 Survey process

The first step in planning and preparation includes defining the aims and purpose of the survey. These will be the basis for selecting the topics and actions in the data collection. They will also guide the decisions on how the EHES standards will be implemented in the national survey. The aims and purposes of the national survey will rely on national and European level health policies, and information needs. National health care systems, previous health surveys and expertise available in the country will also affect the feasibility of different options for carrying out the survey. All decisions need to be made in the context of previous national HIS and HES, as well as other major health surveys in each country. If there are other national surveys, such as surveys on nutrition, lifetyles or health behaviour, or other health interview surveys, the HES needs to be timed and tailored to fit in the national health survey system. An evaluation of already exisiting data sorces is needed to define if the HES is the best way to collect the data. As the national HESs are anticipated to be repeated with regular intervals, the survey planning process needs

to be ongoing with experiences and results on previous surveys leading to the next phase of data collection (Figure 1.1).

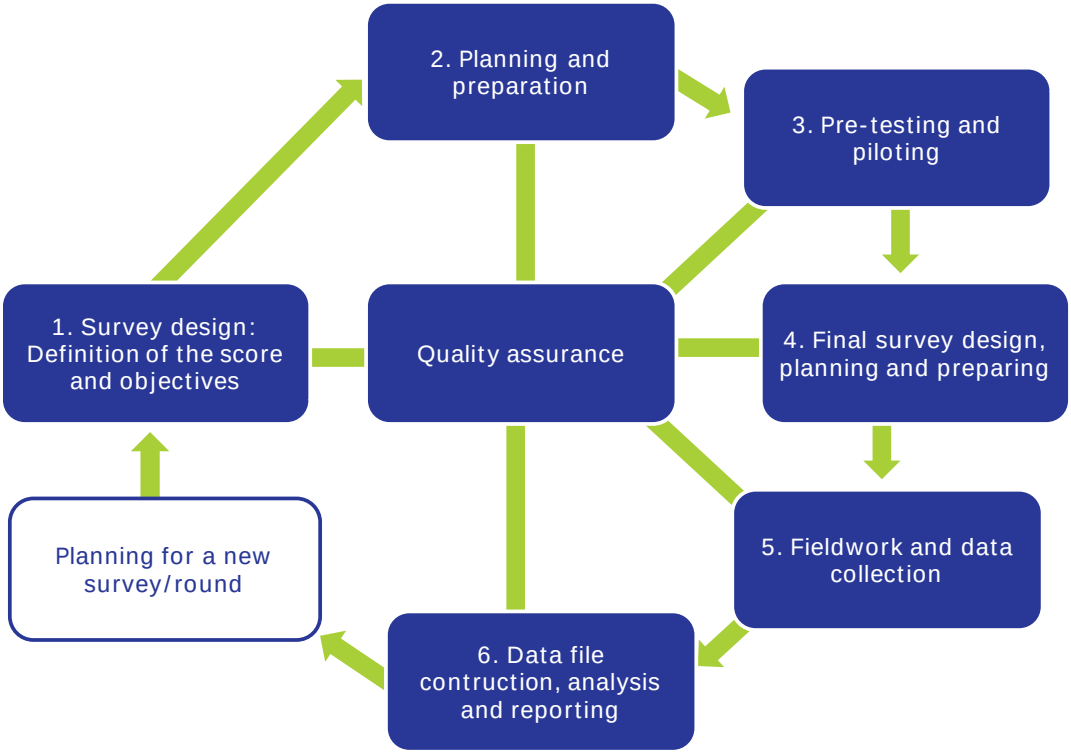


Figure 1.1 Stages in the survey process (stages adapted from Fraklin & Walker 2003, Czaja & Blair 2005)

Six main stages in the survey processa are shown in Figure 1.1. Even though these stages in survey planning, preparation, fieldwork, data processing and analysis as well as reporting proceed after each other, there is a need to return to previous stages througout the survey process to adapt the plans according to experiences and feedback from different experts and stakeholders.

- The output of the stage 1 of the planning of the survey is the first version of the survey proposal. Commitment from key organizations such as the ministry and the national public health institute, national statistical institute and other relevant organizations can be sought based on these preliminary plans and ideas. The survey management structure is also defined as well as a preliminary time schedule for the survey.
- Stage 2 includes the detailed planning of the sampling, survey contents, fieldwork and data collection, data management as well as a preliminary plan for analysis and reporting. The output is a detailed survey plan with the budget and a first draft of the survey manual including the questionnaires, and measurement protocols, and other materials (information leaflets, consent

forms etc). Ethical approval is sought based on the detailed proposal.

- Stage 3 includes pretesting and piloting. After this the proposals and manuals, as well as all survey materials (including the computer programs, survey web-sites as well as cummunication plans) can be finalized.
- Stage 4 sets up the fieldwork and data collection system. Specific attentention should be given to motivation of participants. The fieldwork staff can be hired and trained, first invitations can be launched and first appointments to the interviews and examinations can be scheduled.
- Stage 5 includes the proper fieldwork and data collection. Some chenges and adaptations to original plans may still be needed, e.g. if participation rates in the first weeks are low or if other problems are faced.
- Stage 6 includes finalizing the data sets, documenting data characteristics and quality, the data analysis, as well as reporting and disseminating results.
- Quality assurance is essential throughout the survey process (see Part A, Chapter 11 of the EHES Manual).

1.2 Aims and purpose of the survey

Clearly defined and specified aims and puposes guide the survey planning and fieldwork. Time spent in the development of specific aims is time saved in the design of survey instruments and mesurements (Biemer & Lyber 2003). There are typically interests to include several topics, instruments and measurements in the survey, but all of them are not feasible due to limited time and other resources. The purpose of the survey depends on national needs and uses of health information, e.g. implementation of the ECHI indicators (see www.echim.org). Relevant and valid health information is needed for evidence based health policy, rational planning and evaluation of health promotion and disease prevention programmes, and health services. In each country the objectives of the survey should take into account ongoing or planned national health promotion programmes and key challenges in developing health services to meet the needs of all population groups. Monitoring and forecasting the population's health and health determinants are prerequisites for sound evidence based public health policy, directing and designing health programmes and services as well as social security. HESs can enhance knowledge on health determinants, health needs and population health.The information from a HES is typical used to:

- assess the prevalence of major diseases and their risk factors;
- assess health status and its association with health promotion and disease prevention;
- measure change at an individual (if follow up of the participants is possible) and population level (with regularly repeated surveys);
- predict future health status in the population, based on objective information on major chronic disease risk factors (such as blood lipid levels, obesity);
- analyze equity in health, health care and well being by providing objective data, comparable in all groups in the population;
- estimate met and unmet need for health care, social security benefits and rehabilitation, and to forecast future scenarios concerning the need for health care and social security benefits;
- develop national standards and reference values for the measurements;
- develop a valuable data source for epidemiological studies and health sciences research.

The aims of HESs should be specified and evaluated against other potential sources of health information in each country, such as health interview surveys and administrative registers. This evaluation will show the added potential of HESs to retrieve health information. A HES provides exclusive data on many topics such as disease risk factors not available in any other source. Also, HESs can result in comparable data for many health indicators which are known to differ between countries and between socioeconomic groups. The standardized measurements of health examinations can overcome reporting bias, e.g. the tendency to over-report height and under report weight (Gillium & Sempos 2005, Elgar & Stewart 2008). HESs can also reveal shortcomings in the awareness of risk factors, e.g. having high blood pressure (Kastarainen et al 2009, Ostchega et al 2008). HESs provide population prevalence data also in situations where such data cannot be obtained from routine registers because of limited access and use of health services. For example, routine registers reveal diabetes or cardiovascular disease only in those who have used services and been diagnosed (Gnavi et al 2008, Elo & Karlberg 2009).

The scope of the core EHES is limited to the health of the adult population, as both children and the elderly have their own specific health problems, health risks and protective factors, often requiring specific measurements. Surveys among children and the elderly also have their own challenges in regard to survey

ethics and fieldwork practices, which is why the EHES standards are at first targeted to adult health surveys. The EHES survey can be extended to also cover the elderly as the core measurements are feasible with similar methods among the elderly, but their specific needs should be taken into account (e.g. inclusion of insitutionalized persons, scheduling appointments, and consent among those with cognitive disabilities). Age-group specific measurements and other additions will be developed later and included in the EHES Manual.

1.3 Implementing EHES standards in national surveys

Countries may implement the EHES standards within one of the following options: 1) building a new national HES, 2) synchronizing EHES standards with the existing HES, or 3) incorporating an existing national HIS with the EHES standards (Tolonen et al 2008). There are three alternatives:

1. When building a new national HES without any (recent) prior HESs in the country, the planning and implementation of the survey should be based on the EHES standards. The challenge for the planning and preparation is to set up the survey using the European standards in the national circumstances. National experts need to decide which of the options in this manual are most feasible in their country, taking into account how these choices affect the comparability of data.
2. When synchronizing EHES module(s) and standards with an existing national HES (incorporating the EHES module(s) into the previous national modules), the challenges are balancing the need to follow national time trends and to ensure European comparability. A specific pilot study may be needed to compare results from examinations carried by different protocols. Some measurements and/or questions may need to be administered to the same respondents in two different ways.
3. When combining the EHES with a national HIS, the challenge is in organizing the data collection successfully and minimizing selection bias. This approach may lead to a survey with several phases in the data collection. Everybody in the sample (i.e. not only the respondents of the HIS) should be invited to the HES. There are several examples showing that inviting only the participants to previous phases leads to a diminishing participation rate for HES. The HES phase should

be used for complementing lacking HIS information in HIS nonparticipants.

A HES always includes a questionnaire or interview. Sometimes the questionnaires and interviews may be very extensive and time-consuming. For example, if the survey includes the full EHIS questionnaire, then the HES serves also the needs of EHIS and national HIS.

Other options may also be considered in some countries and the feasibility of these need to be evaluated. Some countries may wish to undertake pilots of collecting information through national health screening services, where a certain age group is invited to screening examinations carried out in primary health care. It can be decided only after the evaluation of the EHES pilots whether such screenings can be standardized to produce data that meets the EHES quality criteria. Key issues in the feasibility of screenings for national health monitoring purposes are their coverage at population level (assuring the representativeness and avoiding selection bias), and standardization of the measurements (e.g. local premises, equipment and training of personnel).

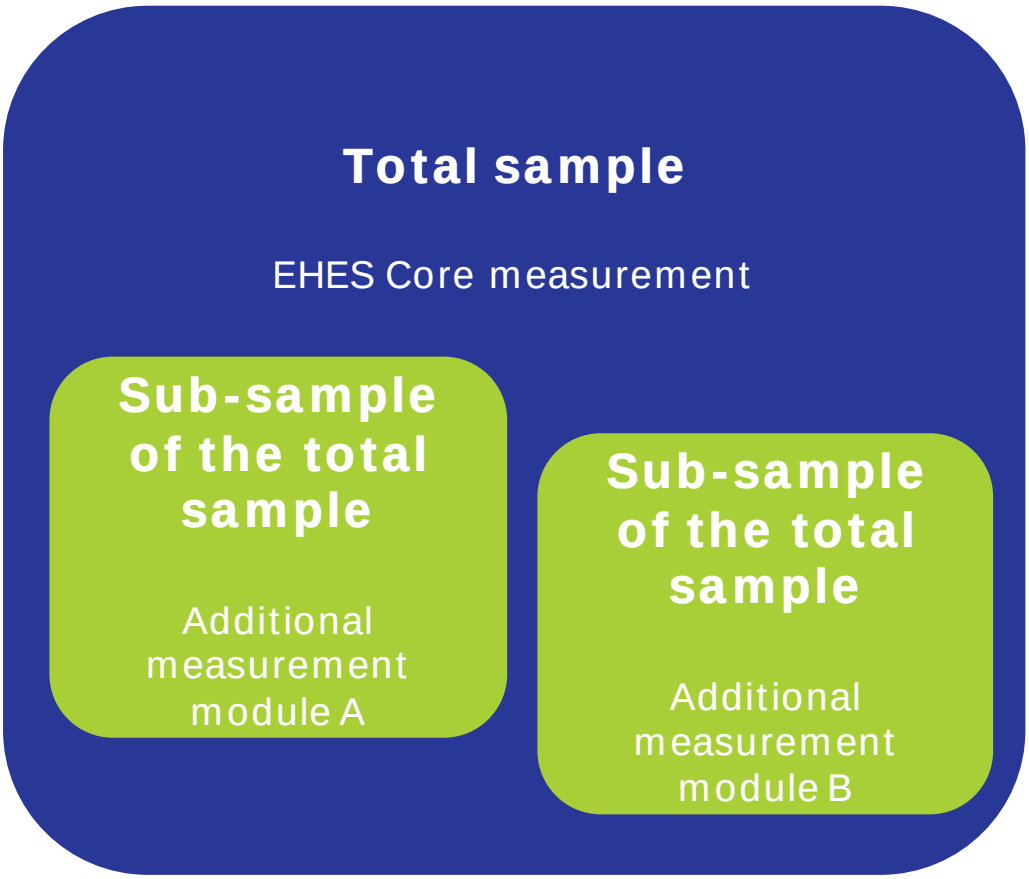


Figure 1.2. Example of a modular structure in the survey

A modular structure can be considered if the survey covers seveal additional topics which are not relevant of feasible to all population groups (Figure 1.2). These modules will need to be taken into account in the survey management and fieldwork logistics. There may be additional measurement modules e.g. an on functional ability for those aged 65 and over. An additional measurement module for a sub-sample may include e.g. a time consuming mental health interview or a dental examination which is not feasible for the total sample due to limited resources.

1.4 Survey management

Often an interdisciplinary survey team is given responsibility for the planning, design, implementation and evaluation of the survey. A core group of key experts is needed to ensure that different aspects are taken into account. In addition, many other experts are needed, and within larger survey organizations their work needs to be organized in different teams, led by members of the core group or others closely involved in the survey. In smaller survey organizations various experts may be consulted without involving them in the actual survey organization. Various types of expertise should include:

- Policy experts to define the needs and use of data for evidence based health policy and to use the results for these purposes;
- Health care and other public service professionals to define the needs of data for planning and evaluating health services and health promotion activities and to use the results for these purposes;
- Scientists in the fields of epidemiology, statistics, public health, other health sciences, social sciences etc. to define the use of the data for scientific research purposes;
- Other experts, such as experts in fieldwork logistics and supervision, laboratory issues, data management, information technology, communication and dissemination etc. to make sure that the data collection runs without problems and to assure high quality data.

It is also recommended to involve different stakeholders such as ministries (e.g. health and research), social insurance organizations, and non governmental organizations to express their interests for the survey, to promote the survey for fund raising and raising interest among the population to participate, and to disseminate the results.

1.4.1. Management structure

The organizational responsibilities of a HES can be divided into (adapted from Tolonen et al 2002):

1. Planning: Definition of the objectives and scope of the survey, planning and preparing the fieldwork and other survey operation.
2. Operation: Implementation and operation of systems for data collection (fieldwork) and data processing.
3. Quality assurance: An authority (if needed independent of the logistics operations) that monitors performance, provides feedback, and ensures that the results are within predefined quality limits.

Planning and operation are most often lead by the same organization, while in some countries e.g. the Ministry or National Public Health Institue are responsible for planning while the organization responisble for the operation is selected from competing organizations. It also needs to be decided in each country if there is a need to carry out the quality assurance by an organization or persons without vested interest in the survey, but with adequate knowledge of the process and methods.

A clear management structure of the survey helps to:

- ensure that the set objects can be met;
- make planning and implementation of the survey more efficient;
- increase the quality of the entire survey;
- decrease the cost of the entire survey.

An example of the management structure of a national HES is given in Figure 1.3.

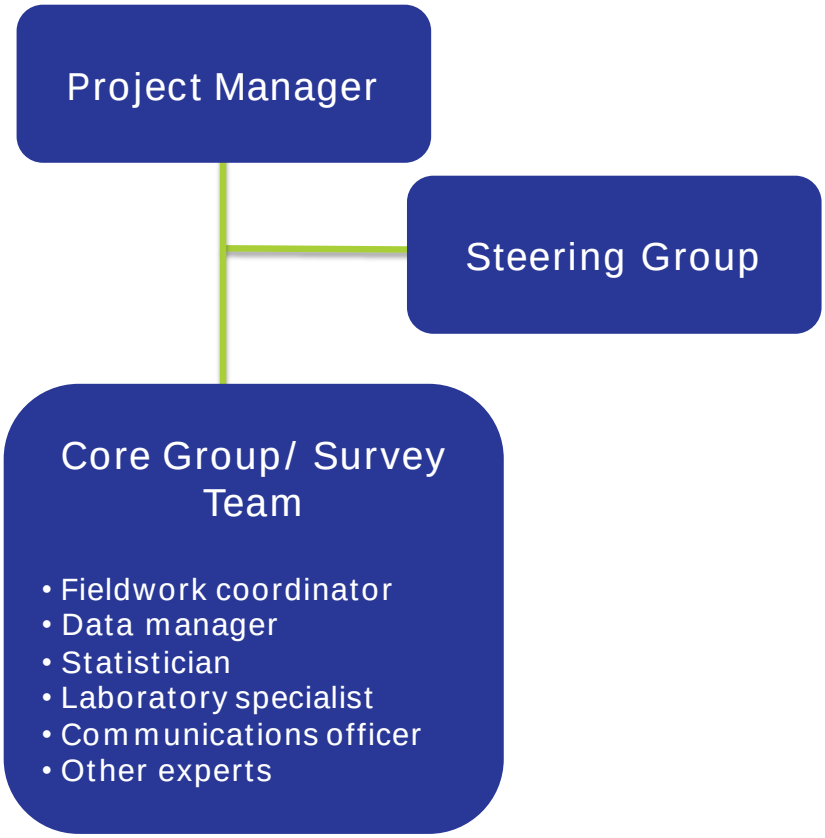


Figure 1.3. An example of a survey project organization in a survey including EHES core measurements and in a more comrehensive survey with several additional measurements

In the example, the different groups and persons have the following tasks:

- The Steering Committee (or a Steering Group) approves the survey objectives, and provides directions and guidelines to meet these aims. It represents is the agency(ies) responsible for the survey and monitors the progress of the survey.
- The Project Manager runs the survey. He/she is directly responsible for the Steering Committee, and his/her responsibilities cannot be shared by other experts. The Project Manager is responsible for:
    - the organization of the survey by allocating responsibilities and resources and by making sure that all areas are covered and that there is no overlap between the responsibilities of different experts;
    - managing the survey process by making decisions, giving guidance, providing and acquiring assistance, motivating team members and solving possible conflicts;

- day-to-day monitoring and evaluation of the survey process, schedules and budget and making adjustments to these when needed;
- reporting to the Seering Committee.

- The Core group assists the Project Manager. It consists of key experts, selected from the Team Leaders, with specific responsibility for coordination of fieldwork, statistical issues, and data management.
- Survey Teams: Different subareas of the survey are planned and implemented in larger surveys by different Survey Teams, led by the Team Leaders. In smaller surveys there may be only single experts in each area, or one expert is covering several areas of expertice. The teams or experts cover different areas of experitice, such as sampling, fieldwork, laboratory issues, communication and quality assurance, as well as different topics of the survey (e.g. blood pressure monitoring, nutrition).

Key experts and tasks in the survey project organization include:

- a fieldwork co-ordinator or fieldwork team. In larger studies a full time fieldwork coordinator is needed to share the workload of the Project Manager. The fieldwork team is led by the fieldwork co-ordinator and the team will prepare the fieldwork logistics, training and day to day data collection activities.
- a data management expert, when needed supported by thee IT team. They are responsible for the computer systems and programs, and the data management:
- a person responsible for the laboratory activities, when needed supported by the laboratory team responsible for the sample collection, analysis and storage;
- in larger studies a quality assurance team may be needed for the quality assurance activities.
- a survey statistician or a team of statisticians with specific expertise on sampling or data analysis;
- a person with expertise in survey ethics and a communications specialist may need to be consulted or invited to the fieldwork team.

Some of the tasks may be carried out under a short-term contract (e.g. computer systems, data entry, printing, mailing) or by contracting out some functions to an external organization. The roles and responsibilities of these persons/teams may vary between countries due to legislation and differences in organization

structures. Legislation in many countries calls for a chief physician in any study classified as medical research. The roles of the fieldwork co-ordinator and chief physician mey be combined.

If the survey team is large and if the survey covers different data collection phases, and/or several topics or modules, it may be useful to have special teams devoted also to each topic area (Figure 1.4). Such topic specific teams should propose questionnaire instruments and measurements for their areas of expertise, participate in the training of the fieldworkers, as well as plan and carry out special studies. All these experts and teams are needed throughout the survey process. When fruitful collaboration is built during the planning and preparation, the members of these expert groups are a valuable resource for e.g. training of the fieldwork staff, quality control during the fieldwork, data analysis and reporting.

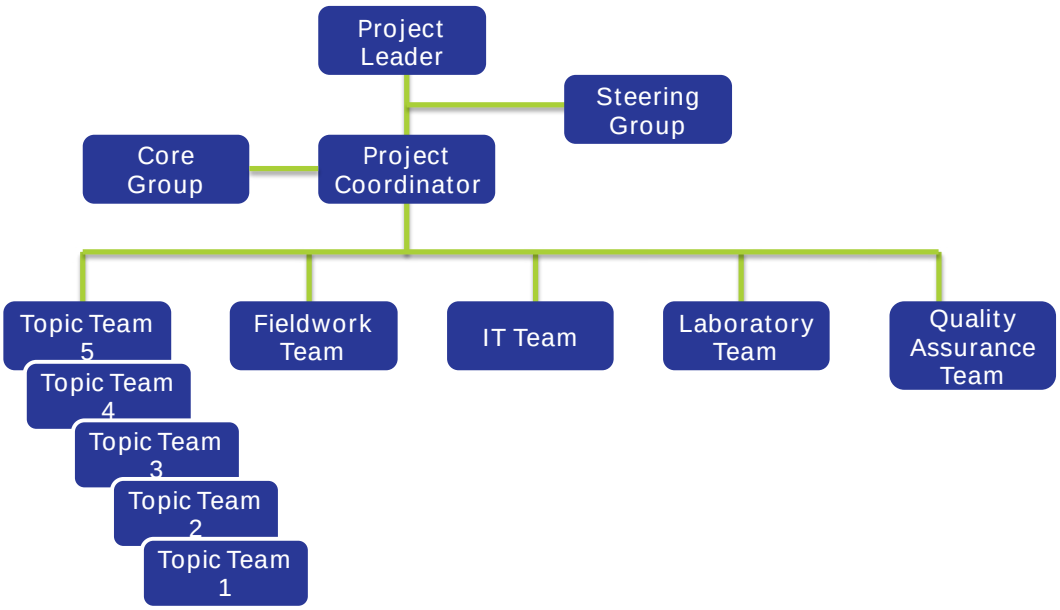


Figure 1.4. An example of a survey project organization in a comprehensive survey including EHES core measurements and several additional measurements

1.4.2. Management tools

It is essential to ensure that there is enough time for different phases of the survey process. The planning and preparation will usually require at least one year before the fieldwork can be started (Figure 1.5). If there is no recent (within last 5-10 years) or only little experience of a previous survey in the country, the planning and preparation for a full scale HES requires a longer period of time.

As collaboration between several organizations and teams is needed, the detailed planning and preparation may benefit from

using specific project planning tools and software to define the project timeline in Gantt Charts and to prepare Critical Path Analysis. Gantt Charts are a type of bar charts that illustrate a project schedule. They show the start and finish dates of the key tasks and activities (Figure 1.5). The Critical Path Analysis (CPA) helps to plan all tasks that must be completed during the survey process (Figure 1.6). CPA acts as the basis both for preparation of a schedule, and of resource planning. It identifies which tasks must be completed on time for the whole survey to be completed on time and identifies which tasks can be delayed if resources need to be reallocated.

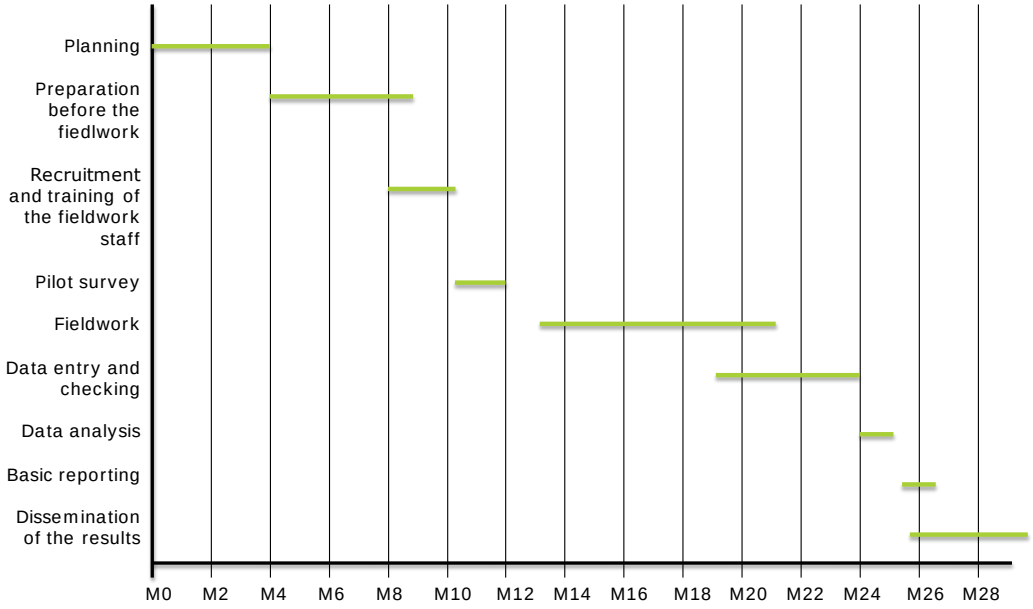


Figure 1.5 Gant chart of the survey, providing the timeframe for the survey

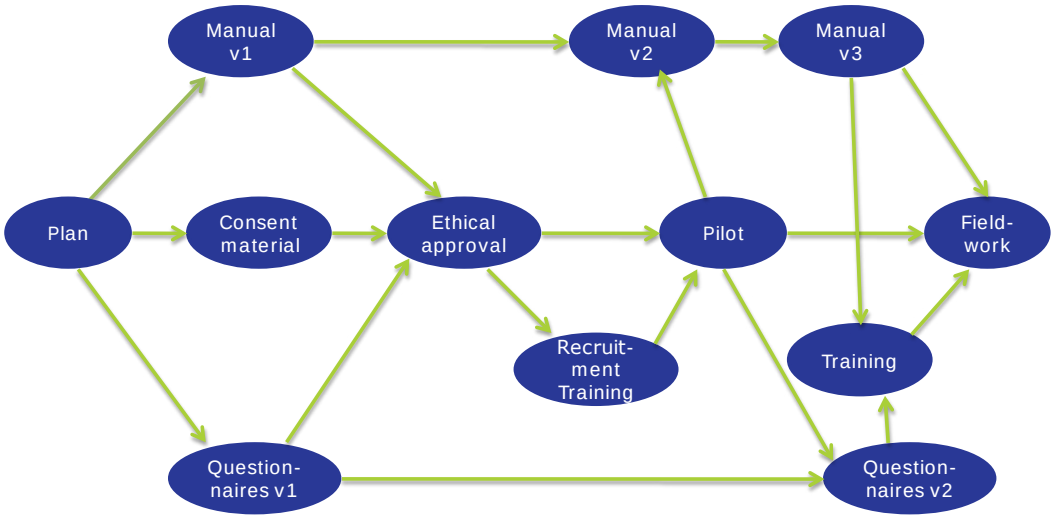


Figure 1.6 Example of a Critical Path Analysis (Perth chart)

One key element in the survey process, to ensure a successful data collection and fieldwork phase, is piloting and detailed evaluation of the pilot process. The countries may consider, if

a small pre-pilot (e.g. fieldwork testing with volunteer participants) is needed before the EHES pilot. These pre-pilots may be needed to test the computer programs, measurement techniques and timing. Specific aims for the national EHES pilots need to be defined during the planning and preparation. A pilot phase is always recommended, but the aims and content of the pilots depend on the previous experience and frequency of the survey. When the schedule of the data collection is planned it needs to be ensured that there is adequate time between the pilot and the actual data collection so that the experiences and results of the pilot are evaluated in detail (see Part A, Chapter 11 of the EHES Manual).

1.4.3 Risk analysis

Risks relate to uncertain events or situations that potentially can adversely affect carrying out a project according to plans. Risk management describes the processes concerned with identifying, analyzing and responding to the risks. The aim is to avoid uncertainties that threaten the goals and timetables of the project, and to take actions in advance to reduce the effect of these risks. Risk analysis should be carried out when planning the project and updated during the process. An example of risk analysis, covering common risks of national HES is presented in Table 1.1.

Table 1.1. Risk analysis in a national HES, examples of potential risks

Risk	Problems caused	Options for avoiding and controlling the risk
Insufficient personnel resources for planning and preparation	Shortcomings in planning and preparation leading to problems during fieldwork, in standardization and quality of data	Careful preparation of the survey organization, and seeking mandate from the ministries (health and research). Seeking specific funding for the planning and preparation, careful budgeting and diverse fund raising (see chapter 16), ensuring that the needed resources are available.

Risk	Problems caused	Options for avoiding and controlling the risk
Shortage of fieldwork personnel	Difficulties in keeping time schedules: problems caused for participants as well as in getting results	Raising interest towards the survey in the ministries and professional organizations, careful piloting and planning for the time schedules, taking potential sick leaves into account when planning the size of fieldwork team(s).
Insufficient time between pilot and actual fieldwork	Not possible to correct errors, specify manuals and training or adapt protocols, problems in standardization	Acknowledging the aims and significance of the pilots. Careful preparation for the time schedule.
Problems in collaboration between different organizations and actors	Difficulties in utilizing all expertise needed, and problems in keeping time schedules	Well defined leadership, building local partnerships throughout the survey process, careful planning for the supervision of the fieldwork teams
Low motivation among the population to participate	Low response, selective participation, biased results	Media campaigns and careful planning of the recruitment process (see chapters 13 and 14)
Violation of personal data protection rules	Loss of confidence	Careful planning and preparation for data management (see chapter 12) and proper training for all survey staff (see chapter 15).
National or local political or ecological crisis situations	Loss of data	Timely data transfer to central national and European data centers.
Epidemics	Absences of fieldwork staff, difficulties in participation	Little possibilities to avoid: infectious disease control at fieldwork settings and offering seasonal flu vaccinations to fieldwork staff.

Risk	Problems caused	Options for avoiding and controlling the risk
Safety risks during fieldwork	Harm caused to staff members or participants	A medical doctor must be available for consultation or present at the fieldwork site. The protocol for needle stick injuries should be easily available to all staff members at all examination sites. Situations with aggressive and violent participants and other safety risks during fieldwork covered in manuals and training. Adequate supervision of field work staff throughout the fieldwork process.

1.4.4 Project evaluation

Project evaluation should be an ongoing task (Table 1.2). It helps to make sure that the survey will be finalised with the resources available and within the timeframe set for the survey. Some parts of the evaluation are directly linked with quality assurance. Indiicators for evaluation should be defined and followed with regular intervals and actions developed if the targets (e.g. numbers of participants) are not met.

Table 1.2. An example of potential evaluation indicators for selected stages in the survey survey process

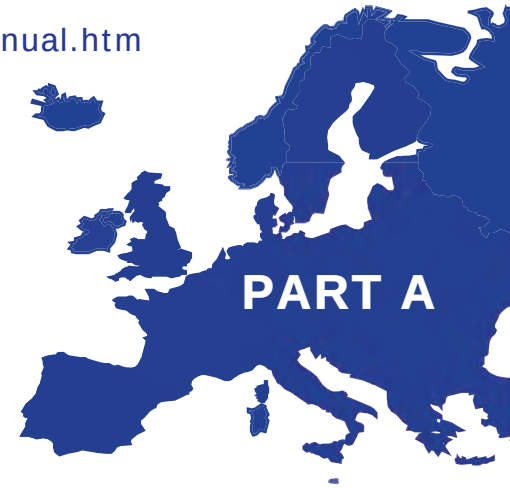
Survey stage	**Process indicators** *program operations*	**Output indicators** *direct results or products of project activities*	**Outcomes indicators** *impacts or changes that can be attributed to the project activities*
Survey design	Organized meetings and seminars	First version of the survey proposal	National consensus on carrying out the HES and timing of the surveys. National HES plans approved by national authorities with at least preliminary decisions for funding for the HES.
Planning and preparation	Number and type of experts involved in the survey planning, personnel resources needed	Detailed survey plan with a budget	Ethical approval

Survey stage	**Process indicators** *program operations*	**Output indicators** *direct results or products of project activities*	**Outcomes indicators** *impacts or changes that can be attributed to the project activities*
Fieldwork during pilot(s) and the actual survey	Training seminars organised for the fieldworkers: hours of training Number of invited persons	Number of fieldwork staff members who participated in the national training (% of all fieldworkers) Number of days for the fieldwork Numbers of participants, those who were found to be ineligible, those who were not contacted and those who refused (by age and gender) Recorded length of examinations per participant – reported average length per participant (minutes/hours) Place of examinations: number of participants examined at the clinic setting/ at home/ at an institution	Participation rate (per age/ gender) Cost of the survey data collection/participant

References

- Biemer PP, Lyberg LE. Introduction to Survey Quality. Wiley Series in Survey Methodology, John Wiley & Sons, New Jersey, USA, 2003.
- Czaja R, Blair J. Designing surveys. A Guide to Decisions and Procedures. 2nd edition, Sage Publications, Inc., California, USA, 2005.
- Elo SL, Karlberg IH. Validity and utilization of epidemiological data: a study of ischaemic heart disease and coronary risk factors in a local population. Public Health. 2009 Jan;123(1):52-7.
- Elgar FJ, Stewart JM. Validity of self-report screening for overweight and obesity. Evidence from the Canadian Community Health Survey. Can J Public Health. 2008 Sep-Oct;99(5):423-7

- Franklin S, Walker C (eds). Survey methods and practices. Statistics Canada. Social Survey Methods Division, 2003. Also available at http://www.statcan.gc.ca/pub/12-587-x/12-587-x2003001-eng.pdf
- Gillum RF, Sempos CT Ethnic variation in validity of classification of overweight and obesity using self-reported weight and height in American women and men: the Third National Health and Nutrition Examination Survey.Nutr J. 2005 Oct 6;4:27.
- Gnavi R, Karaghiosoff L, Costa G, Merletti F, Bruno G. Socio-economic differences in the prevalence of diabetes in Italy: The population-based Turin study. Nutrition, Metabolism and Cardiovascular Diseases, 2008, Volume 18, Issue 10, Pages 678-682
- Kastarinen M, Antikainen R, Peltonen M, Laatikainen T, Barengo NC, Jula A, Salomaa V, Jousilahti P, Nissinen A, Vartiainen E, Tuomilehto J. Prevalence, awareness and treatment of hypertension in Finland during 1982-2007. J Hypertens. 2009 Apr 30.
- Ostchega Y, Hughes JP, Wright JD, McDowell MA, Louis T. Are demographic characteristics, health care access and utilization, and comorbid conditions associated with hypertension among US adults? Am J Hypertens. 2008 Feb;21(2):159-65.
- Primatesta P, Allender S, Ciccarelli P, Doring A, Graff-Iversen S, Holub J, et al. Cardiovascular surveys: manual of operations. Eur J Cardiovasc Prev Rehabil. 2007 Dec;14 Suppl 3:S43-61.
- Tolonen H, Kuulasmaa K, Laatikainen T, Wolf H, for the European Health Risk Monitoring Project. Recommendations for indicators, international collaboration, protocol and manual of operations for chronic disease risk factor surveys. 2002 cited 13 2008; Available from: http://www.thl.fi/publications/ehrm/product2/title.htm
- Tolonen H. Towards the High Quality of Population Health Surveys. Standardization and Quality Control. National Public Health Instiute A27/2005, Helsinki 2005 cited 13 May 2008; Available from: http://urn.fi/URN:ISBN:951-740-546-4
- Tolonen H, Koponen P, Aromaa A, et al. (Eds.) Recommendations for the Health Examination Surveys in Europe. B21/2008, Publications of the National Public Health Institute, Helsinki 2008.Also available from http://urn.fi/URN:ISBN:978-951-740-838-7



2. Target population and sample size

Johan Heldal[1], Susie Jentoft[1]

[1] Statistics Norway (SSB), Oslo, Norway

2.1 Target population and sample size

The target population is the population of individuals which we are interested in describing and making statistical inferences about. EHES RC suggests the following definition for the target population of a country:

1. The core target population is the set of all persons aged at least 25 years and at most 64 years and having permanent residence in the country. (Instructions for more precise definition of age are given in Part A, Chapter 3.)
2. Each country can extend the eligible age group with a lower bound of 18 years and with no limitation for the upper bound.

Some counties have already defined the age ranges for their surveys wider than this. However, the international EHES is a survey of adults and the EHES RC has defined adults as persons of age 18 years and more. This definition describes the ideal target. In order to take a sample from a population one needs a sampling frame from which a sample can be taken. Some countries will have difficulties establishing sampling frames that cover the entire population at a specific date. In a number of the health examination surveys carried out until now, e.g. institutionalized persons have not been available for sampling. Comparisons of survey results across countries should in principle only be done for population groups that are covered in all countries. Every detail of the coverage of each national sampling frame must be well documented. It is not recommended to leave out population

groups difficult to contact or who do not for example speak major languages of the country.

Selection of sampling frames is discussed in Part A, Section 3.2.

2.2 Sample size

A minimum of 4000 persons are sampled to be invited in each country. Each of eight age-sex domains (25-34, 35-44, 45-54 and 55-64 years) should have at least 500 representatives in the sample. The sample size calculation is based on a participation rate of 70 percent, but should be applicable also if the realistic expectation for the participation rate is different. This minimum size relates to the requirements for statistical power when testing differences between countries for age-gender domains. For comparisons between regions or socioeconomic groups within a country, each country will have to set its own standards for accuracy and explain its needs for larger sample sizes. This is not a part EHES.

Sample size relates to the statistical precision of the survey results, whereas bias is the concern related to low response rate. The relative benefit from higher precision, and therefore higher number of participants, is better if the response rate is high. On the other hand, if the expected response rate is low, it will be better to spend resources on increasing the response rate than to increase the total sample size. Specifications and calculations of the minimum recommended sample size are given in the FEHES recommendations (Tolonen 2008). In the pilot we recommend a sample size of at least 200 persons, with at least 25 persons in each age-gender domain.

How to obtain a sample is discussed in Part A, Chapter 3.

2.3 The EHIS definition

For the European Health Interview Survey (EHIS) the Task Force III report on sampling issues suggests the following definition of the target population (Axelson 2009).

For the EHIS, the target population should contain all adults (15 years old and over) living in the country at the place of their usual residence (the place where they mainly live). The sample may not include individuals at a place of residence where they do not mainly live. Such individuals must be treated as not eligible, and the interview stated as terminated.

Apart from defining a wider age range this definition does not specify a reference point in time for the target population. We

have not seen a reference to 'place of their usual residence' as relevant. The EHIS definition may seem to exclude persons not having a residence (homeless people). It is not clear whether this is intentional. However, such persons can be difficult to reach and are not likely to be interviewed or examined anyway.

2.4 People living in institutions

Countries having carried out Health Examination Surveys so far have had different practice with regard to incorporating people living in institutions. This seems to a large extent to be due to practical circumstances. In many countries the main sampling frames only cover persons living in private households. Covering institutions requires special designed sampling frames. Furthermore, it is often difficult to obtain participation in surveys from people living in institutions.

