

Published in final edited form as:

*Exp Gerontol.* 2011 November ; 46(11): 934–945. doi:10.1016/j.exger.2011.08.005.

## Design, recruitment, logistics, and data management of the GEHA (Genetics of Healthy Ageing) project

A. Skytthe<sup>a,\*</sup>, S. Valensin<sup>b</sup>, B. Jeune<sup>a</sup>, E. Cevenini<sup>b</sup>, F. Balard<sup>c</sup>, M. Beekman<sup>d,e</sup>, V. Bezrukov<sup>f</sup>, H. Blanche<sup>g</sup>, L. Bolund<sup>h,i</sup>, K. Broczek<sup>j,k</sup>, C. Carru<sup>l</sup>, K. Christensen<sup>a,m</sup>, L. Christiansen<sup>a</sup>, J.C. Collerton<sup>n</sup>, R. Cotichini<sup>o</sup>, A.J.M. de Craen<sup>d</sup>, S. Dato<sup>p</sup>, K. Davies<sup>n</sup>, G. De Benedictis<sup>p</sup>, L. Deiana<sup>l</sup>, F. Flachsbar<sup>q</sup>, J. Gampe<sup>r</sup>, C. Gilbault<sup>s</sup>, E.S. Gonos<sup>t</sup>, E. Haimes<sup>n</sup>, A. Hervonen<sup>u</sup>, M.A. Hurme<sup>u</sup>, D. Janiszewska<sup>j</sup>, M. Jylhä<sup>u</sup>, T.B.L. Kirkwood<sup>n</sup>, P. Kristensen<sup>v</sup>, P. Laiho<sup>w</sup>, A. Leon<sup>x</sup>, A. Marchisio<sup>l</sup>, R. Masciulli<sup>o</sup>, A. Nebel<sup>q</sup>, G. Passarino<sup>p</sup>, G. Pelicci<sup>y</sup>, L. Peltonen<sup>w</sup>, M. Perola<sup>w</sup>, M. Poulain<sup>s</sup>, I.M. Rea<sup>z</sup>, J. Remacle<sup>aa</sup>, J.M. Robine<sup>c</sup>, S. Schreiber<sup>q</sup>, M. Scurti<sup>b</sup>, F. Sevin<sup>b</sup>, E. Sikora<sup>j</sup>, A. Skouteri<sup>t</sup>, P.E. Slagboom<sup>d,e</sup>, L. Spazzafumo<sup>ab</sup>, M.A. Stazi<sup>o</sup>, V. Tucciaceli<sup>o</sup>, O. Toussaint<sup>ac</sup>, O. Törnwall<sup>w</sup>, J.W. Vaupel<sup>r</sup>, K. Voutetakis<sup>t</sup>, C. Franceschi<sup>b</sup>, and GEHA consortium<sup>1</sup>

<sup>a</sup>Epidemiology, Institute of Public Health, University of Southern Denmark, Odense, Denmark

<sup>b</sup>University of Bologna, Interdepartmental Centre “Luigi Galvani” CIG, Italy <sup>c</sup>INSERM, Health & Demography, France <sup>d</sup>Leiden University Medical Centre, Leiden, The Netherlands <sup>e</sup>Netherlands Consortium for Healthy Ageing, The Netherlands <sup>f</sup>Institute of Gerontology, Kiev, Ukraine

<sup>g</sup>Foundation Jean Dausset, Centre d'Etude du Polymorphisme Humain, Paris, France <sup>h</sup>Beijing Genomics Institute, China <sup>i</sup>Institute of Human Genetics, University of Aarhus, Denmark <sup>j</sup>Nencki Institute of Experimental Biology, Polish Academy of Science, Warsaw, Poland <sup>k</sup>Department of Geriatrics, Medical University of Warsaw, Poland <sup>l</sup>University of Sassari, Sardinia, Italy

<sup>m</sup>Department of Clinical Biochemistry and Pharmacology and Department of Clinical Genetics, Odense University Hospital, Odense, Denmark <sup>n</sup>Newcastle University, Newcastle, United Kingdom <sup>o</sup>Istituto Superiore di Sanità, Rome, Italy <sup>p</sup>University of Calabria, Rende, Italy <sup>q</sup>Institute of Clinical Molecular Biology, Christian-Albrechts-University Kiel, CAU, Germany <sup>r</sup>Max Planck Institute for Demographic Research, Rostock, Germany <sup>s</sup>Catholic University of Louvain, Louvain, Belgium <sup>t</sup>National Hellenic Research Foundation, Athens, Greece <sup>u</sup>Tampere School of Public Health, Tampere, Finland <sup>v</sup>University of Aarhus, Aarhus, Denmark <sup>w</sup>The National Institute for Health and Welfare, Helsinki, Finland <sup>x</sup>Research Innovation s.r.l., Italy <sup>y</sup>IFOM Institute of Milan, Milan, Italy <sup>z</sup>Queens University of Belfast, Belfast, Northern Ireland, United Kingdom <sup>aa</sup>Eppendorf

© 2011 Elsevier Inc. All rights reserved.