There are many types of institutions. Among the most common are nursing homes, elderly homes, children homes, military barracks, jails and monasteries. The Task Force III report on EHIS expresses that *"as concerns EHIS, we are mainly interested in medical institutions or homes for elderly people"*. A reason for that is not given.

It has been decided to create a separate Task Force under EHIS to reflect on the issue on sampling frames to be used to interview people living in institutions, on the way to collect data from those people, and on the measure of the impact on the estimates by including/excluding institutionalised people. The output of this TF will be available early 2010. Moreover, MS should report to Eurostat whether or not institutionalized people are included in their sample.

References

- Tolonen H, Koponen P, Aromaa A, Conti S et al (eds.) Recommendations for the Health Examination Surveys in Europe. Publications of the National Health Institute, Helsinki 2008. B21/2008. Available athttp://urn.fi/URN:ISBN:978-951-740-838-7. Accessed on 22 January 2013
- Axelson M, Bihler W, Djerf K, et al (prepared by) European Health Interview Survey. The Task Force III report on sampling issues. 2009. Available at https://circabc.europa.eu/sd/d/4f89b382-b3e1-4884-94e1-ca9f2e2338b6/Task%20Force%20III%20report%20on%20sampling%20issues%20up%20to%20date%20November%202009.pdf. Accessed 22 January 2013.

3. Sampling procedures

Johan Heldal[1], Susie Jentoft[1]

[1] Statistics Norway (SSB), Oslo, Norway

The goal of a Health Examination Survey (HES) is to produce statistics for clinically measured heath indicators, such as the average state and variation in various health indicators for national populations. This should be done in such a way that the estimates obtained from the survey are as well as possible statistically unbiased for the true averages of these indicators in each participating country. This is required if we want to be able to compare estimates among countries and carry out unbiased tests for these differences. There are many potential sources of bias in a health examination survey. Sampling bias is one of them, but it is also one of the sources of error that can be brought almost entirely within control of the survey taker. At the same time as avoiding bias, purely random errors in the estimates should be made as small as possible.

Control of sampling errors, both systematic (bias) and random, requires a good *sampling design*. Scientific surveys make use of *probability sampling*. This means that every eligible individual or household should have a known probability of being sampled. In probability sampling, randomization techniques and (pseudo) random mechanisms are used to select the individuals to be invited to the survey. Survey sampling is a science which should be carried out or monitored by professional statisticians in each country. The procedures for estimating the health indicators rely on probability sampling.

There are many ways to select a probability sample. Which method to choose will depend on the features of the actual survey and the sampling frames that are available (see Part A, Section 3.2 of the EHES Manual). A two stage sampling design is recommended for health examination surveys in all countries except possibly the smallest ones. Depending on the sampling frame available,

more than two stages might be necessary in some countries in order to reach down to the individual people to be invited to participate in the survey.

The EHES RC has developed a sampling application program to simplify sampling in line with the recommendations given in this chapter. The program is written as an R-package called EHESsampling (Jentoft 2011). R is the statistical software of choice for EHES RC, after consultation with the European Commission. R is freeware and can be downloaded from http://www.r-project.org/. EHESsampling and the user manual can be downloaded from the EHES web site (EHESsampling). EHESsampling carries out the steps described in Part A, Sections 3.3.4 - 3.3.6 of the EHES Manual and contains the recommendations in these sections as default options.

Part A, Section 3.1 of the EHES Manual is a short overview of some general considerations concerning Health Examination Surveys and their implications for design of such surveys compared to Health Interview Surveys.

Part A, Section 3.2 of the EHES Manual discusses some alternative sampling frames for a HES. Section 3.3 treats the design for Stage 1 (sampling PSUs) in further detail and how efficient sample sizes can be calculated. Section 3.4 discusses the design for Stage 2. Section 3.5 considers aspects of using address frames. Section 3.6 deals with documentation and data management and Section 3.7 discusses procedures that are common in use but not recommended here.

3.1 General considerations

3.1.1 Health Examination Surveys

In a two-stage sampling design the sample of individuals to be invited is obtained after two stages of sampling. In a Health Examination Survey where participants are invited to an examination site, it is essential that the distance to the clinic is as short as possible. Short distances are also important when mobile clinics are used to visit invitees closer to their homes (e.g. in rural districts) or when there are home visits by field work staff (See chapter 9, *Guidelines for the selection of survey site* for a more detailed discussion). Both situations call for a clustering of the invitees in a limited number of examination areas that do not cover an entire country but may be selected from a larger set of potential examination areas which do cover the entire country. In agreement with the general terminology of survey sampling, the examination areas will be called *Primary Sampling Units* (PSUs)

in this chapter. The design for sampling of PSUs is Stage 1 of the total sampling design. The individual participants are sampled from the PSUs that have been selected at Stage 1. See figure 1. Within each PSU being sampled at Stage 1 we will sample people, addresses, households or dwellings. These units will be termed *Secondary Sampling Units* (SSUs).

EHESsampling produces R data frames with information on the sampling frame as well as for the sample itself for each stage of sampling in all strata, such as stratum and PSU identification and inclusion probabilities for sampling units at each stage. It is essential that all this information is stored both for those who participate and those who do not along with the data collected in the survey. No information should be discarded. This information is needed for proper analysis, estimation and variance calculations later. All details of the national sampling designs and samples resulting from them must be well documented.

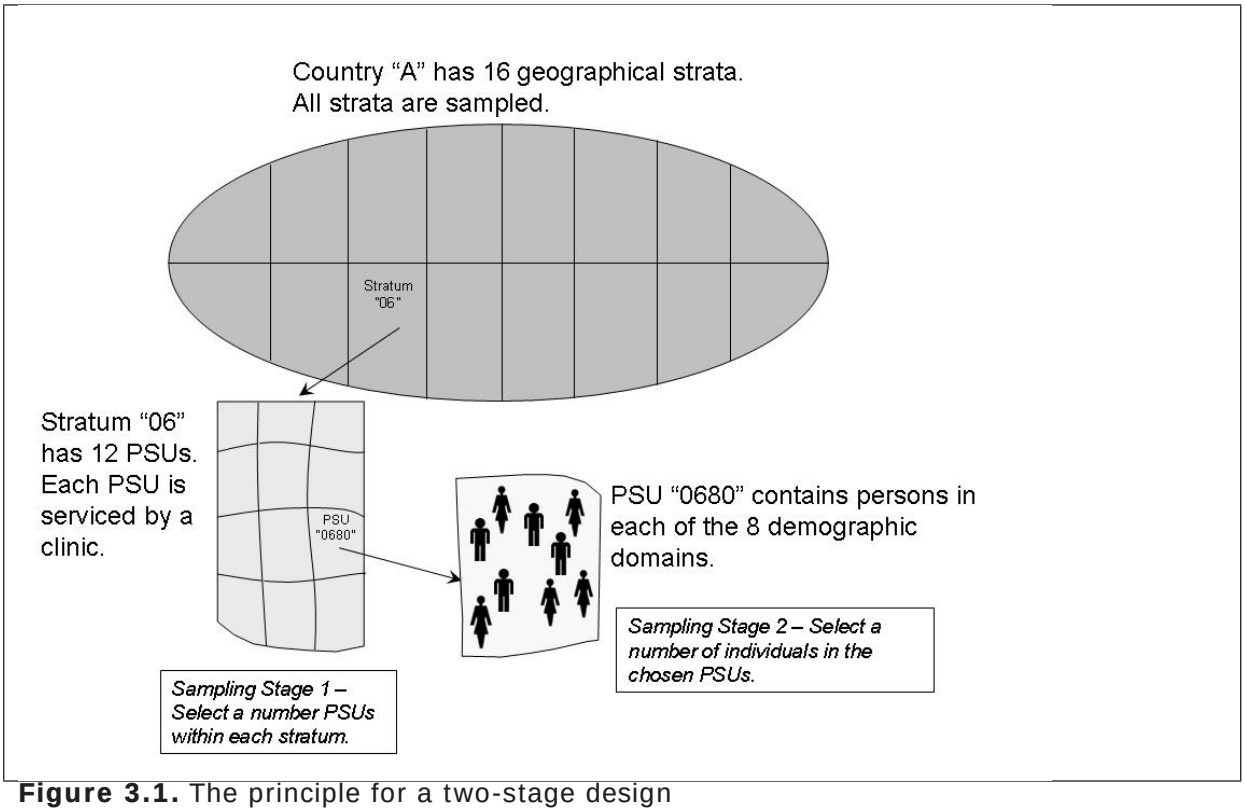


Figure 3.1. The principle for a two-stage design

3.1.2 The recommendations for European Health Interview Survey (EHIS)

For those involved also in EHIS, we describe here its sampling recommendations and the differences from the EHES sampling.

The recommendations for sampling in EHIS are given in the EHIS Task Force III report on sampling issues (Axelson 2009). Since EHIS is an interview survey only, its sampling design does not need to pay attention to such things as "closeness to a clinic"

for the participants. It can even be carried out by telephone or as self-administered survey in which case a two-stage design does not have any advantage. The way HIS surveys have been carried out, as well as survey design, differs between the European countries. The recommendations for sampling in EHIS TF report are therefore not very specific. Basically they recommend that samples should be taken as probability samples where each member of the target population is assigned a non-zero probability of selection.

Some countries have expressed interest in coordinating HES with the HIS, for example conducting HES on a subsample of HIS. Such a strategy requires a sampling design for HIS which is compatible with the requirements for HES. This is not the case for all of the HIS surveys that have been carried out so far. However, it may be possible to coordinate the two surveys in the future.

The EHIS TF III report discusses pros and cons with substitution of non-responding sample units and some practices that occur in the EU Member States. We will not repeat all the details of the discussion here, but in conclusion, it recommends not using substitution of non-respondents in population health surveys. In EHES as well as EHIS we must expect that the state of health of the invited people will often be a factor causing non-participation among the invitees. This will bias the estimates from the survey. Substituting respondents will most likely have of similar health as other respondents and will therefore not reduce this bias. On the contrary, under some circumstances it may increase the bias. Moreover, when substitutions are being used the inclusion probabilities, the probabilities of being selected to the sample, can no longer be exactly calculated. For these reasons we recommend *not* to use substitutions in HES. Observations identified as substitutions will be excluded from the final comparative analyses of the EHES data. Therefore, if a country still chooses to use substitutions these must be identifiable in the data. However, reason for non-participation should be recorded in detail.

Readers interested in more details of former HIS surveys can consult the webpage http://hishes.iph.fgov.be

3.2 Sampling frames

When a survey is carried out in more than one stage, a sampling frame for each stage will be required. The frame for Stage 1 should be a list of all PSUs that can be selected with information on their population sizes (people or households/dwellings). If updated statistics are not available for all PSUs the best available estimates, e.g. the last census, can be used. If feasible, the population sizes should be broken down to at least the core

age by sex groups for which statistics will be published in EHES. There should also be information about which stratum each PSU belongs to. The PSUs and the strata should be equipped with a unique digital number as well as names.

The Stage 2 sampling frame is the list of units (individuals or addresses/households/dwellings) from which a sample of such units can be taken. If the list contains individuals, Stage 2 will be the final stage. If the units are addresses there may be a need for a Stage 3 to select dwelling in a multi dwelling house. Stage 3 sampling will often have to take place in the field and will not be covered in detail in this chapter. A Stage 2 frame should be established at least for all the PSUs selected at Stage 1. It is recommended that the Stage 2 frame is updated as closely as possible to the time when the PSU will be visited by the survey. It can therefore be an advantage waiting as long as possible before taking the Stage 2 sample for a selected PSU. Availability of high quality sampling frames for Stage 2 differs among countries. While some countries have central population registers that can be used in other countries sampling frames for Stage 2 are only available at the local level, e.g. municipalities. Different kinds of frames for Stage 2 are recommended in the following order

1. Whenever legally and practically available, a central file with the most recent and best coverage of the *people* in the target population should be used as the sampling frame. Ideally, this will be a population register. If possible, the main frame can be supplemented with other files to catch parts of the target population not covered by the main frame. Some countries have frames covering individuals but lack for instance non-citizens, homeless or parts of the institutionalized population. The extent of such under-coverage by cause should be estimated. See below. Many countries in Europe are going to have new censuses in 2011. Fresh census data is very useful as a sampling frame and should be considered for the national HES.
2. If a quality frame with individuals is not available, an updated address file or list of housing units can be used as an alternative. However, a postal address can either address a dwelling directly or a house with many dwellings. The two situations require somewhat different approaches to sampling.
3. Countries already carrying out national HES with samples drawn from an established frame may continue to use the same frame in the future. However, all such frames must be compared and evaluated against the general recommendations and standards proposed for EHES.

4. Countries that do not have a sampling frame mentioned in 1 or 2, can construct a Stage 1 sampling frame based on available statistics for the units chosen for Stage 1 and the sampling of such units can be carried out in the same way as for countries covered by item 1 or 2. Some countries have local population registers which can be frames for Stage 2 sampling when the Stage 1 sample has been selected. If these kinds of local frames are not available, a local frame must be constructed. It may be necessary to sample in more than two stages. The strategies may differ between urban and rural areas. In cities, street maps which identify city blocks may be useful. The number of dwellings in each block must be mapped and some of them sampled. Dwellings can then be sampled within each selected block. In may be better in rural areas, to use areal squares as PSUs. The number of houses in each square should be counted and a sample of the inhabited squares selected. Either a sample or all houses in the sampled squares, is included in the sample. This is called an area frame. The National Health and Nutrition Examination Survey in the USA and the Canadian Health Measure Survey use this kind of strategy. See the NHANES Analytic guidelines (NHANES 2006) and (http://www.statcan.gc.ca/pub/82-003-s/2007000/article/10363-eng.pdf) for descriptions. Each country needing this kind of frame must adapt a procedure that fits the national structures.

The FEHES Review Report (Tolonen 2008) provides a list for accessible sampling frames in each country. However, the list may not be complete. If no acceptable frame seems to be available, the national statistical institute or other national institutions, public or private, regularly carrying out national sample surveys in other fields should be consulted for assistance.

The target population is defined in Chapter 2 of Part A as all individuals of an eligible age, living in the country. Whenever the general Stage 2 frame does not cover all residents that should be eligible the various kinds of under-coverage should be explained and the size of the under-coverage estimated, preferably by sex and age. If for instance parts of the institutionalized population are not covered by the main frame for the survey, an overview of such institutions by category should be constructed. If feasible this should be done in such a way that this overview can be used as a supplementary Stage 1 frame for institutions although it will rarely be possible to take samples from the institutions.

3.3 Sampling design for Stage 1

3.3.1 Creating the PSUs

Whatever sampling frame for Stage 2 will be used, the sampling frame for Stage 1 (the PSUs) should be established approximately as follows.

Partition the geographical area of the country into a set of disjoint areas, the PSUs. Each PSU should be small enough to be served by one examination site and with acceptable travel distances to the site for all people living in the PSU or for field work team and mobile units to travel between the homes of all potential invitees. The PSUs should be areas for which statistics for total population sizes (number of persons) or the number of postal addresses or dwellings are accessible. What alternatives for PSUs are available may vary among countries, but small census tracts, municipalities, electoral districts and post code areas are examples. Most National Statistical Institutes in Europe have detailed population statistics by sex and age for all administrative units and sometimes also for smaller units defined for statistical purposes. For many countries this information is freely available on their websites and can be downloaded as excel files. If more detailed statistics is needed the statistical offices should be contacted.

From a statistical point of view it is desirable that PSUs, at least those within the same stratum, are statistically as similar as possible so that which PSUs are actually selected will affect influence on the survey results as little as possible. As a PSU will have to be a contiguous area that will have to meet practical constraints there will always be limits to how similar it is possible to make them. They should however not be smaller than necessary to meet the practical demands since small PSUs will often tend to be internally more homogenous and therefore less similar to their neighbours. The sizes of the PSUs can vary within the same stratum, but not “too much”.

3.3.2 Measure of size for the PSUs

A measure of size should be established for all PSUs. This will usually be the number of SSUs in the PSU according to the Stage 1 frame, people or household addresses, but if such up-to-date are not immediately available, cruder measures of size, e.g. old census counts, should be used.

If the SSUs are people the size should be the number in eligible age living in them. If the distribution by sex and age is available this information should be taken into the file defining the Stage

1 frame. Age should be recorded by the groups that will be used for publication and comparison among countries, at least the age groups 25-34, 35-44, 45-54 and 55-64.

If the SSUs are households, dwellings or postal addresses, their numbers in the PSUs should be used as the measure of size. If the number of dwellings at each postal address is known, it will be better to use the number of dwellings than the number of addresses as the size measure of the PSU. If it is feasible to select dwellings directly rather than addresses the need for Stage 3 to select dwellings at multi-dwelling addresses can be avoided or reduced. In households with a large number of eligible individuals, it may be necessary to limit the number of participants. Techniques for doing this (Kish Grid, last birthday etc.) will not be discussed in this document. For an example, see the Health Survey of England (Craig 2008).

If neither a frame based on individuals nor addresses or dwellings is available for Stage 2, frames for further sampling must be established within the selected PSUs. An example of such a frame is the National Health and Nutrition Examination Survey (NHANES 1994).

3.3.3 Stratification of the PSUs

The PSUs should be stratified by grouping together relatively similar PSUs, e.g. urban PSUs versus rural PSUs, PSUs having similar age distribution by taking into account the social or demographic profile of the PSUs. Good stratification increases the precision of the survey estimates. Although the PSUs in a stratum do not need to be geographically contiguous, geography is also important. There is often interest in comparing regions within a country with respect to various health indicators. It is therefore desirable that these regions consist of complete strata. When considering how many strata to create, think about how many PSUs it is natural to select. Generally to facilitate variance estimation in a two-stage design, two PSUs should be selected per stratum. To be able to measure uncertainty in the estimates is important when comparing estimates from different countries or regions within a country. However, other considerations can justify selecting only one. Detailed stratification may reduce sampling variance but make unbiased estimation of the variance infeasible. Sometimes other considerations makes it is natural establish some small strata where selecting more than one PSU is difficult. PSUs that are very large in population can be strata alone (i.e. metropolitan areas). Very large PSUs can either be cities where it can be seen as appropriate to sample in one stage or are those that are assigned a probability larger than one according to the formula. EHESsampling will automatically select PSUs that are

too large compared to other PSUs in the same stratum with probability one. Such PSUs will be treated separately at Stage 2. As a basic rule, to be able to select two PSUs in a stratum it should contain at least four PSUs.

EHESsampling can calculate the number of PSUs to be sampled in each stratum and the anticipated costs doing the survey with this stratification. Using the software these calculations should be carried out for alternative stratifications as a tool to find the best ones, the one that gives the lowest variance for a given cost.

3.3.4 Sample sizes at Stage 1

For each stratum, the number *m* of PSUs to be selected at Stage 1 and the number *p* of SSUs to be invited within each PSU at Stage 2 can either be decided directly or be established based on cost-variance considerations. It will be demonstrated in this section and in Section 3.4.2 that if the PSUs are selected with Probability Proportional to Size (PPS), which is recommended, the Stage 2 sample size *p* should be the same within every sampled PSU in the same stratum.

How to calculate a cost-variance optimal value of *m* and *p* will be demonstrated below.

For a given stratum, let C_{PSU} be the average cost of sampling a PSU (i.e. setting up an extra site) in the stratum and let C_{SSU} be the average cost of inviting an (extra) SSU (person or dwelling) to the survey. Let *m* be the number of PSUs to be selected in the stratum and let *n* be the number of SSUs. A model for the expected variable survey cost in the stratum is then

$$C = C_{PSU} m + C_{SSU} n \tag{3.1}$$

Let *Y* be a survey variable. For a given stratum, let Y_{ij} be the value of this variable associated with SSU no. *j* in PSU no. *i*. Let N_i be the size (no. of main frame units) of PSU no. *i*. Let

$$\begin{aligned} \mu_i &= \sum_{j=1}^{N_i} Y_{ij} / N_i = \text{ average of } Y \text{ within PSU } i \text{ and} \\ \sigma_i^2 &= \sum_{j=1}^{N_i} (Y_{ij} - \mu_i)^2 / N_i = \text{ variance of } Y \text{ within PSU } i. \end{aligned} \tag{3.2}$$

Let V_{Among} be the (weighted) *variance of* μ_i *(across PSUs) of the within* PSU *averages* and let V_{Within} be the (weighted) *average (across PSUs) of the within* PSU *variances* σ_i^2, that is

$$N = \sum_{i \in S} N_i = \text{ total number of frame units in all PSUs in stratum S}$$
$$\mu = \sum_{i \in S} N_i \mu_i / N = \text{ average of } Y \text{ in stratum S}$$
$$V_{Among} = \sum_{i \in S} N_i (\mu_i - \mu)^2 / N = \text{variance among the PSUs of their averages}$$
$$V_{Within} = \sum_{i \in S} N_i \sigma_i^2 / N = \text{ average of the within PSU variances} \quad (3.3)$$

We wish to sample *n* units within a stratum to estimate the average μ of *Y* in that stratum using a two-stage sample where every final unit has the same probability of being selected. When the PSUs are selected with probability proportional to size as described in Section 3.3.5 such an equal probability sample is obtained by allocating the sample size *n* equally with *p* = *n*/*m* units to each of the *m* sampled PSUs. The variance of the simple sample mean estimator $\hat{\mu} = \sum_{ij \in sample} Y_{ij} / n$ is then

$$Var(\hat{\mu}) = \frac{V_{Among}}{m} + \frac{V_{Within}}{n} \quad (3.4)$$

The number of individuals *p* to be invited within a PSU and the number *m* of PSUs to be drawn to minimize the $Var(\hat{\mu})$ (given *n* and *C*) is given by the formulae

$$p_{opt} = \sqrt{\frac{V_{Within}}{V_{Among}} \frac{C_{PSU}}{C_{SSU}}}, \quad m = \frac{n}{p_{opt}} \quad (3.5)$$

The value of p_{opt} obtained with formula will be different for different *Y*-variables. If formula is to be made operational, V_{Within} and V_{Among} must be calculated or estimated for some compromise *calculation variable*. Using a variable available in the frame is best. Age is a recommended variable for this purpose since health in general depends strongly on age.

Note that p_{opt} does not depend on the total number of units (*n*) to be sampled in the stratum, but *m* does. Formula says that *if the sample size n of SSUs is to be increased within a stratum then this should be done by taking a larger sample of PSUs, not by selecting more SSUs within each PSU*. Note also that p_{opt} may be calculated larger than *n.* Since μ must be 2 or more, formula will only play a role if p_{opt} is calculated to be less than roughly 40 percent of *n*. *m* and *p* also need to be rounded to integers. This is done in the program EHESsampling.

Assessing the costs C_{PSU} and C_{SSU} is a part of the budgeting of the survey. The variances will differ among the variables and some compromised calculation values must be established to apply formula . Note however that it is basically the ratios C_{PSU}/C_{SSU} and V_{Within}/V_{Among} that are needed to calculate p_{opt}. If nothing else can be assumed, C_{PSU}/C_{SSU} can be set equal for all strata. If information about age distribution within the PSUs comes with the sampling frame it is possible to establish values for V_{Within}/V_{Among} based on the age distribution within and across the PSUs. An example on calculation of V_{Within} and V_{Among} based on age distribution is given in section 3.6.1 and in the manual for EHESsampling.

Before calculating *m*, the number *n* of units (individuals or addresses) to be sampled in the stratum must be set. In an equal probability sample the total sample size should be allocated to the strata proportional to the number of units in the frame and this should be the basis for calculating *m*.

R-program EHESsampling does the calculations for *m* and *p* based on the given values for C_{PSU}/C_{SSU} and V_{Within}/V_{Among} and rounds the calculated sample sizes to integers in such a way that the total national sample size is maintained.

3.3.5 Inclusion probabilities for the PSUs

The sampling of Primary Sampling Units at the first stage should be done with Probability Proportional to Size (PPS-sampling). This means that the probability π_i for selecting PSU no. *i* in the stratum is

$$\pi_i = m\frac{N_i}{N}. \qquad (3.6)$$

Note that if *m* is large and there is significant variation in the sizes of the PSUs, formula can assign the largest PSUs probabilities greater than one. This should be avoided if possible, and setting a minimum size for the PSUs aims at that. PSUs that are 'too large' in the sense that produces a probability larger than 1, are automatically assigned $\pi_i = 1$. EHESsampling will select them and calculate π_i correctly among the rest. Specific examples will be provided in the manual for EHESsampling. EHESsampling will also at this stage calculate the Stage 2 inclusion probabilities and anticipated sample sizes to be used in each PSU if the PSU is selected at Stage 1. How the Stage 2 inclusion probabilities and sample sizes are calculated is described in Section 3.4.2.

3.3.6 Sampling

When all the preparations described have been completed, sampling can proceed using for example the R-program EHESsampling package. *PPS*-sampling is rather technical and there is a host of methods. EHESsampling uses an R-package called *sampling* developed by Yves Tillé and Alina Matei (2009).

Having chosen the desired number of PSUs to be selected, EHESsampling calculates the number of SSUs to be sampled at Stage 2 within each PSU if that PSU is being selected. If the sampling units are addresses or if age-sex stratification is not used (see Section 3.4.1), the sample sizes will be the same in all PSUs selected with probabilities less than one. There may be deviations of no more than one due to rounding. PSUs selected with probability equal to one will have larger sample sizes than the other PSUs in the same stratum. If age-sex stratification is used, the number of persons to be sampled to the same age-sex domain will vary among the PSUs in the same PSU-stratum. If EHESsampling is used, the sample of PSUs along with the Stage 1 and 2 inclusion probabilities, PSU-size and anticipated sample sizes for Stage 2 is stored in an R data frame.

3.3.7 Distribution of PSUs over time

A national Health Examination Survey may be carried out over a long period of time, often a year, sometimes more. Teams may travel and visit each sampled PSU, one at a time. The examination site will be in operation for a limited period, from one day to a couple of weeks and then the team will move to another PSU. The order in which the PSUs are being visited is not indifferent. It is well known that there are seasonal variations in people's health caused by varying temperatures and weather conditions. If the teams starts operating in one part of the country, for instance in the southern part and then move gradually north to finish the survey in the northern part, the effects of season and geography on health variables which should be distinguishable in the data, will be confounded. It will be impossible to estimate them separately.

This should be considered when timing the survey. Ideally, a randomization of the order in which the sampled PSUs are visited is recommended but can be difficult to implement. If teams have to move across the country in a completely random order carrying a heavy cargo of equipment, the cost of travelling can be high both in terms of time and money.

For further discussion, see Part A, Chapter 6 of the EHES Manual, Length and time of year for the fieldwork.

3.4 Sampling design for Stage 2

Simple random sampling is proposed for sampling persons or addresses within each PSU selected at Stage 1. If age-sex stratification is used, simple random sampling should also be used within each age-sex domain in each selected PSU. However, age-sex stratification complicates the issue of sample sizes.

3.4.1 Stratification by age-sex domains

Results from national HESs will be compared across countries within age-sex domains. The four age domains 25-34, 35-44, 45-54 and 55-64 years will be crossed with sex to form eight domains. For this reason it is desirable to guarantee all eight domains a minimum sample size. With a minimum total sample size of 4 000 this means at least 500 persons in each domain. This can be obtained by a kind of stratification with respect to the eight domains where one sample is taken for each domain at Stage 2.

The eight age-sex domains will intersect the PSUs. Stratification with respect to domains that intersect the PSUs in a two-stage design is not common, but can be carried out so that every person in the same domain in the same stratum (or country) has the same inclusion probability. This is recommended when the total sample size for the survey is not large enough to by itself warrant the minimum sample size for each domain. When age-sex stratification is used, then the sample sizes will not be quite fixed. This is discussed in Section 3.4.2.

Age-sex stratification cannot be applied if the SSUs are addresses, dwellings or households.

3.4.2 Sample sizes at Stage 2 - with and without age-sex domains

We first consider the simplest case, without age-sex stratification. The goal is for all units within a stratum to have the same selection probability after two stages. This will be achieved if the Stage 2 sample sizes are calculated as:

$$n_i = n\frac{N_i}{N\pi_i} \qquad (3.7)$$

If all PSUs are selected with probability less than one, the same sample sizes n_i at Stage 2 will all be equal:

$$n_i = \frac{n}{m} \tag{3.8}$$

for all *i* (PSUs) selected at Stage 1. The selection probability at Stage 2 is then

$$\varphi_i = \frac{n_i}{N_i} = \frac{n}{mN_i} \tag{3.9}$$

If the largest PSU, say PSU no. 1, is selected with probability $\pi_1 = 1$, then PSU no. 1 will have a different sample size that is calculated first:

$$n_1 = n\frac{N_1}{N} \tag{3.10}$$

where N_1 is the total number of sampling units in PSU no. 1. In this case $\varphi_1 = n/N$. The n_is calculated are usually not integers. This does not matter. When sampling, each PSU gets a sample size which is n_i rounded either up or down in such a way that they sum to *n* and no PSU is statistically favoured.

If individuals are sampled and age-sex stratification is used at Stage 2, we recommend a solution where *every person in the same age-sex domain in the same stratum has the same selection probability after two stages whichever PSU the person belongs to in the stratum*. But the inclusion probabilities for the PSUs at Stage 1 have been set based on the total population in each PSU, not the population in each age-sex domain. Therefore, the total sample size for an age-sex domain within a stratum will depend somewhat on which PSUs have been selected at Stage 1 and so will the total sample size across all age-sex domains and PSUs in a stratum. Different domain definitions in different strata cam be allowed, but different domain definitions in different PSUs in the same stratum is not allowed.

Starting at the top we define a *desired* sample size for an age-sex domain within a stratum, say n_d^* for domain *d in the actual stratum*. For instance, if you want a total sample size *n* for a stratum, and there are eight domains and you want to have approximately equal sample sizes for each domain you can choose $n_d^* = n/8$. Or if

you want every member of domain *d* in your country to have the same probability of being selected, n_d^* should be set by allocating the nationally desired sample size proportional to the stratum size of domain *d*. In order to obtain equal probability sampling within the domain, the sample size for domain *d* within PSU no. *i* if PSU no. *i* is being selected must be

$$n_{id} = \frac{N_{id} n_d^*}{N_d \pi_i} \tag{3.11}$$

If all $\pi_i < 1$ this amounts to

$$n_{id} = \frac{n_d^*}{m} \frac{N}{N_d} \frac{N_{id}}{N_i} \tag{3.12}$$

Here N_d is the size of the age-sex domain *d* in the stratum, N_i is population size of PSU *i* across all domains and N_{id} is the population in domain *d* in PSU no. *i in the Stage1 frame*. The Stage 2 inclusion probabilities for all eligible individuals in domain *d* in PSU *i* is

$$\varphi_{id} = \frac{n_{id}}{N_{id}} \tag{3.13}$$

The sample sizes n_{id} are calculated for all PSUs before sampling at Stage 1. Their sum over the *selected* PSUs constitutes the *actual* sample size n_d in domain *d* in PSU no. *i* and depends on which PSUs have been selected. The actual sample size is therefore random and usually different from n_d^*, but is equal to n_d^* in expectation.

The variation of n_{id} should be kept as small as possible. Variation in age distribution and to some extent the sex distribution across the PSUs within strata will contribute to variability of the domain sample sizes. This is a strong argument *for considering age distribution when stratifying the PSUs*, whether age-sex stratification is being applied or not. We must expect to find considerable variation in age distribution across PSUs in all countries. Typically, recently established housing areas and areas with considerable immigration have a much younger population than districts where there is emigration and a declining population. It is most often the young population that moves.

The R-data frame produced by EHESsampling when taking the Stage 1 sample contains the anticipated sample sizes for Stage 2 in each selected PSU as shown in Table 3.3. The sampling variation in the domain sample sizes should be assessed for each age-sex domain, for instance with variance calculations. At this stage this has not been implemented in EHESsampling but might come in a future version. Studying the variation by simulating many Stage 1 samples is better than theoretical variance assessments and this option can also be implemented in a future version.

In practice in many countries the Stage 2 samples will often be selected later that the date when the Stage 1 frame was constructed and one PSU by one based on local registers. These local registers will show up PSU-sizes and domain-sizes that are different from the sizes N_i and N_{id} that we used when taking the Stage 1 sample and calculating φ_{id}. Correctly done, φ_{id} will be applied to the local registers to do the actual Stage 2 sampling. The real sample sizes will result from this process and they will be somewhat different from n_{id} in formulae (3.11) - (3.12). Never the less, the method will provide a sample where the selection probability $\pi_i\varphi_{id}$ is the *same for all individuals in the same age-sex domain in the same stratum* whichever PSU (*i*) the person belongs to in the stratum. This is the procedure offered in EHES-sampling

The effect of age-sex stratification is illustrated in Table 3.1 with an example from the test population used in developing EHES-sampling. It creates an overrepresentation of the smallest age-sex domains and an underrepresentation of the larger domains and thus a more even distribution among them in the sample, but not exactly equal sample size.

Table 3.1 Illustration of the effect of age-sex stratification on sample size

Domain	F25_34	F35_44	F45_54	F55_64	M25_34	M35_44	M45_54	M55_64	Total
Total sample	1318	1305	1317	1308	1338	1309	1295	1310	10500
Pct. of sample	12,55	12,43	12,54	12,46	12,74	12,47	12,33	12,48	100,00
Total in testpop	206329	238764	209751	187288	211359	248490	216226	187932	1706139
Pct. of testpop	12,09	13,99	12,29	10,98	12,39	14,56	12,67	11,02	100,00

A simpler procedure for age-sex domain stratification at Stage 2 can be selected.

1. Decide the total sample size at Stage 2 for every (selected) PSU, say n_i = 200

2. Divide this size by the number of domains, e.g. n_{id} = 200/8 = 25.

And take 25 as the sample size for all domains in the PSUs. (3.13) will still be valid. The stratum domain sample size n_d will be exactly as desired. However, this method will *not* yield $\pi_i \varphi_{id}$ the same for every individual in the same age-sex domain in the same stratum whichever PSU (i) the person belongs to in the stratum.

3.4.3 When the Stage 1 frame is approximate

The sampling frame used to establish the Stage 1 frame and to do the calculations described in Section 3.4.2 may be approximate in relation to the actual sampling process. The selection of the PSUs will have to take place well before the survey is carried and may be based on population statistics that are not up-to-date. A Stage 2 sampling frame with the desired SSUs may also exist only locally and not centralised at national level. Furthermore, since the survey will have to take place over a year or more, those selected at the beginning of the survey period may have died or moved from the PSU at the time when the survey team establishes a clinic there. Therefore, it is desirable to take the Stage 2 sample for a PSU as close as possible to the time when the PSU will be visited. This may mean that the Stage 2 samples have to be taken at different times for different PSUs. When considering eligibility and age-sex domains, the age of a person included in the Stage 2 frame should then be taken as the *age at the middle of the data collection period in the actual PSU*. Then the sizes of the PSUs (N_i) and the age-sex domains (N_{id}) may have changed slightly and the anticipated Stage 2 sample sizes calculated in Section 3.4.2 will be adjusted somewhat according to that.

When adjusting the sample sizes we wish to maintain the Stage 2 inclusion probabilities φ_i (or φ_{id}) and over all inclusion probabilities $\pi_i \varphi_i$ (or $\pi_i \varphi_{id}$) calculated in Section 3.4.2. This means that the sample sizes for Stage 2 will have to be recalculated based on new population counts. The Stage 2 sampling procedure described in Section 3.4.4 will automatically handle this.

3.4.4 Taking the Stage 2 sample

To prepare for sampling at Stage 2, select from the main sampling frame all individuals or addresses that belong to the PSUs

selected in Stage 1. The sampling procedure described below is documentation for the interested reader. Other readers can rely on EHESsampling which has a module that does it all.

1. In the data frame consisting of all Stage 2 sampling units in the selected PSUs, associate the Stage 2 inclusion probabilities φ_i (φ_{id}) to each record.
2. For each record, generate a random number *u* between 0 and 1.
3. Sort the data frame by PSU by age-sex domain by *u.*
4. Create a new variable *a* by aggregating the φ_i (φ_{id}) successively up to the previous record in the file and $b = a + \varphi_i$.
5. Generate a new random number *r* between 0 and 1.
6. Find the record in the frame for which $a \leq r < b$. This record is selected.
7. Set *r* = *r* + 1.
8. Go to step 6 and repeat the procedure until you are through all records in the file.

This provides equal probability random samples in all domains by PSUs. If the calculated sample size in a selected PSU is say 23.84, then the actual sample size by this algorithm will be 24 with probability 0.84 and 23 with probability 0.16.

3.5 When using address frames

3.5.1 Multi-dwelling houses

If the main sampling frame is addresses and each dwelling at multiple dwelling addresses cannot be identified in the frame, a third sampling stage may be necessary. It is not desirable to include all dwellings at an address with many dwellings.

If the *number* of dwellings at an address (*k*) is known, copy the record for the address so that there are *k* records for the address in the frame. Take the sample as described previously and apply a unique rule for assigning a physical dwelling to the selected records. The advantage of this is that all dwellings at such multiple dwelling addresses will still have the same probability of being selected as single-address dwellings.

If the number of dwellings at an address is not known, the selection of dwellings should occur when the address is visited for the first time. The dwellings must be mapped and the number of dwellings to be selected must be decided. Sampling dwellings

at an address can be seen as a third stage in the design. If the addresses have been sampled with equal probability, taking a sample of the dwellings at multi-dwelling addresses means that such dwellings will have lower inclusion probabilities than single-address dwellings. This is a disadvantage which will have to be corrected by weighting. To reduce this disadvantage at least two dwellings should be sampled at each such address. On the other hand, sampling many dwellings at multi-dwelling addresses will increase the total number of dwellings and people in the sample.

3.5.2 Selection of individuals within a dwelling

When using address frames, the participant invitation must take place when visiting the address and the selected dwelling for the first time. In both cases, the selection must be random. Random selection of individuals within a dwelling can be seen as a third (or in some cases fourth) stage of sampling and will affect the actual sample sizes, inclusion probabilities and the sampling weights to be used later in estimation.

Basically, everyone in the core age group (25-64 years) living in a selected dwelling should be invited to the survey. This gives every person the same probability of being selected, independent of household size. If the number of eligible people in the dwelling is very high, a maximum should be set. Defining this limit is a national decision, but should not be less than three. To select participants, all eligible individuals in the dwelling must be listed and a random sample must be taken from that list. The selection probabilities for people living in such dwellings will be lower than for those living in “take all” dwellings by a factor equal to the fraction of eligible persons selected in the dwelling. This must be corrected for by proper weighting at the estimation stage. For selection of individuals within a dwelling one can use a (modified) Kish grid. For references to Kish grid techniques see for instance Kish, (1949, 1965) and Nemeth (2001, 2003).

3.5.3 Other situations

When neither of the kinds of sampling frames mentioned in item 1 or 2 in section 3.2 are available, it should be possible to carry out Stage 1 in the sampling design much like when individual or postal-address frames are available. But small area population counts from censuses or other sources must be available at a suitable level. Detailed discussions of such cases will not be done here, but will be taken with the countries which it concerns.

3.6 Documentation and data management

An overview and details for data management are described in Part A, Chapter 12 of the EHES Manual. This section will consider the documentation of the sampling design and the samples produced by that design.

3.6.1 Reporting the sampling at Stage 1

The documentation for Stage 1 sampling must describe the sampling frame for Stage 1, which kind of units are being used for PSUs, how many they are, their stratification and how the PSUs have been selected within each stratum. The documentation must contain two files/tables with a minimum set of columns described below. If EHESsampling is being used for organizing the sampling and selecting the sample, it will produce R data frames with all the information that we ask for and more that if desired can be exported to other formats. The preferred format when submitting the files to EHES RC is semicolon separated ascii (CSV) text file. In each file, the first row should be for the variable names, also separated by semicolons. The main features of the files are described below.

A. Stratification file

A file that describes the stratification. This file must have a name with the format EHES_*CC*_*SC*_stratification. Here *CC* represents the EU's two-letter *C*ountry *C*ode and *SC* represents a two digit *S*urvey *C*ode that identifies different EHES-surveys within the same country. See chapter 12.2. The file must contain one row for each stratum. Below is the list of variables for this file. The variable names are typed in **ARIAL (bold)**.

1. **COUNTRY**. Character (2) *C*ountry *C*ode *CC*.
2. **SURVEY.** Character (2). The *S*urvey *C*ode *SC*.
3. **STRATUM_ID.** Character (max 3). A stratum identifier (code).
4. **STRATUM_NAME.** Character (max 20). Common name of stratum
5. **STRATUM_SIZE.** Integer. The size of the stratum. The total number of SSUs, (*N*)
6. **DOMAINS.** Integer. The number of age-sex strata in Stage 2 sampling. = 1 if no age-sex stratification is used.
7. **ST1_ANT_SSU**. Decimal (2). The anticipated number of SSUs to be selected within the stratum (*n*)

8. **ST1_NO_PSU.** Integer. The number of PSUs in the stratum (M_{PSU})
9. **ST1_SEL_PSU.** Integer. The number of PSUs to be selected in the stratum (m)
10. **ST1_CV.** = 1 if **ST1_SELNO_PSU** has been calculated using cost-variance optimization (section 3.3.4). = 2 otherwise.

    The following items are only relevant if **ST1_CV** = 1.

11. **ST1_CPSU.** Integer. The average cost of establishing a PSU in the stratum (C_{PSU})
12. **ST1_CSSU.** Integer. The average cost of inviting SSU in the stratum (C_{SSU})
13. **ST1_WITHIN.** Decimal (4). The average within PSU variance of the calculation variable (V_{Within})
14. **ST1_AMONG.** Decimal (4). The variance of the PSU means for the calculation variable (V_{Among})
15. **ST1_COST.** (Optional). Integer. The total cost of carrying out the survey in the stratum as calculated by formula (*Cost*)

Table 3.2 shows an example stratification file. The correspondence between the variable names in the formulae and the variable names in the file is shown in the two-line heading. The variable **DOMAINS** has been set to 8 for all strata indicating that age-sex stratification with eight domains is used I all strata. The variable **ST1_SEL_SSU** (n) has in this example been calculated based on a proportional allocation of 9000 sampled individuals but with sample sizes less than 400 adjusted up to a minimum of 400 persons in each stratum. Notice that n has been calculated by an allocation formula which usually does not produce an integer result and is therefore given with decimals. In EHESsampling the rounding to an integer will take place in the sampling process at Stage 2 (see Section 3.4.4). **ST1_VWITHIN** and **ST1_VAMONG** (V_{Within} and V_{Among}) have been based on age coded with 1 = '25-34 years', 2 = '35-44 years', 3 = '45-54 years' and 4 = 55-64 years' and calculated using formula (3.3). This is sufficient accuracy for the purpose although 'Age' could have been used more directly. The values of **ST1_CPSU** and **ST1_CSSU** (C_{PSU} and C_{SSU}) in this example are not real costs, but 'raw guesses' made up for testing purposes. The variable **ST1_COST** (*Cost*) is the total cost of carrying out the survey in the actual stratum calculated according to formula (3.1) in Section 3.3.4 with the values of m, n, C_{PSU} and C_{SSU} given in the table. The variable **STRATUM_NAME** provides a common name for the stratum in addition to its code.

As already mentioned EHESsampling produces a table with the variables that we ask you to report.

Comment: In stratum 03 in the example the value for V_{Among} is large compared to the other strata. This results in a low value for p_{opt} and a high value for m. The high value of V_{Among} and the slightly low value of V_{Within} express a large variation of average age among the PSUs in that stratum which may be typical for cities. In such cases the precision of the survey would benefit from splitting the stratum in two strata, one stratum for the PSUs with average age less than the median (or mean) and one for the PSUs with average age above the median for the PSUs. This would also result in having to select a smaller number of PSUs and a lower cost for that stratum. An alternative is to combine PSUs to larger and less homogenous PSUs as long as they do not become too large to be suitable for the survey.

A stratum file will have to be produced before the actual Stage 1 sample is taken. A file that is used as the input to EHESsampling can have ready made columns for sample sizes (n and m), possibly two or three alternatives in which EHESsampling can do calculations. Or it can contain no columns for n and m and let EHESsampling calculate them. However, the stratification file to be reported should only contain the sample sizes for the design actually used.

B. Primary Sampling Unit (PSU) file

The PSU-file to be reported is a file that describes the *selected* PSUs only, the Stage 1 sample. This file is described below. But before taking the Stage 1 sample EHESsampling will establish a PSU file similar to the one we ask you to report but with all PSUs in the frame. Input for establishing this file is a file with the more basic variables from which the remaining new variables are calculated. The input file is described in the EHESsampling manual. The Stage 1 sample will be a sample of PSUs from this file. The file name should have the format EHES_*CC*_*SC*_PSUSAMPLE where *CC* and *SC* follow the same standard as for the stratum file A. *If age-sex stratification will not be used at Stage 2 the PSU-file must contain one record for each selected PSU. If age-sex stratification is to be used the file should contain one row for each PSU by age-sex domain in each selected PSU.* Each row should contain the variables

Table 3.2 Example Stratification File (EHES_NO_CC_stratification)

				N		*n*	M_{PSU}	*m*		C_{PSU}	C_{SSU}	V_{Within}	V_{Among}	*Cost*
COUNTRY	**SURVEY**	**STRATUM_ ID**	**STRATUM_ NAME**	**STRATUM_ SIZE**	**DOMAINS**	**ST1_ ANT_ SSU**	**ST1_ NO_ PSU**	**ST1_ SEL_ PSU**	**ST1_ CV**	**ST1_ CPSU**	**ST1_ CSSU**	**ST1_ VWITHIN**	**ST1_ VAMONG**	**ST1_ COST**
NO	01	01	Østfold	72970	8	400.00	43	2	1	10000	50	1.1531	0.0068	40000
NO	01	02	Akershus	259795	8	1370.44	99	9	1	8000	45	1.0905	0.0091	133670
NO	01	03	Oslo	175955	8	928.17	139	22	1	8500	60	1.0936	0.0908	242690
NO	01	04	Hedmark	65801	8	400.00	34	3	1	7000	55	1.1644	0.0060	43000
NO	01	05	Oppland	68775	8	400.00	42	2	1	8000	50	1.1563	0.0054	36000
NO	01	06	Buskerud	101992	8	538.01	56	4	1	9000	60	1.1610	0.0104	68281
NO	01	07	Vestfold	70045	8	400.00	43	3	1	9500	70	1.1313	0.0078	56500
NO	01	08	Telemark	32045	8	400.00	18	3	1	8500	65	1.1592	0.0070	51500
NO	01	09	Aust-Agder	35719	8	400.00	21	3	1	7500	45	1.1788	0.0100	40500
NO	01	10	Vest-Agder	73670	8	400.00	37	3	1	8000	50	1.1594	0.0099	44000
NO	01	11	Rgaland	182525	8	962.83	89	10	1	10000	55	1.1446	0.0235	152956
NO	01	12	Hordaland	169316	8	893.15	108	10	1	9500	40	1.1529	0.0358	130726
NO	01	14	Sogn og Fjordane	26729	8	400.00	12	2	1	8000	35	1.1471	0.0074	30000
NO	01	15	Møre og Romsdal	63492	8	400.00	39	3	1	8500	50	1.1636	0.0099	45500
NO	01	16	Sør-Trøndelag	112555	8	593.74	59	9	1	7500	55	1.1729	0.0347	100155
NO	01	17	Nord-Trøndelag	47782	8	400.00	22	2	1	6500	50	1.1630	0.0038	33000
NO	01	18	Nordland	63237	8	400.00	35	2	1	8000	45	1.1505	0.0042	34000
NO	01	19	Troms	58134	8	400.00	33	4	1	9000	60	1.1535	0.0183	60000
NO	01	20	Finnmark	25602	8	400.00	15	3	1	9500	75	1.1591	0.0072	58500

1. **COUNTRY**. Character (2). *C*ountry *C*ode *CC*.
2. **SURVEY.** Character (2). The *S*urvey *C*ode *SC*
3. **STRATUM_ID.** Character (max 3). A stratum identifier (code).
4. **PSU_SN.** Character (max 4). A PSU serial number (maximum four digits) which *replaces* the real PSU ID (e.g. postcode, municipality code etc.) that has to be used nationally to identify the PSU for data collection. The purpose of the PSU serial number is to tell which individuals or households belong to the same PSU. This is important for proper analysis of sampling variance. It will not be necessary or even desirable for the Reference Centre to know which real PSU is represented by a PSU serial number and for confidentiality reasons that information should not be transferred. We recommend that the serial numbers run across strata since this will distinguish PSUs without using the stratum variable. A link between the PSU serial number and the real PSU ID should be maintained by the national survey organizer only.
5. **PSU_SIZE.** Integer. The size of the PSU (N_i)
6. **ST1_PROB.** Decimal (4). The Stage 1 inclusion probability (π_i) used in sampling
7. **ST2_PROB.** Decimal (4). The Stage 2 inclusion probability or for the PSU (φ_i) or domain (φ_{id}) (if **DOMAINS** > 1 in the stratification file). Notice that if age-sex stratification is used this probability will be different for different age-sex domains.
8. **ST2_ANT_SSU.** Decimal (4). Optional. Anticipated sample size within the PSU or age-sex domain (n_i or n_{id})

   If age-sex stratification is used (**DOMAINS** > 1 in the stratification file) in at least one stratum it should also contain

9. **DOMAIN_ID.** Character (max 10). A domain identifier specifying the age-sex domain for the record
10. **DOMAIN_SIZE_PSU**. Integer. The number of people in each age-sex domain in the PSU (N_{id}).
11. **DOMAIN_SIZE_STR**. Integer. The number of people in each age-sex domain within the stratum (N_d).

Table 3.3 presents an excerpt from an example PSU file without age-sex stratification and Table 3.4 present an excerpt from an example with age-sex stratification. In these files, postcode is used as the PSU identifier.

In Table 3.3 **ST1_PROB** (π_i) is calculated from formula (3.6) and **ST2_ANT_SSU** (n) is calculated from formula (3.8), Section 3.4.2 and **ST2_ANT_SSU** is the same for all PSUs in the same stratum and will be rounded to an integer in the Stage 2 sampling. **ST2_PROB** (φ_i) is calculated by formula (3.9) and the final selection probabilities after two stages $\pi_i\varphi_i$ will be the same for all PSUs (and SSUs) in the same stratum.

Table 3.3 Excerpt from a Primary Sampling Unit file without age-sex domains

				N_i	π_i	φ_i	n_i
COUN-TRY	**SURVEY**	**STRA-TUM**	**PSU_SN**	**PSU_ SIZE**	**ST1_ PROB**	**ST2_ PROB**	**ST2_ ANT_ SSU**
NO	01	01	098	1886	0.0517	0.1060	200.00
NO	01	01	131	1433	0.0393	0.1396	200.00
NO	01	02	001	2549	0.0883	0.0597	152.27
NO	01	02	007	5681	0.1968	0.0268	152.27
NO	01	02	013	2767	0.0959	0.0550	152.27
NO	01	02	017	1958	0.0678	0.0778	152.27
NO	01	02	119	2364	0.0819	0.0644	152.27
NO	01	02	026	2999	0.1039	0.0508	152.27
NO	01	02	054	7316	0.2534	0.0208	152.27
NO	01	02	067	2869	0.0994	0.0531	152.27
NO	01	04	018	6983	0.3184	0.0191	133.33
NO	01	04	037	2099	0.0957	0.0635	133.33
NO	01	04	083	1422	0.0648	0.0938	133.33
NO	01	05	037	1244	0.0362	0.1608	200.00
NO	01	05	081	2217	0.0645	0.0902	200.00
NO	01	06	105	1426	0.0559	0.0943	134.50
NO	01	07	022	1700	0.0728	0.0784	133.33
NO	01	07	024	1059	0.0454	0.1259	133.33
NO	01	07	067	1414	0.0606	0.0943	133.33

In Table 3.4 **ST2_DOMAINS** = 8 and each PSU is represented by eight rows, one row for each age-sex domain, labelled by the variable **ST2_DOMAIN_ID**. **ST1_PROB** (π_i) is calculated in the same way as in Table 3.3. **ST2_ANT_SSU** (n_{id}) is calculated from (3.12) and varies over the domains in the same PSUs, but their sums over all domains are the same as in Table 3.3. **ST2_PROB** (φ_{id}) has been calculated using formula (3.13). The final selection probabilities after two stages $\pi_i\varphi_{id}$ will be the same for the same age-sex domain in all PSUs (and SSUs) in the same stratum but differs across domains.

Countries that do age-sex stratification by simply taking the same number of persons in each age-sex domain in each PSU (see end of Section 3.4.2) should report that number for **ST2_ANT_SSU. ST2_PROB** (φ_{id}) should still be calculated using formula (3.13), but since **ST2_ANT_SSU** (n_{id}) will be different the final selection probabilities after two stages $\pi_i\varphi_{id}$ will *not* be the same for the same age-sex domain in all PSUs (and SSUs) in the same stratum.'

Table 3.4 Example Stratification File (EHES_NO_CC_stratification)

COUNTRY	SURVEY	STRATUM	PSU_SN	N_i PSU_SIZE	π_i ST1_PROB	DOMAIN_ID	N_{id} DOMAIN_SIZE_PSU	N_d DOMAIN_SIZE_STR	φ_{id} ST2_PROB	n_{id} ST2_ANT_SSU
NO	01	02	103	3957	0.1371	M25_34	377	26141	0.0478	18.02
NO	01	02	103	3957	0.1371	M35_44	663	39975	0.0313	20.73
NO	01	02	103	3957	0.1371	M45_54	601	35218	0.0355	21.33
NO	01	02	103	3957	0.1371	M55_64	337	28111	0.0445	14.98
NO	01	02	103	3957	0.1371	F25_34	386	27438	0.0455	17.58
NO	01	02	103	3957	0.1371	F35_44	704	40418	0.0309	21.77
NO	01	02	103	3957	0.1371	F45_54	554	33824	0.0369	20.47
NO	01	02	103	3957	0.1371	F55_64	335	28670	0.0436	14.60
NO	01	02	029	4133	0.1432	M25_34	304	26141	0.0458	13.91
NO	01	02	029	4133	0.1432	M35_44	690	39975	0.0299	20.65
NO	01	02	029	4133	0.1432	M45_54	610	35218	0.0340	20.72
NO	01	02	029	4133	0.1432	M55_64	437	28111	0.0426	18.60
NO	01	02	029	4133	0.1432	F25_34	368	27438	0.0436	16.05
NO	01	02	029	4133	0.1432	F35_44	685	40418	0.0296	20.28
NO	01	02	029	4133	0.1432	F45_54	629	33824	0.0354	22.25
NO	01	02	029	4133	0.1432	F55_64	410	28670	0.0417	17.11
NO	01	02	052	1698	0.0588	M25_34	112	26141	0.1114	12.48
NO	01	02	052	1698	0.0588	M35_44	249	39975	0.0729	18.14
NO	01	02	052	1698	0.0588	M45_54	255	35218	0.0827	21.09
NO	01	02	052	1698	0.0588	M55_64	218	28111	0.1036	22.58
NO	01	02	052	1698	0.0588	F25_34	176	27438	0.1061	18.68
NO	01	02	052	1698	0.0588	F35_44	250	40418	0.0721	18.01
NO	01	02	052	1698	0.0588	F45_54	220	33824	0.0861	18.94
NO	01	02	052	1698	0.0588	F55_64	218	28670	0.1016	22.14

3.6.2 Reporting the sampling at Stage 2

The sample resulting from Stage 2 must be reported to the EHES RC in full. *All sampled units must be included in the file, even if they were later found to be ineligible to the sample or they did not participate in the survey*. The filename format should be EHES_*CC*_*SC*_st2sample. There should be *no direct identifiers* such as ID number, names or addresses on the file. The file should contain

1. **COUNTRY**. Character (2). The two-character *C*ountry *C*ode *CC*.
2. **SURVEY.** Character (2). The two digit *S*urvey *C*ode *SC*.
3. **STRATUM_ID.** Character (max 3). A stratum identifier, max 3 characters.
4. **PSU_SN.** Character (max 4). A PSU serial number (maximum four digits), the same as in the PSU-file.
5. **SERIAL.** Character (max 12). Serial number that uniquely identifies a person within the survey. When the SSUs are individuals, the number can be assigned immediately after the sample has been selected. EHESsampling provides option for this. If households are used as SSUs the serial number must be assigned when the household is visited. Must not contain information that identifies the person in the population. See section 12.2. Assigned after visit of household if **HOUSEHOLD_UNIT** = 1.
6. **ST2_DOMAIN_ID.** Character (max 10). A domain identifier specifying the age-sex domain for the record. Only relevant if **DOMAINS** > 1 in the stratum file. Then equal to **DOMAIN_ID** in that file.
7. **ST2_SEL_SSU.** (Integer). Number of SSUs actually selected within the PSU or domain. Must be calculated when the Stage 2 sampling has taken place. All who were selected should be counted here, also those who were later found to be not eligible to the sample and those who did not respond.
8. **HOUSEHOLD_UNIT.** = 1 if addresses/households/dwellings are used as SSUs. = 2 otherwise
9. **HOUSEHOLD_SN.** Character (max 5). Relevant only if **HOUSEHOLD_UNIT** = 1. An address or *H*ousehold *S*erial *N*umber (*HSN*) with maximum five digits (code "88888" on all records if addresses or households are *not* used as sampling units). The number can be assigned within or across the PSU serial numbers.

10. **ST3_SAMPLING.** Only relevant if **HOUSEHOLD_UNIT** = 1. =1 if there is probability sampling within households. = 2 otherwise.
11. **ST3_PROB.** Decimal (4). Stage 3 inclusion probabilities. If **ST3_SAMPLING** = 1 then the stage 3 incusion probability. Otherwise ST3_PROB= 1.0000.
12. **ALL_PROB.** The overall inclusion probability. **= ST1_PROB * ST2_PROB * ST3_PROB.**
13. **SAMPLING_WEIGHT.** Decimal (4). **= 1/ ALL_PROB**.

With reference to Part A, Section 12.2, a SERIAL NUMBER must be given to everybody selected to the sample (i.e. not only for example those eventually found eligible or to the survey participants). This serial number must be unique to every person within the survey. Whenever the sampling units are individuals this serial number should be assigned immediately after sampling.

However, when addresses or households are used as sampling units, *SN* can be completed only at the stage when the household is visited and all eligible subjects at the address or in the household have been mapped.

Table 3.5 Excerpt from a Stage 2 sample (with age-sex stratification)

COUNTRY	SURVEY	STRA-TUM	PSU_SN	SERIAL	HOUSE-HOLD_ UNIT	DOMAIN_ID	ST2_SEL_ SSU	HOUSE-HOLD_ SN	ST3_ SAMP-LING	ST3_ PROB	ALL_PROB $\pi_i\varphi_{id}$	SAMPLING_ WEIGHT $1/\pi_i\varphi_{id}$
NO	01	03	004	00261	2	M25_34	4	88888	2	1.0000	0.0036	277.189
NO	01	03	004	00262	2	M25_44	4	88888	2	1.0000	0.0036	277.189
NO	01	03	004	00263	2	M25_44	4	88888	2	1.0000	0.0036	277.189
NO	01	03	004	00264	2	M25_34	4	88888	2	1.0000	0.0036	277.189
NO	01	03	004	00265	2	M35_44	5	88888	2	1.0000	0.0045	221.691
NO	01	03	004	00266	2	M35_44	5	88888	2	1.0000	0.0045	221.691
NO	01	03	004	00267	2	M35_44	5	88888	2	1.0000	0.0045	221.691
NO	01	03	004	00268	2	M35_44	5	88888	2	1.0000	0.0045	221.691
NO	01	03	004	00269	2	M35_44	5	88888	2	1.0000	0.0045	221.691
NO	01	03	004	00270	2	M45_54	3	88888	2	1.0000	0.0065	154.031
NO	01	03	004	00271	2	M45_54	3	88888	2	1.0000	0.0065	154.031
NO	01	03	004	00272	2	M45_54	3	88888	2	1.0000	0.0065	154.031
NO	01	03	004	00273	2	M55_64	4	88888	2	1.0000	0.0083	120.753
NO	01	03	004	00274	2	M55_64	4	88888	2	1.0000	0.0083	120.753
NO	01	03	004	00275	2	M55_64	4	88888	2	1.0000	0.0083	120.753
NO	01	03	004	00276	2	M55_64	4	88888	2	1.0000	0.0083	120.753
NO	01	03	004	00277	2	F25_34	6	88888	2	1.0000	0.0036	277.189
NO	01	03	004	00278	2	F25_34	6	88888	2	1.0000	0.0036	277.189
NO	01	03	004	00279	2	F25_34	6	88888	2	1.0000	0.0036	277.189
NO	01	03	004	00280	2	F25_34	6	88888	2	1.0000	0.0036	277.189
NO	01	03	004	00281	2	F25_34	6	88888	2	1.0000	0.0036	277.189
NO	01	03	004	00282	2	F25_34	6	88888	2	1.0000	0.0036	277.189
NO	01	03	004	00283	2	F35_44	4	88888	2	1.0000	0.0051	196.920
NO	01	03	004	00284	2	F35_44	4	88888	2	1.0000	0.0051	196.920
NO	01	03	004	00285	2	F35_44	4	88888	2	1.0000	0.0051	196.920
NO	01	03	004	00286	2	F35_44	4	88888	2	1.0000	0.0051	196.920
NO	01	03	004	00287	2	F45_54	3	88888	2	1.0000	0.0068	146.550

3.7 Some common designs – a discussion

Two stage sampling is a complicated matter. There are some common ways of doing two-stage sampling that we have not recommended. One of these is to sample the PSUs with equal probability within each stratum at Stage 1, perhaps only one or two PSUs. This will produce a sample with many more of the smaller PSUs compared to the large ones than a probability proportional to size (PPS) design. Another strategy is to sample the same proportion of SSUs in each selected PSU at Stage 2. Depending on which combination of strategy for Stage 1 and Stage 2 one chooses one will get equal or unequal selection probabilities after two stages, over or under representation of SSUs in small PSUs versus large ones or fixed or random total sample size. Random sample sizes results in an unpredictable number of invited participants. This is disadvantageous both from a statistical, cost and administrative point of view since both will depend heavily on the sample size. The four combinations and their respective advantages and disadvantages are depicted in Table 3.6. Notice the only design recommended in this chapter requires both fixed sample sizes and equal selection probabilities for every secondary sampling unit (SSU) after two stages. The table assumes that the Stage 1 design provides a fixed number of PSUs in the Stage 1 sample.

Table 3.6 Combinations of sampling strategies for Stage1 and Stage2

Combinations of sampling strategies for Stage 1 and Stage 2		
Stage 2	**Stage 1**	
	A. Euqal probabilities selection	**B. Probabilities proportions to size**
1. Same sampling proportions in all PSUs	Equal probabilities for all people/-households after two stages within stratum. Small samples in small PSUs. Large samples in large PSUs. Final total sample size unpredictable and depends on which PSUs are selected at stage 1.	Large samples in large PSUs and small samples in small PSUs in the same stratum. Unequal selection probabilities after two stages: People/households in large PSUs over represented Final total sample size unpredictable and depends on which PSUs are selected at stage 1.

Combinations of sampling strategies for Stage 1 and Stage 2		
Stage 2	**Stage 1**	
	A. Euqal probabilities selection	**B. Probabilities proportions to size**
2. Same sample sizez in all PSUs	Fixed sample sizes after two stages. Smaller stage 2 probabilities (sampling proportions) for people/-households in large PSUs than small PSUs resulting in unequal selection probabilities after two stages: People/households in large PSUs are under represented compared to people/households in small PSUs.	Fixed sample size after two stages. Equal selection probabilities after two stages. **RECOMMENDED DESIGN**

Combination of stratification of PSUs by size and strategy A2 is common. A strategy seen in some countries has been to first stratify the PSUs by regions and within each region by three sizes. If the same number of PSUs is selected in each size-stratum, people and households in large PSUs will be underrepresented in the total sample. It is possible to compensate for this by selecting more PSUs in the strata for large PSUs, but it will be difficult to establish exact equal probability samples that way.

Selecting only one or two PSUs per stratum may save the costs of establishing a large number of examination clinics, but may lead to larger sampling variances than selecting a larger number of PSUs and less people within each of them, in particular if the PSUs are not very similar with respect to relevant characteristics (V_{Among} is large). To find a good balance between cost and variances is the purpose of allocation formula .

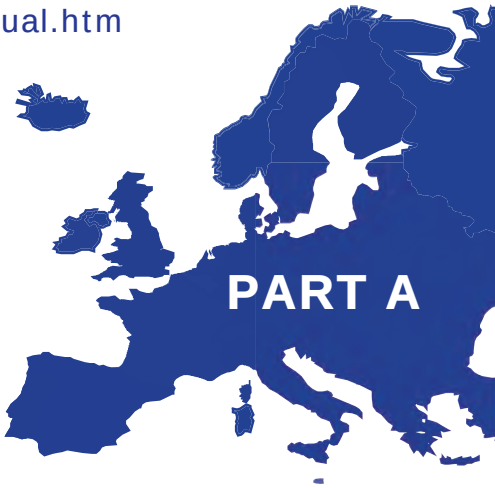
On the other hand, if it is possible to make a detailed stratification with homogenous strata (V_{Among} is small), selecting only two PSUs per stratum may be optimal. This is shown in Table 3.2 in Section 3.6. Even selecting one PSU per stratum may be cost-variance optimal. But this will render unbiased estimation of the Stage 1 component of the variances infeasible in these strata since no variation among the PSUs will then be visible in the data. For this reason EHESsampling always selects at least two PSUs per stratum.

References

- Craig R, Shelton N, (eds). Health Survey for England 2007. Volume 2: Methodology and Documentation. Leeds: The NHS Information Center; 2008.
- Axelson M, Bihler W, Djerf K, et al (prepared by) European Health Interview Survey. The Task Force III report on sampling issues. 2009. Available at https://circabc.europa.eu/sd/d/4f89b382-b3e1-4884-94e1-ca9f2e2338b6/Task%20Force%20III%20report%20on%20sampling%20issues%20up%20to%20date%20November%202009.pdf. Accessed 22 January 2013.
- Jentoft S, Heldal J. RcmdrPlugin.EHESsampling v. 2.0. [Software Package]. Comprehensive R Archive Network (CRAN); 2011.
- Kish L. A procedure for objective respondent selection within the houshold. Journal of the American Statistical Association. 1949 Sep;44(247):380-7.
- Kish L. Survey Sampling. New York: Wiley; 1965.
- Nemeth R. Respondent selection within the houshold: A modification of the Kish Grid. 2001; Available from: https://www.stat.aau.at/Tagungen/Ossiach/Nemeth.pdf. Acceessed on 22 January 2013
- Nemeth R. Sampling design of Health Surveys: Household as the sampling unit. 2003.
- NHANES. National Health and Nutrition Examination Survey (NHANES). Vital and health statistics. Plan and operation of the third National Health Examination Survey 1988-94. 1994.
- NHANES. National Health and Nutrition Examination Survey (NHANES). Analytic and reporting guidelines. 2006.
- Tolonen H, Koponen P, Aromaa A, Conti S, Graff-Iversen S, Grøtvedt L, et al. (eds.) Reveiw of Health Examination Surveys in Europe. (2008) Publications of the National Public Health Institute. Helsinki. B18/2008. Available from http://urn.fi/URN:ISBN:978-951-740-843-1. Accessed on 22 January 2013

Additional literature on sampling

- Tolonen H, Koponen P, Aromaa A, Conti S, et al (eds.). Recommendations for Health Examination Surveys in Europe. (2008) Publications of the National Public Health Institute. Helsinki. B21/2008. Available from:http://urn.fi/URN:NBN:fi-fe201204193358. Accessed on 22 January 2013.
- Lethonen R, Pahkinen E. Practical Methods for Design and Analysis of Complex Surveys. New York: Wiley; 2004.
- DEGS. Operationshandbuch. Berlin: Robert Koch Institut; 2010.
- Rao PSRS. Sampling Methodologies with Applications. Chapman & Hall; 2000
- Särndal CE, Swensson B, Wretman J. Model assisted Survey Sampling. New York: Springer-Verlag; 1992. (Springer Series in Statistics).
- Tillé Y. Sampling Algorithms. New York: Springer; 2006. (Springer Seriers in Statistics).
- Thompson SK. Sampling. 2nd ed. Wiley; 2002
- Valliant R, Dorfman AH, Royall RM. Finite Population Sampling and Inference – A Prediction Approach. Wiley; 2000



4. Legal and ethical aspects

Susanna Conti[1], Mark Kanieff[1], Grazia Rago[1]

[1] Istituto Superiore di Sanità (ISS), Rome, Italy

In any research involving humans, ethical conduct is a fundamental concern. This means that the research must be performed so that participants are protected not only from risks to their physical and mental health but also from risks to their privacy and from receiving misinformation. Although performing a HES does not pose any serious risk to health, the safeguarding of privacy and acquiring informed consent are crucial ethical aspects. In this chapter, we describe a series of recommendations related to the legal and ethical aspects of performing a HES in Europe. These recommendations are based, in part, on a survey of how Member States addressed these concerns in previous HESs or similar studies (Tolonen 2008). In particular, we provide some general recommendations on the ethical conduct of a HES, with specific reference to the safeguarding of privacy (or "data protection"); we also provide and discuss a model of an informed consent form, which is intended as a guide for creating such a form for HESs in Europe.

4.1. Legislation and guidelines

National HESs must be conducted according to ethical standards, which, for all research on humans, are regulated by national legislation and national and international guidelines, for example:

1. national acts regulating the status and/or rights of patients
2. national medical research acts
3. other national ethical principles of research involving humans
4. international biomedical research guidelines, such as:

- the Declaration of Helsinki, "Ethical Principles for Medical Research Involving Human Subjects", which is considered to be the pillar of ethical standards (WMA 2008)
- the Belmont Report ("Ethical Principles and Guidelines for the Protection of Human Subjects of Research" (NIH 1979), and
- the Recommendation of the Committee of Ministers No. R(90) 3 concerning medical research on human beings (Committee of Ministers 1990); and
- the Oviedo Convention on Human Rights and Biomedicine. (Oviedo 1997)

4.2 Role of ethics committees

An ethics committee is a body that is responsible for evaluating research proposals from an ethical standpoint. In particular, this committee, which can be local, regional, or national, evaluates the given proposal in terms of its compliance with national legislation and regulations. The evaluation covers not only the performance of the research itself (i.e., that the patient will not be harmed or placed at risk) but also the contents of the informed consent and how it is obtained, the safeguarding of privacy, and the use of data and biological materials, both for the research being conducted and any future purposes.

The approval of the ethics committee is needed not only for full-size national HES but also for the pilot studies. It must also be considered that obtaining ethical approval can be a time-consuming process; in some countries or circumstances it may take up to one year. Therefore, the procedures for obtaining approval need to be started as early as possible, during the beginning of the planning phase.

The general steps for obtaining ethical approval are illustrated in Figure 4.1, though the detailed procedures may vary by country. The first step is to identify the appropriate ethics committee and the documentation that this committee requires for applying for approval. In preparing this documentation, it is recommended that experts in ethical issues be consulted. Once the proposal is submitted for approval, the ethics committee may request modifications if the proposal does not fulfill the established criteria. The HES cannot be started before approval is obtained.

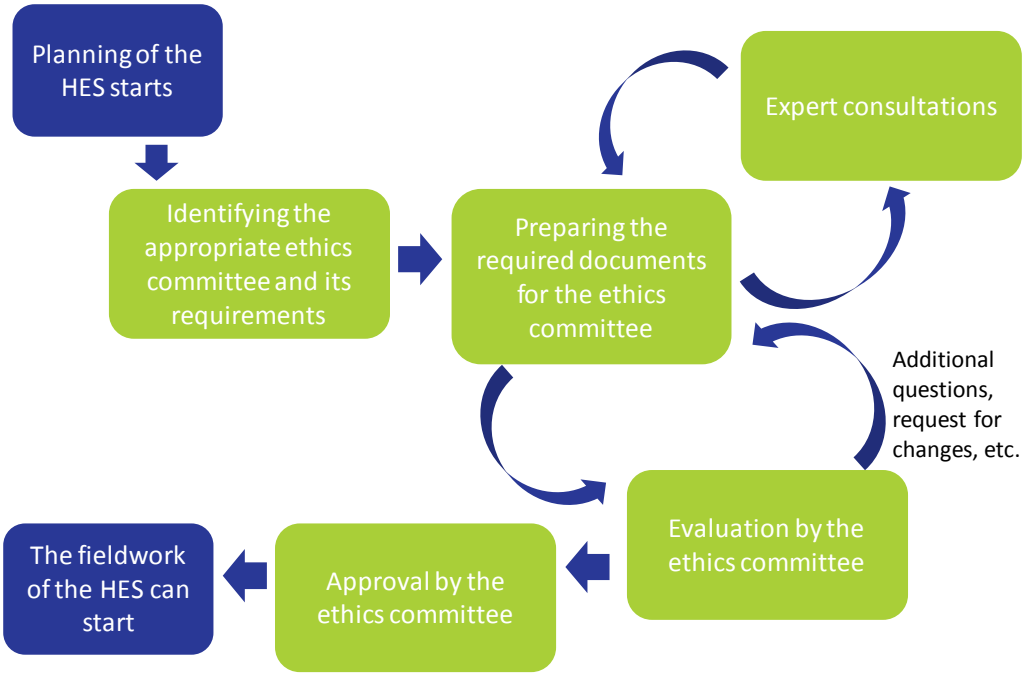


Figure 4.1 Process for obtaining ethical approval

4.3 Data protection

The Declaration of Helsinki states, *"Every precaution should be taken to respect the privacy of the subject and the confidentiality of the patient's information..."*. This issue has become increasingly important in light of the progress made in information technology and the consequent ease of access to data. That privacy is safeguarded is ensured through legislation (generally a "Data Protection Act").

Performing a HES includes collecting individual level data which are also personal data (i.e., sensitive data regarding health). For this reason, the HES protocol must comply with the given country's Data Protection Act and cover all aspects of data protection, in particular: access to data, the exchange of data, record linkage, and anonymisation procedures (more detailed information on methods for ensuring data security are provided in Part A, Chapter 12). In Europe, the most important document regarding data protection is: "Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data" (WMA 2008). The issue of ensuring data protection is also of extreme importance in developing informed consent material.

To understand better the concept of data protection, some commonly used terms are defined below (more detailed definitions are provided in the above-mentioned Directive).

- *Personal Data* - information regarding an identifiable person, that is, one who can be directly or indirectly identified, in particular by reference to an identification number or to factors specific to his/her physical, physiological, mental, economic, cultural or social identity
- *Processing of Personal Data* - any operation (automatic or not) performed on personal data, for example, collection, storage, adaptation or alteration, retrieval, linkage, destruction and dissemination
- *Controller* - the person or entity that determines the purposes and means of the processing of personal data
- *Processor* - the person or entity that processes personal data on behalf of the controller
- *Personal Data Act (or Data Protection Act)* - legislation for protecting the privacy of natural persons in the processing of personal data
- *Sensitive Data* - personal data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, trade-union membership, criminal convictions, and data concerning health or sex life
- *Right of Access* - the right of a human subject to consult the data collected on him/her
- *Duty of Notification* - the obligation of the controller to notify the data protection authorities of the intention to perform data processing, including a description of the processing

4.4 Informed consent

4.4.1 Objectives of informed consent

Before performing any kind of research involving humans, informed consent must be obtained. The objective of informed consent is to allow a person to make a truly informed decision as to whether or not to participate in the HES. In other words, obtaining informed consent goes beyond getting an individual to sign a form: it is a process of communication between an individual and the HES personnel. Its goal is to ensure that the individual fully understands the scopes of the study, the methods adopted, and how the data will be used.

The first step in obtaining informed consent is to provide the study candidate with information. Given that the ultimate goal is to ensure that participants are truly informed, it is fundamental that this information be complete and clear. This is also important in terms of the HES participation rate, in that unclear or

poorly written material (or even material that relies too heavily on "scientific" terms, which could be intimidating) could result in a candidate's decision not to participate.

4.4.2 Means of providing information for informed consent

In this section, we discuss how to provide the HES candidate with information on the HES and how to obtain written informed consent. These procedures basically consist of three main activities: i) making the candidate aware of the HES; ii) providing the candidate with a clear understanding of what participation involves; and iii) obtaining the candidate's signature attesting to his/her consent to participate. This is done using what is called an "informed consent form". This form contains all of the information required by the HES candidate for understanding the HES and what participation entails and a space for the candidate's signature, attesting to the fact that he/she has understood the information and agrees to participate. In some cases, the information on the HES is contained on a separate document (referred to as an "information notice" or "information leaflet"). Furthermore, an "invitation letter" can also be provided, which is used as an introduction to explain in general what the study is about, its importance, and how and when the candidate will be contacted (If used, the invitation letter should be brief yet "appealing"; for more information on the invitation letter, see Part A, Chapter 13, "Recruitment of participants".).

One advantage to using separate documents is that the information notice and invitation letter can be provided some time before the HES candidate provides a signature, so that he/she has sufficient time for reading and understanding the information before agreeing to participate. The choice of the information material's format also depends on such factors as the general organization of the study and the laws and regulations regarding privacy or related issues. For example, in some countries, the ethics committee explicitly requires that the informed consent form consist of a single document that includes both the necessary information and the signature. In developing informed consent material, setting up telephone help-lines for answering candidates' questions and providing clarifications can also be considered, as can the translation of material into other languages. A web-site dedicated to the HES could also be created, with all of the information about the study, including the information notice itself.

EXAMPLES OF INFORMED CONSENT MATERIAL AND WHEN TO PROVIDE IT

Given below are three examples of formats for the informed consent material and when to provide it. The extent of the information in each document can be modified, depending on specific needs. For example, if the information notice is “extensive” (i.e., if it contains most of the necessary information), then the invitation letter and the informed consent form can be brief. Whatever format is chosen, the documents should be complementary, that is, there should not be excessive overlap, to avoid burdening the candidate with too much reading.

Example 1

Invitation letter + extensive information notice, sent together some days before the candidate presents for the HES:
⇨ brief informed consent form, provided when the candidate presents for the HES

Example 2

Invitation letter + brief information notice, both sent some days before the candidate presents for the HES:
⇨ extensive informed consent form, provided when the candidate presents for the HES

Example 3

Invitation letter sent some days before the candidate presents for the HES + extensive information notice published on a web-site, with toll-free line provided to ask questions:
⇨ written information notice (the same published on the web-site) and informed consent form provided and explained to the candidate when presenting for the HES.

4.4.3 Recommendations for creating an informed consent form

This section is intended to help you to create an informed consent form for the HES in your country. If you feel that another format would be more suited to your HES, then you are free to make any changes deemed necessary. For example, in the present form, candidates are asked to provide a single signature which indicates consent to participate in all parts of the study. However, in some countries it may be required (or preferable) that the candidate provide a separate consent and signature for

each individual activity that he/she will undergo (e.g., blood taking, linking of data to other databases). Many of the statements on this form are followed by a comment (in italics) that provides suggestions or considerations which may help you to adapt the form for use in your HES.