\*Corresponding author at: Epidemiology, Institute of Public Health, University of Southern Denmark, J B Winsløvs Vej 9B, DK 5000 Odense C, Denmark. Tel.: +45 6550 3034; fax: +45 6550 3682., askytthe@health.sdu.dk (A. Skytthe).

<sup>1</sup>A list of all contributors is found at the end of the paper.

### Contributors

The following persons contributed to the recruitment process of the GEHA project (for affiliation, see title page):

Andersen S<sup>a</sup>, Bemer E<sup>d</sup>, Berardelli M<sup>p</sup>, Blauw GJ<sup>d</sup>, Bødker F<sup>a</sup>, Börn S<sup>q</sup>, Castaldi M-H<sup>c</sup>, Celani L<sup>b</sup>, Christensen GB<sup>a</sup>, Christiansen SE<sup>a</sup>, Clavel A<sup>c</sup>, De Groot-Nys M<sup>d</sup>, De Man LW<sup>d</sup>, De Rango FP, Debaq-Chainiaux F<sup>cc</sup>, Dewit F<sup>c</sup>, D'ippolito C<sup>o</sup>, Ervasti M<sup>u</sup>, Facchini R<sup>o</sup>, Fagnani C<sup>o</sup>, Faurobert J<sup>c</sup>, Ferrari D<sup>o</sup>, Frackowiak B<sup>j</sup>, Frapsauce A<sup>c</sup>, Frölich M<sup>d</sup>, Granger B<sup>c</sup>, Hirviniemi S<sup>u</sup>, Hughes J<sup>n</sup>, Jussila M<sup>w</sup>, Kcholin V<sup>f</sup>, Kirik V<sup>f</sup>, Kløjgaard S<sup>a</sup>, Knudsen S<sup>a</sup>, Knuutila H<sup>w</sup>, Kupisz-Urbanska M<sup>j-k</sup>, Lecomte C<sup>s</sup>, Locuratolo CN<sup>o</sup>, Maenpaa K<sup>u</sup>, Marais M<sup>c</sup>, Mari V<sup>p</sup>, Marras MS<sup>l</sup>, Martin B<sup>c</sup>, Martino CP, McNerlan S<sup>z</sup>, Meucci S<sup>o</sup>, Milbradt K<sup>q</sup>, Mooijekind IG<sup>d</sup>, Muradian K<sup>f</sup>, Murphy A<sup>z</sup>, Nielsen L<sup>a</sup>, Nikolaus S<sup>q</sup>, Nisticò L<sup>o</sup>, Palmas MG<sup>b</sup>, Panourgia M<sup>p</sup>, Pascal T<sup>cc</sup>, Penna L<sup>o</sup>, Pes G<sup>l</sup>, Peters U<sup>q</sup>, Phelipot M<sup>c</sup>, Pini E<sup>b</sup>, Podkówka-Sieczka R<sup>j</sup>, Poletayeva K<sup>f</sup>, Polyakov A<sup>f</sup>, Polyakova I<sup>f</sup>, Prokopenko N<sup>f</sup>, Pulciani S<sup>o</sup>, Purcheddu S<sup>l</sup>, Rasmussen KM<sup>a</sup>, Repo J<sup>u</sup>, Ricciardi E<sup>l</sup>, Romieu I<sup>c</sup>, Rønne H<sup>a</sup>, Saarela K<sup>u</sup>, Salemi M<sup>o</sup>, Skaaning J<sup>a</sup>, Slynchak P<sup>f</sup>, Snoek N<sup>c</sup>, Stelmach T<sup>f</sup>, Tango F<sup>o</sup>, Timonen E<sup>u</sup>, Tomarevskaya E<sup>f</sup>, Tuikkala A<sup>u</sup>, Tuovinen J<sup>u</sup>, Ustimenko A<sup>f</sup>, Vaiserman A<sup>f</sup>, Vallone AP, Van Boxxel T<sup>d</sup>, Van der Breggen R<sup>d</sup>, Van der Star MGHF<sup>d</sup>, Vandervelpen G<sup>s</sup>, Vavrik G<sup>f</sup>, Verkkonen K<sup>u</sup>, von Eller-Eberstein H<sup>q</sup>, Westendorp RGJ<sup>d</sup>, Widén E<sup>w</sup>, Yli-Hukkala K<sup>u</sup>, Aalto A<sup>u</sup>.

Array Technologies, Belgium <sup>ab</sup>Istituto Nazionale Ricovero e Cura Anziani, Italy <sup>ac</sup>University of Namur, Narillis, URBC, Namur, Belgium

## Abstract

In 2004, the integrated European project GEHA (Genetics of Healthy Ageing) was initiated with the aim of identifying genes involved in healthy ageing and longevity. The first step in the project was the recruitment of more than 2500 pairs of siblings aged 90 years or more together with one younger control person from 15 areas in 11 European countries through a coordinated and standardised effort. A biological sample, preferably a blood sample, was collected from each participant, and basic physical and cognitive measures were obtained together with information about health, life style, and family composition.