Given that the ultimate goal is to ensure that participants are truly informed, the information provided must be complete and clear. You should thus use terminology that is simple and easy to understand, avoiding scientific terms when possible. Moreover, excessively complex or long descriptions can confuse or intimidate study candidates.

The protocol for conducting the HES in your country, including the informed consent form, will have to be approved by your national, regional, or local ethics committee, so as to ensure that it complies with national legislation and ethical standards. Many of the sections in this form may have to be modified to be consistent with the legislation in your country (e.g., that regarding access to data and storage of samples in biobanks).

4.4.4 Model of an informed consent form to be used in European HESs

Below is provided a model of an informed consent form. The model includes an introduction which explains its purpose and provides recommendations for those who will be responsible for this aspect of the HES. In the model, the information notice and the signature form are a single document, yet as mentioned above, these can be two separate documents provided to study candidates at different times, together with an invitation letter.

In the model, asterisks indicate items that are “mandatory”, that is, those that should always be included. The other items depend on the specific characteristics of the HES. For example, if data linkage is not performed, then it is not necessary to ask the candidate for consent.

a) Introductory information on the HES

To the survey candidate,

You are invited to take part in a National Health Examination Survey (or “HES”).

> ***Comment:*** *If the HES includes minors or persons not capable of providing informed consent, the word “you” should be substituted with “your child” or “your legal ward” or with the actual name of the study participant. To this regard,*

the issue of providing results to persons not capable of providing informed consent will also need to be considered.

A HES is a study carried out for obtaining information on general health by interviewing individuals and measuring certain indicators that can be important to health, such as weight, blood pressure, and cholesterol level. This information is used to acquire knowledge on the health status of the population, which can be important in promoting and improving the health of all.

> ***Comment:*** *Information on health concerns that are important in the specific country and for which a HES could be beneficial can be added here. For example "In Italy, obesity is becoming an increasingly important health concern, yet there is little information on what percentage of the population can be considered as obese." If the candidate feels that the study would be socially useful, then the chances of him/her participating could increase.*

The HES is being conducted by (specify name of organization conducting the HES in your country) among a sample of (specify expected number of participants individuals in specify study area, such as the town or province). Your name was chosen from (specify source of the person's name and the area to which it refers)

> ***Comment:*** *This sentence should specify how the individual was chosen (e.g., from electoral rolls, social insurance registers, population registers), so that he/she is aware of how the research personnel obtained his/her name.*

All aspects of this study have been approved by the Ethics Committee of the (specify the name or level of the ethics committee , e.g., France's National Ethics Committee). The present form includes important information about the study and a description of what will be asked of you if you decide to participate. In order to participate, you will need to carefully read and sign this form. If any part of this form is not clear to you, please feel free to ask the person/s obtaining the informed consent.

> ***Comment:*** *The wording of this sentence may change according to who is available for providing clarifications or depending on whether or not information aids, such as telephone help-lines, are provided.*

Your participation is important to us, but please be assured that it is voluntary, that you may leave the study at any time, and that your data will be kept confidential.

b) Collection of personal data

During the survey, you will be asked to answer questions on ...

> ***Comment:*** *Specify the topics that the questions will cover. If an interview is not conducted (e.g., if a self-administered questionnaire is used), the wording of this section should be modified accordingly).*

Measurements of your height, weight, waist circumference and blood pressure will be taken; blood/urine/saliva samples will also be taken.

> ***Comment:*** *If the HES comprises additional modules, then modify accordingly.*

These samples will be tested for ...

> ***Comment:*** *To be modified in accordance with the specific objectives of the HES.*

> ***Comment:*** *It is important to assure study candidates that the samples will not be used to test for other purposes (e.g., HIV testing, drug testing); examples could be provided. If DNA testing is performed, this should be explicitly declared.*

To perform the interview and the physical examinations and collect the samples needed for the survey, approximately ___hours of your time will be needed. These activities will be performed in __visits.

> ***Comment:*** *Specify the total time in hours, number of visits, and the amount of time per visit. This is an important consideration for candidates in deciding whether or not to participate, in that an excessive amount of time could discourage participation, though the time needed should not be underestimated.*

c) Information on risks

The only health risk to participating in this survey is for the taking of a blood sample, yet the risk posed is minimal.

> ***Comment:*** *Given that the risk associated with the taking of blood samples is minimal, this section can be eliminated, although in certain countries it may be necessary to make such a statement. However, if any activities that may pose a risk are added to your HES, the potential risks must be disclosed.*

Comment: *To reassure the candidate, the following sentence may be included: "All examinations are conducted by qualified and specially trained operators; they are also trained to react competently to unforeseen situations".*

Comment: *If insurance coverage is provided for the duration of the stay of the participant at the study centre, then this should be stated.*

d) Compensation

For your participation in this survey, you will receive....

Comment: *If no compensation is to be provided, then it is possible to write "You will not be paid for taking part in this study." or to eliminate this statement. If instead it is provided, the description of compensation must be clear. Payment or other forms of incentive may not be allowed in certain countries.*

e) Use of results

Would you like to receive the results of the tests performed on the blood/urine/saliva/etc. samples taken from you?

☐ Yes
☐ No

Comment: *It is assumed that in the HES no information that could be potentially upsetting to participants will be collected (e.g., HIV test results), though it should nonetheless be considered whether or not the participant could be upset by such information as, for example, obesity. If the participant's general practitioner is involved in collecting information for the study and is responsible for providing the results to the participant, then this should be specified. It may also be a good idea to specify an approximate time frame for providing the results (e.g., "The results of the tests will be provided to you in approximately 6 months".*

The data collected from you will be used by the (specify name of institution conducting the HES).

These data will also be provided to other institutes collaborating on the survey, possibly in other countries. However, all of the data provided to other institutions will remain anonymous.

Comment: *Given that data protection laws may vary by specific country, the institutes with access to data may differ. For example, in some countries it may be legal to*

provide data on individuals to general practitioners or insurance companies. It is important that the candidate be aware of who will have access to his/her data.

Your data will also be stored in a computer database at (specify name of institution conducting the HES). Only the persons conducting this survey will have access to this database. The data will also be combined with data from the HES conducted in other European countries in a centralized database.

f) Record Linkage

The data may also be combined (or "linked") with other data from different sources. For example, if you have a specific health condition and the data on your condition have been recorded in another database, then the researchers may combine these data with the data collected in the present survey, so as to study causes and relationships for certain health conditions, which is important in determining the population's health status.

Do you consent to having this done?
☐ Yes
☐ No

Comment: *The databases that are to be linked, if known, should be indicated.*

g) Confidentiality/Privacy

The data collected will be kept strictly confidential. They will be stored, analysed and handled in accordance with legislation on Data Protection and Privacy. No information that could be used to identify you will be provided to third parties. The results of this study could be published in an article, presented at a scientific meeting, or placed on a specific website, but they would not include any information that would let others know who you are.

Comment: *Describe procedures that will be followed to keep subject information and specimens secure and confidential. For example: "To ensure that the data collected from you remain confidential and that your privacy is protected, records will be kept in a separate research file that does not include names or other information that could be used by anyone but the researchers to identify you." Your country's specific laws or regulations on Data Protection and Privacy could also be provided here.*

In any case, your name or any data that could possibly be used to identify you will only be known to the specify institution conduct-

ing the HES. If you withdraw from the study, you may decide that your data and the samples will not be used / will be eliminated.

> ***Comment:*** *Whether or not data from persons withdrawing from the study must be discarded depends on the specific legislation in the given country.*

The person/entity responsible for safeguarding privacy in this study is [specify]

Permission to perform this study has been provided by [specify Data Protection Authority].

> ***Comment:*** *depending on the Data Protection Act, it might be necessary to notify or request permission from the Data Protection Authority.]_At any point during or after the study, if you are concerned about a possible violation of your privacy, you can contact [specify name and contact information of the person/entity responsible for privacy in this study.*

h) Long-term storage

Your samples may be stored at the (Specify name of organization conducting the HES) or in what is referred to as a "biobank" (that is, a long-term storage facility for biological materials) and used at a later time for other health studies). However, as mentioned above, these samples will not be tested for [specify tests NOT to be performed; see comment above].

> ***Comment:*** *In the given country, there may be legal limitations regarding the storage (including duration) and use of biological materials. Keep in mind that the term "biobank" may be intimidating for some and that terms such as "long-term storage" may be more suitable.*

i) Additional Studies/Follow-up

After this survey is complete, we may want to re-contact you for more questions and other examinations; do you agree to be re-contacted (please note that this could even be in a few years)?

☐ Yes
☐ No

j) CONTACT INFORMATION ABOUT THE PRESENT STUDY

For any questions or concerns, you can contact the researcher(s) listed below.

Comment: *It is important that the participant be provided with the possibility to speak with someone for any questions or doubts that he/she may have. Not only can this be reassuring for the study candidate or participant, but it might also increase the participation rate.*

Principal Investigator: specify name of Principal Investigator

Comment: *The person available for providing clarifications may change according to how your HES is organized.*

E-mail:
Mailing Address:
Telephone:

Consent

Participant:

I understand the information printed on this form. I understand that if I have more questions or concerns about the survey or my participation, I may contact the person(s) listed above.

Comment: *This section can be modified to emphasise the interactive aspects of informed consent, for example: "I have read and understood all of the information regarding this study, which has also been verbally explained, and all of my questions have been adequately answered."*

Signature of participant: _____
Date: _____
Name (Print legal name): _____
Participant ID: _____

Legal Representative (if applicable):

Comment: *If persons unable to fully consent for themselves are included in the HES, this section should be filled in by the person's legal guardians.*

Signature of person legally authorized to give consent

_____
Date: _____
Name (Print name): _____

Relationship to participant:
☐ Parent
☐ Spouse
☐ Son/Daughter
☐ Sibling

☐ Legal Guardian
☐ Other: _____

Reason participant is unable to sign for self:

_____

Person receiving the informed consent:

I have received the informed consent of (name of participant).

> ***Comment:*** *A sentence can be added to emphasise the interactive aspects of informed consent, for example:"I have informed the participant of the objectives and conduct of this study and of its compliance with data protection procedures, both verbally and in writing."*

Signature of person receiving informed consent:

_____
Date: _____
Name (Print legal name): _____

4.4.5 Template of an informed consent form

National Health Examination Survey (NHES)

To the survey candidate,

You have been selected to take part in a National Health Examination Survey (or "NHES"). The NHES is a study carried out for obtaining information on general health by interviewing individuals and measuring certain indicators that can be important to health, such as weight, blood pressure, and cholesterol level. This information is used to acquire knowledge on the health status of the population, which can be important in promoting and improving the health of all.

The NHES is being conducted by *(name of organization conducting the HES)* among a sample of *(number of participants)*. Your name was chosen from *(source of the person's name)*. All aspects of this study have been approved by the ethics committee of *(specify ethics committee)*.

The present form includes important information about the study and a description of what will be asked of you if you decide to participate. In order to participate, you will need to carefully read and sign this form. If any part of this form is not clear to you, please feel free to ask the person/s obtaining the informed consent. Your participation is important to us, but please be assured

that it is voluntary, that you may leave the study at any time, and that your data will be kept confidential.

During the survey, you will be asked to answer questions on your health status, lifestyle and background. Measurements of *(specify measurements)* will be taken; a *(specify type of sample)* sample will also be taken.

The samples will be analyzed *(specify analyses to be performed)*.

To perform the interview and the physical measurements and to collect the samples needed for the survey, approximately *(amount of time required)* of your time will be needed. These activities will be performed in *(number of visits)* visit(s).

The only health risk to participating in this survey is for the taking of a blood sample, yet the risk is minimal.

For your participation in this survey, you will receive *(specify any compensation or reimbursement offered)*.

Would you like to receive the results of the tests performed on the *(specify samples to be taken)* samples taken from you?

☐ Yes
☐ No

With regard to who will be provided with, or have access to, the data collected from you, these data will be used by the *(specify the institutes with access to data)*. These data will also be provided to other institutes collaborating on the study, possibly in other countries, yet no information that can be used to identify you will be provided to these institutions.

Your data will also be stored in a computer database at the *(specify location of database)*, which can only be accessed by the researchers conducting the study. The data will also be combined with data from the HES conducted in other European countries in a centralized database.

The data may also be linked to other databases containing health data, such as *(specify databases)*. Do you consent to having this done?

☐ Yes
☐ No

The information collected will be kept strictly confidential. The data will be stored, analysed and handled in accordance with legislation on *(specify type of legislation, such as "legislation on Data Privacy")*. No information that could be used to identify you will be provided to third parties.

In *(specify country)*, data confidentiality is guaranteed by *(specify legislation)*; the provisions of this law have been adopted in this study.

The person/entity responsible for safeguarding privacy in this study is *(specify person/entity responsible)*.

Permission to perform this study has been provided by *(specify data protection authority)*.

At any point during or after the study, if you are concerned about a possible violation of your privacy, you can contact *(specify name and contact information of the person/entity responsible for privacy in this study)*.

Your (type of sample) samples will be stored at *(specify where samples are stored)* and used at a later time for *(specify future uses)*.

After this survey is complete, we may want to re-contact you for more questions and other examinations; do you agree to be re-contacted (please note that this could even be in a few years)?

☐ Yes
☐ No

For any questions or concerns, you can contact the researcher(s) listed below.

Principal Investigator:
E-mail:
Mailing Address:
Telephone:

Consent

Participant:

I understand the information printed on this form. I understand that if I have more questions or concerns about the study or my participation, I may contact the person(s) listed above.

Signature of participant: _____
Date: _____
Name (Print legal name): _____
Participant ID: _____

Legal Representative (if applicable)

Signature of person legally authorized to give consent _____
Date: _____

Name (Print name): _____

Check relationship to participant:
☐ Parent
☐ Spouse
☐ Son/Daughter
☐ Sibling
☐ Legal Guardian
☐ Other: _____

Reason participant is unable to sign for himself/herself:
_____

Person receiving the informed consent:

I have received the informed consent of (name of participant).

Signature of person receiving informed consent:
_____
Date: _____
Name (Print legal name): _____

References

- Committee of Ministers (1990) Recommendation No. R (90) 3 of the Committee of ministers to member states concerning medical research on human beings. Available at https://wcd.coe.int/wcd/com.instranet.InstraServlet?command=com.instranet.CmdBlobGet&InstranetImage=569941&SecMode=1&DocId=590274&Usage=2. Accessed on 6 June 2011
- NIH(1979) The Belmont Report. Ethical Principles and Guidelines for the protection of human subjects of research. Available at http://ohsr.od.nih.gov/guidelines/belmont.html. Accessed 6 June 2011
- Oviedo (1997) Convention for the protencion of Human Rights and dignity of the human being with regard to the application of biology and medicine: Convention on Human Rights and Biomedicine. Available at http://www.portaledibioetica.it/documenti/001316/001316.htm. Accessed on 6 June 2011
- Tolonen H, Koponen P, Aromaa A, Conti S, et al (eds.) Recommendations for the Health Examination Surveys in Europe. Publications of the National Institute for Public Health, Helsinki 2008. B21/2001. Available at http://urn.fi/URN:ISBN:978-951-740-838-7. Accessed on 22 January 2013.
- WMA (2008) Declaration of Helsinki - Ethical Principles for Medical Research Involving Human Subjects. Available from http://www.wma.net/en/30publications/10policies/b3/. Accessed on 6 June 2011

5. Selecting the measurements

Johanna Mäki-Opas[1], Kari Kuulasmaa[1], Päivikki Koponen[1], Hanna Tolonen[1]

[1] National Institute for Health and Welfare (THL), Helsinki, Finland

EHES collects data through physical and clinical measurements, questionnaires, and analysis of biological samples. This chapter outlines those measurements as well as the importance and rationale of them. Measurements have been divided into core and additional measurements. Core measurements are a minimum set of measurements which should be included to every national HES. When the country has more experience, funding and national need for information, additional measurements can be added.

5.1 Criteria for selecting the measurements

The selected measurements should be based on objectives of the survey and research questions as well as analysis plan (de Bruin 1996). It is important to review each measurement carefully to make sure that they are really needed and that they provide required valid information for selected indicators. One measurement may contribute to several indicators (Tolonen 2005). As a guideline for selecting the measurement the criteria in Table 5.1 should be used. The EHES core measurements meet these criteria, but also the additional measurements should be evaluated against them.

Table 5.1 Criteria for selecting the measurements for national HES (modified from Primatesta et al. 2008, Tolonen 2005)

The criteria for selecting the measurements	Rationale and importance of the criteria
Availability of international standards	Internationally standardized measurement protocols are recommended to be used when ever possible; this ensures comparability of the results between countries and in time.
Clear interpretation of the results	The measurements need to reliable.
Practicality, easy to administrate	The measurements need to be feasible at population level.
Interesting for participants	It is recommended to have at last one measurement that motivates people to take part. This may increase the participation rate. Personal results can also be interpreted and used to estimate needs for care and preventive activities for the individual.
Acceptability to the participant	The selected measurements should not be too time consuming, causing extra burden, pain or discomfort for the participants.
Ethical acceptability	Measurements have to be ethically approved and safe for the participants, as well as accepted by health care professionals. If deviations from normal values are identified, access to care and preventive activities needs to be assured.
Costs	Costs of measurements and available funds need to be in balance. Selecting one expensive measurement may drop out several cheaper ones.
Public health importance	Selected measurements should address key public health problems.

5.2 Measurements

The core EHES physical and clinical measurements, analyses of blood samples and questionaire items collect data mainly on major chronic diseases (e.g. cardiovascular diseases and diabetes), and their risk factors (e.g. obesity, high blood pressure and high serum cholesterol) which are preventable at both individual and community level (Vartiainen 2009).

5.2.1 The core physical and clinical measurements

Physical measurements are needed because self-reported data is not sufficiently reliable to follow population level trends or to

make comparisons between populations. These selected measurements are also the ones that have been measured in previous national HESs conducted in Europe (Tolonen 2008).

The core physical measurements are:

- Height
- Weight
- Waist circumference
- Blood pressure

Body Mass Index (BMI) is a widely used indicator of obesity. It is defined as body weight divided by the square of height. The increase of obesity and overweight among the population is one of the most important public health issues in developed countries. Overweight and obesity represent a high risk factor for diseases of the circulatory system, diabetes and other chronic diseases (Malnick 2006). The evolution of the way of life and food consumption in the EU member States is characteristic by low physical activity and energetic food intake which increases the body mass index (EHIS: Background and rationale of the questions).

Waist circumference is used as an indicator of abdominal obesity. Since increasing evidence has shown that waist circumference reflects the accumulation of visceral fat better than waist-to-hip ratio, the waist circumference is the preferred measure in population studies (Seidell 2001). Waist circumference is significantly associated with the risk incident of CVD events and type 2 diabetes (de Koning 2007).

Measuring blood pressure gives a prevalence of actual and potential hypertension. Single-occasion blood pressure measurement has been shown to be a strong indicator of coronary and cerebrovascular risk (MacMahon 1990). However, the diagnosis of hypertension requires follow-up and observed high blood pressure on several occasions.

5.2.2 The core biological samples

The EHES surveys include the collection and analysis of biological samples. The core blood samples are:

- Non-fasting blood samples
  - Total cholesterol
  - HDL cholesterol
- Fasting blood sample (8-14 hours)
  - Glucose

High serum total and HDL cholesterol are major risk factors of cardiovascular diseases. Increased glucose level in blood may indicate insulin deficiency or insulin resistance which indicates risk for diabetes. (Emerging Risk Factors Collaboration 2009.)

Because of potential difficulties in requiring fasting from all participants, the glucose may cover only a sub-sample of the survey. It should be noticed that the fasting should not last more than 14 hours.

In consideration for the future:

The classification and diagnosis of type 2 diabetes has relied on the measurement of fasting plasma glucose concentrations or oral glucose tolerance tests. Interpretation of non-fasting glucose values is difficult, if not impossible for classification of diabetes in large health surveys. On the other hand, obtaining blood samples from adequately fasted participants is often impractical in health surveys.

Glycated haemoglobin, HbA1c, reflects the time-averaged blood glucose concentration during the previous 2-3 months. There is a close relationship between HbA1c and glucose. Therefore, it has been proposed to substitute plasma glucose with HbA1c not only for following the effectiveness of diabetes treatment but also for classification of type 2 diabetes (The International Expert Committee 2009). Its superiority over plasma glucose, especially for health surveys, lies in that its measurement does not require a fasting blood sample.

In the past, measurement of HbA1c has been hampered by the measurements not having been standardized to a sufficient level. Recently, however, a consensus statement on the worldwide standardization of the HbA1c measurement has been published (Hanas 2010). It is foreseen that in the very near future HbA1c could replace plasma glucose as a core measurement. Therefore, measurement of HbA1c is strongly recommended already now.

5.2.3 The EHES core questions

The EHES core questionnaire items which are based on the EHIS questionnaire should be administered as recommended in EHIS. The following documents should be reviewed while preparing the national version of the EHES questionnaire and when training the fieldwork personnel:

- EHIS Background and rationale of the questions.
- EHIS Conceptual translation cards and guidelines.

Guidelines and quality criteria for EHIS questionnaire administration have also been documented in Davidsson et al 2009. The standard EHIS questions and the above mentioned documents are available on the CIRCA website of the Commission: http://circa.europa.eu/Public/irc/dsis/health/library?l=/methodologiessandsdatasc/healthsinterviewssurvey/2007-2008_methodology&vm=detailed&sb=Title

The EHES core questions are mostly questions that are necessary for the reporting and interpretation of the data from the physical measurements and biological samples. Whenever possible, the EHIS questions should be used. In the EHES core questionnaire some questions have been slightly modified from EHIS and there are also a few other than EHIS questions. Note, however, that there are plans to revise parts of the EHIS questionnaire based on the experiences from the first round. This revision is expected to be completed by 2014. Most likely the EHIS questions selected as EHES core questions will not be revised, but for example previous EHIS questions on alcohol and physical activity will most likely be revised. The EHES RC is following the revision of the EHIS questionnaire and will update the EHES Manual accordingly.

The EHES core questionnaire includes questions on:

- Household size
- Sex
- Age
- Marital status
- Socioeconomic status
- Height and weight
- General health
- Chronic diseases
- Use of medication
- Smoking

All the questions can be found from Part B, Section 5.7.

Age and sex enable reporting of the HES results by sex and age group and the age-adjustment of the results for comparison between populations. Education, occupation and household income are needed for the estimation of socioeconomic differences in the population. Some countries may have the possibility to obtain these demographic and socio-economic data through linkage with registry data so there is no need to ask them at all. However, if such data linkage is not possible in the country or if

the coverage of the registry data is incomplete, it is important to include these questions into the survey.

Even though height and weight will be measured, they should also be asked. Sometimes the questionnaire is the only data source. This enables for example the analysis of non-participant's BMI in a case that participant fills in only the questionnaire but does not take part to the physical examinations. Asking height and weight also enables to estimate the differences of measured and self-reported height and weight between the countries, and by sex, age ans socio-economic status.

The three questions on general health form the Minimum European Health Module (MEHM), which is expected to be included in all European social and health surveys, in order to link results among surveys according to these standard health characteristics of the population. The structural indicator Healthy Life Years is calculated on the basis of questions of MEHM. More generally the three questions are used for the calculation of the prevalence of perceived health, self-reported longstanding illnesses or health problems and long-term activity limitations. (EHIS: Background and rationale of the questions.)

The questions on chronic diseases measure the main public health concerns, which are also a major reason for using the health care services. Measuring chronic morbidity is useful for overall evaluation of health and health status. It is also useful for the study of health care systems in terms of evaluation, policy formulation and assessment of need for health care. (EHIS: Background and rationale of the questions.) The answers to the questions on specific diseases are needed from the same person as the physical measurements are taken. For example the question on use of medication on hypertension is commonly combined with the results of blood pressure measurement to see how well hypertension is treated and controlled in the population. Also the cholesterol and glucose levels in blood are often combined with the related questions.

Smoking is an important factor for lung diseases and cancer, other cancers and diseases of the circulatory system. Lung, trachea and larynx cancer is the type of cancer with the higher standardized death rate among men in EU. In addition, important policy activities are developed at EU level in order to limit tobacco consumption and many of the Member States are in the process to forbidding smoking in working places and public areas. For these reasons it is a major determinant of health outcomes. (EHIS: Background and rationale of the questions)

5.3 Additional measurements

In addition to the core measurements, countries may include other physical measurements and questions into the national HES. When choosing the additional measurements, the criteria shown in Table 5.1 should be kept in mind. Countries with little experience from earlier HESs are recommended to keep the number of additional measurements low to allow adequate planning and preparation for all measurements and fieldwork procedures. Experienced countries may include a wide range of additional measurements to the survey if they are confident that they can manage the survey process and they have sufficient funding. Additional measurements can be added to the survey as modules that are relevant for example to specific subgroups of the population, such as certain age groups, ethnic groups or other sub-populations of regional/local interest.

When selecting additional measurements, the countries should consider their implications to the survey administration, the time taken for training, by administrating the questionnaire and carrying out the physical measurements as well as the costs and the periodicity of the survey. If the survey will be repeated frequently, different additional modules can be considered for each round of data collection. When the survey will be carried out less frequently, it may be feasible to build a more comprehensive survey covering several health topics. There is a commonly used target that the physical/clinical measurements should be limited to take one hour. Some evidence suggests that longer surveys are less acceptable to respondents. But there are also experiences (e.g. the Health 2000 survey in Finland) where a more comprehensive survey with long examinations has been attractive to the participants as it gives more information on their own health.

5.3.1 Additional physical and clinical measurements

For many of the additional measurements, no EHES recommendation is currently available. Procedures for such measurements will be added to Part B of this manual after there is an agreement on the standards. Meanwhile, the countries planning to include measurement for which there is no standard available in the EHES manual should inform the EHES RC and other countries about their plan, so that the countries interested in the same measurement can collaborate on preparing their procedures. In this way, also unintentional use of different procedures in countries can be avoided.

Potential additional physical measurement are for example:

- Hip circumference
- Lung function test
- Physical function tests
- Vision and hearing tests
- Cognitive function tests
- ECG
- Bone density
- Dental examination

5.3.2 Additional biological samples

It is recommended that countries collect more blood samples than are needed for the core analyses. Once suitable blood samples have been collected in the survey and stored properly, they can be used for various measurements in the future, if ethical approval and participants' consents for the storage and future analysis are obtained.

From additional blood samples, following issues can be considered:

- Many countries may want to assess serum triglycerides, which are an indicator of cardiovascular risk. Furthermore, triglycerides, together with total and HDL cholesterol can be used to estimate LDL cholesterol, a major risk factor for coronary heart disease. The measurement of triglycerides is complicated by the fact that fasting will be required before blood sampling.
- The measurement of apolipoproteins A1 and B are under consideration for core measurements. They are correlated with HDL cholesterol and LDL cholesterol respectively, and there are indications that they predict cardiovascular diseases better that HDL and LDL cholesterol. These measurements are easier to standardize than HDL cholesterol and much easier than LDL cholesterol. Furthermore, fasting is not needed.
- Countries may also want to collect samples of whole blood for DNA. This will increase the future research potential of the survey, as today the availability of large population studies with DNA is a major limitation of genetic research. The DNA collection will imply additional ethical requirements for the survey.
- Many other measurements, such as nutritional biomarkers and possible new emerging measurements can be done on the stored samples

- Glycated haemoglobin (HbA1c) may in the future replace plasma glucose since it is a sensitive biomarker of excessive glucose and is not affected by meals (see Part A, Section 5.2.2).

There has also been discussions with the European Centre for Disease Control (EDCD) about the measurement of some infectious diseases, and the with the European Commission, DG Environment about the inclusion of biomonitoring markers.

5.3.3 The additional questions

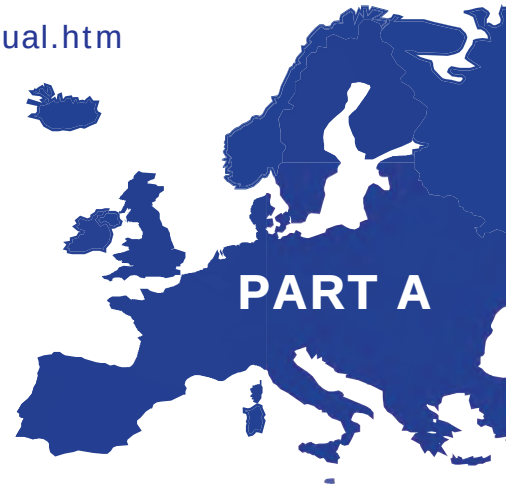
The questionnaire can also include additional questions. These are the questions that are related to the additional physical measurements or biological samples or otherwise collect information needed to meet survey aims and purposes. EHIS questions are recommended to be used when suitable questions are available. Potential topics for additional questions are:

- Physical activity
- Alcohol consumption
- Use of health care services
- Social support
- Fruit and vegetable consumption

References

- European Health Interview Survey (EHIS). 1st round 2007-2088. Background and rationale of the questions. European Commission. Available at https://circabc.europa.eu/sd/d/150a71c0-7a17-407d-8d20-2fdc284b0bf7/EHIS%20-%201st%20round%202007-2008%20-%20Background%20and%20rationale%20of%20the%20questions.pdf. Accessed on 22 January 2013
- European Health Interview Survey (EHIS). 1st round 2007-2008. Conceptual translation cards and guidelines. European Commission. Available at https://circabc.europa.eu/sd/d/671d5958-7928-471c-a7b9-0194c34b-9cb8/EHIS%20conceptual%20card%20and%20guidelines.pdf. Accessed on 22 January 2013
- Emerging Risk Factors Collaboration, Di Angelantonio E, Sarwar N, Perry P, Kaptoge S, Ray KK, Thompson A, Wood AM, Lewington S, Sattar N, Packard CJ, Collins R, Thompson SG, Danesh J. Major lipids, apolipoproteins, and risk of vascular disease. JAMA. 2009 Nov 11;302(18):1993-2000.
- Davidsson G, Demarest S, Tolonen H, De Vitiis C, Primatesta P, Agafitei L. European Health Interview survey Task Force II. Report on guidelines and quality criteria for population health survey design and methods. Eurostat, 2009. Available at https://circabc.europa.eu/sd/d/0d4d38ff-fb62-4bf6-ba53-ed64063fc193/Task%20Force%20II%20report%20on%20survey%20design%20%28November%202009%29.pdf. Accessed on 22 January 2013

- de Koning L, Merchant AT, Pogue J, Anand SS. Waist circumference and waist-to-hip ratio as predictors of cardiovascular events: meta-regression analysis of prospective studies. European Health Journal. 2007:28(7):850-56.
- Hanas R, John G, Consensus Committee. 2010 Consensus statement on the worldwide standardization of the hemoglobin A1C measurement. Diabetes Care 2010;33:1903-4.
- MacMahon S, Peto R, Cutler J, Collins R, Sorlie P, Neaton J, Abbott R, Godwin J, Dyeer A, Stamler J. Blood pressure, stroke, and coronary heart disease. Part 1. Prolonged differences in blood pressure: prospective observational studies corrected for regression dilution bias. Lancet. 1990:335:746-74
- Malnick SD, Knobler H. The medical complications of obesity. QJM. 2006;99(9):565-79.
- Primatesta P, Chaudhury M. Core Module and additional topics. In Tolonen H. et al. (eds.) Recommendations for the Health Examination Surveys in Europe. Kansanterveyslaitoksen julkaisuja 2008:B21/2008. Available at http://urn.fi/URN:ISBN:978-951-740-838-7. Accessed on 22 January 2013
- Seidell JC, Després J-P, Bouchard C. Waist and hip circumference have independent and opposite effects on cardiovascular disease risk factors: The Quebec Family Study. ACJN 2001;74:315-21.
- The International Expert Committee. International expert committee report on the role of the A1C assay in the diagnosis of diabetes. Diabetes Care 2009;32:1327-34.
- Tolonen H. et al. (eds.) Review of Health Examination Surveys in Europe. Publications of the National Public Health Institute 2008:B18/2008. Available at http://urn.fi/URN:ISBN:978-951-740-843-1. Accessed on 22 January 2013
- Tolonen H. Towards the high quality of population health surveys. Standardization and quality control. Publications of the National Public Health Institute 2005:A27/2005. Available at http://urn.fi/URN:ISBN:951-740-546-4. Accessed on 22 January 2013
- Vartiainen E, Laatikainen T, Peltonen M, Juolevi A, Männistö S, Sundvall J, Jousilahti P, Salomaa V, Valsta L, Puska P. Thirty-five-year trends in cardiovascular risk factors in Finland. Int J Epid 2009: 1-15.



6. Timing of the fieldwork and order of measurements

Päivikki Koponen[1], Hanna Tolonen[1], Johan Heldal[2]

[1] National Institute for Health and Welfare (THL), Helsinki, Finland
[2] Statistics Norway (SSB), Oslo, Norway

This chapter describes general issues related to timing of the fieldwork and order of measurements that need to be taken into account when planning the survey. For example seasonal and diurnal variation in symptoms, morbidity, body functions and health behavior need to be taken into account. Timing of the examinations will also affect participation rates (see Part A, Chapter 13). General principles for the order of measurements need to be considered when estimating the time needed to carry out the fieldwork. These issues will all have an effect on personnel resources and other survey costs (see Part A, Chapter 16). Details for timing specific measurements will be given in Part B of the EHES Manual.

6.1 Periodicity

The recommenation is to repeat the national HES with the EHES core measurements about every five years, while some additional measurements may be repeated less frequently (e.g. every 10 years). More frequent surveys do usually not reveal interpretable changes for most of the measurements. They can be considered on ad hoc basis if there is a need to closely follow trends related to potential effects of specific health promotion activities.

An alternative is to build a system of continuous data collection. In such surveys data from different years can be aggregated to provide precise estimates for indicators that require larger samples. Continuous data collection also allows keeping permanent fieldwork staff, which may decrease staff recruitment and training costs. Examples of HESs with continuous data collection are the Health Syrvey for England and the National Health and Nutrition Examination Survey of USA. In a continuous survey the permenent core survey content can be kept brief while varying

additional measurements can be introduced yearly or every second year. However, a more comprehensive survey content will allow more possibilities to study how different health topics are related to each other at individual level. The feasibility of a short or a longer, more comprehensive examination may vary in different countries.

6.2 Length and time of year for the fieldwork

High seasonal variation has been identified in several health determinants as well as in biological measures. For example seasonality impacts physical activity patterns (Merchant et al 2007, Pivarnik et al 2003), food consumption (Fowke et al 2004) as well as quality of life (Jia et al 2009). In countries and regions with cold winters, leisure-time physical activity is more likely during summer and spring than during winter and fall (Merchant et al 2007, Pivarnik et al 2003). People may be more likely to eat fruits (Fowke et al 2004) and to report better quality of life (Jia et al 2009) during summer than during winter. Significant summer-winter differences have been identified in blood pressure, fasting plasma glucose levels, blood lipid levels, body mass index and waist circumference, with lowest risk factor levels in the summer (Chen et al 2006, Visscher & Seidell 2004).

Therefore, it is important for the estimation of trends that repeated surveys are carried out at the same time of the year. For best international comparability, the surveys in each country should cover evenly all seasons, which means that the fieldwork should last at least a year. Seasonal variation may differ between countries and regions, depending on the climate. If the survey covers only a part of the year, it is essential to evaluate the potential effect of weather and national/regional climate and other issues related to fieldwork timing (e.g. common flu epidemics) to measurement results.

A short survey duration usually needs a relatively large temporary staff, whereas long or yearly repeated surveys allow a more stable employment of the core staff. When the survey lasts more than a few months, particular attention needs to be paid to regular quality control, re-testing and re-training of the fieldwork staff.

It will be desirable to compare the survey results between men and women, between age groups and possibly between regions of the country. Such comparisons will be possible only if the days of examination of all such population subgroups are distributed evenly over the whole survey period. For sex and age this usually becomes taken care of if all sex and age groups are examined

simultaneously. Regional comparability needs to be taken into account when scheduling visits to the different examination sites. A safe option is to order the examination sites of the country in random. Alternatively, a systematic ordering, where each region is visited evenly in all seasons can be used. The latter alternative can usually lead to lower travelling costs for the fieldwork personnel and less time needed for moving from one examination site to another.

6.3 Weekdays and time of day

To allow easy access for participants and to minimize the effect of timing to measurement results, morning, day and evening appointments should be available, as well as several weekdays (see Part A, Chapter 13). Also weekends should be used, if it is feasible to schedule these from the point of view of cost and availability of premises and staff, and if they are preferred by the participants. Measurements that require overnight fasting may be organized only in the mornings and may therefore be feasible only for a subsample (see Part A, Chapter 10 on Blood sample collection). Experiences from previous surveys in several countries have shown that the working age population prefers early mornings (before working hours) or late afternoons and early evenings (after work) for their examinations during the week. Fridays also seem to be less often preferred than other days of the week. In some surveys additional options for an appointment during weekends (Saturdays) have been used to raise willingness to participate.

Many of the HES measurements, such as blood pressure and some blood analyte concentrations, have diurnal variation. For the lipids it is difficult to dissociate changes in their concentration from the effects of a meal. THis is best adjusted for by spreading the measurements throughout the day, and in any case the time of the day needs to be recorded (see Part B of the EHES Manual).

6.4 Order of measurements

The order of measurements has often constraints because of logistical requirements, such as composition of the survey staff and fieldwork teams, subject flow, costs and examination duration. However the following requirements need to be taken into account to ensure valid measurements and comparability between surveys.

6.4.1 Clinical measurements

The order of measurements should be determined as much as possible by (adapted from Tolonen et al 2002, Tolonen et al 2008):

1. Importance of the measurement; most important measurements should be made early in the session, in case the participant is unable to follow the full examination protocol (time constraints, limitations in functional capacity etc.). The EHES core measurements should be conducted first, before additional measurements.
2. Sensitivity of questions; uncontroversial questionnaires should occur early in the interview to allow participants to become relaxed and comfortable with the procedures.
3. Stressfulness of procedure;blood pressure measurement should precede venapuncture and other potentially (mentally or physically) stressful tests/interviews.
4. Order in previous surveys; unless good reasons exist for change, it is suggested to maintain the former order of measurements.
5. Other effects on measurement results; blood pressure and blood samples should be taken before physical fitness tests or tests of physical function.

For similarity of the order of the EHES core measurements between countries the following order is recommended: first blood pressure, then anthropometric measurements, third blood samples, and all additional measurements after these.

6.4.2 Questionnaires and interviews

The selection of self-administered questionnaires and interviews is described in chapter 8. The decisions on when the questionnaires or interviews will be administered should be based on the following:

1. Before the examination (mailed with the invitation to examinations, administered at the examination site before the clinical measurements or interviews before ther examination)

   When the self-administered questionnaires/interviews are completed before the examination, the responses are not affected by the examination. Self administered questionnaires should be checked during the examination. It is recommended that the questionnaire data collection before examination is restricted to most important key

questions and not to include questions that can be considered as too sensitive. The core EHES questions (presented in Part A, Chapter 5 of the EHES Manual) are recommended to be administered before the examinations.

The time lag between the questionnaire data collection and the examination may cause problems with linking the measurement and questionnaire data (e.g. the effect of acute infections on clinical measurements). The questionnaires have to be short and easy to fill in, not to discourage participating in the examinations. However, the interviewers can be trained to motivate participation to examinations and to book time for the examination that best suits the participant.

2. During the examination (between measurements)

   When the self-administered questionnaires/interviews are completed during the examinations, the responses can be affected by the examination (learning that the participant has e.g. high blood pressure, knowing that smoking behaviour can be detected from blood/saliva cotinine etc.). It is recommended to ask questions on acute symptoms and current medication during the examination as these may affect the clinical measurements.

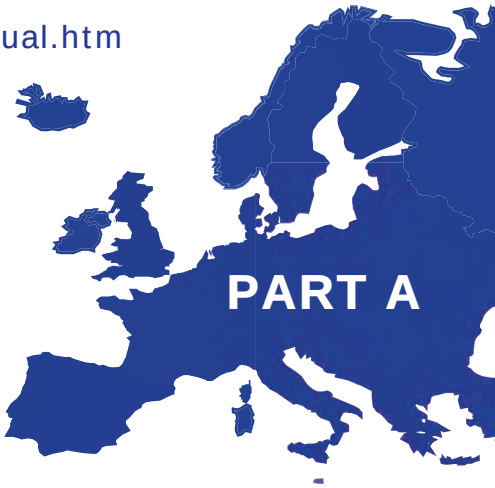
3. After the examination

   When the self-administered questionnaires/interviews are completed after the examinations the non-participation can be highest. It is recommended to collect after the examination information on sensitive questions and questions that are less important from the point of view of the top key aims of the survey, but which potentially raise new issues for research and health policy/health care development. The use of sensitive questions after the examinations should be explained to the participants during the examination to make sure that their final impression of the survey participation is positive.

To avoid respondent burden, the questionnaires may be split into different parts administered before, during and after examinations. In this case, the selection of questions to different phases of data collection should based on the principles described above: questions that are effected by clinical mesurements, EHES core questions and non-sensitive questions before measurements, questions on acute symptoms during the measurements, and less important and sensitive questions after the examination.

References

- Chen S-H, Jen I-A, Chuang S-Y, Lin K-C, Chou P. Community-based study on summer-winter differences of component of metabolic syndrome in Kinmen, Taiwan. Preventive Medicine 2006;43:129-135.
- Fowke JH, Schlundt D, Gong Y, Jin F, Schu XO, Wen W, Liu DK, Gau YT, Zheng W. Impact of season of frood frequency questionnaire administration on dietary reporting. Ann Epidemiol 2004 14(10):778-785.
- Jia H, Lubetkin EI. Time trends and seasonal patterns of health-related quality of life among U.S. adults. Public Health Rep 2009;124(5):692-701.
- Merchant AT, Dehghan M, Aktar-Danes N. Seasonal variation in leasure-time physical activity among Canadians. Can J Public Health 2007:98(3):203-208.
- Pivarnik JM, Reeves MJ, Rafferty AP. Seasonal variation in adult leasue-time physical activity. Med Sci Sports Exerc 2003 35(6):1004-1008.
- Tolonen H, Kuulasmaa K, Laatikainen T, Wolf H, for the European Health Risk Monitoring Project. Recommendations for indicators, international collaboration, protocol and manual of operations for chronic disease risk factor surveys. 2002 [cited 13 2008]; Available from: http://www.thl.fi/publications/ehrm/product2/title.htm. Accessed on 22 January 2013
- Tolonen H, Koponen P, Aromaa A, et al. (Eds.) Recommendations for the Health Examination Surveys in Europe. B21/2008, Publications of the National Public Health Institute, Helsinki 2008. Also available from http://urn.fi/URN:ISBN:978-951-740-838-7. Accessed on 22 January 2013
- Visscher TLS, Seidell JC. Time trends (1993-1997) and seasonal variation in body mass index and waist circumference in the Netherlands. International Journal of obesity 2004;28:1309-1316.



7. Selecting the examination site

Johanna Mäki-Opas[1], Hanna Tolonen[1], Päivikki Koponen[1], Kari Kuulasmaa[1]

[1] National Institute for Health and Welfare (THL), Helsinki, Finland

This chapter outlines the possible examination sites and their advantages as well as disadvantages. The selection of the survey site has to be based on general requirements, national practices and cultural factors. Examination site may have an impact on the quality of the data and participation rate.

Potential examination sites are:

- Participant's home;
- Temporary examination site outside health care organizations, for example school premises;
- Examination site within existing health care premises, such as a health centre or GP office (with the regular staff of with specific survey teams allowed to use the premises);
- Mobile examination site, for instance a bus equipped for examination.

7.1 Requirements for examination site

The HES can be conducted either at participant's home or in any other place arranged and equipped for the survey purpose. When physical examinations take place somewhere else than in the participant's home, the following issues should be considered:

- Participants should have easy access to the examination site. The maximum distance to the examination site varies between countries and even between areas in countries. In urban areas, people are not necessary

willing to travel to another side of the city, but in rural areas longer distances can be considered acceptable.

- The availability of public or in another way organized transportation to the examination needs to be assured.
- Access of participants with limited functional ability;
- Handling and storage of the blood samples;
- Requirements for the EHES core measurements;
  - Privacy;
  - Quietness;
  - Comfortable room temperature;
- Requirements for the additional measurements, e.g.
  - Enough room for functional ability tests;
  - Sound proof environments for audiograms.

The only way to be sure that the examination site is suitable for carrying out physical measurements is to visit the place before selecting it. This requires adequate time and personnel resorces during survey preparation.

7.2 Requirements for home visit

Home is a private place and when the examinations takes place at participants home some special issues should be taken into account.

These can be for example:

- Acceptability of home visits among the population, e.g. are people used to home visits by the primary health care personnel;
- Only measurements that do not need equipment that is heavy or otherwise difficult to transport, or have other special requirements for the environment, can be conducted;
- Special attention to the safety of the fieldwork staff should be paid;
- It might be difficult to guarantee the privacy during the examinations/interview;
- Standardization and calibration of the equipment and following the measurement protocol may be challenging;
- Restrictions for handling and storage of the blood samples may compromise the quality of the samples;

- Challenges for data transfer and data confidentiality.

When other examination sites are used, it should be considered if it is feasible and useful to offer home visits to those who are not willing or able (e.g. due to limited fuctional ability) to come to the examination site.

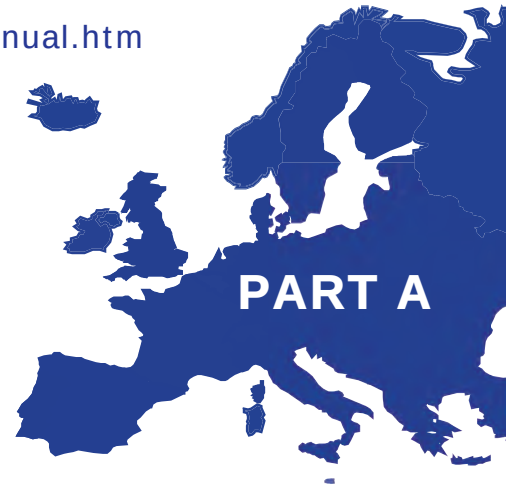
7.3 Advantages and disadvantages of examination sites

In all examination sites has their advantages and disadvantages. When selecting the site issues In Table 7.1 some of these has been reviewed.

Table 7.1 Advantages and disadvantages of the examination sites

	Participant's home	**Temporary examination site**	**Examination site within existing health care premises**	**Mobile examination site**
Access by participants	Easy access	Requires effort	Requires effort	May be easy if mobile examination site can be taken close to the participants
Cost for participants	None	Travel costs	Travel costs	Some travel costs
Environment				
Atmosphere	Relaxed	Some tension	Some or a lot of tension	Some tension
Privacy	Limited privacy if other family members at home	Can be controlled	Can be controlled	Can be controlled
Temperature	Cannot be controlled by the survey team	Can usually be controlled	Can usually be controlled	Can be controlled
Quietness	Cannot be controlled by the survey team	Can usually be controlled	Can usually be controlled	Can be controlled
Safety of the field work staff	Limited and cannot be controlled	Can be controlled	Can be controlled	Can be controlled
Travel cost of field work staff	Expensive	Some	Some	Some
Traveling for field work staff	Lot of traveling	Some	Some	Some

	Participant's home	Temporary examination site	Examination site within existing health care premises	Mobile examination site
Restriction to measurements	Only measurements for which devices can be transported easily and which do not have specific environmental requirements	Generally none	Generally none	Generally none, sometimes a lack of facilities for specific measurements may come up (e.g. limited space)
Calibration/standardization of the measurements	Difficult	Can be done	Can be done (but if used equipment of health care centre, standardization and calibration may be difficult)	Can be done
Acceptability	Some people are not willing to let the survey team into their home	Generally accepted	In some countries may not be highly valued among some people	Generally accepted
Time and cost for setting up an examination site	Minimal	Time consuming	Takes some time (depends on used equipment, if equipment from the health centre are used, carefull calobration before fieldwork is needed, otherwise like temporary examination site)	Time consuming and costly
Cost of the maintenance of the examination site	None	Some costs	Some costs (depends on agreements with the local health care administration)	Costly



8. Questionnaire design and administration

Päivikki Koponen[1], Johanna Mäki-Opas[1], Hanna Tolonen[1]

[1] National Institute for Health and Welfare (THL), Helsinki, Finland

This chapter considers issues that need to be taken into account when preparing the national HES questionnaire and planning the questionnaire administration. The questionnaire design has an impact on participation rate and validity of the obtained data. The questionnaire administration mode has an effect on survey budget but it may also affect on participation rate, item non-response and validity of the answers.

8.1 Questionnaire design

Every national HES should also include a questionnaire to collect information which is needed e.g. to interpret the measurement results. The questionnaire design affects the participation rate as it gives the participant an impression of how easy, convenient and time-consuming it is to take part in the survey. It also affects the reliability and accuracy of the information obtained by the questions. Therefore, enough time and resources for planning the questions and preparing the questionnaire should be allocated (Tolonen 2005). Often, after the pilot survey, the questionnaire is at least slightly modified and improved based on the experiences obtained during the pilot.

Language, wording of the questions, selection of the response alternatives, formulation of sensitive questions, recall bias, order of questions, jump rules and the length of the questionnaire are the main elements of questionnaire design.

8.1.1 Language and wording

The proper wording of the questions is essential; the questions should be simple and straightforward. This ensures that respondents understand the questions correctly. When formulating the

questions, effort must be devoted to avoiding ambiguity in the wording. Professional or highly technical terms, slang, abbreviations or words which may be considered as insulting, should be avoided. In each question only one issue should be addressed. All questions should be available in the native language of the respondent. (Rea 2005.) In many European countries, several language versions should be considered. The translations should be prepared with a careful validation process. The EHES core questionnaire (see Part B, Chapter 5) consists many questions from the European Health Interview Survey (EHIS). The EHIS questions have been usually translated to the national language(s) by the national Statistics Office and at least for some sets of questions, cognitive validation has been done. These should be used when ever possible.

8.1.2 Recall bias

When formulating the questions it is good to remember that people tend to forget events. It is usually easier to remember things that happened recently than for example a year ago. When the recall period is longer the accuracy is often worse. Recall can become a source of bias (de Bruin 1996). Therefore recall of the events should be assisted by adding aids to the questionnaire and by ordering of the questions. For example holidays and national festivals can be used to refer to a certain time period, or the respondents can use a calendar. (Tanur 2004)

8.1.3 Order of the questions

The order of the questions in the questionnaire is also important. A poorly organized questionnaire may confuse the respondent, bias the responses, has an effect on response rate, as well as on the willingness to answer sensitive questions. (Rea et al. 2005, Tanur 2004, Biemer et al. 1991.) The questionnaire should start with the easy questions. When more difficult questions are placed at the end of the questionnaire and if the respondent stops answering, at least some data for earlier questions have been collected. During the interview asking the easy questions first may help to build trust between the interviewer and the respondent so the respondent may be more willing to answer more difficult questions in the end. All the questions should be grouped by the topic. This makes answering easier. Also filtering questions should be used. This reduces the respondents burden. Use of jump-rules in the questionnaire avoids respondents answering irrelevant questions. Also the order of the response alternatives can greatly influence the results (Biemer et al. 1991).

Each national HES should include at least the EHES core questions (see Part B, Chapter 5). If the national questionnaire in-

cludes several additional items, it is recommended to keep the EHES core questions early in the questionnaire to make sure that the participants give valid responses to all of them. However, the structure of the whole questionnaire needs to be taken into account.

8.1.4 Length of the questionnaire

The length of the questionnaire affects the response rate as well as reliability of the data. A short questionnaire increases the response rate but may lack important questions for the indicators. With the longer questionnaire the respondents often get careless towards the end and the reliability or the answers suffers (Biemer et al. 1991). The ideal length for filling in a self-administered questionnaire is 15 minutes and for the face-to-face interview 30 minutes. In practice, questionnaires which are designed for these lengths, may require about 15 minutes longer for most respondents. (Rea et al. 1997)

8.1.5 Layout of the questionnaire

Issues to be considered when using paper questionnaires include e.g.:

- font size and font layout feasible for persons with problems in visual capacity (especially for the elderly);
- number of questions in each page;
- number of pages needed;
- if some questions can be skipped, clear advice to jump to next questions;
- using colors and pictures.

Issues to be considered when using web questionnaires include e.g.:

- font size and font layout;
- number of questions visible at each screen;
- jump rules to be followed (controlled by the program, not the respondent);
- using colors and pictures;
- downloading the file in respondent's own computer not too time consuming.

Issues to be considered when using computer aided questionnaire administration (CAPI or CATI or CASI) include e.g.:

- possibilities for layout when using different programs;
- visibility of instructions for the interviewer;

- jump rules to be followed (controlled by the program, not the interviewer).

8.2 Questionnaire administration

Survey questionnaires can be filled in either by the respondent (i.e. self-administration) or by an interviewer. Both self-administration and interview have several alternatives how then can be organized and all of them have advantages and disadvantages, see Table 8.1. (Franklin & Walker 2003, Czaja & Blair 2005, Tolonen 2005). The questionnaire administration mode may effect participation rate and the accuracy and reliability of the responses. It is recommended that the core EHES questions are collected through face-to-face interview. Other administration modes can be considered for additional questions and when the person does not respond to the first contact attempt or refuses the examination. Use of mixed-mode data collection and several phases of questionnaire administration may avoid participant's burden and selection bias.

Table 8.1 Comparison of different questionnaire administration methods (adapted from Franklin & Walker 2003, Czaja & Balir 2005, Tolonen 2005)

Aspect	Self-administered paper questionnaire	Self-administered electronic questionnaire	Interview Telephone	Interview Face-to-face
Cost	Low	Low*	Medium	High
Length of questionnaire	Medium	Short	Medium	Long
Complexity of questionnaire	Must be simple	May be complex	May be complex	May be complex
Control of question order	Poor	Fair	Very good	Very good
Use of visual aids	Good	Very good	Not possible	Very good
Sensitive topics	Good	Poor	Fair	Fair
Control of response situation	Poor	Poor	Fair	Good
Language version	Poor	Very good	Good	Very good
Socially desirable answers	No	No	Yes	Yes
Item non-response	High	High	Medium	Low
Response rate	Low	Low	Medium	High
Needed literacy level	High	High	Low	Low
Verifying the respondents identity	Low	Low	Medium	High

* Setting up the electronic questionnaire may be costly but after that costs of data collection are low

8.2.1 Self-administration

Self-administration of the questionnaire is cost effective but assumes that participants are not visually impaired and have a good literacy level. The self-administered questionnaire should be relatively short and all questions need to be completely self-explanatory; format and question wording must be simple, without complex skip patterns. Self-administration eliminates the interviewer effect but may result in missing data as a result of uncertainty about the question. The self-administered questionnaire can be either a paper form or an electronic version. Paper forms also require separate data entry. The electronic questionnaire can be at the internet or on stand alone software on computer. The electronic questionnaire can be more complex (with skip patterns) than the paper format. The computer program can have built-in checks for responses (e.g. upper or lower limits for response categories).

A self-administered questionnaire can be delivered to the participant before the clinical examination. In this case, the questionnaire is filled in at home before the examination and checked by field work staff at the examination site. The possibilities to motivate participation to examinations are poor if questionnaires are mailed before examinations. It is also known that response rates tend to be low when self-administration is used. Alternatively, the questionnaire can be given to the participant when he/she arrives to the examination site and he/she fills in the questionnaire at the examination site. In this case, the participant can ask help from the field work staff if he/she has any problems with the questionnaire. Also in this case, the completed questionnaire should be checked by the field work staff for completeness before the participant leaves the site.

Self-administration provides more privacy for the respondent and is particularly suitable for sensitive questions (e.g drug use, sexual behavior, income). The questionnaire can contain printed reference materials and pictures (visual aids), e.g. pictures can be useful for showing portions in questions on alcohol intake and food consumption/diet. But when web based questionnaires are used for self-administration sensitive questions may become problematic, because the respondents do not always trust in data security.

Web-based questionnaires can be considered as one form of self-administration which may be easy for certain groups in the population. However, they may result in participant selection because it requires easy access to Internet. Therefore it should in most European countries be used as an alternative to the traditional paper forms, rather than as an exclusive mode of data collection.

The use of web-based questionnaires also requires extra efforts to ensure privacy.

8.2.2 Interviews

Interviews are time consuming and carry additional personnel costs, but they eliminate the issues of literacy level and visual impairment and they provide an opportunity for clarifying the questions if needed. These clarifications have to be described in the manual and training for the interviewers and/or in the questionnaires to avoid biased responses and to ensure standardization of questionnaire admission. Interviews can be conducted either by telephone or face-to-face. In both modes the questionnaires can be quite long and complex, when skip patterns and jump rules are used and followed by the interviewer and/or controlled by the computer program and not by respondent. This reduces the burden of the respondent. Automatic built-in checks for responses and data entry by the interviewer may reduce errors in computer assisted interviews.

Face-to-face interviews are usually the most expensive mode for questionnaire administration. However, it has many advantages: the interviewer has a possibility to check the personal records (e.g. medication), personal contact may increase the response rate and the use of printed reference materials (visual aids) is possible. Telephone interviews are less expensive but provide no control over the environment in which the interview is conducted. Question wording needs to be simple and it requires good hearing capacity from the respondent. There is a risk that interviewers introduce bias by not asking the questions verbatim, modifying the questions or by incorrect prompting. This risk can be reduced, but not fully eliminated, by proper training and quality control. Sensitive questions may be problematic in interviews, because the respondent may reply according to what is socially most acceptable.

8.2.3 Mixed method

When there are several additional topics and many questionnaire items, a mixed method should be considered: e.g. a short self administered questionnaire mailed before examinations, interview during examinations, and another questionnaire given to be filled in later at home. Several modes of data collection can also be used for the same questionnaire to obtain better response rates, e.g. self-administered questionnaires are mailed as a paper version to all subjects, but a possibility to fill this in as a web-based questionnaire is given in the cover letter. In addition, interviews during the examination may be offered to those who have been unable to fill in the questionnaires by themselves. In this case

the mode of questionnaire admission should be recorded to allow comparison of responses by different administration modes.

8.3 Use of proxies

In EHES data collection, proxy use during the interviews is only allowed when the selected person him/herself is unable to respond due to major limitations in communication skills and/or cognitive ability. The reason for proxy use (why the selected individual was unable to respond on his/her own behalf) and type of proxy (spouse, child or other relative/significant other, or nurse for e.g. institutionalized persons) should always be recorded. When the use of proxy is considered, special attention should be paid for the decision if the person him/herself is capable to provide informed consent (see Part B, Chapter 2). Proxy use can be avoided by proper resources during data collection and scheduling adequate time to contact all selected persons.

There is a lot of evidence that the use of proxies introduces systematic biases, affecting national disability estimates and the incidence of several chronic conditions as well as their trends in repeated surveys (Shields 2004, Todorov & Kirchner 2000). Proxy responses and self-reports differ significantly depending on the type of questions, age and gender of both the proxy and the selected person, and relationship between the selected person and the proxy (Neumann et al 2000, Toldrov & Kirchner 2000, Shields 2004, Snow et al 2005). For younger persons there is evidence on proxy respondents under-reporting chronic conditions, disability and medication use, while for older persons the bias may be opposite, proxies reporting more impairment than self-respondents.

References

- Biemer PP, Groves R, Lyberg LE, Mathiowez N, Sudman S, editors. Measurement errors in surveys. New York: John Wiley & Sons, Inc.; 1991.
- Czaja R, Blair J. Designing surveys. A Guide to Decisions and Procedures. 2nd edition, Sage Publications, Inc., California, USA, 2005.
- de Bruin A, Picavet HSJ, Nossikov A. Health interview surveys. Towards international harmonization of methods and instruments. WHO Regional Publications. European Series. 1996: No 58.
- Franklin S, Walker C (eds). Survey methods and practices. Statistics Canada. Social Survey Methods Division, 2003.
- Neumann PJ, Araki SS, Gutterman EM. The use of proxy respondents in studies of older adults: lessions, challenges, and opportunities. J Am Ger Soc 2000; 48(12):1646-54.
- Rea LM, Parker RA. Designing and conducting survey research. A comprehensive guide. San Francisco: Jossey-Bass; 2005.
- Shileds M. Proxy reporting of health information. Health Reports 2004; 15 (3): 21-33.

- Snow LA, Cook KF, Lin P-S, Morgan RO, Magaziner J. Proxies and Other External Reters: Methodological Considerations. Health Services Research 2005;40(5):1676-1693.
- Tanur J. Questions about questions. Inquiries into the cognitive bases of surveys. New York: Russell Sage Foundation; 1992.
- Todorov A, Kirchner C. Bias in Proxies' Reports of Disability: Data from the National Health Interview Survey on Disability. American Journal of Public Health 2000; 90(8): 1248-1253.
- Tolonen H, Kuulasmaa K, Laatikainen T, Wolf H, for the European Health Risk Monitoring Project. Recommendations for indicators, international collaboration, protocol and manual of operations for chronic disease risk factor surveys. 2002. Available at http://www.thl.fi/publications/ehrm/product2/title.htm. Accessed on 22 January 2013
- Tolonen H. Towards the High Quality of Population Health Surveys. Standardization and Quality Control. National Public Health Institute A27/2005, Helsinki. Available at http://urn.fi/URN:ISBN:951-740-546-4

9. Selection of the fieldwork staff

Päivikki Koponen[1], Hanna Tolonen[1]

[1] National Institute for Health and Welfare (THL), Helsinki, Finland

Competent and motivated fieldwork staff is a key to successfull data collection. Characteristics of the staff members can influence nonresponse as well as validity and reliability of survey data. es-pecially in public health surveys, which cover sensitive issues and topics prone to socailly desirable responding (Davis et al 2010). In HES the selection of fieldwork staff has to be based on general requirements and competences needed to carry out the clinical measurements. Differences in the national health care systems as well as national guidelines for the responsibities of different health professionals need to be considered.

9.1. General principles and criteria for recruitment

Interviewer and measurer effects have to be considered when selecting and recruiting fieldwork personnel. Existing literature on interviewer race and ethnicity effects fails to conclude wheth-er respondents feel more comfortable with, trust, prefer or pro-vide more accurate data to interviewers of their own race, sex and ethnicity (Davis et al 2010). However. it is clear that these effect should be taken into account and evaluated. The General principles for the selection abd recruitment of fieldwork staff are (adapted from Tolonen et al 2008):

1. Legislation concerning medical practice and nursing in each country as well as the EU directives for the rec-ognition of professional qualifications have to be taken into account.
2. The personnel should be motivated to strictly follow the survey protocols to ensure reliability and accuracy of the survey results.

3. General appearance (non-provocative, calm and neutral appearance and good manners), friendliness, respect, empathy, encouragement and interest shown towards participants may affect participation and the results of the measurements. Age, gender, and ethnicity of the fieldwork personnel need to be taken into account in respect to the national and the participants' culture. It is recommended that the fieldwork teams consist of personnel with a variety of backgrounds. For example a similar ethnic background of the nurse and the participants may help to build trust and understanding among participants from ethnic minority groups. Similarly male nurses may not be accepted to carry out measurements requiring light clothing for women.
4. Willingness and possibility to travel around the country with the survey team may be needed depending on survey logistics. For example, this may be a problem for persons with small children.
5. Professional competence of the staff members and service given to participants may also be an important factor affecting survey response. Feedback given to the participants during and after the measurements needs to be considered also in the selection of survey staff. For example physiotherapists may be better qualified than nurses to carry out some physical functioning tests, while registered nurses may be better qualified than nurse assistants to carry out blood pressure measurements.
6. Fluency in national language(s), and if needed, languages of the major migrant groups.

Fieldwork staff may be recruited specifically for the survey. An altarnative is to use personnel from the local health care organizations (e.g. primary care units or health centers or hospitals) in the selected survey sites. It is usually easier to ensure standardization of measurements, if fieldwork staff is recruited specically for the survey. When permanent personnel of the local health services are trained to carry out the survey fieldwork they may be temped to follow their regular practices instead of the survey protocols. This may happen especially if they also have their regular tasks during the survey, and are only part time carrying out the survey fieldwork. In any case the use of the local personnel in each survey site increases substantially the time and efforts needed for training. The use of regular health service personnel may also effect survey results by the differences in willingness of the survey participants to disclose their personal issues to the practitioners they are familiar with. This familiarity may both enhance and restrict open communication.

A combination of the two gropus of personnel may be considered. Specially recruited personnel travel from survey site to another is trained to carry out the measurements that are most challenging to standardize, such as blood pressure measurements. These specific survey staff members may also supervise the local personnel responsible for other tasks. The local personnel may also be more efficient in recruiting participants.

9.2. Professional groups

From the point of view of these general requirements and implications to the survey budget different professional groups have both benefits and disadvantages (Table 9.1). The professional groups which should be considered for most of the measurements are physicians, nurses and other health care professionals. It is recommended that registered nurses carry out the EHES core measurements. The person performing the blood collection should be a certified phlebotomist. In most countries, this certification is offered through national accrediting agencies for clinical laboratory sciences. Employing a certified phlebotomist for the invasive blood collection procedure provides not only a measure of safety for the participant but also some medical-legal protection for the survey organizers, in case something should go wrong. A medical doctor is needed for back-up. Especially the person who takes the blood samples should know whom they can contact in case something happens with the participant during or after the blood drawing.

9.3 Fieldwork teams

When estimating the number of survey personnel needed for the fieldwork, potential sick leaves and other absences need to be anticipated. In most cases it is recommended to train a few extra persons for substitutes to ensure that time schedules are kept, and the participants are served as well as possible. Especially when the fieldwork period lasts for several months and the examinations are carried out by a team consisting of specific personnel for each measurement, the possibility to rotate duties between staff members should be considered. Such rotation of duties helps to minimize measurer effects and to motivate the staff members to follow the standards. This requires staff members with broader competence, who can also substitute other team members in case of absences (e.g. sick-leaves).

Table 9.1. Requirements, benefits and disadvantages of different professionals in survey fieldwork

Professional group	Specific requirements	Benefits	Disadvantages
Physicians (or dentists if oral health is measured)	Needed if clinical or diagnostic examinations are carried out and if physician's presence is required for clinical measurements. This may depend on national regulations.	May increase participation based on higher professional respect/ regard among the population. Better readiness for acute situations during the fieldwork, and in interpreting test results and informing participants about their test results (better service to participants may affect willingness to participate).	High salary level (effect on survey costs). Higher tendency to adapt survey protocols, not follow standards (Graves & Sheps 2004) and make independent decisions (also in conflict with survey protocols). Higher "white coat"/observer effect on some measurements, such as blood pressure (Graves & Sheps 2004, Labinson et al 2008**).**
Nurse	Registered nurse generalists with training according to the Directive 2005/36/EC are recommended for most measurements	Better adherence to follow standards (Graves & Sheps 2004) in survey protocol than physicians. Lower salary level and lower survey costs compared to physicians.	Differences in professional independence and respect among the population in European countries.
Certified phlebotomists	Recommended for blood sample collection	In-depth qualifications for blood sample collection.	Differences in basic professional training in European countries.
Interviewers	Specific interviewer training recommended if personal (face-to-face) or telephone interviews are used.	Standardized interviewing techniques.	Interviewers with medical/nursing background are more qualified for asking questions on medical conditions and medication. Lay interviewers may get more valid answers to questions on health behavior, as people may tend to give more socailly acceptable answers to professionals.

Professional group	Specific requirements	Benefits	Disadvantages
Other professional groups such as medical-technical assistants, nutritionists, dental assistants or physiotherapists	Depending on selected measurements	In-depth qualifications for specific measurements	Restricted roles/tasks. Having larger fieldwork teams with specific professionals for different measurements is challenging for the fieldwork logistics.

Support and supervision needed from the survey organizers or from the survey core group ("the central survey office") need to be ensured. This is particularly important if different teams cover different parts of the country. Well-defined leadership within the team is also essential. Each fieldwork team should have a specified supervisor/leader who follows the work progress and adherence to standards among all team members. In addition to the medical back-up required for drawing the blood samples, physicians may be needed to interpret measurement results or to give medical advice when abnormal measurement results, which may need urgent consultation, are found. When physicians are not part of the field teams their availability for consultation has to be organized in another way.

In case of home visits, the fieldwork teams seldom include more than two persons (interviewer and a nurse) and therefore multiprofessional fieldwork teams are not feasible. Instead, the personnel needs to be well trained generalists who have specific professional training for making home visits and whom the public easily accept to visit their homes. Typically public health nurses or health visitors are used for home visits. For surveys carried out in clinic settings the professionals selected for the fieldwork teams may vary. Two examples are presented here.

9.3.1 Team for a survey in clinic environments and with the EHES core measurements

Nurse 1, tasks: reception of the participants, obtaining informed consent, short health interview or checking the self administered questionnaire.

Nurse 2, tasks: Blood pressure measurement, height, weight and waist circumference measurement.

Phlebotomist, tasks: drawing and processing blood samples

Nurse 1 can be selected with less professional competence and with lower salary level (e.g. medical receptionist, medical-tech-

nical assistant). However, if nurse 1 and nurse 2 are both registered nurses rotation of tasks e.g. with monthly intervals and substitution of the other nurse in case of is sudden absences is possible. A survey physician may be needed as a back-up person (on call), easily available for consultation This consulting physician can cover several fieldwork teams working in different locations.

9.3.2 Team for a survey in clinic environments and with several additional measurements

Nurse 1, tasks: reception of the participants, obtaining informed consent, short health interview or checking the self administered questionnaire.

Nurse 2, tasks: Blood pressure measurement, height, weight and waist circumference measurement, lung function test (spirometry).

Phlebotomist/bioanalyst, tasks: drawing and processing blood samples

Nurse 3, tasks: diagnostic mental health interview (e.g. CIDI)

Physiotherapist, tasks: hand grip strength measurement, test of standing balance and timed Chair stand test

Physician, tasks: clinical medical examination with e.g. auscultation of the heart and lungs, interpreting previous measurement results (e.g. spirometry), and diagnostic assessments

In this team it is possible to rotate tasks between several team members if the bioanalyst is trained also to cover the tasks of nurse 2 and nurse 2 is also certified/qualified to draw blood samples. Nurse 1 and nurse 3 can easily be trained for both tasks. The last professional whom the participants meet at the end of the examination is the physician who is checks all measurement results and may advice the participants to seek further medical help when needed.

References

- Davis RE, Couper MP, Janz NK, Calwell CH, Resnicow K. INterviewer effects in public health surveys. Health education research 2010;25:14-26.
- Directive 2005/36/EC of the European Parliament and of the Council of 7 September 2005 on the recognition of professional qualifications. *Official Journal of the European Union* 2005:22-142.

- Graves JW, Sheps SG. Does Evidence-based medicine suggest that physicians should not be measuring blood pressure in the hypertensive patient? Am J Hypertension 2004;17:354-360.
- Labinson PT, Giacco S, Gift H, Mansoor GA, White WB. The importance of the clinical observer in the development of a white-coat effect in African-American patients with hypertension. Blood Press Monit 2008 Jun;13(3):139-142.
- Tolonen H, Koponen P, Aromaa A, et al. (Eds.) Recommendations for the Health Examination Surveys in Europe. B21/2008, Publications of the National Public Health Institute, Helsinki 2008. Available at http://urn.fi/URN:ISBN:978-951-740-838-7. Accessed on 22 January 2013

10. Blood samples and laboratory analyses

Georg Alfthan[1], Jouko Sundvall[1]

[1] National Institute for Health and Welfare (THL), Helsinki, Finland

This chapter gives guidelines and describes requirements for selection of laboratories, blood sampling, processing and storage, and the laboratory analysis of the samples. Also procedures on quality requirements of analytical laboratories and guidelines for the standardization of methods are are described.

10.1 Selection of analytical laboratories

It is recommended that all analyses pertaining to the core measurements of a country should be performed at the same laboratory, hereafter called the National HES Laboratory (NHESL). The most important criteria for selection of the laboratory should be based on its performance in external quality assessment (EQA) programmes. Whenever possible, the laboratory should be accredited by a national organization.

Accreditation

A prerequisite for a laboratory to become accredited is to have a documented quality management system. The usual contents of the quality manual follow the outlines of either the ISO/IEC 17025 for Testing and Calibration Laboratories or the ISO 15189:2007 for Medical Laboratory Standards.

Laboratories use the above standards to implement a quality system aimed at improving their ability to consistently produce valid results. This is the basis for accreditation from a national Accreditation Body. Since the standard is about competence, accreditation is simply formal recognition of fulfillment of that competence.

In these instructions we assume that the NHESL will be responsible also for the long-term storage of all samples. If this is not the case in the country, it should be taken into account in the national HES Manual.

It is hoped to establish a Central EHES Reference Laboratory (EHES RL) for support and EQA of the National HES Laboratories. For the time being, the NHESLs should keep in contact with the EHES Reference Centre (EHES RC) for the survey procedures and quality assessment.

10.2 Selection of measurements on the blood samples

Total cholesterol, HDL cholesterol and fasting glucose are core measurements, which should be included in all surveys. It is recommended that all countries collect more blood samples than are needed for the core measurements. This will make it possible to do various additional measurements on the samples in the future. There is more discussion on potential additional measurements in chapter A5.Selecting the measurements, where the following additional measurements are listed:

- Infectious disease antibodies, possibly to be measured in collaboration with the European Centre for Disease Control (ECDC);
- Environmental biomarkers, possibly to be measured in collaboration with DG Environment of the European Commission;
- Triglycerides;
- Apolipoproteins A1 and B;
- Glycated haemoglobin (HbA1c);
- DNA (depending on the participant's consent and legal restrictions in the country).

For the time being, instructions for the analytical laboratory covers mainly the core measurements.

10.3 Blood collection

The collection of blood samples for the analysis of the core measurements is described here.

10.3.1 Core blood measurements

Total cholesterol and high density lipoprotein (HDL) cholesterol:
These lipids should be measured from serum. Fasting is not necessary.

Glucose:
Plasma glucose should be measured from fluoride-citrate plasma. 8-14 hours fasting is necessary. Because of potential difficulties in requiring fasting from all participants, the measurements may cover only a subsample of the survey.

- In case fluoride-oxalate or another agent is used as an anticoagulant/inhibitor, a 5% lower glucose concentration per each 30 min may be expected before separation of red cells.

Additional measurements are considered in Part A, Chapter 5 of the EHES Manual.

10.4 Critical issues of the blood collection

10.4.1 Fasting before the sample collection

The serum samples for total cholesterol and HDL cholesterol can be taken at any time of the day, with the subject non-fasting. If measuring fasting glucose, lipoprotein fractions and triglycerides, the samples should be collected after a fasting period of at least 8 hours and at most 14 hours (excessively long fasting causes major changes in energy metabolism, with implications for blood triglycerides). In practice, this means that fasting must be overnight and that the samples can only be taken in the morning and can only be expected from persons who are invited to undergo the examination in the morning. In all cases the length of time from the last meal in full hours should be documented.

10.4.2 Position of the subject

All blood samples should be drawn with the subject in a sitting position preceded by a 10-15 min rest. Preferably, blood should not be collected from the arm that is used for blood pressure measurement, (i.e., blood should usually be drawn from the left arm).

10.4.3 Use of a tourniquet

Prolonged venous occlusion can cause changes in concentrations of blood constituents. Therefore, the use of a tourniquet should be minimized. If a tourniquet is used to search for a vein, it should be released before withdrawal of blood begins. In any case, the use of a tourniquet should be limited to less than one minute.

10.4.4 Effects of seasonal variation

Diurnal effects on analyte concentrations are varied. For the lipids it is difficult to dissociate changes in their concentration from the effects of a meal. Studies suggest that cholesterol concentrations are higher in autumn and winter than in spring and summer.

10.4.5 Effect of physical training

Excessive physical training may cause dehydration resulting in raised serum electrolytes and several enzymes of muscle origin. Except for dehydration, other effects are difficult to estimate. Therefore, abstaining from heavy physical training for 8 hours preceding phlebotomy is recommended.

10.5 Equipment for drawing of blood samples

10.5.1 Choice of type and order of blood tubes

The number and type of blood collection tubes depend on the core and other anticipated measurements on the samples. The blood collection kit, including all tubes and equipment needed for the procedure, needs to be planned and prepared in such a way that all parts are compatible. An example of a kit, covering the measurements specified in section "Selection of measurements" above is provided in Table 10.1.

Table 10.1. A recommended kit including all supplies for blood collection, processing and storage

a -	plain serum gel tube (9/8 ml) used for core measurements, e.g. lipids, lipoproteins (serum)
b -	fluoride-citrate (5/3 ml) used for glucose, clotting factors, adhesion molecules (plasma)
c -	EDTA tube (9 ml) used for DNA extraction (whole blood)
d -	EDTA tube (9 ml) used for e.g. vitamins, antioxidants (plasma)
e -	EDTA tube (3 ml) used for HbA1c (whole blood)

f -	tube holder
g -	needle
h -	Plastic tubes (short-term storage - 20°C)
i -	Cryogenic vials (long-term storage -70°C)
j -	storage boxes (for whole blood tubes, plastic and cryogenic tubes and vials)
k -	Sheet of bar code labels (for blood and storage tubes and storage boxes)
l -	tourniquet, skin cleaner, pipettes, tips, skin tape, etc.

Items *a* to *g* must be supplied by the same national supplier and all tubes are evacuated. Items *h* to *j* should be compatible with the systems of both NHESL and EHES RL. Eg. the cryogenic tubes must be straight-walled in order to enable reading of the bar code. Please contact the EHES RC for the requirements of the EHES RL.

10.5.2 Other equipment for handling, transfer and storage

- centrifuge, capable of 3,000g. If gel tubes are used, the centrifuge should have a swinging bucket rotor.
- racks for tubes
- special boxes for tube transfer and storage. The storage boxes should fit the freezer racks.
- set of (bar code) labels with identification codes or other method to mark the tubes (note that these should not be vulnerable to freezing)
- freezer (as required).

10.6 Processing, storage and transport of blood tubes

10.6.1 Processing

- Immediately after venipuncture, bar coded labels are placed on those blood collection tubes which have been successfully filled with blood.
- Centrifuge tubes at room temperature (20-25°C) for 10 min at 2000 g.
    - Plain serum gel tubes (a) are centrifuged within 30 - 60 min from venipuncture. Adherence to the time range and room temperature is necessary for complete clotting.

- Plasma tubes (b,d) are centrifuged together with the plain serum tube within 60 min from venipuncture. Simultaneous centrifugation of both serum and plasma tubes ensures the identification and aliquoting of samples from the same subject.
- Tubes c and e are NOT to be centrifuged
- The caps should not be removed before centrifugation

- Immediately after centrifugation remove the caps. Place bar code labels on serum and plasma storage tubes. The labels should be fixed upright (see Figure 10.1).
- Transfer with pipette serum or plasma into the storage tubes according to a prefixed scheme, example in Figure 10.2. When using gel-containing tubes, it is convenient to pool the serum before pipetting into aliquots, as shown in Figure 10.2. Cap the tubes tight.