From 2004 to 2008 a total of 2535 families comprising 5319 nonagenarian siblings were identified and included in the project. In addition, 2548 younger control persons aged 50–75 years were recruited. A total of 2249 complete trios with blood samples from at least two old siblings and the younger control were formed and are available for genetic analyses (e.g. linkage studies and genome-wide association studies).

Mortality follow-up improves the possibility of identifying families with the most extreme longevity phenotypes. With a mean follow-up time of 3.7 years the number of families with all participating siblings aged 95 years or more has increased by a factor of 5 to 750 families compared to when interviews were conducted. Thus, the GEHA project represents a unique source in the search for genes related to healthy ageing and longevity.

## Keywords

Longevity; Healthy ageing; Nonagenarian sib pairs; Genetics; Europe

## 1. Introduction

The number of men and women reaching 90 years of age and beyond is rapidly increasing in Europe and other developed countries. The proportion of this age group that survives for many years without major health problems is also increasing (Christensen et al., 2009). This development has led to an increase in research aimed at understanding the nature of the ageing process.

In 2003 a consortium was established to undertake a multi-national research project titled Genetics of Healthy Ageing (GEHA). The GEHA consortium has 25 partners, of which 23 are from the European Union, one is from Ukraine, and one is from China (Franceschi et al., 2007). The GEHA consortium brings together expertise from many different fields that are relevant to the study of ageing, including molecular biology, genetics, clinical medicine, epidemiology, demography, statistics, and ethics.

The primary aim of the GEHA project is to identify genes involved in healthy ageing and longevity, using approaches first proposed by Schächter et al. (1993). Previous research has established that, in the general population, genetic factors account for approximately 25% of the heritability of human longevity (Herskind et al., 1996; Hjelmborg et al., 2006). Various studies on the possible association of candidate regions of nuclear and mitochondrial DNA with human longevity have been conducted, but many of these studies report contradictory results (Beekman et al., 2006; De Benedictis et al., 2001; Dato et al., 2004; Christensen et al., 2007). Consistent evidence for association has so far been described for the

apolipoprotein E (APOE) locus and the forkhead box 3A (FOXO3A) gene (Christensen et al., 2007; Willcox et al., 2008; Flachsbart et al., 2009). There is great interest in identifying the molecular processes underlying genetic factors that increase the chance of an individual's reaching advanced old age in good cognitive and physical function and in the absence of major age-related diseases.

To address this aim, the GEHA project has established an unprecedented sample consisting of more than 2500 sib pairs aged 90 years or more recruited from 15 regions in 11 European countries. In addition, a younger person aged 50–75 years, preferably with the same gender as the oldest sibling in the sib pair, was recruited as a control to each sib pair. The function of the controls, who have a reasonably low probability of sharing any genetic predisposition to healthy longevity with the sibling pairs with which they are matched, was to provide reference data.

Biological material was collected from each participant for the purpose of genetic analysis. In addition, a structured interview was carried out with each participant to obtain information on his or her health status and basic physical and cognitive function. The purpose of the structured interview was to permit classification of the health and functional status of the old siblings. Additional family information about parents and other siblings was obtained with the further purpose of characterising patterns of familial longevity.

In this paper, we describe the basic design, the procedures employed and the logistics of recruitment of the participants in the GEHA project both in general and country-specific terms.

## 2. Material and methods

### 2.1. Study design

The basic unit of the study is the nonagenarian sib pair, i.e., two or more siblings aged 90 years or older, together with a younger control person who is unrelated to the sibling pair but from the same geographical area (Fig. 1). This basic unit is termed a '*trio*'. The family to which the nonagenarian siblings belong is termed a *GEHA-family*.

The study design enables several types of genetic analyses: (i) utilising the affected sib pair design, linkage analyses can be performed to identify candidate regions of interest in the genome and (ii) case–control studies can be conducted to test whether allelic variation is associated with longevity, either in targeted candidate genes or in genome-wide association studies (GWAS).

The structured phases of the project were conceived as forming a linked process. The first phase was the recruitment of subjects (90+ sib pairs and younger unrelated controls) all over Europe; this phase included the collection of information on their phenotype (functional status) as well as of biological samples (blood and/or cheek swab). The second phase of the project involved the extraction of DNA from the collected biological samples, its quality control and shipment to the GEHA partners who will conduct the genetic analysis. The third phase consists of ongoing genetic analysis. The fourth phase will be the further analysis of data by means of new analytical methods; these methods will be developed as part of the GEHA project using state-of-the-art mathematical models.

This paper focuses on the first phase of the project and outlines the process from identification of eligible subjects to the collection of biological material and phenotypic information.

Table 1 lists the countries involved and their recruitment regions. Countries were chosen in order to cover the diversity in ethnicity, culture, and socio-demographic conditions among European populations.

Because none of the participating recruitment centres had pre-existing lists available, one of the first and main challenges associated with the GEHA project was the identification of potential participants.