Note: g =(relative centrifugal force, RCF) is calculated from the formula:

$rpm = 1000 \times \sqrt{(RCF/(11{,}17 \times r))}$, where

- r = radius, distance from tip of tube to center of rotor (cm),
- rpm=rotations per minute.

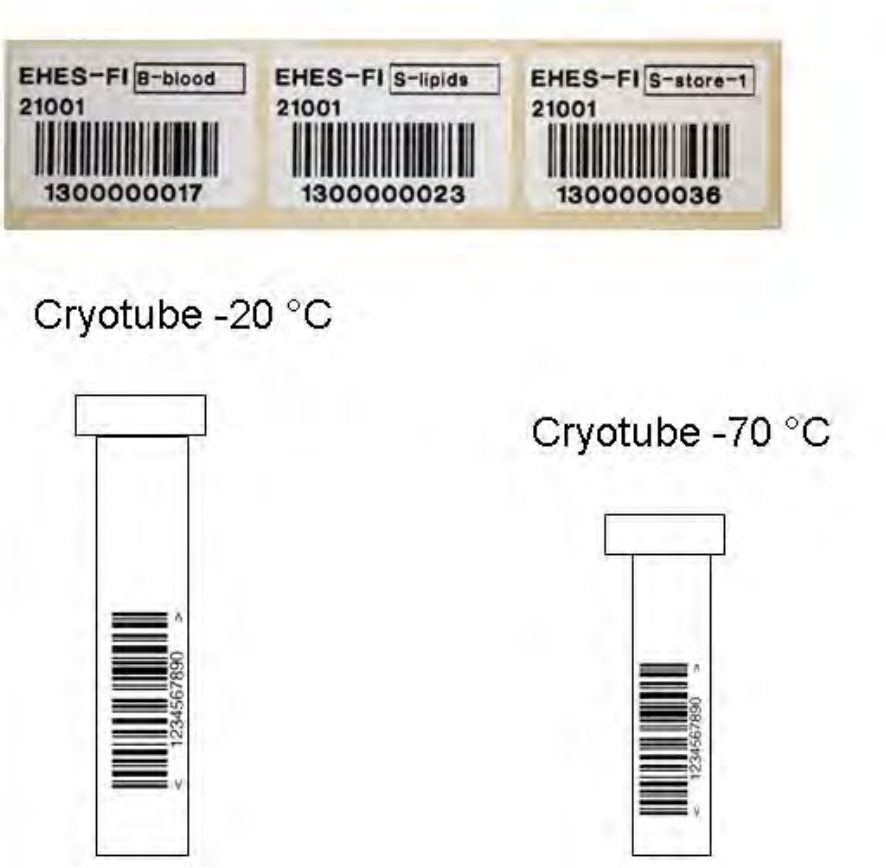


Figure 10.1 Labelling of storage tubes

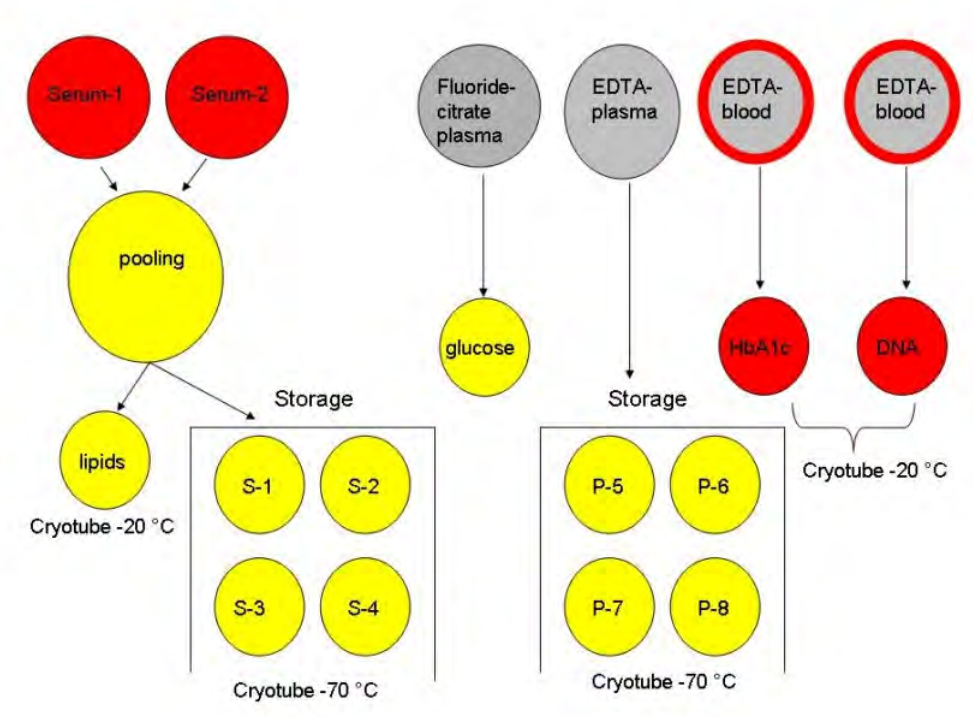


Figure 10.2 Example of blood processing

10.6.2 Storage of whole blood, serum and plasma tubes

Only a few aliquots of serum and plasma will be used for the core measurements. Samples frozen at -20°C should be analyzed within six months. These are typically reserved for the core measurements. For long term storage reserved for additional measurements and future use, the samples must be frozen at -70°C. Note that tubes intended for core measurements should not be discarded after analyses, but should be returned to the original storage temperature.

The storage boxes should be labelled with their appropriate bar code label BEFORE placing them in the freezer. Otherwise the labels will not stick. Place tubes upright in their designated boxes without delay and keep the boxes in a - 20°C freezer. An inventory of all stored specimen must be documented daily at the examination site.

10.6.3 Transport of specimen from the examination centre to the NHESL

The frozen samples should be sent in suitable batches during or at the end of the fieldwork. An inexpensive temperature check is to place a pre-frozen tube, half-filled with water, upside down in each batch. The boxes are transported to the NHESL with an adequate amount of dry ice.

The transportation should be organized well in advance (with a courier company). A courier company will provide the service including necessary paper work for “door-to-door” transport. Before sending the shipment, please inform the contact person of the receiving laboratory of date, courier and tracking details by email. The receiving laboratory should acknowledge the receiver shipments by e-mail or phone. The examination centre should keep a log book of all shipments, where also the acknowledgments are recorded.

10.7 Guidelines on laboratory performance

10.7.1 Performance of laboratories

Concerning the core measurements, the golden standards are values determined by the Centres for Disease Control (CDC, Atlanta). Data on the three following levels of accuracy (bias) per-

formance ascertainment for core measurements will be monitored by the central EHES RL:

- Bias in EQA programmes of NHESLs related to the core measurement methods during the preceding year.
- Bias between NHESLs and EHES RL
- Bias between EHES RL and Centres for Disease Control (CDC, Atlanta).

This accumulated data will be reported annually.

10.7.2 Standardization of analytical data

1. A pilot calibration between the NHESLs and EHES RL is carried out before the beginning of the survey. Therefore, the NHESL should contact the EHES RL at least 6 months prior to the planned starting date of the national HES. If the pilot calibration is satisfactory, proceed to step II. If it is not satisfactory, continue calibration pilots until results agree.

   The calibration consists of a series of reference samples having target values.
2. Depending on the analyte and number of survey participants, a random 5 or 10% of actual survey samples are transported and reanalyzed at the EHES RL.

10.7.3 Recommendation for analytical methods

No recommendation to use a specific method is given. However, only validated methods should be used, and the procedures should be documented. The documentation for each method used in the survey should be available at the NHESL and the EHES RL.

10.7.4 Quality Control

Data on precision of methods within and between days (series) must be kept with a computer-aided protocol. The goal and acceptable precision between days(series) for the core and additional analytes are shown in Table 10.2. The precision limits provide guidelines for the performance of a mehtod. The limits are based in part on data from instrument manufacturers and experience of the EHES RL.

10.7.5 Accuracy (bias) and external quality assessment (EQA)

Data and documentation on accuracy of the methods are provided by participating regularly, >1 times per year in national or international EQA programmes.

It is recommended to check and document the performance of all instruments (clinical chemistry analyzer, photometers, balances and pipettes, eg.) once a year. As with all components of a method, the traceability of calibrators should be documented in order to ensure high quality.

The recommended goal and acceptable bias values for the core and additional analytes are shown in Table 10.2. The data on bias take into account the biological variation of the analyte. The values are modified from the reference: www.westgard.com/europe.htm.

Table 10.2. Recommended goals for bias and precision of methods

	Bias (%)		CV (%)	
Core analytes	**Goal**	**Acceptable**	**Goal**	**Acceptable**
Serum total cholesterol	3	5	2	3
Serum HDL- cholesterol	5	10	2	3
Plasma glucose	4	8	1	2
Additional analytes				
Apo A-I	5	12	2	3
Apo B	5	12	3	5
Serum Tg	7	15	3	11
Blood $HbA1_c$	2	3	2	3

10.7.6 Corrective measures

If the bias of a method exceeds the acceptable value, corrective measures should be performed. Likely errors stem from erroneous calibrators, change of reagents (kits), malfunctioning instruments, wrong type of sample (serum-plasma-whole blood) and reagents not compatible with the instrument. An unexpected shift in bias may be observed, for example, when new calibrators are introduced. Also change of the method (technique) may provide remedy. According to good quality criteria of a laboratory, all NHESLs must document all changes done to methods or procedures and report promptly to the central EHES RL.

10.8 Safety and laboratory quality procedures

Guidance on issues regarding safety procedures of laboratory personnel and laboratory quality assessment, eg. equipment are given in Part B, Section 5.5.

11. Quality assurance

Kari Kuulasmaa[1], Päivikki Koponen[1], Hanna Tolonen[1]

[1] National Institute for Health and Welfare (THL), Helsinki, Finland

Quality assurance of a heath examination survey (HES) refers to the measures that are undertaken to ensure a good quality survey. Well planned quality assurance is essential in order to obtain high quality data, which will be comparable between countries and in particular over time, so that reliable long-term trends can be calculated form the data in the future. The basic components of quality assurance are:

- **Good overall management** of the survey.
- **Agreement on survey procedures** that ensure standardized stable measurements. These, together with all other requirements for the national survey management, should be described in the manual of operations.
- **Training** of the survey personnel on using the standard procedures.
- **Piloting** the fieldwork phase.
- **Quality control**, which refers to measures taken to monitor the survey process, so that any problems can be detected at an early stage. The term 'quality control' also includes the action taken to correct the detected problems. In the ideal case, the problem will be detected early enough so that it can still be remedied. We will also use term Quality assessment, which refers to the monitoring and documenting the quality, but does not include the corrective action. Accordingly, a part of quality control is quality assessment.
- **Evaluation of the achieved quality.** This step is necessary in order that the results of the survey can be interpreted correctly, taking into account the limitations

of the survey quality. This step is also important for the documentation of the experiences, so that similar problems can be avoided in future surveys.

A separate section will be devoted to each of these components below.

Quality assurance should be seen as an integral component of all phases of a HES (Figure 11.1).

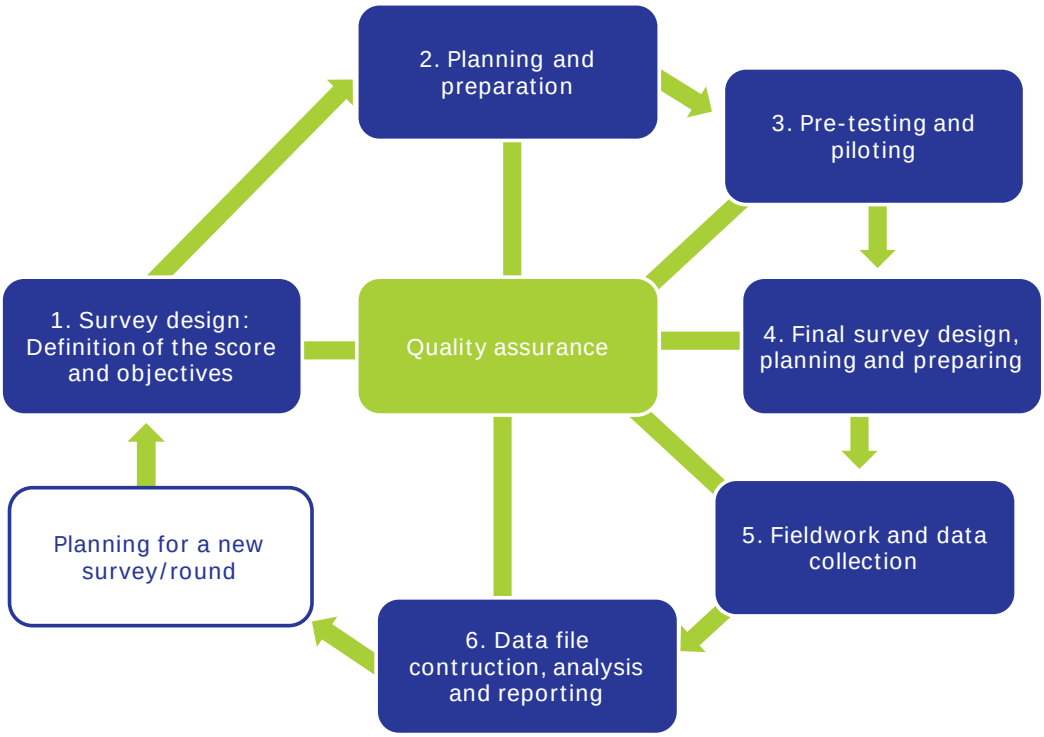


Figure 11.1. Rome of quality assurance in a HES

11.1 Good overall management

Survey management is considered in Part A, Chapter 1. From the point of view of quality assurance, it is important that there is a well defined survey organization. This includes a management structure, clearly defined responsibilities of the survey personnel and professional coordination.

A careful risk analysis helps the survey management to anticipate and prevent many problem situations, which could otherwise have serious implications to the quality of the survey. Risk analysis, which is considered in Part A, Chapter 1 should be done in the planning stage of the survey and reviewed periodically during the survey.

11.2 Agreement on survey procedures

European guidelines and standards for the survey procedures are described in this EHES Manual. The national HES manual should describe the details of the procedures to be used in the national HES, also for issues where the EHES Manual can only give alternatives or general guidelines.

Whenever a country considers a procedure which deviates from the European recommendation, the issue should be discussed with the EHES Reference Centre. The decided deviations together with the justifications should be documented in the national HES manual. They will also be posted in the EHES Extranet, where they serve as examples for the other countries. Countries are also encouraged to discuss such issues in the EHES extranet with other countries which may face a similar situation.

Templates to facilitate the preparation of the national manuals are available in the EHES Extranet, where also the English version of the national HES manuals will be posted. These English translations are needed for the communication with the EHES Reference Centre, external quality assessment and survey evaluation. They are also examples for others preparing their national HES manuals. A national language version of the manual will be needed for the local purposes.

11.3 Training

Training is needed for acquiring the skills to follow the survey procedures. Each country is responsible for training the national survey personnel. Specific training is necessary in particular for the persons performing survey measurements in the field. The training and certification procedures for each measurement are described in Part B, Chapter 5 under the each measurement procedure.

If the field work takes more than few months, re-training sessions are usually needed. Measurement practices of the survey personnel often have a tendency to change over time. Re-training will reinforce the use of the standard procedures.

The contents of the Europe wide and national training programmes are considered in Part A, Chapter 15. The national training programme should be specified in the national HES manual of each country.

11.4 Piloting the HES

Piloting the HES and detailed evaluation of the pilot is crucial to ensure a successful data collection and field work phase of the HES.

Each country should carry out a pilot survey prior to the full-size HES. The purpose of the pilot survey is to evaluate the entire survey process and to obtain additional information for the planning of the actual survey. The pilot survey also helps in familiarizing the fieldwork personnel with potential practical problems. The sample size of the pilot HES should be estimated in such a way that it will lead to about 200 participants.

The aims and content of the pilot survey depend on the contents of the survey, previous experience and frequency of surveys. At least the following issues need to be considered (adapted from Primatesta et al. 2007, Biemer et al. 2003):

- Identifying need for further quality assurance activities, such as further specification of the recruitment, measurement and training procedures. This facilitates planning the training of the survey personnel and finalizing the survey manuals.
- Getting feedback from the invited participants. This may concern the willingness to participate, recruitment process and information leaflets, informed consent and experiences on the measurements. It is also useful to observe the respondents' reactions. The feedback is needed to develop different ways to motivate participation in the population.
- Recording timing and calculating average duration of interviews and examinations per participant. This is needed to estimate the need for personnel resources (which has implications to budgeting) and potential burden to participants.
- Testing the use of equipment, computer programmes, data management, and the processing, transfer and storage of blood samples. This is needed to avoid problems in data and sample collection and management, and to assess the need for storage space, equipment and logistics.
- Familiarizing staff with potential practical problems. This helps to avoid problems during fieldwork, and may give rise to refine practices and add further specifications to the national HES manual.

The questionnaire is evaluated in regard to

- the length of recall period,
- clarity of concepts and definitions,
- the question wording and the response alternatives,
- the sensitivity of topics,
- the questionnaire layout,
- the choice of administration mode, and
- the respondents' burden (i.e. how long it takes to complete the questionnaire).

An optimal timing for the pilot survey is 3-6 months before the full-size HES, so that there will be sufficiently time to evaluate the pilot and to make the necessary adjustments to procedures before the full-size survey. The pilot survey should be conducted by the same personnel who would do the full-size survey.
The countries should also consider the need for a small pretesting before the actual HES pilot. Such a pretesting may be needed to test the computer programmes, measurement techniques and timing with a small number of volunteer participants.

11.5 Quality Control

The term "quality control" refers to the measures taken to monitor the survey process so that any problems can be detected at an early stage. Well planned and conducted quality control will save resources because it will minimize the time and resources needed for detecting and solving problems and for repeating tasks. It will also minimize the need to reject survey data because of loss or poor quality. Quality control is also needed to convince the users of the survey data about the good quality.

11.5.1 Quality control of the planning of the HES

The main quality control activities of the planning stage are:

- to check that the plan for the preparatory phase of the survey covers all relevant aspects with sufficient detail (see Part A, Chapter 1); and
- to monitor the the time schedule of the planning and preparatory phase.

The national survey organizers will be assisted by the EHES RC, which is responsible for monitoring the progress of the planning and preparation of the national HESs and reviewing the national HES manuals. Therefore, each country will have to provide the EHES RC with:

- a schedule for planning the national HES as early as possible, and preferably one year prior to the beginning of the field work of the survey, and
- an English translation of the national manual preferably six months before the planned start of the survey.

The EHES RC will have to review and comment on the proposed schedule and the manual without delay, and in any case within three months after receiving them.

11.5.2 Organizing quality control for the survey procedures

Most of the quality control is carried out by the national survey team. This is called internal quality control. In addition, it is important to have the survey observed and assessed by an independent outside body. This is called external quality assessment.

11.5.2.1 Internal quality control

In principle, quality control is relevant for all phases where data are transferred from one form or place to another. Therefore, quality control is relevant for:

- the interview and measurement instruments and procedures,
- data and sample handling and transfer, and
- the data management.

Procedures for the internal quality control for each of these are specified in the respective Chapters of Part B (Fieldwork procedures).

The corrective action must always be thoughtful, to make sure that it really corrects the problem. A wrong correction may add a new component to the measurement bias. For measurements which involve a subjective component, such as interview or blood pressure measurement, the way of approaching the measurer needs to be planned carefully in order to prevent over correction. The best action can be for example a routine retraining of all personnel doing that measurement.

The activities of internal quality control should be documented in a log book, together with any concerns detected and the action taken to correct problems. Examples of such log books are given in the Part B, Section 5.6.

The implementation of the internal quality control in each country should be described in the national manuals.

11.5.2.2 External quality assessment

External quality assessment is never a substitute for the internal quality control. It complements the internal quality control by providing an independent review of the performance, checking that the national standards are similar between the countries and over time, and overseeing that the internal quality control functions as planned.

The EHES RC coordinates and carries out external quality assessment in EHES. Individual countries may use additional sources of external quality assessment, but they should keep the EHES RC informed of this.

The external quality assessment carried out or coordinated by the EHES RC includes:

- Monitoring the progress of the planning and implementation of the national surveys.
- Review of the English versions of the national manuals.
- Site visits during the national surveys to assess the survey and quality control procedures.
- Assessment of the data obtained from the surveys. This will be described in more detail in subsection “Evaluation of the achieved quality” below.
- External laboratory quality assessment. This will be described in more detail in Part A, Chapter 10: Biological samples and laboratory analyses.

11.6 Evaluation of the achieved quality

The evaluation and documentation of the achieved quality involves analytical investigation of:

- the actual survey procedures used,
- the data generated in the surveys, and
- the data and information generated though external quality control.

The evaluation report is an essential prerequisite for the analysis and correct interpretation of the survey data.

For the overall conduct of the survey and for each of the core measurements, the evaluation will be carried out by the EHES RC, with the help of national survey organizers. For any national additions to the survey, such evaluation should be carried out nationally.

As part of the planning of the survey, there is a need to ensure that the data collection in the survey includes all data needed for the evaluation. For the EHES core measurements, the required data items are described in Part B of the EHES Manual, under each measurement protocol. For additional measurements, countries should define the required information before they start their field work.

References

- Primatesta P, Allender S, Ciccarelli P, Doring A, Graff-Iversen S, Holub J, et al. Cardiovascular surveys: manual of operations. Eur J Cardiovasc Prev Rehabil. 2007 Dec;14 Suppl 3:S43-61.
- Biemer PP, Lyberg LE. Introduction to Survey Quality. Wiley, 2003

12. Data management

Ari Haukijärvi[1]

[1] National Institute for Health and Welfare (THL), Helsinki, Finland

A well-organized data management is an essential part of health examination survey (HES). It ensures that:

1. the data will be available for analyses, and that the available data are
   - *complete*. No data collected from the survey subjects are lost.
   - *correct*. There are only justified differences between the values which were originally measured and the values in the final data storage.
   - *verifiable*. The relationship between the original data collected and the data in the final data storage can be described.
2. the data analysis will be
   - done using the correct data and other information which are relevant for the data analysis.
   - done without errors.
   - documented in such a way that the whole analysis or a part of it can be repeated later. If the documentation is not done properly, it may be difficult or impossible to reach the same results when similar analyses are repeated in other situations.
3. the confidentiality of the data is secure.

Point one above involves data collection, checking of data, error correction, data transfer from the field to the final storage (database), documentation and back-up of the data. Point two concerns analysis of the final data to obtain survey results. *Separation of the data management into these two stages is recom*

mended. If the survey data are not completed and the quality of the data are not documented before the data are analyzed, it is likely that the analysis will reveal problems in the data, many of which could have been detected earlier. This in turn can result in much longer delays in the final analysis than if more care had been taken during the preparation of the data. Furthermore, use of unchecked and uncorrected data will lead to incorrect results. Well-planned data management facilitates good quality and availability of the data for analysis.

Therefore it will be necessary to create a detailed plan for the data management including all phases of the survey. Planning of the survey data management should be part of the general planning of the survey from the beginning. The things that need to be focused are:

- detailed data flow in the survey
- transfer, storage and security issues during each phase
- rules for data correction (data correction procedures)

12.1 Basic work and data flow

Below the following topics are considered from the point of view of organizing national HES data management:

- sample selection and recruitment
- appointment scheduling
- collecting the survey data
- error checking, correction, and documentation of the data
- transfer and storage of the data

It is assumed that each country establishes a database for their survey data and maintains it locally in the country. A database should be prepared to store individual level data on the national HES measurements (including the questionnaire part), information on the quality of the data, as well as sampling data of each survey respondent. The national HES database serves as the local repository for the data used for evaluation of the surveys and for analyzing the survey results. An example of the work and data flow in the national HES is depicted below in Figure 12.1.

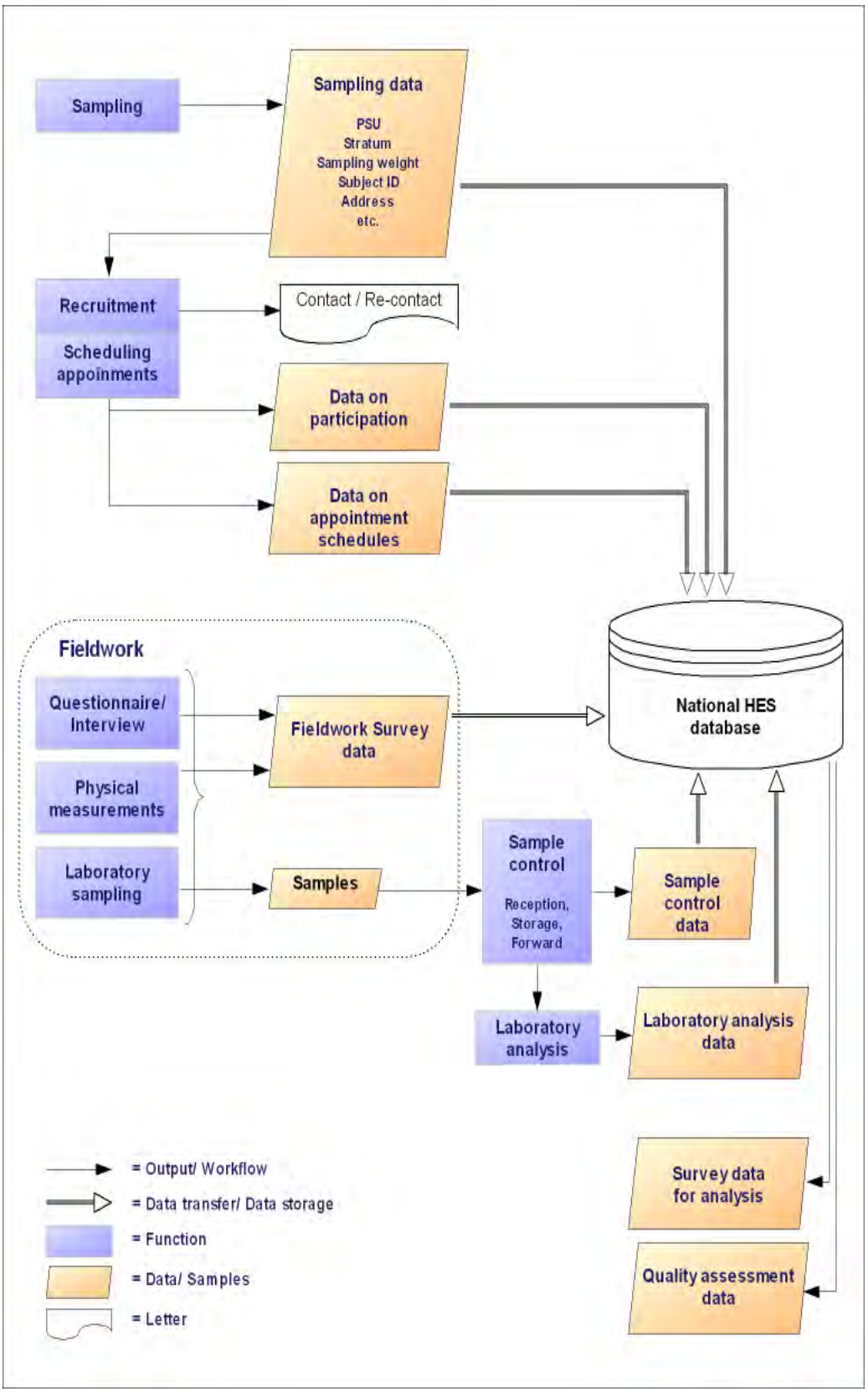


Figure 12.1. An example of HES work and data flow

There are several good methods to implement the stages of survey data management ranging from manual methods to computerized ones. In each phase modern information technology can

be utilized. The choice of the methods will depend on local facilities, existing practice and the expertise available.

12.2 Subject Identification

The subject records in EHES will be identified using four levels of codes:

- **EHES:** This specifies that the survey is a part of the EHES framework.
- **COUNTRY:** This identifies the country of the survey, using the EU-coding (see http://publications.europa.eu/code/en/en-370100.htm). It is the same as the two-letter ISO 3166-1 alpha-2 code with two exceptions: Greece is EL (not GR) and United Kingdom is UK (not GB).
- **SURVEY:** This two digit number code identifies different EHES-surveys in the country. It is assigned by the national survey organizers but shoud be confirmed with EHES Reference Centre before applying.
- **SERIAL NUMBER:** This number identifies the different persons selected to the sample. It is assigned by the national survey organizers, and is subject to the following principles:
  - SERIAL NUMBER is given to everybody selected to the sample (i.e. not only for example those eventually found eligible or to the survey participants.)
  - It is unique within the survey. Only one person selected to the survey can have the same SERIAL NUMBER. However, surveys in different countries or different surveys within same country (identified by different SURVEY codes) can use same SERIAL NUMBERS. When addresses or households are used as sampling units, the part of the serial number identifying the subject may be completed only at the stage when the household is visited.
  - The maximum length of the SERIAL NUMBER is 12 characters.
  - Because errors in the SERIAL NUMBER usually lead to a loss of the record, it is strongly recommended that the SERIAL NUMBER includes a check digit (or check digits). An example of a convenient single character numeric check digit, which detects all one digit errors and all transpositions of adjacent digits has been described by Gumm (1986).

- The serial number should not include information which makes it possible to identify the person in the population.

12.3 Data sources

12.3.1 Sample selection, recruitment and appointment scheduling

12.3.1.1 Sample selection

The first major data management issue relates to sample selection and recruitment (see Part A, Chapters 2, 3 and 13). As a minimum, the following information is to be recorded for every subject selected for the sample:

- *SERIAL NUMBER* (see Part A, Section 12.2) should be given to everybody selected to the sample and used to identify the subject throughout the survey and data management.
- *Sampling information*, as specified in Part A, Section 3.8.
- Address information, such as the person's name, address and any other information will be needed to contact the person.
- Additional information, such as sex and age are also often available from the sampling frame.

The *subjects selected to the sample form the basis for the control of the data completeness* through the data management process. The survey history of every subject should be verifiable from the final database.

12.3.1.2 Recruitment

At the recruitment stage, attempts made to contact each subject need to be monitored, and the eventual success of the recruitment needs to be recorded. Also the contact information may need to be updated. For each person invited to participate in the HES, it is necessary to keep a record of the following information:

- Eligibility status
- The number and type of contact attempts
- Contact status
- Participation status

- The reason for non-participation. For the subjects who did not participate the survey examinations, the reason should be recorded using the classification listed in Part A, Section 13.3.

12.3.1.3 Appointment scheduling

Organizing the appointment schedule is necessary for a successful fieldwork examination and is closely linked to the recruitment phase. For example in a case when an appointment was fixed, but the subject is nevertheless unable to participate the survey, this changes not only the appoinment schedule, but also the participation status (or even eligibility status) in the recruitment data. Here at least the following information should be logged:

- The subject identification (see Part A, Section 12.2)
- The contact information (name, address, phone number)
- The appointment information (time, place)
- The recruitment information (the record of participation)

Regarding the future surveys it might also be important to keep a record of changed appointment times.

There are several commercial applications for scheduling (e.g. patient scheduling software used by hospitals and clinics). Some of them are web-based scheduling services, while others are standalone client software. The usefulness of these software depends on how they can be customized to manage the necessary data. One possibility is to build a dedicated HES database application for the survey project to serve both the recruitment data and the appointment scheduling. The approcation should preferably interact with the national HES survey database. Change of experience between countries will be useful when planning or selecting an application for this purpose.

12.3.2 Survey data

12.3.2.1 Sources

Recording the survey measurements and getting data from different examination sites to the common database are essential parts of the national HES data management. These include:

- completion of self-administered questionnaires (if self-administered questionnaires are used)

- interview
- recording the values of physical measurements
- biological sampling, processing and transfer to the laboratory
- recording the laboratory results
- transfer of paper forms and/or electronic records from the examination site or laboratory to the common database
- getting the data into electronic format

Three main challenges for the data management during these steps are to ensure that

1. no errors are made in recording the results;
2. the data records are complete;
3. no records are lost or different persons' records are not mixed up.

Errors and incompleteness of the records can be prevented by good design of the record forms and by routine checking of the forms and the data. The earlier the errors will be detected, the easier their correction is. When feasible, detection of errors should be done when the subject is still in the interview or examination site.

Relevant data which were not obtained from the subject should not be left blank, but a specific code for missing data should be used. Subsequently, the incompleteness of the data can be detected as blanks in the data forms.

To prevent the loss of records, it is important that the subject identification becomes correctly recorded at all stages and the Subject Identification code will be used. If feasible, laboratory samples (biological samples, storage tubes and storage boxes) should be labelled with bar codes with a reference to the correct Subject Identification code.

All steps where data are transferred from one form to another or from one place to another require specific attention when data management of the survey is being planned.

12.3.2.2 Forms of data collection

The procedure of collecting the survey data will be different according to whether the data are collected directly to computers or on paper forms.

- A computer-assisted data collection has the advantage of reducing the number of manual data transfers and facilitates extensive data checking at an early stage. However, such a system should be used only if it has been tested in the field and found reliable. Otherwise there will be an increased risk for losing records or delaying the examination schedule due to breakdown of the system. Paper forms should always be on hand as back up in case of power failure or other problems with computer devices.
- The use of paper forms has proven to be reliable, but they have the problem that on-site data checking is more difficult. If paper forms are used, the typing of the data into electronic format needs to be done carefully. (In this case the traditional double typing method by different persons is worth considering). Also optical character recognition (OCR) can be used to convert scanned images of handwritten or printed text into electronic files. This again sets up challenges for design of the forms and data error checking, i.e. validating the OCR converted data.

Table 12.1. Computer-assisted interviewing methods

Method	Description
Computer-assisted personal interview (CAPI)	Interviewers meet respondents and conduct a face-to-face interviews using a computer. There may be an online connection to an external database from the computer or the data are sent to a central computer after the interview (either via Internet or by sending data disks by mail).
Computer-assisted telephone interview (CATI)	The interviewer sits at a computer and asks the questions appearing on the screen. The respondents are on the telephone. The respondent's answer is typed by the interviewer. Supervisors are present for quality control and to assist with specific problems.
Computer-assisted self interview (CASI)	In computer-assisted self interview or self-administered web-survey the respondents themselves read the questions on the screen and enter the answers. There is no interviewer; the interviewing program guides the respondent through the questionnaire. This procedure can appear also as part of a computer-assisted personal interview session where the interviewer hands over the computer to the respondent for a short period, but remains available for instructions and assistance. This is equivalent to the procedure used in traditional face-to-face interviews where an interviewer might give the respondent a paper questionnaire containing sensitive questions.

The computer-assisted data collection may include both the interview and the measurement phase. Computer-assisted interviewing methods are described in Table 12.1. In computer-assisted data collection automatic built-in checks for responses become possible, and data entry by trained fieldworkers reduces errors. A computer-assisted interviewing system and survey processing tool (such as Blaise software) can include features to define questionnaires, data validity and range checks, conditional error handling, etc. which facilitate both the questionnaire design and data entry, and can help to prevent printing errors.

In general when computerized data collection is used during the fieldwork, it will be necessary to store the survey data and measurements as they are gathered - either locally or directly to the national HES database. Storing the final data once, as soon as possible after the measurement will make it possible to have only one recording round for the data, enhance data security, and ensure that the data will not be forgotten or lost.

Good practices in recording the data are:

- Minimize the number of times the same data are recorded or stored
- Immediate recording/storage of the data (the data will not be forgotten or lost)
- The data are recorded by the producers of the data (i.e. the members of the fieldwork team). In consequence:
  - there is a well-defined responsibility in recording the data,
  - the security of the data will be ensured

12.3.2.3 Preparation of the fieldwork

The preparation of the fieldwork phase includes:

- Planning of the data management and data transfer system for the fieldwork (i.e. computers, network, software, and other equipment needed) and testing these
- Arranging the training, responsibilities and support of fieldwork teams

When necessary computer equipment and network are considered things that require particular attention are data security, data transfers, and back-up of the data to avoid any losses of data in case of a system flaw. Questions arising are:

- Shall a private LAN/WLAN be established for each field-work team?
- What computer equipment will be needed?
- What software will be needed?
- Planning software update procedures during the field-work
- Planning data back-up equipment and procedures
- Planning the transfer of the data
- Data security (locking of computers, usernames and passwords to access the computers, encrypting data on computers, how to store paper forms on the field, etc.)

Regarding the software all workstation computers are recommended to be identical to each other and thereby easy to replace if broken. Server computers (if any) in the fieldwork can be designed to be stand-alone, independent portable computers, easy to relocate or replace in case that one of them is broken. A backup mirror can be implemented between servers, in which case the data on the first server are copied to the second one.
The equipment and data transfer and storage systems should be tested thoroughly prior to the training of fieldwork personnel. Time needs to be reserved both for testing and analysing the test results.

12.4 Error checking, correction and documentation of the data

After the data collection the data should be checked as soon as possible for

1. strange *values*, i.e. for values which have not been defined, and also for values which are possible but rare;
2. *consistency* between the values of different data items;
3. *completeness*, i.e. that all data items have been recorded and no records have been missed.

A visual checking of the key items can be done at the interview or examination site even if paper forms are used. An extensive checking should take place as soon as the data have been computerized. When potential errors are detected, they should be investigated for correctness, and corrected only if it is found that they really are errors. It is advisable to authorize only those who have made the errors to correct them, since they are usually in the best position to tell if there really is an error, and are often

the only ones who know the correct value. Each error and its possible correction should be documented.

The frequency of errors, which were not possible to remedy should be documented. The same concerns the results of the quality control during the data collection, any deviations from the survey protocol, and any other information which may be relevant for the interpretation of the results. Knowledge of these issues is essential to those who analyze the data and interpret the results.

Each data transfer and import into the central national HES database should include at least a routine check for each data variable.

Examples of routine error checking criteria and documentation of the quality of the data in a multinational setting can be found e.g. in the WHO MONICA quality assessment reports, which are available at http://www.ktl.fi/publications/monica.

12.5 Transfer and storage of the data

12.5.1 Data transfer and interface to import the data

Data transfer and import into the national HES database depends on whether the data are collected using computer-assistance or manually by using paper forms. Whenever data or samples are transferred from one place to another, it is important that the data transfer is logged properly. All data transfers should be traceable whether they are computerized or manual: The recipient of the data or samples should be able to check that he or she has received exactly the same records which were sent, and the person sending the data should make sure that everything was received. The transfer of the data into the central data storage, the national HES database, should be done regularly and via a secure data transfer medium.

Good ways to transfer computerized data are through Internet via a secure connection and data encryption or by storing data on disks delivered via a secure mail. For a user interface to import the data into final data storage there exist primarily following options:

- Web application - a dedicated web software designed to interact with the national HES database or other database from which the data are further transferred into

the national HES database. This kind of an application typically funtions on clients' web browser.

- Standalone client PC application, operating on dedicated computers.
- Other - e.g. direct import from data files. The data are delivered to the HES coordinating centre on a dedicated medium from which data files are uploaded into the central HES database.

An example of a possible HES database system architecture is depicted below in Figure 12.2.

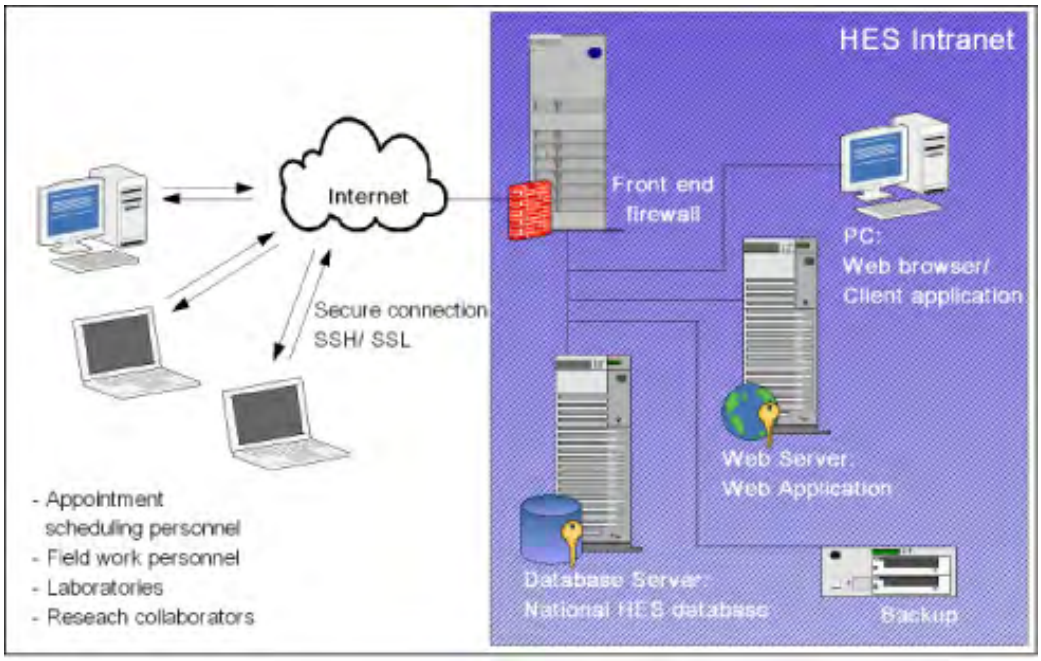


Figure 12.2. An example of HES system architecture

12.5.2 Data security

The core principles of information security are data confidentiality, integrity and availability regardless of the form the data may take: electronic, print, or other forms. (Table 12.2)

Table 12.2. The principles of data security

Security principle	Description
Confidentiality	Ensuring that information is accessible only to those authorized to have access. This is necessary for maintaining the privacy of the people whose personal information the system holds and to prevent the disclosure of information to unauthorized individuals or systems.

Security principle	Description
Integrity	Safeguarding the accuracy and completeness of information and processing methods from intentional, unauthorized, or accidental changes. Maintaining data integrity is essential to the privacy, security, and reliability of the data. Data integrity can be compromised by malicious users, hackers, software errors, computer virus infections, hardware component failures, and by human error in entering or transferring data. Therefore, the access to protected information should be done only through proper identification and authentication of the users.
Availability (the degree to which the system is operable)	Ensuring that authorized users do have access to information when required, i.e. the information is available. This can be accomplished utilizing data back-up plans and continuity/recovery plans.

To ensure data confidentiality and integrity it is necessary to use technical controls - e.g. passwords, network firewalls, access control lists, and/or data encryption - to monitor and control access to the computing systems and collected data. Some standards for these are described below in Part A, Section 12.7. For example during the fieldwork it is necessary to protect data on local computers' hard drive, or when transferring data over a network.

It is essential that the information connecting the survey data to the personal identification of the subject will be available only to persons who have authorized access to such data. Only authorized persons should have access to the data and all of them must understand the importance of the confidentiality of the data. After data collection, the information from which a person can be identified and the code which connects this information to the subject identification of the survey records, should be stored separately from the survey data, and maintained e.g. on encrypted hard drive. Normally only few people need access to the person identification information, whereas the rest of the survey data will need to be accessed by all who analyze the data. Specific precautions should be defined for the handling and storage of paper forms in the examination site and elsewhere when these are used.

Regarding the administrative controls, approved written policies and guidelines on data transfer and confidentiality, see Part A, Chapter 4.

12.5.3 Back-up

All data in electronic format should be backed up routinely for accidental breaks of the storage devices, failures in data transfer and unintentional deletion of data files. Especially during the fieldwork phase it will be important to back-up the data on local computers' hard-drives against accidental losses. Common situations where important data have been lost, although some back-up was in place, are:

- Loss of data during data collection or data processing because of absence of back-up at these early stages.
- Accidental loss of the back-up data together with the original data since the two were stored together, or the broken device or system which destroyed the original data was used to open the back-up data.
- The complete back-up data had already been replaced by the incomplete data before the loss of data had been detected. This may happen if the system for long-term back-up is incomplete.
- There was a back-up, but there were insufficient documentation on its location or on the procedures needed to retrieve the data from the back-up. This problem could arise due to unforeseen changes in personnel.

Technically several storage media can serve as back-ups. Here the primary options are:

- Magnetic tapes (LTO-4 provides up to 1.6 TB of capacity per cartridge)
- External hard disks
- Optical disks (CD, DVD)
- Another computer dedicated to back-up purpose

Back-ups are needed not only for data in electronic format, but also for important paper documents, such as log books of the survey examinations.

12.6 Recommended standards, techniques and tools

12.6.1 Database

The national HES database can be structured in different ways depending on the available facilities. The database should be cre-

ated using a well-established database management platform, designed for scalability and extensibility.

Recommended standards for database construction are:

- Well-established relational database management systems, such as Oracle (http://www.oracle.com), PostgreSQL (http://www.postgresql.org), MySQL (http://www.mysql.com) or Microsoft SQLserver (http://www.microsoft.com/sqlverver).
- Language to implement database structures and logic: ANSI SQL (http://en.wikipedia.org/wiki/SQL#Standardization).
- Standard database connection interfaces, such as ODBC (http://en.wikipedia.org/wiki/ODBC) and JDBC (http://en.wikipedia.org/wiki/JDBC).

12.6.2 Development tools, use of statistical software and XML

Several programming languages and development environments, as well as dynamic web content technologies exist to be used to implement an appropriate application logic and user interface for the national HES database ranging from Java (http://www.oracle.com/us/technologies/java/overview/index.html) to .NET (http://www.microsoft.com/net) solutions. The choice will depend on the local facilities, existing practice and the expertise available.

It is recommended to use the XML-based general standards when implementing dynamic web content output and/or interchanging data over the Internet. In the survey data analysis it is recommended to produce analysis and reports by a well-established statistical software, such as R (http://www.r-project.org/) or SAS (http://www.sas.com/). SDMX (Statistical Data and Metadata eXchange, http://sdmx.org/) technical standards provide technical specifications for the exchange of statistical data and metadata based on a common information model.

12.6.3 Data encryption

The system should enforce security through data access control and auditing:

- *Access control* to restrict access to the data.
- *Auditing* to log the actions and changes which have been performed, when and by whom.

Data encryption may be necessary, for example to protect data on local computers' hard-drive during the fieldwork, or when transferring the data over network.

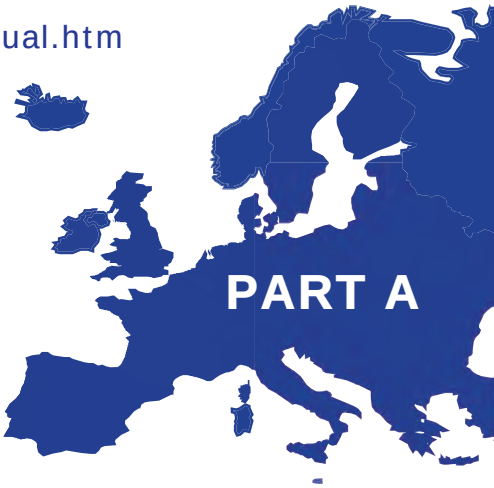
- To secure and encrypt data connection e.g. the following techniques can be used:
  - SSH (Secure Shell network protocol, http://en.wikipedia.org/wiki/Secure_Shell)
  - SSL (Secure Sockets Layer, http://en.wikipedia.org/wiki/Secure_Sockets_Layer) or TLS (Transport Layer Security, http://en.wikipedia.org/wiki/Secure_Sockets_Layer, http://en.wikipedia.org/wiki/Advanced_Encryption_Standard)
  - AES (Advanced Encryption Standard)
  - Blowfish (Fast block cipher, http://en.wikipedia.org/wiki/Blowfish_(cipher))
- A web application can be built on an information server with SSL/ TLS support to ensure encrypted connections to the server. Recommended protocols for data transfer between client and web server are HTTP/ HTTPS (by SSL/ TLS) and FTP through SSH.

12.7 Local data management and the EHES Reference Centre

The collection of anonymous individual level data from each country to a centralized database at the EHES Reference Centre is necessary for data quality assessment and for assessing the success of the standardization and documentation of country-specific characteristics of the data. This is described in Part C of the EHES Manual.

References

- Gumm HP. Encoding of numbers to detect typing errors. Int J Appl Engng Ed. 1986;2:61-65.



13. Recruitment of participants

Sanna Ahonen[1], Päivikki Koponen[1], Hanna Tolonen[1], Kari Kuulasmaa[1]

[1] National Institute for Health and Welfare (THL), Helsinki, Finland

This chapter provides general guidelines for recruitment process, recruitment methods, definition of participation rate and non-participant data. The strategy and methods for recruitment have to be determined by each country based on national and regional feasibility and legislation, the survey budget, and cultural norms.

A high participation rate is fundamental to the reliability and validity of the survey. The participation rate depends directly on the success of recruitment. Proper recruitment is also necessary for the HES to be ethically acceptable. Description of the recruitment process is a key element in the research proposal reviewed by the ethical committees. The recruitment process needs to be carefully prepared and piloted.

13.1 Recruitment process

The purpose of the recruitment process is to ensure as high participation rate as possible. The recruitment process includes all stages, where selected participants are contacted to provide information and to make appointments for examination visits. The recruitment process varies between countries and should be planned according to what is the most feasible way in each country. In some countries, there are also legal restrictions regarding to contacting potential participants.

13.1.1 First contact attempt

IIt is important to obtain as high participation rate as possible already with the first contact attempt. A successful first contact, without the need for additional attempts, saves costs. The first contact attempt can be made for example by an invitation letter

and information leaflet and then be followed by a phone call in order to schedule an appointment. The written material that is used in recruitment, can be divided as follows:

- **Information leaflet:** A leaflet that contains key information on the survey in a concise form, targeted to selected persons or also targeted more widely to stakeholders. It should be visually attractive, but easily distinguishable from advertising materials. It typically addresses:
  - Objectives of the survey
  - Brief description of the measurements
  - Importance of the survey for improving public health
  - Importance of participation
  - Benefits for the participant
  - Information on receiving personal results and how the survey results will be reported
  - Information on partners and financial support
  - Name and signature of the leader of the survey
  - Strict confidentiality of survey data
  - Website address for more information and possibly for scheduling appointment
  - Contact information (toll-free number for more information, e-mail)
  - See an example of the information leaflet (at the end of this Chapter).

- **Invitation letter:** This is a personal invitation to participate in the survey. The invitation letter can be short, if other relevant information is given in an attached information leaflet.
  - Information on how the person was selected.
  - Pre-scheduled appointment time (with contact information for rescheduling) or instructions how to schedule the appointment.
  - Name and signature of the survey leader (or other important/respected person)
  - See an example of the invitation letter without pre-scheduled appointment time (at the end of this Chapter) and with pre-scheduled appointment time (at the end of this Chapter).

It is also possible to combine the information leaflet and the invitation letter.

- **Instructions to participant:** Instructions on practicalities regarding the participation.
  - Includes information on examination (fasting, instruction to the examination site etc.) (see Part B, Section 2.1. Instructions to the participants).
  - These instructions may also be included in the invitation letter.

- **Information sheet**
  - Provides the necessary information to participant before obtaining informed consent (see Part A, Chapter 4. Legal, ethical and data confidentiality issues.)

Invited participant's response to the first contact attempt highly depends of the contents and the format of the written materials. The format of written materials should be informative, but also easy to understand, even by participants with a slight linguistic or cognitive impairment. The format, length and wording of the invitation could be modified according to the age of the participants. The material should be translated into all relevant languages.

13.1.2 Re-contact attempts

Figure 13.1 shows different responses to contact attempts. Regardless of how successful the first contact attempt is, at least 1-3 re-contacts should be made if feasible and not restricted by national legislation. In all types of surveys, incorrect addresses and difficulties to obtain telephone numbers are well known problems. Many of the persons who do not show up after the first contact attempt, probably simply did not receive the invitation. If feasible, the envelope of the invitation letter could include a note for the post office informing that for recipients who have moved, the letter should be returned with information on the new address rather than being forwarded. For re-contact attempts, the accuracy and the recentness of the contact information must be checked, if possible.

The re-contacts may consist of a letter (with or without the questionnaire), phone calls or home visits, depending on the cultural acceptability. Among persons below a certain age, the most suitable means of contact is usually a mobile phone, where as fixed telephone-lines are often the best way to reach older adults, although this is country specific. A personal approach is usually more effective than a second letter of invitation and allows the scheduling of the appointment to be "tailored". If a second let-

ter of invitation is sent, it should be modified (e.g. introduction, signature, format) in respect with the first letter of invitation. The hours in which the measurements are taken can be made more flexible (early mornings, evenings, weekends, drop-in visits). Home measurements or visits to institutions (e.g. hospitals, nursing homes, prisons) may be offered if the person is unable (e.g. health condition) to participate otherwise. Reimbursements, incentives or small gifts additional to those used in the first invitation should be considered. If a selected person refuses to participate in the survey, it should be respected and recruitment attempts should end at that point. However a short non-participant questionnaire may be offered to those who refuse.

Substitution of a non-contact with, for example, a neighbour or a person with similar characteristics (e.g. sex and age), is not acceptable (see also Part A, Chapter 3. Sampling techniques). Obtaining information from proxies for the interview components of the HES is generally not acceptable (e.g., information on health issues provided by the spouse for a person working abroad). However, if there is a separate questionnaire to collect information on non-participants,this may be answered by proxy if the selected person cannot be reached or is otherwise incapable to answer it (see an example of the non-participant questionnaire at the end of this Chapter). In addition parts of the interview can be answered by a proxy if the selected person is unable to answer due to e.g. limited cognitive functions (see Part A, Chapter 8).

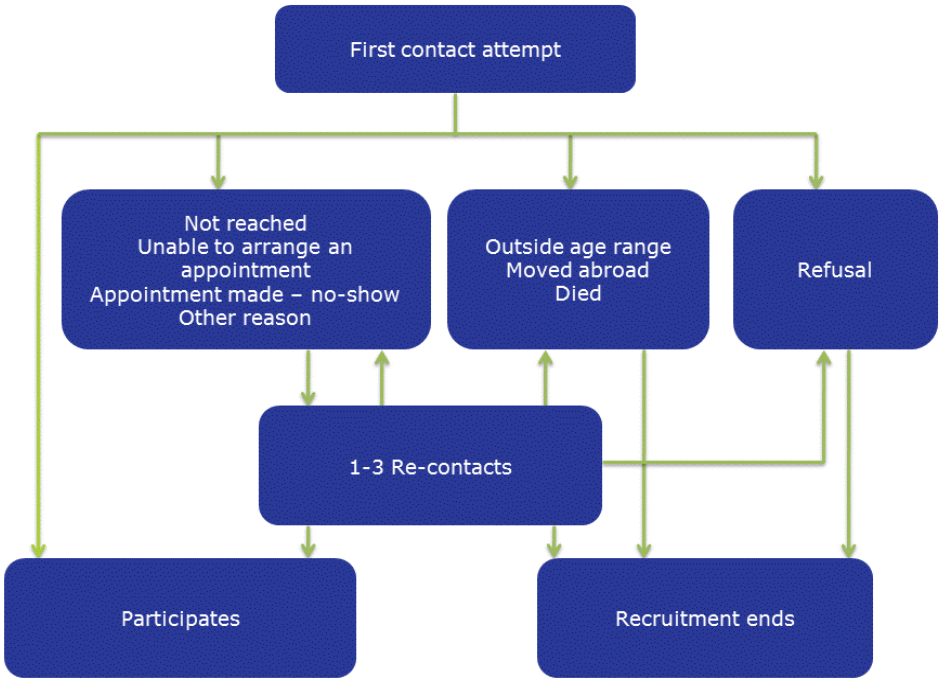


Figure 13.1. Responses to contact attempts

13.2 Participation rate

Participation rates should be calculated separately for the interview/questionnaire information and examinations whenever feasible.

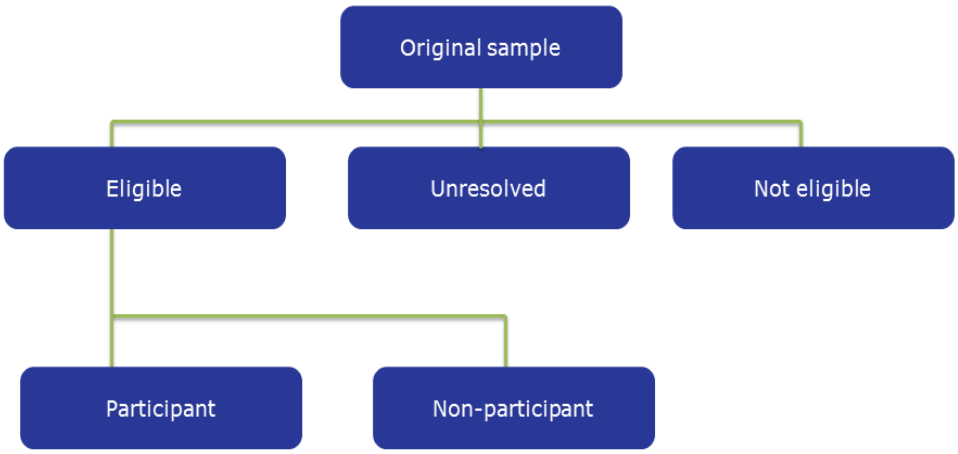
13.2.1 Definition

Figure 13.2 shows the classification of the original survey sample. The definitions are:

- **Eligible:** A person is coded as eligible, if she/he belongs to the target population (see Chapter 02. Target population and sample size).
  - **Participant:** An eligible person is coded as participant if she/he has at least one valid examination measurement, such as height and weight, in addition to some questionnaire results.
  - **Non-participant** is a person, who refused or otherwise did not participate after the invitation was assumed to have been received or other contact was established. (Tolonen 2005, Wolf 2005)

- **Not eligible:** A person selected to the sample is coded as not eligible if she/he does not belong to the target population. This includes over-coverage of the sampling frame (i.e. possible persons who are in the sampling frame although they do not belong to the target population, e.g. not within the age limits) and persons who died or moved out of the primarily sampling unit (PSU) prior to the scheduled examination. The reason for being not eligible should be recorded. Persons who were temporarily absent during the survey period because of work, studies, tourism, hospitalization, or for other reasons are part of the target population and therefore eligible.
- **Unresolved:** There may be persons whose eligibility status cannot be resolved. In a typical case,
  1. the invitation letter was returned to the survey administration indicating that there is no such person in the address; AND
  2. other contacts were not possible or not successful; AND
  3. no information was available to assess the eligibility status.

Although it may be likely that the person does not belong to the target population, there is no certainty about this. The number of unresolved persons is usually small, but may be substantial in some countries, where good sampling frames are not available.



The formula to calculate participation rate (PR) and its fractions co-operation rate and contact rate are shown in Table 13.2.

Table 13.2 Calculating participation rates

Participation rate = (number of participants) / (number of eligible AND unresolved)
Co-operation rate = (number of participants) / (number of eligible)
Contact rate = (number of eligible) / (number of eligible AND unresolved)
Note that Participation rate = Contact rate x Co-operation rate

13.2.2 Target participation rate

Recruitment efforts should be geared towards obtaining the highest possible participation rate so that the sample will represent the target population. The target rate of participation should be at least 70%, but preferably higher. If there is indication that participants differ from non-participants on important variables such as health factors, the rate should be closer to 80% or over (Tolonen 2005). It is known that non-participants are more often young, men and from lower socio-economic class when compared to participants (Shahar 1996, Jackson 1996, Tolonen 2005, Eaker 1998). Non-respondents have also worse health profile, more psychological disorders (van der Akker 1998, Shakar 1996), are more often smokers (Tolonen 2005, Barchielli 2002, Macera 1990) and have higher total and cause-specific mortality than participants (Cohen 2002, Hara 2002, Harald 2007, Jousilahti 2005).

Previous HESs have shown great variations between participation rates among European countries. Only a few surveys have

reached participation rates of 70% or higher during the last few years (HIS/HES database). This is why special attention should be given to developing actions which may help to obtain the highest possible participation rates.

13.2.3 Ways to increase participation

13.2.3.1 Selection and training of personnel

Competent and motivated survey personnel play an important role during the recruitment process. The selection of fieldwork personnel has to be based on general requirements and competences needed to carry out the fieldwork tasks, as stated in Part A, Chapter 9. Selection of fieldwork staff. Good social skills, especially good communication skills, are prerequisites when selecting survey personnel. After selecting competent personnel, sufficient training must be provided (See Part A, Chapter 15. Training programme). It is important, that the personnel responsible for recruitment is familiar especially with the following issues:

- Understand the importance of a high participation rate to survey quality
- Ways/ actions how to motivate participation
  - Know the correct answers to frequently asked questions
  - Know what options can be offered in case of difficulties in scheduling a visit (e.g. weekend and evening hours, drop in or home visits)

If there is a need to motivate the personnel responsible for recruitment, a bonus may be considered to be offered, if feasible, for high participation rates in districts or age groups where participation is expected to be lower.

13.2.3.2 Factors affecting participation rate

There are several actions that can be used to reach the target participation rate. In addition to the importance of the survey for serving public health and research, potential participants are also interested in personal benefits. For some participants a possibility to receive information on their own health status and risks may be an important reason for participation. Therefore, inclusion of additional examinations which offer more information to the participants on their health status should be considered. Also, the use of incentives may be an important motivator to participate in the survey, especially for population groups that are challenging to recruit otherwise.

The feasibility of personal contacts differs between countries, e.g. due to differences in the availability of telephone/mobile phone numbers and acceptability of home visits. Contacting potential participants by phone or mail may be challenging due to people's negative attitudes caused e.g. by numerous telemarketing calls and junk mail (Samanic 2003, Sinicrope 2009). Using media and different personal contact methods such as telephone calls, home visits, and reminders before appointment (phone call/ text message reminders) may help to raise the participation rate (Heistaro 2008). Home visits are usually efficient in recruiting persons who are unable or unwilling to participate otherwise (Heistaro 2008). Participation should be facilitated through flexibility: rescheduling of an appointment, prolonged opening hours, offering appointments also on weekends, possibility for selected persons to drop in without an appointment, and easy access to the examination site. Factors that may affect the participation rate are gathered in Table 13.1. It should be taken into account that the effect of some actions varies between cultures and population groups and also within countries.

Table 13.1. Factors that may affect participation rates

Factor	Possible effects to participation rate
Pre-notification	Pre-notification prior to invitation to participate in the survey usually raises the participation rate (Phillips 2002, Spry 1989).
Phone call	Phone contact is an effective way of increasing participation rate (Heistaro 2008).
Multiple contacts	Multiple contacts significantly increase participation rates (Porter 2004).
Flexibility in scheduling appointment	Offering evening and weekend times, drop in visits and different locations for measurements increases participation especially among busy people (Heistaro 2008).
Relevance and importance	Survey relevance and importance to the survey recipient is an important factor when designing surveys and key messages. Highly relevant surveys raise the participation rates. (Porter 2004, Phillips 2002)
Personal fulfilment	Feeling valued and appreciated increases the willingness to participate (Phillips 2002). Signature or introduction in the invitation letter written by a respected person may increase the feeling of being valued.
Statements of confidentiality	Loss of privacy when providing biologic specimens can be a major concern affecting participation rate. This is why it is important to explain confidentiality issues to the participants (Samanic 2003).
Requests for help	People with personal appeal to altruism tend to follow a norm of social responsibility and may be more willing to take part in the survey, if a phrase "it would really help us..." is used in the invitation (Porter 2004, Sinicrope 2009).
Sponsorship	Surveys sponsored by academics or governmental organizations have higher participation rates in general than surveys sponsored by commercial organizations (Porter 2004).
Mass media campaigns	Raising public awareness about the survey: the importance in national, community and individual levels.

Factor	Possible effects to participation rate
Home visits	Home visits raise the participation rate if a person is unable (e.g. difficulties in functional capacity) or unwilling to participate otherwise (Heistaro 2008) , or when people prefer home visits in their health services.
Domestic vs international use of research samples	Participants may be more willing to allow samples to be used for domestic rather than international studies (Tupasela 2009).
Several languages	Using several languages helps in recruiting ethnic minorities (Sproston & Mindell 2004)
Incentives	The use of compensation or small “thank-you gifts” for participation (financial or other) may be considered. Prepaid incentives (paid with the survey itself) raise participation, while postpaid (paid after the survey) usually don’t (Porter 2004). Long survey with incentives can make it achieve the same participation rate as a shorter survey without incentives (Groves 1999). The effect of incentives may depend on cultural norms.
Survey environment and background	Economic and social environments may affect by lowering or raising the participation rate; e.g. lower socio-economic groups tend to have lower participation rates (Harald 2007, Porter 2004).
Feedback from focus groups	Discussions in focus groups (small groups with representatives of potential participants) may produce important information for planning leaflets and invitations in a way that they raise interest to participate (Sinicrope 2009, Samanic 2003).
Internet survey vs. paper survey	Participation rate may be even higher in web survey compared to paper survey, but it depends on the population and the design of the web survey (Porter 2004). Typically web surveys can be used as an additional data collection method, as there are currently only few populations where most sampled persons can be reached with web surveys. (see also Part A, Chapter 8. of the EHES Manual.)
Length of a questionnaire form	Long questionnaire forms (several pages) may have lower response rates than short forms (1-2 page), but only moderate effect (Porter 2004).
Deadline	Deadlines (giving respondents a deadline) haven’t shown important effects on either increasing or decreasing the participation rate (Porter 2004).

13.2.3.3 Partnership for enhancing participation

Partnership and collaboration with local organizations, professionals and communities help to raise awareness of the importance of the survey, and to arrange easy access to the examinations.

- The employers of the participants can be encouraged to allow their employees to participate in the survey during working hours.
- Cooperation with regional or local hospitals, non-governmental organizations, research centres and universities may increase the interest in participation.
- National and local health authorities and health professionals must be informed prior to the survey.

- Local community leaders need to be notified to ensure the community's understanding and support.
- The public should be notified using mass media around the same time that the invitations are sent. (See Part A, Chapter 14. Dissemination and publicity)

13.3 Non-participation

In order to assess the non-participation bias, it is important to collect information on non-participants to evaluate potential biases in estimates (Harald 2007, Jousilahti 2005, WHO MONICA Project 1997). This is important even when the participation rate is high. Some key information, such as age, sex and possibly some aspects of social status can in most countries be obtained already from the sampling frame or other registries through record linkage. Other key information, and also these if not otherwise available, should be asked using a non-participant questionnaire (see example of the non-participant questionnaire at the end of this Chapter). The questionnaire may be sent by mail or e-mail, or it can be filled in during a telephone interview or home visit. If the invited person is not available (by phone, e-mail or other means), proxy information may be used for completing the short non-participant questionnaire.

13.4 Data to be recorded

It is necessary to keep a record of the participation status of each person invited to participate in the HES. The number and type of contact attempts should be recorded. If the person was contacted, it should be recorded if the person participated, refused or dropped out after having agreed to participate. Information on completed and not completed examinations should be recorded. If the person refused, the reason should be recorded, if this information can be obtained. Reasons for not being examined are listed below (modified from the Health 2000 survey in Finland). Some of these reasons should not be easily accepted, and it should be attempted to convince the person that his/her participation is highly valued.

Reasons for not being examined:

- Refused: no reason given
- Refused: lack of time
- Refused: personal principle
- Refused: health problem (e.g. disability restricting access to the examination site or is hospitalised)

- Refused: feeling healthy (therefore thinks that there is no reason to participate)
- Refused: survey topic (is not interested in health issues or considers this too personal)
- Contacted: not able to schedule an appointment (e.g. participant could only attend during evening hours or week ends)
- Contacted: no show (does not come to the scheduled visit, and the visit cannot be re-scheduled)
- Not contacted: not reached (no address/phone number available, outdated information)
- Not eligible: moved abroad
- Not eligible: moved out of the primary sampling unit (PSU)
- Not eligible: age out of survey range
- Temporarily unavailable: e.g. holiday
- Language problems
- Not eligible: died
- Impossible to examine for other reason (this reason should be specified, if feasible)

References

- 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 Evaluation and Models. Helsinki: National Public Health Institute; 2003.
- Barchielli A, Balzi D. Nine-year follow-up of a survey on smoking habits in Florence (Italy): higher mortality among non-respondents. Int J Epidemiol 2002 Oct; 31(5): 1038-42
- Cohen C, Duffy JC. Are nonrespondents to health survey less healthy than respondents? Journal of Official statistics 2002;18(1):13-23
- Eaker S, Bergstrom R, Bergstrom A, Adami HO, Nyren O. Response rate to mailed epidemiologic questionnaires: a population-based randomized trial of variations in design and mailing routines. Am J Epidemiol 1998 Jan 1; 147(1): 74-82
- European Health Interview & Health Examination Surveys Database. 2010 [cited 7 April 2010];Available from: https://hishes.iph.fgov.be
- Groves RM, Singer Eleanor, Corning AD, Bowers A. A laboratory approach to measuring the effects on survey participation of interview length, incentives, differential incentives, and refusal conversion. Journal of Official Statistics 1999;15(2):251-268.
- Hara M, Sasaki S, Sobue T, Yamamoto S, Tsugane S. Comparison of cause-specific mortality between respondents and nonrespondents in a population based prospective study: ten-year follow-up of JPHC study cohort I. Japan Public Health Centre. J Clin Epidemiol 2002 Feb; 55(2):150-6

- Harald K, Salomaa V, Jousilahti P, Koskinen S, Vartiainen E. Non-participation and mortality in different socioeconomic groups: the FINRISK population surveys in 1972-92. Journal of epidemiology and community health. 2007 May;61(5):449-54.
- Heistaro S. Methodology report. Health 2000 Survey. KTL-National Public Health Institute, Finland. Helsinki 2008.
- Jackson R, Chambless LE, Yang K et al. Differences between respondents and nonresponsents in a multicentre community-based study vary by gender ethnicity. The Atherosclerosis Risk in Communities (ARIC) Study Investigators. J Clin Epidemiol 1996 Dec; 49(12): 1441-46
- Jousilahti P, Salomaa V, Kuulasmaa K, Niemela 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. Journal of epidemiology and community health. 2005 Apr;59(4):310-5.
- Kaila-Kangas L, ed. Musculosceletal Disorders and Diseases in Finland. Results of the Health 2000 Survey.
- 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, et al., eds. Health surveys: evaluation and recommendations Subproject reports of Phase 2 of the project Health Surveys in the EU: HIS and HIS/HES Evaluation and Models. Helsinki: National Public Health Institute 2003.
- Macera CA, Jackson KL, Davis DR, Kronenfeld JJ, Blair SN. Patterns of nonresponse to a mail survey. J Clin Epidemiol 1990; 43(12): 1427-30
- Phillips M, Woodward C, Collins D, O'Connor W. Encouraging and maintaining participation in the families and children survey: Understanding why people take part. 2002. Working paper number 6. National Centre fo Social Research. Available from: http://research.dwp.gov.uk/asd/asd5/WP6.pd
- Porter AR. Raising Response Rates: What works, New Directions for Institutional Research 2004;121:5-21
- Samanic CM, Colt J, Campbell W, WH Chow. Attitudes towards participation in a population based epidemiologic study: Results from focus groups. Annals of Epidemiology 2003;13(8):559.
- Sharar E, Folsom AR, Jackson R. The effect of nonresponse on prevalence estimates for a referent population: insights from a population-based cohort study. Atheresclerosis Risk in Communities (ARIC) Study Investigators. Ann Epidemiol 1996 Nov; 6(6): 498,506
- Sinicrope PS, Patten CA, Bonnema SM, Almquist JR, Smith CM, Beebe TJ, Jacobsen SJ, Vachon CM. Healthy women's motivators and barriers to participation in a breast cancer cohort study: a qualitative study. AEP 2009;19(7):484-493.
- Sproston K and Mindell J (eds). Health Survey for England 2004. Volume 2: The health of minority ethnic groups. Methodology and documentation. The Information Centre, 2006a.
- Spry VM, Hovell MF, Sallis JG, Hofstetter CR, Elder JP, Molgraad CA. Recruiting respondents to mailed surveys: Controlled trials of incentives and prompts. American Journal of epidemiology 1989;130(1):166-172.
- Tolonen H. Towards the high quality of population health surveys. Standardization and quality Control. National Public Health Institute A27/2005, Helsinki 2005 [cited 1 January 2010]; Available at http://urn.fi/URN:ISBN:951-740-546-4. Accessed on 22 January 2013
- Tolonen H, Dobson A, Kulathinal S. Effect on trend estimates of the difference between survey respondents and non-respondents: results from 27

populations in the WHO MONICA Project. European journal of epidemiology. 2005;20(11):887-98.

- Tolonen H, Kuulasmaa K, Laatikainen T, Wolf H. Recommendations for indicators, international collaboration, protocol and manual of operations for chronic disease risk factor surveys. EHRM Project. 2002 [cited 1 January 2010]; Available a http://www.thl.fi/publications/ehrm/product2/title.htm. Accessed on 22 January 2013
- Tupasela A, Sihvo S, Snell K, Jallinoja PA, Aro AR and Hemminki E. Attitudes towards biomedical use of tissue sample collections, consent, and biobanks among Finns. Scand J Public Health 2010; 38; 46 originally published online Nov 11, 2009;DOI: 10.1177/1403494809353824.
- van den Akker M, Buntinx F, Metsemakers FK, Knottnerus JA, Morbidity in responders and non-responders in a register-based population survey. Fam Pract 1998 Jun; 15(3): 261-3
- WHO MONICA Project. MONICA Manual, Part III: Population Survey, Section 1: Population survey data component. 1997 [cited 01 January 2010]; Available at http://www.thl.fi/publications/monica/manual/part3/iii-1.htm. Accessed on 22 January 2013
- Wolf HK, Kuulasmaa K, Tolonen H, Sans S, Molarius A, Eastwood BJ. Effect of sampling frames on response rates in the WHO MONICA risk factor surveys. European journal of epidemiology. 2005;20(4):293-9.

Thank you for your help with this survey. Your co-operation is very much appreciated.

For further information, please contact:

Name of the contact person
Address
Toll-free phone number
E-mail address
Survey web page

EHES
Measuring the Health of Europeans

EHES

Health Examination Survey

Examples of the information leaflet, invitation letter and non-participant questionnaire

What is this survey about?

The intention of this health examination survey is to receive up to date health information of the adult population of x (*country*). The information gathered will be used for planning health care as well as assessment of the prevalence of diseases, their causes, and care. The survey is being carried out by x and y (*name of the organization/partners*).

Why am I selected?

We have invited 4000 people to take part in the survey. You have been randomly selected from the national population register.

Why is it important to participate?

This survey is important for improving public health. As you have been selected to the sample, your participation is very important. As the selection has been made by random, it is not possible to replace a selected person by anyone else.

Do I benefit from the survey?

Yes, you will receive important information about your health. During the examinations you have a change to get feedback on results and talk to the health professional. After the examination you will get a report of your results in mail. The health examination is free of charge.

All participants will receive xx (*if incentives are used*).

What measurements are included?

The measurements include height, weight, waist circumference, blood pressure and x (*list additional measurements*). Also a blood sample will be taken to measure total and HDL-cholesterol and glucose.

The measurements are safe and are made by specifically trained and qualified personnel.

Is the survey confidential?

All survey data is confidential and protected by legislation (Data Protection Act). This means that survey results will not be presented to reveal your identity at any point.

Is the survey compulsory?

Participation is completely voluntary. The success of the survey relies on the co-operation and goodwill of those asked to take part. The more people take part, the more useful the results are. You may withdraw from the survey at any point.

Whom can I contact to ask further questions?

We will help you with any questions or concerns you may have. Please call us at xx-xx-xxx (*toll-free phone number*). The survey website at http://www.hes.xx also has more information.

28 April 2011 *Study ID*

Mr./Ms. First name Last name
Street address
City

Dear Mr. /MS. Last name,

We are inviting you to participate to the Health Examination Survey of country x (*substitute with the survey name)*. This survey studies the health of population in country x (*replace with your country)*. You have been selected from national population register to represent 25-64 years old people of the country (*replace with your country)*.

In the survey, an interview will be conducted and your height, weight, waist circumference, blood pressure will be measured and blood sample collected.

Representativeness and usefulness of the results of the survey depend on people we contact to get involved. It takes 30-45 minutes to go through the interview and measurements. You cannot be replaced by anyone else. Your participation is voluntary.

All information collected during the survey, will be handled confidentially. You can find answers to the questions regarding the survey from attached leaflet. You can also call on Monday-Friday at 9:00-16:00 to TOLL-FREE-PHONE-NUMBER if you have any questions.

Our survey team will contact you within next few days to arrange the appointment time for you.

The HES survey team thanks you for your collaboration.

Sincerely,

Mark Model, Dr.
Project Leader

Susie Super, PhD
Head of Department

28 April 2011 *Study ID*

Mr./Ms. First name Last name
Street address
City

Dear Mr. /MS. Last name,

We are inviting you to participate to the Health Examination Survey of country x (*substitute with the survey name)*. This survey studies the health of population in country x (*replace with your country)*. You have been selected from national population register to represent 25-64 years old people of the country (*replace with your country)*.

In the survey, an interview will be conducted and your height, weight, waist circumference, blood pressure will be measured and blood sample collected.

Representativeness and usefulness of the results of the survey depend on people we contact to get involved. It takes 30-45 minutes to go through the interview and measurements. You cannot be replaced by anyone else. Your participation is voluntary.

All information collected during the survey, will be handled confidentially. You can find answers to the questions regarding the survey from attached leaflet.

We have booked you an appointment for the examination clinic (*provide address of the clinic*) on

6 May 2011 at 8:30.

If this time is not suitable for you, please call on Monday-Friday at 9:00-16:00 to TOLL-FREE-PHONE-NUMBER to schedule new appointment.

Please, read the instructions to the participants leaflet attached to this invitation before coming to the examination clinic.

The HES survey team thanks you for your collaboration.

Sincerely,

Mark Model, Dr.
Project Leader

Susie Super, PhD
Head of Department

Example

Non-participant questionnaire

Version: 22 March 2011

Identification

Participants identification code:

Background information

Sex

- ☐ Man
- ☐ Woman

Date of birth (dd.mm.yyyy)

Age (in full years)

Educational level

What is the highest education leaving certificate, diploma or education degree you have obtained? (Please, include any vocational training)

- ☐ No formal education of below ISCED 1
- ☐ Primary education (ISCED 1)
- ☐ Lower secondary education (ISCED 2)
- ☐ Upper secondary education (ISCED 3)
- ☐ Post-secondary but not-tertiary education (ISCED 4)
- ☐ First stage or tertiary education (ISCED 5)
- ☐ Second stage of tertiary education (ISCED 6)

Health status

How is your health in general?

- ☐ Very good
- ☐ Good
- ☐ Fair
- ☐ Bad
- ☐ Very bad

Diagnosed diseases

Do you have or have you ever had any of the following diseases or conditions, diagnosed by a medical doctor?

Myocardial infarction	☐ Yes	☐ No
Coronary heart disease (angina pectoris)	☐ Yes	☐ No
High blood pressure (hypertension)	☐ Yes	☐ No
Evelated blood cholesterol	☐ Yes	☐ No
Stroke	☐ Yes	☐ No
Diabetes	☐ Yes	☐ No

Height and weight

How tall are you without shoes? (cm)

How much do you weight without clothes and shoes? (kg)

Smoking status

Do you smoke at all nowadays?