Two main methods for the identification and enrolment of participants in the GEHA project were employed. In the first method, a list with identified suitable sib-ships was created before potential participants were contacted. If it was not possible to create such a list, a second method was followed in which an older person was contacted to establish whether an eligible sibling(s) was also alive.

In Appendix A, the methods used by each recruitment centre to identify sib-ships and younger controls are described.

## 2.2. Protocol

Despite differences in the methods used to identify potential participants, all recruitment centres adopted agreed-upon inclusion criteria and a uniform study pathway, both of which are described subsequently.

Participants were required to fulfil the following conditions to be enrolled in the project: i) they should be at least 90 years of age (a nonagenarian); ii) they should have a living sibling who also is a nonagenarian; iii) they should be able to understand the scope of the project; iv) they should consent to participate in it, subject to a carefully developed and ethically robust consenting process.

Once identified, individuals aged 90 years or older were contacted by mail with a letter of introduction and an information booklet outlining the purpose of the project and giving details of the interview and the collection of blood samples. If an individual expressed an interest in participating, an appointment was made to visit the individual in his or her usual place of residence.

During the visit, the interviewer explained the project in detail; if the individual agreed to participate, the interviewer obtained a signed informed consent statement after ascertaining that the individual had the capacity for free and informed consent.

Upon consent, a structured interview was conducted based on a standardised questionnaire. The interview included questions about parents' age, the vital status of other siblings, socio-demographic conditions, self-reported physical function (Katz Activity of daily living index (ADL)) (Katz et al., 1963), cognitive function (Standardized Mini-Mental State Examination (SMMSE)) (Molloy et al., 1991), and past and present medical history. Physical function was examined by chair stand test (including 5 chair rises), and hand grip strength was measured using a hand-held dynamometer (Smedlys' dynamometer, Scandidact, Kvistgaard, Denmark). A measurement of height to the nearest centimeter was obtained, and weight was measured to the nearest kilogram. If the participant was bedridden, self-reported measures were obtained. The total time for the home visit was limited to approximately 1 h so as not to overburden the participants.

Finally, blood samples consisting of three 7.5 ml EDTA tubes of blood were collected from each participant. If the participant refused blood sampling, biological material was obtained by means of a cheek swab. However, recruiters were instructed that cheek swab sampling should only be used as a 'back up' method because of a lower yield of genetic material

compared with blood samples. Samples were stored at  $-20^{\circ}\text{C}$  until shipment to the National Institute for Health and Welfare in Helsinki.

To ensure a standardised procedure for the extraction and handling of DNA, sample processing was conducted by one central facility, the National Institute for Health and Welfare, Department of Chronic Disease and Prevention, Public Health Genomics Unit, Helsinki, Finland (THL). One sample tube of blood from each participant was sent to THL, Helsinki for extraction of DNA, while the remaining tubes were kept at the recruitment centres. If the extraction of DNA from the first tube failed, a second tube was shipped to Helsinki. In three recruitment centres (Bologna, Tampere, Calabria), which had the expertise and facilities for separation, lymphocytes and granulocytes were isolated according to specific protocols (Bøyum, 1976) from one sample tube immediately after collection for the analysis of mitochondrial DNA (mtDNA) heteroplasmy. The lymphocytes and granulocytes were stored in the local laboratory at  $-20^{\circ}\text{C}$  and later sent to THL for extraction of DNA.

Genomic DNA was extracted from whole blood or cell fractions using Gentra's automated extraction instrument Autopure LS; Gentra's Puregene Salt Precipitation kit was used to manually extract DNA from cheek swabs. After extraction, the DNA samples were subjected to preliminary concentration measurement and quality control using UV-spectrometry. The extracted DNA is stored at THL and in France at the Centre d'Etude du Polymorphisme Humain (CEPH), Paris and is distributed from these locations to genotyping centres on request.

The following inclusion criteria were applied to recruitment of younger controls: i) the participant should be between 50 and 75 years of age; ii) he or she should be genetically unrelated to the old siblings to whom he/she was assigned as control; iii) he or she should be of the same ethnicity and geographical background as the old siblings; iv) whenever possible, he or she should be of the same gender as the oldest sibling in the sib pair or group.

Two main methods of recruiting younger controls were used. In cases where it was possible to contact the spouse of a child of an old sibling, he/she was asked to participate. Alternatively, control volunteers were recruited from the same geographical area as the nonagenarians. This process was achieved by random selection and by approaching individuals in the population or social centres or by approaching persons already known to the recruitment team.

After obtaining written consent from the younger control, a shorter interview that included questions about the individual's socio-demographic conditions, health status and the origin of parents and grandparents was conducted. Origin of parents and grandparents was reported as place of birth and was recorded at the level of country and region, corresponding to a population of 800,000 to 3 million (Eurostat, 2004). Finally, biological material was collected following the same protocol as for the old siblings.