- ☐ Yes, daily
- ☐ Yes, occasionally
- ☐ Not at all

Reason for non-participation

Why did you not participate to the survey?

- ☐ Not interested
- ☐ No time
- ☐ Not able to get suitable appointment time
- ☐ I'm healthy, no need to participate
- ☐ I'm too ill to participate
- ☐ Don't participate to any surveys

14. Dissemination and publicity

Hanna Tolonen[1], Jari Kirsilä[1]

[1] National Institute for Health and Welfare (THL), Helsinki, Finland

Words dissemination, communication and publicity are closely linked to each other and ofter used in overlapping meeting. Here we define these terms as follows:

- Dissemination is a a transmission of the information to the public without direct feedback from the audience. Dissemination can be for example a seminar presentation, newsletter or newspaper article.
- Communication is activity of conveying meaningful information in the way that there is specified sender of the information, message and also intended receiver for the message. Difference between dissemination and communication is blurred.
- Publicity is the deliberate attempt to manage the public's perception of the issues. Promotion of the survey can also be considered as one form of publicity.

Dissemination, communication and publicity are all needed and important aspects of health examination survey (HES). With these, survey organizers keep all stakeholders informed about the survey plans, progress and outcomes.

Dissemination of the information has four purposes (Nelson 2009):

- to increase knowledge,
- to instruct,
- to facilitate informed decision making, and
- to persuade.

14.1 Why dissemination and publicity is needed?

Dissemination and publicity are needed to

- Supporting the decisions to carry out the HES by effectively communicating its benefits
  - Support the fund raising
- Supporting the collaboration with local authorities by constructing active lines communication
- Motivating the invitees
  - letter of invitation and background materials
  - General public - information about the HES, health topics to media and public discussion
  - NGOs (employer organizations, trade unions, GPs, etc.)
- Disseminating and marketing the results to
  - participants (ethical approval for the provided information - link to ethical chapter)
  - Funders
  - General public
  - Policy makers
  - Health care authorities
  - NGOs
  - Scientific community

14.2. Target groups for dissemination

A national HES has number of target groups for dissemination who need to be informed about the importance of the national HES, its progress and results. The important groups, key stakeholders whose support or consent is essential to carring out the HES successfully varied between countries but may include (in alphabetical order):

- funders,
- general public,
- health care authorities,
- health care personnel,
- invitees, their relatives, intimates,
- media,

- non-governmental organizations (NGOs),
- policy makers, and
- research community.

Funders are needed in order to be able to conduct a national HES. There can be different kinds of funders for the survey. Often, the core part of the survey, which is mainly used for the health monitoring in the country is funded through the national ministries but other additional modules of the survey may have other funders, like heart foundations, insurance companies, etc. For potential funders, it is important to know about the survey plans well in advance so that they can make funding decisions in time in relation to the survey planning. It should be noted that nowadays some funders require a dissemination plan for the projects they fund as part of the funding application. This plan is also used in evaluation of the applications and outcomes of the funded projects.

General public is a key partners in the HES. Without collaboration from general public, it is not possible to obtain reliable and representative survey results. Therefore, informing general public about the survey may help to increase the awareness about the survey and through that, the participation to the survey.

Collaboration with local health care centres, general practitioners/physicians etc. is needed when conducting a national HES. Local physicians should be kept well informed about the survey. People invited to the survey may contact their own physician to ask more about the survey and should they participate. Especially people, who see they doctor in regular basis for the treatment of chronic condition may be in contact with they doctor. In case like this, it is important that the doctor is aware about the survey and can tell that measurements on the survey are meant to estimate population's health not the health of the individual and that they won't replace the regular visits to the doctor.

Also in come countries, survey measurement results are communicated back to the participant through their own doctor. If this is national practice, some times required by law, the involvement of the local physicians is needed for the success of the survey.

Media is not directly gaining from the HES but they are delivering information about HES to other stakeholders.

NGOs like national heart foundations and diabetes associations, etc. are stakeholders of the surveys for number of reasons. They can be partial funders, but also they usually have good channels to distribute the information about survey to the general population and use survey results to prevention activities. Additional to

that, among NGOs there may be needed knowledge to formulate correct questions to survey questionnaires, to select the most reliable and accurate measurement devices and also to obtain persons doing the analysis from the collected data.

Policy makers can be in different levels. For example national health policy makers, regional health authorities, or heads of local health care centre. All of them need valid and reliable information about the health behaviors and health status of the population as bases if their policy making, evidence-based policy making. As the HES will provide representative and reliable health information, policy makers in all levels are considered end users of the results. Often policy makers, like ministries of health are also funders of the HES.

Research community can have valuable input to the contents of the survey as well as methodological input. Also after the survey has been completed, research community is needed to analyze the results and disseminate the scientific finding to the research community for further use.

14.3 Message to disseminate

What the message to be disseminated includes is dependent on the target group and is the message distributed during the planning phase of the survey, when the survey is already of the field and after the survey has been finished and results are disseminated. In articulating the messages, they should be easy to remember and repeat in various situations so that members of the national organizing team are all providing a coherent message.

Therefore the letter of invitation should give all the essential information the invitees need to give their consent: what is the aim of the survey, to what uses will the results be put, why are they important, what will the invitees personally gain, what are the possible risks (even minor ones), how well has privacy been protected and how to cancel one's participation.

A leaflet can easily be annexed to the letter of invitation, and it is a useful way to give basic information to your stakeholders whenever we meet them:

- Why is this survey important?
- Who is doing the survey?
- Who is asked to participate?
- How will the examinations be conducted?
- Does the selected person have a choice about participating in the survey?`

- What kinds of measurements and questions will be done/asked?
- What happens to the results?
- Although your participation is voluntary, your cooperation is needed to ensure that the results are reliable and accurate.
- “You have been randomly selected to represent others in your country; your unique contribution is valuable to the study.”
- We are taking every precaution to ensure that the physical measurements and collection of samples is safe for participants, etc.

14.3.2 Examples of the messages

The national key messages may for instance be:

- Health surveys are vital for understanding the health situation and the behaviours of the population, and they provide an evidence-base for health policies.
- Identifying health differences between population groups is a prerequisite in the work to narrow down health inequalities
- To support healthy aging we need to know the current state of health of adults and children.
- The national HES is conducted by a reliable public health authority, the methods are secure and science based, and the results do not serve any other interests but the public benefit.
- Participating in the survey will give participants a free-of-charge opportunity to receive up-to-date information on their own health.
- Information about people’s health is vital to building an efficient health care system geared to our health needs and that of our families. Each individual’s contribution is important in making the study representative.
- The physical health examination survey will verify and complement data collected through other health questionnaires and registries.

14.4 Brand building

To help the dissemination and increasing the awareness about national HES, it is important to pay attention to how the survey is represented. This includes all materials prepared and distributed

about the survey as well as messages given out from the survey. This can be considered as a brand building of the national HES. The brand of the survey is used to identify the survey from others.

The brand building intents to provide a uniform and professional message and image about the survey. This includes issues like:

- name of the survey, acronym
- logo of the survey
- slogan
- colors to be used in survey materials
- fonts to be used in survey materials
- template for the presentations (e.g. PowerPoint)
- template for other published materials
  - letters
  - posters
  - questionnaires
- web site layout
- image of the fieldwork team (name tags, possibly clothing)
- signs how to come to the examination site
- promotion materials
  - pens
  - notepads
  - cups
  - etc.

14.5 Means of dissemination

After defining what we want to say and to whom, we need to choose the most appropriate means of dissemination. There is number of ways to disseminate the information about HES. It can be done through:

- face-to-face discussion,
- seminars,
- letters,
- leaflets and brochures,
- Newsletters,
- press releases or press conferences,

- web sites,
- social media,
- posters,
- reports,
- TV and radio advertisement,
- information desks on public places,
- road shows, and
- promotion materials.

In a public health examination survey the availability of the basic information about the survey is critical. Easy access to good quality information will bring about confidence in the survey, as well as focus the attention of stakeholders. Thus, the quality of the most important outputs of your dissemination needs to be maintained. For example:

- the letter of invitation
- a leaflet about HES
- the national websites
- press materials

Quality in dissemination is measured by how clear and relevant our messages are for our target groups, using language that is accessible and mindful of questions that may arise, i.e. identifying our target groups interests and concerns.

Where computers and the internet are widely available in a country or area, a website is one of the best ways to give basic information on the HES 24/7. Consideration tought to be given to the design of the website. If the website comes across as poorly developed, the impression could easily be that the whole survey is not being done professionally.

The existing websites of co-operating organizations in the host country could serve as an easy solution to communicate to different target audiences. Publishing on the net needs to take account of layout, navigation and architecture of the site, maintaining simplicity wherever possible. Keep pertinent information together in one place, i.e.:

- who is doing the survey, and what are the funding and organizing bodies
- the aims of the survey
- who are to be the participants, how are they invited
- instructions to the participants

- results
- etc.

Pages should be visually effective: use pictures or graphical elements to illustrate your messages. Information should be kept up to date with the project progress, e.g. news section. Consultation with communications experts will help in finding available resources and ensuring the accessiblility of the web pages. Communications like any other human interaction differs from one culture and country to an other. Therefore it is important that you are able to consult national experts in communication.

The criteria for how media outlets pick up their stories and act as partners may sometimes be confusing. The media culture differs from country to an other, even in Europe, and it's important that you have a skilled communication officer in your team. However, reporters usually want to serve their audiences and to keep an eye on authorities, wether national or international. So, it is essential to state the benefits to the home country of the HES clearly in press releases so that the concrete benefits are visible and relevant to the average citizen.

Likewise, possible negative press should be anticipated so as to protect the message from easy distortion by possible critics. Think what is your main message to the media, what might be of interest them (stories), and put that idea right into the headline. Do consider organizing a press conference, if your think that the survey can catch media attention or if it appears to have been misunderstood or misrepresented.

If these means of communication are inadequate, further means can be considered, e.g.:

- telephone information service to the participants
- radio broadcastings, advertisements or supporting programmes
- roadshows, information desks in public places, etc.

14.6 Dissemination plan

All the above mentioned elements should be put into a concrete time scale for the project, saying what will be done, and when and by whom. For instance, will we organize an information seminar for health policy makers to support their decision-making and fundraising, or will we meet them personally? When and what kinds of information materials will the participants receive? Will you inform the local or national media a few days before the participants receive the letters of invitation? How and when do

you communicate the results, by personal letter to invitees, on the national website to policy makers, in a press conference to media? What information material you need in these actions? Do you need national website for the project?

A pre-prepared dissemination plan helps to ensure that the dissemination is done properly, in a timely manner and without giving rise to constant concerns. The survey management or other person defined for this task has to follow up the implementation of the dissemination plan, where a responsible person for each action has been designated.

The outline of the dissemination plan could be:

1. The role and main objectives of dissemination in the national HES. What we want to accomplish with dissemination.
2. Key messages. What is our message to be disseminated in different points of time to different target groups.
3. Key target groups for dissemination.
4. Means of dissemination.
5. Time scale and concrete actions.
6. Organization of dissemination. Who will do, what, when and how.

Dissemination is a strategic part of the survey and may contribute to the success or failure of the entire project. Successful project leaders are fully committed to the goals of dissemination. It is usually very beneficial, that the national survey team includes a skilled communication officer who has an overview of the dissemination strategy of the HES. Her/his duties may include:

- preparation of the dissemination plan;
- follow-up of the implementation of the plan; and
- contacts with the media.

14.6.1 Active vs. passive publicity

Decision about the publicity strategy should be made at the same time when preparing the dissemination plan. Are we targeting actively for publicity or do we have more passive strategy where we wait and see what comes out.

14.6.2 Preparing for unintended publicity

In case of unintended publicity, which is quite often also negative, there should be a plan how to react. This plan should include information who from the team has the main responsibility to respond to the unintended publicity. This responsibility may be divided to several persons by topic. In all communication, the response should come openly, from your own initiative and timely after the unintended publicity. Do not start hiding the issue but confront it calmly with facts. Make sure that the message is same all the time regardless who from the team provides it and on which forum.

References

- Nelson DE, Hesse BW, Croyle RT. Making data talk. Communicating public health data to the public, policy makers, and the press. Oxford University Press, 2009

15. Training programme

Päivikki Koponen[1], Hanna Tolonen[1], Kari Kuulasmaa[1]

[1] National Institute for Health and Welfare (THL), Helsinki, Finland

This chapter outlines the training programme developed by the EHES Reference Centre for the EHES pilots and presents key issues which should be considered when planning and preparing the future EHES surveys and the national training programmes for all staff members who take part in the data collection. It is essential to outline the national training at early stages of the planning process, as this will affect both budgeting (training costs) and timing of the data collection. Training is a key element of standardization and quality assurance (see Part A, Chapter 11).

15.1 EHES training

The EHES training includes two dimensions:

1. Europe wide training seminars for the persons responsible for the planning and organizing the surveys and for those resposible for training of the national survey personnel, and
2. Outline for the training of the national survey personnel actually conducting the survey. This training is conducted nationally.

The EHES training programme also aims to promote the use of e-learning methods and materials targeted to the national survey organizers and trainers of the national survey teams. (http://www.ehes.info/training_programme.htm)

Similar training should be organized periodically, for other countries planning and preparing their first national HES using the EHES standards, and for all EHES pilot countries before their next

survey. Key aims for the EHES training are to ensure standardization and to share experiences between countries.

15.1.1 EHES training programme

The EHES training progamme includes three trainig seminars. Two of these were organized for the EHES pilot and their details are available at http://www.ehes.info/rc/training_seminar/training_seminars.htm. A need for the third seminar has been identified during the EHES pilot project.

Training seminar covering issues relating to the **planning and preparing** for the European Health Examination Survey (EHES) at the national level.

The target group for this seminar are those who plan and prepare national surveys in EU Member States and EFTA/EEA countries. The objective of the seminar is to train the participants for planning a national HES according to the EHES standards. Other objectives of the seminar are to raise awareness on EHES in all European countries, to receive feedback from the participants on the EHES standards and the EHES Manual, and to discuss possible national adaptations. The seminar will support preparing the national manuals and finalizing national study plans.

Training seminar covering issues relating to the **field work** of the national health examination surveys.

The target group for this seminar are those who will train the national fieldwork team members in each country. The objective of the seminar is to promote the use of the standard EHES training materials and to otherwise ensure that the training for the fieldworkers will be organized following the EHES standards. The focus is on the core measurements but also additional measurements can be included, when feasible. The seminar will support finalizing the national manuals and training programmes.

Training seminar focusing on **data analysis, reporting and dissemination of results**.

The target group of this seminar are stastisticians, researchers and survey organiziers reaponsible for the data analysis, reporting and dissemination of the results. The aim is to promote comparsion of the national results, to develop European level reports and to support both national and European dissemination of the results.

15.1.2 EHES training materials

All training materials for EHES will be available at the EHES website http://www.ehes.info. The national survey organizers and national trainers are encouraged to translate, use, develop further and adapt these materials for their national purposes. However, they should keep in mind that the key contents and methods for the national training should be standardized to assure the international comparability.

15.2 National training programme

All members of the national survey team, both those working at the central office and all fieldwork staff members should participate in the national training programme. It is essential for the quality of the survey that everyone, including secretaries and assistants working at the central survey office, those who contact the selected persons, send the invitations and schedule the visits, data managers, statisticians and all field work staff members know and understand the aims of the survey and the whole data collection process.

The key contents of the national training should be similar in all countries, but some parts will depend on how the fieldwork is organized and which additional measurements are carried out in addition to the EHES core measurements.

15.2.1 Outline for the national training seminars

The training should include both general issues for all staff members, general fieldwork skills and practices for the fieldwork staff, and specific training for each selected measurement (Figure 15.1). If the staff members have experience from previous surveys some parts of the general training may be only short refresher lectures. Practical measurement sessions are needed also for the experienced staff members to ensure that the standards are followed correctly.

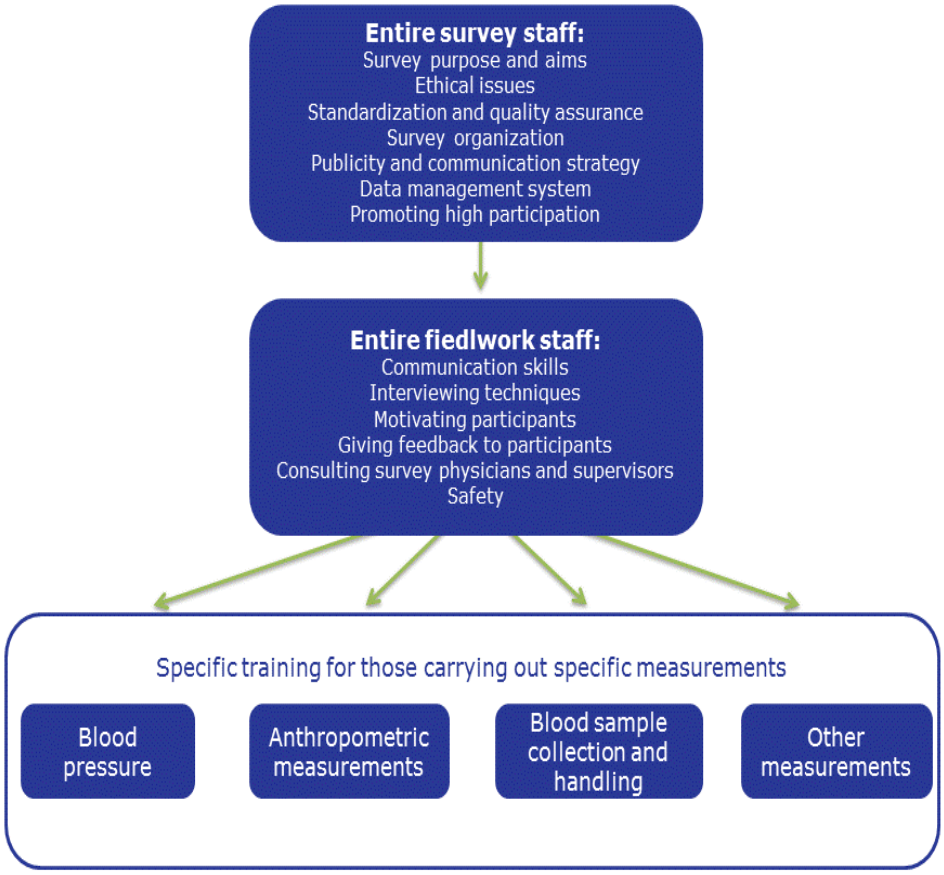


Figure 15.1. Training process for survey staff members

The training should include at least the following topics for all staff members:

- Purpose and aims of the survey: It is important that all staff members understand the importance of the survey and are able to describe the aims and purpose of the survey to the participants in a standard way;
- Ethical issues and confidentiality: What is data confidentiality and how it is assured by all staff members, why an informed consent is needed, what is meant by the informed consent, and how the informed consent should be obtained;
- Random samples and the importance of high participation rates: How people are selected, and why all selected persons are equally important regardless of their health status or other characteristics, how participation can be encouraged and motivated;
- The importance of standardization and quality assurance: Understanding the aims of audit visits and qual-

ity assurance, the role of the survey manuals, the importance of consulting supervisors when needed;

- Survey organization: roles and responsibilities of each staff member at the central office and in the fieldwork teams;
- Communication skills, including similarities and differences in professional conduct during survey data collection and clinical practice in normal health care settings;
- Working with the local health care professionals e.g. to build and maintain good collaboration, so that they encourage their patients to participate in the survey, and referring participants with abnormal measurement results to their GPs or other local health care professionals;
- How the survey results will be reported and published, publicity rules and working with local media during fieldwork;
- The data management system and IT skills for data entry, handling and reporting.

As in most cases all fieldwork staff members have at least some interview questions to ask before or after the clinical measurements, all of them will need training in general interviewing skills, including (adapted from Czaja & Blair 2004):

- rules for accepting proxy responses;
- reading questions verbatim;
- using non directive probes (when allowed);
- asking all questions;
- recording answers correctly, especially in case of open ended questions.

The training for the members of the field work teams who are carrying out the measurements should include at least the following topics:

- Specific procedures for each interview module or instrument;
- Specific measurements: rationale why they are measured, measurement techniques, including practical training and certification if needed;
- Giving feedback to participants concerning measurement results;

- Consulting survey physicians and local health care professionals when needed;
- Safety of the fieldwork team members (e.g. actions needed in case of needle stick injuries, violently acting and aggressive participants).

For example, the personnel responsible for collecting blood samples should be familiarized with the part of the protocol that pertains to blood collection. The safety instructions for protecting the participant and the nurse or technician during the blood sample collection should be reviewed. Similarly those who will carry out the blood pressure measurements need specific information on why standardized blood pressure measurements are needed, what are the key steps in the measurement protocol, how the results are recorded and how the results are explained to the participants. The practical training will include e.g. carrying out adequate number of measurements observed by supervisors and feedback sessions. Detailed guidelines for the training and certification needed for each measurement will be provided in Part B of the EHES Manual.

15.2.2 Selection of the national trainers

The national trainers should:

- have participated in the EHES training seminars targeted to national trainers:
- be well informed both on the aims and purposes of the national survey as well as on EHES standards (e.g. members in the national survey project teams), and
- have specific expertise in the subject area (e.g. survey ethics, blood pressure measurements).

The supervisors and persons with experiences from previous surveys can act as training assistants to train the other team members. They are needed in the practical training and in the role playing sessions.

15.2.3 Use of training materials and different training methods

The traineers should be encouraged to read the survey manuals before the training sessions, during and/or after the training. The survey manuals form the basis for all training. The EHES training materials will usually require national translations and adaptations. The training materials may include standard presentations, videos on interviewing and measurement techniques, and web-based education tools. Giving material to watch and read later at

home will support learning. Newspaper articles and reports from previous surveys (if available) may help to see the importance of the survey and understand how the data will be utilized. An effective training programme will emphasize participatory exercises over lectures (Czaja & Blair 2004). If the fieldwork staff members do not practice their skills in a training session, they will practice them with real participants, which may lead to poor quality of data during the first days or even during the first weeks of the proper fieldwork. Role playing can be used in the participatory exercises, where the staff members take turns in playing different roles of the field work member (interviewer, measurer) and the participant.

In the role playing sessions those who play the role of the participants should be encouraged to vary their behaviour and to challenge the fieldwork member e.g. with asking several questions on the purpose of the survey, acting to be very busy, shy, fearful, reluctant or aggressive. If possible, practical training sessions can be recorded. Watching own own performance helps to understand the purpose of standardization. Getting direct feedback during the practical sessions is important. Time needs to be allocated to discuss encountered difficulties and solutions directly after these exercises. The final step of the training should be to carry out the examination of a actual survey participant with supervisor observation.

Placing all training material and learning tasks and keeping a common discussion forum in the Internet (e.g. a specific survey training extranet site) will help to make sure that all staff members have up-to date information available throughout the fieldwork period. Open discussions between all field work members and other survey staff members should be encouraged during the training sessions. During the fielwork, meetings with the supervisors, audit visits and feedback sessions will support learning and point out the importance of standardization.

15.2.4 Duration and timing of the training

The EHES core measurements will require at least two or three training sessions, depending on the previous survey experience of the selected staff members.When the blood pressure is measured using the auscultation method, at least one week of training is required to ensure that all measurers have the same level. Each additional measurement will increase the duration of the training.

Training should be organized just before the fieldwork will be started, but some additional sessions may also be needed during the fieldwork. To allow substitution of other fieldwork team members when needed and rotating tasks (see Part A, Chapter 9) it is

recommended that each team member will be trained to handle several measurements, even if the measurements are carried out by teams where the staff members have different tasks. Retraining during fieldwork should be organized if the fieldwork lasts for more than two or three months to ensure that the standards are kept. Retraining is essential also if observer effects or non-adherence to survey standards are observed during audit visits or by other forms of quality control during the fieldwork.

15.2.5 Certification

Certification for specific measurements is needed at least for the most challenging measurements requiring strict adherence to detailed protocols, such as blood pressure and waist circumference measurements and drawing blood samples. Certification is given after observed competent performance in practice and proven theoretical knowledge on measurement techniques and standardization.

References

- Czaja R, Blair J. Designing surveys. A guide to decisions and procedures. Sage Publications Ltd, Oine Firge Press, California, USA, 2004.

16. Preparation of the survey budget

Hanna Tolonen[1], Päivikki Koponen[1], Kari Kuulasmaa[1]

[1] National Institute for Health and Welfare (THL), Helsinki, Finland

Conducting a national health examination survey (HES) requires resources, which include personnel costs and materials as well as funds for travel, accommodation, rent, transport of materials, etc. The type and amount of resources needed depends strongly on the number of persons to be examined, the measurements to be done and the setting of the surveys. The preparation of the budget has to include the entire survey process (see Part A, Chapter 1) to ensure adequate resources for the planning and preparation, fieldwork as well as for the data analysis and reporting.

This chapter will provide guidelines for estimating the costs of the different phases of a national HES. It should be noted that these are just guidelines and have to be adjusted for the local situation. An Excel template (http://www.ehes.info/tc/tools/time_cost.xls), which may assist in preparation of the national HES budget, is also provided.

16.1 Purpose of the survey budget

The survey budget gives an estimate of the amount of money needed to carry out the planned survey components. With a well prepared survey budget, the work can be carried out without major surprises in the actual costs. The budget can also be used in discussions with the collaborators when possibilities to include additional measurements are negotiated. Adding a new measurement to the survey protocol will increase the total survey cost more than just the required equipment, trough longer examination times per person which affects the costs of survey personnel and survey site, and also through training, data management as well as handling, quality control and reporting costs.

The funding available for the survey is always limited, and the survey budget has to be adjusted to the available funds. This may mean limiting the number of included measurements or the number of persons to be examined from what was initially planned. Also the selection of the survey mode has to be considered in light of the funding available.

16.2 Components of the survey budget

From the national HES, 12 stages, which affect the survey budget can be identified:

1. Planning and preparation
2. Coordination
3. Training of personnel
4. Dissemination (PR-activities)
5. Piloting
6. Sampling
7. Recruitment of participants
8. Field work of the full-size HES
9. Laboratory analysis and sample storage
10. Data entry and cleaning
11. Quality assurance
12. Analysis and reporting

When the measurements to be included in the national HES have been selected and the survey setting has been decided, the time needed to examine one survey participant should be estimated. For example, let us assume a survey setting where the participants come to the fixed examination site and we measure height, weight, waist circumference, and blood pressure and draw blood samples for total and HDL cholesterol, and for fasting glucose measurements. Additional to that, the participants have to fill in the survey questionnaire at home, which is checked and completed at the examination site after the informed consent is explained to and signed by the participant. We can estimate that the checking of the questionnaire and obtaining informed consent will take 15 minutes, anthropometric measurements 10 minutes, blood pressure measurement 15 minutes and drawing the blood sample 15 minutes. This sums up to 55 minutes per participant.

The total time required to measure the entire sample can be calculated by multiplying the time per participant with the sample size. In practice, the number of survey participants to be meas-

ured will be less than the sample size. This should balance out the time needed for setting out the examinations sites, mandatory breaks of the fieldwork staff, re-training, etc.

For example, if we have a sample of 4000 persons and the time to measure one participant is 55 minutes, a total time required to measure the entire sample is 220,000 minutes = 3667 hours. If each field work day lasts 8 hours, this would mean 459 days. Depending on the number of parallel field work teams, the length of the field work period can be calculated. If we have 4 field work teams, the field work would take 115 days and in case where the field work is conducted only during the working days from Monday to Friday, this would mean that the field work lasts 23 weeks. All this assumes that participants are examined at the fixed examination site one after other, without any overlap and without time needed for traveling (as in case of home visits). If the examinations can be organized with a field work team so that while one member of the team is examining one participant, the other one is at the same time examining the other one, i.e. there is overlap, the needed examination time decreases.

16.2.1 Planning and preparations

The planning process is described in Part A, Chapter 1. The main resource needed for the planning and preparation stage is personnel. The expertise of different professionals is needed for the planning. Each planning and preparation team should include or consult at least following experts:

- Project leader, who has the main responsibility of the survey.
- Survey coordinator, who will organize the practicalities and will monitor the progress of the work.
- Senior researchers, who will provide epidemiological and public health perspective to the selection of the measurements and to the preparation of the survey questionnaires and manuals.
- IT expert, who will plan and prepare the IT infrastructure of the survey.
- Survey statistician, who will be consulted on sampling and analysis of the results.
- Press officer, who will be consulted on promotion of the survey as well as on dissemination of the results.
- Experts on the measurements, who will provide information on practical points of each measurement included in the survey.

- Laboratory experts, who will plan the blood sampling, sample processing, storage, transport and analysis.
- Expert on legal and ethical issues, who will be consulted in the questions relating to the data confidentiality, and ethical issues. This also includes obtaining the ethical approval for the survey, informed consent, etc.

There may also be need for a person with special knowledge on survey logistics. Survey logistics, scheduling of the examinations, and transfer of personnel and materials can have a major impact on the survey budget. In some countries, translation of the survey questionnaires, etc. to different languages may be needed.

16.2.2 Coordination

The coordination activities of the survey are described in Part A, Chapter 1. The main resource needed for the coordination is the personnel but also some basic equipment and other resources are needed.

For the personnel, at least following is needed:

- Project leader, usually a senior researcher, who has the main responsibility for the survey.
- Survey coordinator, who will organize the practicalities and will monitor the progress of the work.
- Fieldwork supervisor, who will take care of the personnel management.

Often also an assistant is needed to assist with various practicalities, like recruitment of survey personnel, ordering the equipment and materials, etc.

The coordination team needs at least computers with internet connection, telephones/mobile phones, printers, software licenses, and office materials. Also premises for the coordination office are needed, although they are often provided by the organizing institute. In many cases, the coordination team will also travel to the survey sites to promote the survey and to monitor the progress of the work. The travel and subsistence allowances need to be budgeted for this. Also some money has to be budgeted for the recruitment of field work personnel (newspaper advertisements, etc.).

16.2.3 Sampling

The definition of the sample size is described in Part A, Chapter 2 and the sampling process is described in Part A, Chapter 3. De-

pending on the local situation, the actual sampling can be done by a survey statistician hired to the survey team, or it can be bought as a service from a statistical institute or other sampling frame owner. In any case personnel and equipment are needed as well. Sampling has to be made for both the pilot survey and the full size survey.

Regardless of the way the actual sampling is done, at least following personnel is needed:

- Statistician to determine the adequate sample size, to sort out available sampling frame(s) and to design the sampling.
- Database manager to form a database where the sample information is stored.

Computers, and software licenses to establish a database for the sample are needed. Often the survey database is not established on a PC but on a server of the institute organizing the survey. Depending on the institute, the use of server space and database platforms may or may not cost separately for the projects.

16.2.4 Training

The training programme is described in Part A, Chapter 15. Salaries and travel costs for the trainers and the field work staff to be trained, equipment and some other costs like preparation of the training materials have to be budgeted.

As trainers, at least following expertise is needed:

- Trainers for each measurement included in the survey. Depending on the qualifications of the trainers, one person can train the measurement protocols for several measurements or each measurement may need to be trained by different persons.
- IT-support and/or data management person(s) to train how to use the IT programmes in the field work and how the data management is organized.
- Press officer to tell about the promotion activities for the survey.
- Statistician to tell about the sample selection as participants on the field may ask from the field work personnel how just they got selected to this survey.
- Legal and ethical expert to explain the importance of data confidentiality and how to obtain the informed consent.

Equipment needed for the training depends on the included measurements (Part A, Chapter 5). A full set of equipment for each included measurement needs to be available during the training. A list of the required equipment for each recommended core measurement is given in Part B, Chapter 2. It is good to provide each trainee with a folder, which includes training material together with the local survey manual.

In addition to the personnel and equipment, also premises for the training are needed. For the field work staff to be trained and for the trainers, travel expenses and subsistence may also have to be paid.

16.2.5 Dissemination of information

The dissemination activities are described in Part A, Chapter 14 and Part C, Chapter 7. Dissemination of the national HES includes the promotion of the survey before and during the field work as well as the dissemination of the survey results to the survey participants, various stakeholders, general public and the scientific community. The resources needed depends on the dissemination strategy of the specific survey, but some personnel resources and other costs should always be budgeted.

Regardless of the dissemination strategy, at least a press officer is needed to plan and supervise the national dissemination strategy. There may also be need for a graphical designer to prepare promotional material for the survey. Often this service is bought from a service provider.

Depending on the planned promotion activities, there may be printing costs of promotional leaflets, advertisement costs for newspapers, radio and TV, and costs of press conferences and other promotion events. Costs from press conferences and promotion events may include payments to the publicly known persons who come to promote the survey, travel and subsistence costs, refreshments, etc.

16.2.6 Piloting

The piloting process is described in Part A, Chapter 11. This requires both personnel resources, equipment and other resources.

For the personnel resources, at least following is needed:

- Full field work teams are needed to conduct examinations of the pilot sample.

- IT support to assist with the computer programs in use on the field.
- Laboratory personnel to process, transfer and storage of the blood samples and to carry out the laboratory analysis of the collected samples.
- Data manager to handle the incoming data.
- Statistician to assess and analyze the pilot survey data.

Examination equipment have been obtained for the training but additional sets may be needed for the pilot phase.

Other costs that may occur during the pilot are the printing costs of the invitation letters, phone bills, rents of the examination sites, travel costs and subsistence of the field work and coordination personnel and transportation of the material to and from the field. Additional costs for the pilot phase come from the evaluation of the success of the pilot and from possible needs to change the survey protocol, to correct the computer programs, and to re-train the field work personnel.

16.2.7 Recruitment of participants for the full-size HES

The recruitment process and ways to increase participation rate are described in Part A, Chapter 13. The used recruitment strategies are survey specific and need to be adjusted for the national situation.

For the recruitment of participants, at least following personnel is needed:

- Assistant, who will prepare the invitation letters and mail them out or in case the first or re-contact is through telephone, personnel who make the required phone calls.
- A designated person, who is named as a contact person and can provide more information about the survey when ever any person selected to the survey requires that.
- Data manager or assistant, who will make sure that the status of the recruitment and contacts is also recorded to the survey database.

Computers and software licenses are needed to prepare invitations. It is highly recommended to establish a toll-free telephone number to which survey participants can call to change their ex-

amination times or ask additional questions. If a survey web site with for example web-based appointment scheduling system is established, personnel, equipment, and other costs relating to that have to be budgeted.

16.2.8 Field work for the full-size HES

All participants of the survey will be examined during the field work. The selected survey setting (Part A, Chapter 7) will have a marked effect on budget. If questionnaires are filled in during the interview and are not self administered (Part A, Chapter 8), the cost of interviewers has to be included to the budget. Also in case where the examinations are done at the participant's home, the travel expenses for the field work staff have to be adjusted for this setting. It is good to remember that the coordination goes through out the survey process and is an essential part of the field work.

For the field work, at least the following personnel is needed:

- Full field work teams to carry out the measurements (Part A, Chapter 9 and Part B, Chapter 2).
- IT support to assist in the use of computer programs used in the field work (Part A, Chapter 12 and Part B, Chapter 3).
- Data manager to handle the incoming data (Part A, Chapter 12).

There may also be need to have a medical doctor on call for consultation.

Most of the needed equipment (Part B, Chapter 2) have already been obtained during the training and pilot phase but additional sets of equipment may be needed for the field work teams.

Additional to the personnel and equipment, at least costs from transport of the materials to and from the field examination site, transport of the personnel, accommodation and subsistence of the personnel during the field work, storage of the material in the field, rents of the examination sites, and data transfer, phone and internet connection have to be included in the budget, depending on the survey setting. The cost of the feedback to the participants about their results (e.g. letters with laboratory results) has to budgeted as well.

16.2.9 Laboratory analysis and sample storage for the full-size HES

The issues related to the laboratory analysis and sample storage are described in Part A, Chapter 10. Decisions on what is analyzed immediately and how much of the samples are stored for the future analysis effect the budget.

For the laboratory analysis, at least the following personnel is needed:

- Laboratory personnel with specific qualifications to handle the analysis.
- Data manager to handle the incoming data from the laboratory analysis.

A medical doctor should also be available for consultation if abnormal results are discovered.

Assuming that laboratory facilities are provided by the organizing institute, the following laboratory consumables need to be budged: aliquot tubes, pipettes, storage track and reagents. If the laboratory analysis is bought from the laboratory, these consumables should be part to the contract prize.

For the long time storage, the specific tubes which endure -70^{o}C, storage boxes and deep freezers of -70^{o}C are needed. For the deep freezers there has to be a security system to ensure that the samples do not melt during the power breaks and which will alarm if freezers get broken. Also log book of the stored samples is needed and has to be established if this is not already used in the laboratory.

16.2.10 Data entry and cleaning

The data management issues relating to the field work are described in Part B, Chapter 3. The collected data have to be entered to the database and checked for completeness and correctness.

For data entry and cleaning, at least following personnel is needed:

- Data manager to maintain and update the survey database.
- Data entry persons in case data is entered manually.
- Statistician to work on data checking and cleaning.

Also transferring data forms to a data entry company and the data from there to the database will create costs. In case data forms are scanned, the scanning costs have to be included in the budget.

If computers and software licenses and server space needed for the maintenance of the database is not already budgeted during the planning phase, it should be included in budget here.

16.2.11 Quality control

The quality assurance of the survey process is described in Part A, Chapter 11 and the quality control of the data in Part C, Chapter 4.

For quality control, at least following personnel is needed:

- Senior researcher/epidemiologist who know the survey protocol well and can detect deviations from it. Number and specific qualifications depend on the contents of the specific survey.
- Data manager who can detect possible systematic errors in the data already when they are entered to the database.
- Statistician to do basic checking of the data entered to the database.
- Laboratory personnel to perform both internal and external quality control of the laboratory analysis, and costs for taking part in the laboratory standardization program.

Equipment needed for the quality control of the measurements are specified together with the measurement procedures in Part B of this manual. In addition to these, computers and software are needed. An important part of the quality control is observing the work. There may be need to budget some travel and subsistence costs for the personnel conducting the quality control during the field work.

There may also be costs of the external quality control, such as for laboratory measurements.

16.2.12 Analysis and reporting

Issues relating to the analysis of the survey data and reporting are described in Part C, Chapters 6.
For the data analysis and reporting, at least the following personnel is needed:

- Data manager to maintain and share the survey data.
- Statistician to conduct the statistical analysis.
- Researchers to specify the research questions to be analyzed from the data and to interpret and report the results.

Additional to the personnel, also computers and software licenses are needed. In case the basic results are published in a form of book, the layout and printing costs need to be budgeted. Often results are also published on the web and a web manager may be needed to prepare the reports for the web.

16.3 Template for budget calculations

There is an Excel template at the EHES Web site under EHES Reference Centre Tools (http://www.ehes.info/rc/tools/time_cost.xls), which can be used while preparing the survey budget. The template is only a helping tool and each component has to be evaluated in the national situation. Instructions for using the Excel template are given in the first worksheet of the Excel template.

ISBN (pdf) 978-952-245-842-1

ISSN (pdf) 2323-4172

http://urn.fi/URN:ISBN:978-952-245-842-1

http://www.ehes.info

Reference:

Tolonen H (Ed.) EHES Manual. Part A. Planning and preparation of the survey.
National Institute for Health and Welfare, 2013, Directions 2013_001.
URN:ISBN:978-952-245-842-1,
URL: http://urn.fi/URN:ISBN:978-952-245-842-1