A trio was considered complete only when interviews and blood samples had been completed for at least two older siblings and a younger control. Interviews with siblings were conducted as closely together in time as possible; however, in a number of cases, the interview of the second sibling was not completed because of either the intervening death of the second sibling or the refusal of that sibling to participate in the study.

In families that contained more than two surviving nonagenarian siblings, all siblings were eligible for participation. In these families, sets with three, four and five siblings that participated plus a younger control were formed. In cases where participating nonagenarians had siblings who had not reached 90 years of age at the time of initial recruitment but would do so within the recruitment period, details were noted, and an agreement was made to

return to them when they reached their 90th birthdays. In some cases, a sibling aged 89 years was interviewed before his or her 90th birthday; however these siblings were only included in the project if they reached 90 years of age.

A number of standardised documents were prepared in English and translated into the native language of the participating countries. These documents included a letter of introduction, informed consent forms, and questionnaires.

The interviews were conducted by trained nurses, biologists, biotechnologists, students and, in some cases, medical doctors. Prior to recruitment a training course was held in Bologna in which procedures for obtaining initial contact, informed consent, standardisation of interviews and tests, collection and handling of biological material were discussed and coordinated.

Most of the recruitment for the study was carried out between November 1, 2004, and October 31, 2008. The actual recruitment period in the various recruitment centres varied somewhat because of different procedures for obtaining ethical approval of the study. The data and samples from a small pilot study conducted prior to the main recruitment period were also included in the main study (Table 2).

### 2.3. Verifying age

An important issue in ageing studies is verification of the ages of the participants, in particular the ages of the nonagenarian siblings. Therefore, a central part of the recruitment of old participants was to collect evidence of their age.

Three types of evidence were considered valid for age verification: i) registration of birth in an official birth register (e.g., in Denmark); ii) registration of birth in an official population register (e.g., in Italy); iii) an official document, preferable a birth certificate, in which the date of birth was stated (e.g., in The Netherlands). In some countries (e.g. Poland) in which records of population registration were lost during wartime, other types of personal documents served as valid evidence (e.g., documents issued prior to the Second World War). For all participants, care was taken to ensure that valid evidence was used.

In most cases, older participants or their relatives reported the ages of other family members such as parents and non-participating siblings. However, information about age of other family members was generally not verified against official sources.

### 2.4. Mortality follow-up

It is not known to what extent a GEHA-family represents a long-lived family, only that at least two siblings have lived to age 90. To identify extreme long-lived families, repeated follow-up on vital status and, thereby, calculation of attained age for the old siblings has been factored in at regular intervals.

The first follow-up included vital status up to September 1, 2008; successive vital status follow-up was completed every year by January 1. However, as with the recruitment procedure, the existence of different conditions in the 11 countries that participated in the recruitment necessitated the use of a variety of methods. In some countries, it was possible to perform a central follow-up by querying official population registers either nationally or locally; however in others, the only possible way of checking vital status was by contacting the participants or their relatives. The different methods adopted by the recruitment centres are described subsequently.

In Denmark, Finland, The Netherlands, Belgium and Poland the vital status for all participants was obtained by linkage to a central population register, which in the event of death returned date of death, otherwise if persons were alive, the date of linkage was used in the calculation of attained age. In England, vital status updates were obtained through registration of the project with 'The Information Centre for Health and Social Care'. This service allows the 'tagging' of consenting participants' medical records and the reporting of significant events, including death. In addition, information on participants' vital status was often reported by relatives and caregivers.

In other recruitment centres, the follow-up procedure involved contact with many different sources from which information on vital status could be obtained. In Italy, follow-up was performed by contacting the local population registers (the 'anagrafe') and requesting information for each participant, and when necessary, by contacting the participants or their relatives.

In Northern Ireland, vital status was obtained either by contacting the local GP surgeries and/or the participants and their relatives. In Germany, Greece, and Ukraine, vital status follow-up was made by contacting the old siblings or their relatives. In France, vital status follow-up was initially made by contacting the nonagenarians or their relatives, but for the 2010 follow-up, the procedure was changed, and information on vital status was obtained from the Répertoire National d'Identification des Personnes Physiques (RNIPP) (National Register for Identification of Individuals). In all countries, follow-up of vital status proved to be quite costly and time-consuming, and, for those countries in which follow-up required re-contacting individuals, the risk of loss to follow-up was high.

## 2.5. Data management

Data management was centred on methods for handling and sharing data from two different sources in a way that fulfils both practical and ethical purposes.

**2.5.1. Phenotype data**—Data from interviews of participants were entered into a central database hosted at USD, Odense, Denmark. Each recruitment centre entered data in a structured form using EpiData (Lauritsen, 2000-2008). The data were transferred to an Oracle database from which each recruitment centre could download data originating from their own recruitment area.

**2.5.2. Genetic data**—Genetic data from the participants include single nucleotide polymorphism (SNP) data obtained from nuclear DNA. Genotype data, which are used for candidate gene studies, GWAS and linkage analysis, are stored in a genotype database hosted in Kiel, Germany. Data from analyses of mitochondrial DNA are stored in a mitochondrial genotype database hosted in Tampere, Finland.

No phenotype data are kept in the genetic databases; apart from gender and the data necessary for quality control of genetic analyses, only basic data on samples and conditions of analysis are stored in these databases.

**2.5.3. Analyses**—Researchers within the consortium who wish to analyse pooled data from several recruitment centres and/or combined phenotypic and genetic data are required to submit a request to the Max Planck Institute for Demographic Research (MPIDR), Rostock. Researchers outside the consortium may approach members of the consortium for collaboration, including the use of the collected data.

The request must include a description of purpose and specification of data. The request is distributed among the other consortium members; if no consortium member objects to the request within two weeks, MPIDR creates and delivers a specific data set to the researcher.

## 2.6. Ethics

Ethical issues have played a central role throughout the planning of the GEHA study. Recommendations and suggestions from a dedicated ethical working group that consists of internal as well as external members of the GEHA consortium have guided the production of documents and the establishment of procedures for conducting interviews and handling data. Special attention has been given to the key issues of recruitment, informed consent and confidentiality and data access, exchange, and storage.

The introduction letter made it clear that participation in the project was voluntary and that it was possible to withdraw from the project at any time without being asked to provide a reason.

Before each individual consented to participation, utmost care was taken to assure that he or she fully understood the purpose of the project. During the first personal contact, a series of questions was asked to be sure that the participant had the capacity to give informed consent. Participants were also informed of the possibility of completely withdrawing from the GEHA project and were informed that withdrawal would include destruction of the participant's biological samples and deletion of the data obtained in the interview. So far, no participant has requested to withdraw from the project.

The identity of the participants is known only to the central team within local recruitment centres. No identifiable information (i.e., name, address or personal identification numbers) was transferred to the central databases.

Two different identification numbers were used for the phenotypic information and the genotypic information collected in the study. This method was followed to protect the identity and respect the privacy of the participants. The link between the two identification numbers resides in only two places: THL, Helsinki, where the genetic identification number was issued, and MPIDR in Rostock, where combined data sets of phenotypic and genetic data for analyses are created and, at the same time, where a new analytic identification number is created for each participant.

Each recruitment centre obtained approval for the study from local ethical committees. The conditions for obtaining approval varied greatly among the participating countries (see Table 2). Most recruitment centres received their approval by February 2005, at which time the recruitment could commence.

## 3. Results

### 3.1. Recruitment

At recruitment centres where families with at least two siblings aged 90 years or older could be identified, the contact effort was targeted to only those families. In other recruitment centres with no means of identifying such families, the contact effort was extended to all individuals aged 90 years or older.

Table 3 shows the number of persons contacted and the resulting number of eligible families at recruitment centres where family status was not known prior to contact. Not surprisingly, a total of almost 86,000 older persons were contacted in the process of identifying a total of 1902 families with at least two living nonagenarians.



For recruitment centres in which the status of a family was ascertained as eligible before contact, the number of old persons contacted were lower; however, a considerable number of persons still had to be approached to reach the allotted number of families (Table 4).

The lists detailing eligible families were subject to attrition because of the time that passed from the creation of the list to the time at which contact was actually made. Because death of one old person resulted in exclusion of the sib pair, mortality gradually reduced the number of eligible families.

Table 5 displays the results of the recruitment effort at each recruitment centre. In total, 5418 old participants originating from a total of 2634 families were visited and interviewed during the recruitment period. In 99 families, only one old participant was interviewed at the age of 90 years or above. In some of these cases, an interview was conducted with a sibling younger than 90 years (41 interviews, not included); in other cases, the second sibling died or changed his/her mind before the visit took place and did not want to participate after all. Thus, a total of 5319 old siblings from 2535 GEHA families with at least 2 old siblings were enrolled in the GEHA project.

A total of 2548 younger control persons consented to participate in the study and were interviewed.

The central DNA extraction centre in Helsinki received a total of 11392 samples belonging to 7918 participants from the recruitment centres. The majority of the samples were whole blood samples; samples consisting of isolated blood cells also constituted a significant fraction of the samples. Although not preferred, almost 1000 cheek swab samples were received. The number of old siblings providing cheek swab samples was considerably higher than the number of younger controls, probably as a result of younger people being more willing to have blood drawn and also because of a later start in recruitment of younger controls at some recruitment centres. Three recruitment centres (Bologna, Calabria, Tampere) also shipped additional aliquots of lymphocytes and granulocytes, locally separated from blood samples, to Helsinki for mitochondrial DNA analysis.

In a number of cases, more than one sample of each type was shipped. In addition, more than one sample type was taken from some participants; in some cases, up to three different sample types were obtained. One participant provided only a lymphocyte sample. Thirty-two participants provided both a blood sample and a cheek swab sample (cheek swab sample were taken if the amount of blood drawn was less than expected).

For a number of reasons, some samples (n=244) shipped to Helsinki could not be used for the formation of trios. The main reasons included low yield of DNA in the sample and failure of collection of blood samples from both nonagenarians due to the death or refusal to participate of one of them.

The total number of participants who provided biological samples and who were considered in the formation of trios was 7674. Of these participants, 2392 were younger controls, while the remaining 5282 were nonagenarian siblings.

The number of complete trios from each recruitment centre is shown in Table 6. A total of 2249 complete trios with blood samples from at least two old siblings and the younger control could be formed. It is worth noting that 171 GEHA-families had 3 older siblings donating blood, 18 GEHA-families had 4 old siblings, and 3 GEHA-families consisted of 5 old siblings aged 90 years or more, all donating blood samples. These families represent the extreme longevity phenotype.

If the participants with cheek swab samples are also included in the formation of trios, it is possible to form 99 additional trios consisting of at least two old siblings and a younger control person. The use of cheek swab sampling was more prominent at some centres (Belgium, France, Ukraine, and Greece).

For 211 GEHA-families, a trio could not be established, either because only one old sibling participated or a younger control was not linked to the GEHA-family (if neither or only one old sibling donated a blood sample, a younger control was not linked to the GEHA-family).

### 3.2. Characteristics of participants

The age at interview of old siblings ranged between 90.0 years and 107.1 years (Table 7). The median age of all old siblings was 92.99 years, while the median age of younger controls was 62.47 years, ranging from 50.00 to 75.92 years.

Although the intention was to match the gender of the younger control person to the gender of the oldest sibling, the recruitment strategies employed in some recruitment centres made it difficult to follow this recommendation. In 414 out of 2354 pairs of oldest sibling and younger control (18.0%), the gender of the oldest sibling and the younger control was not the same.

Table 8 shows the distribution of gender within the participating GEHA-families. Of the 2315 GEHA-families with 2 old siblings, there was a clear dominance of sib pairs with a female sib; only 1 in 8 consisted of two brothers.

Fig. 2 shows the proportion of men among the old participants. With the exception of two recruitment centres, Greece and Ukraine, the average proportion of recruited men is approximately 25–30%. The very high proportion of men in Greece (68%) is probably due to the recruitment method, which exploited lists of possible old participants obtained from the Agricultural Organization. In contrast, among the sib pairs recruited in Ukraine, only 8% of the old participants are men.

The proportion of men who represent the oldest participant in a GEHA family does not differ from the proportion of men among all old participants (Fig. 2).

By including follow-up on vital status, a longitudinal dimension is introduced into the GEHA project. This dimension provides the advantage that the number of concordant families in which both sibs have reached a certain age increases. At interview, the number of families in which both sibs aged 95 years was 150, and in 3 families, there were 3 old siblings aged 95 years or more. However, at the most recent follow-up (January 2010), the number of families with all sibs aged 95 years or more increased to 801, of which 5 families had 4 old siblings aged 95 years or more, 46 families had 3 siblings aged 95 years, and 750 families had two siblings aged 95 years. Thus, according to the most recent follow-up, which had a mean follow-up time of 3.7 years, the number of families in which both sibs has reached 95 years of age had increased by a factor of 5 compared with the time of interview.

## 4. Discussion

Through a massive effort, the GEHA consortium has managed to establish a cohort of more than 2500 sib pairs of 90 years of age or more and a similar number of younger persons aged 50–75 years of same ethnicity and with the same geographical background as the sib pairs. Interviews were made with more than 5400 old participants, and biological samples were collected from almost all individuals, usually in the form of a blood sample. Not only were sib pairs included, but almost 9% of the included families had three or more old siblings as

participants. To our knowledge, this is the largest familial study in Europe that has been initiated with the aim of studying the genetic background of longevity and healthy ageing.

With 15 recruitment centres in 11 countries, the logistics of organising a common framework for recruitment were challenging. The process of setting up a common questionnaire for the interview required achieving a balance between the scientific questions to be investigated and the endurance of the older participants. In addition, it was important to respect cultural differences in the populations of participants. The existence of different types of legislation regulating permissible methods for identifying families with living siblings aged 90 years or more added further to the complexity of the recruitment process.

The conditions for carrying out a recruitment of the expected size were quite different among the recruitment centres. The effort to identify and obtain consent from members of a trio varied from having a list with GEHA families to having no prior knowledge of the family status when making contact with a nonagenarian.

A general problem in the recruitment of old sib pairs was that of identifying nonagenarian siblings living within a reasonable distance of each other. The size of the recruitment area at different centres varied, but several centres performed recruitment throughout their respective countries. Another general problem was the necessity for two individuals to agree to participate in the study. If one old sibling in an identified GEHA family of two old siblings refused to participate, the GEHA family could not be included. Thus, a participation rate of 60% at the individual level may correspond to a participation rate of 36% at the family level; this rate is close to the observed participation rate among the recruitment centres with knowledge of familial status before contact.

It should be kept in mind that the recruited participants do not represent a random sample of nonagenarians, nor do they comprise the complete population of families with at least two nonagenarians alive at the same time. The recruited sample represents a selection of different quotas of families from the recruitment partners in the GEHA project; once a given recruitment centre reached its quota, recruitment ceased.

On the basis of the foregoing, the old siblings in the GEHA project do not necessarily represent the total population of nonagenarians. Comparisons with other studies of nonagenarians must be made to determine whether the health status and other characteristics of the participants in this study are comparable with those of participants in other studies.

Despite these qualifications, the data and biological samples collected in this study comprise a unique resource in the search for genes associated with longevity and healthy ageing. A number of studies based on the collected material will emerge within the coming years. The inclusion of trios allows us to apply combined linkage and association studies for this purpose. Furthermore, a combined analysis of mitochondrial and nuclear DNA will be possible through the combination of data from the genetic database in Kiel with data on mtDNA from a new database created for the storage, retrieval and analysis of data. In addition, the mtDNA database will represent one of the largest collections of mtDNA sequence data from a single study, and by adding other already published sequences, it will constitute one of the largest mtDNA databases in the world.

By continued follow-up of the old participants in this study, the families with the most extreme longevity phenotype will be identified. This information will increase the power of the GEHA project for identification of longevity genes and other genes related to healthy ageing compared with cross-sectional studies.

## Acknowledgments

### Funding

The work described in this paper was funded mainly by the EU GEHA Project contract no. LSHM-CT-2004-503-270.

The work has additionally been supported by the following programs and agencies:

The Competitive Research Funding of the Tampere University Hospital and Academy of Finland (Tampere);

United States National Institute of Aging (PO1-AG08761) (Odense);

The Innovation Oriented Research Program on Genomics (Senter-Novem IGE05007), the Centre for Medical Systems Biology (CMSB) and the National Institute for Healthy Ageing (NCHA 05060810), all in the framework of the Netherlands Genomics Initiative (NGI)/Netherlands Organisation for Scientific Research (NWO) (Leiden);

The Institute for Ageing and Health and the UK NIHR Biomedical Research Centre for Ageing and Age-related disease award to the Newcastle-upon-Tyne Foundation Hospitals NHS Trust (Newcastle);

Fondation Caisse d'Epargne Rhône-Alpes Lyon CERAL (2004–2007) (Montpellier);

Regione Autonoma della Sardegna (Sassari)

## Appendix A Identification of families with old siblings

Two principal methods were employed for the identification and enrolment of participants in the GEHA project. In one method a list with identified sib-ships with old participants was created, before contact with possible participants was established. In the other method, such a list could not be obtained, and it was necessary to contact an old person and inquire about the existence of a living old sibling, before it could be established whether an old sib pair could be formed. In this appendix the method employed at each recruitment centre is described. The population data are from Eurostat, 2004 (<http://epp.eurostat.ec.europa.eu>).

### Odense, Denmark

The recruitment area covered all of Denmark. The population is 5.4 million inhabitants.

In 2004, a list of old sib pairs was established based on information from the central national population register and local birth registers. First, all living persons in Denmark born before 1918 were extracted from the central nation population register. The information extracted included date and place of birth (parish) as well as current and historic name. Using a matching algorithm, individuals with the same historical surname who were born in the same parish were combined in groups of up to 4 members. Approximately 20,000 groups were created.

Subsequently, birth registers in each parish were scrutinised to identify persons with the same parents listed, thus comprising an old sib pair. As part of this procedure, information about date of birth, and thus age, was verified for all potential old participants in GEHA. The total number of families with old sib pairs identified in this manner was 3644. Sib pairs were considered eligible for enrolment in the GEHA project when at least two siblings were 90 years of age or older. In total 1229 GEHA families were contacted with 476 families giving a positive response.

The younger control persons were selected from the central national population register as a random sample of individuals aged 50-75 years and with residence on the island of Funen. A sufficient number of persons were invited to take part in the GEHA project. An appointment

was made for those who agreed to participate, either as an in-home visit or at the Institute of Public Health at the University of Southern Denmark, Odense.

## Tampere, Finland

The recruitment area covered all of Finland. The population is 5.2 million inhabitants.

The majority of the old sib pairs were identified as a result of advertisements and the appearance of the Finnish GEHA coordinator in TV, radio, and newspapers. A minor fraction of old sibs were found through census lists.

In total, 325 families were contacted, of which 234 gave a positive response. The remaining families did not participate because of death (31 families), dementia (9 families) or refusal of participation by one or more members of the old sibling pair (51 families). One family was excluded because a sibling was untraceable.

Younger control persons were recruited in the same manner as old participants.

## Belfast, Northern Ireland

Recruitment took place in all parts of Northern Ireland. The population is about 1.6 million inhabitants.

Identification of old participants was mainly made through general practitioners. A request for help in identifying persons aged 90 years or older was made of all general practitioners in Northern Ireland. If the persons were willing, the GP provided names and addresses so that a formal invitation letter could be sent to the potential participants. A follow-up was made by a research nurse to determine whether there was a living sibling over the age of 90; if so, an invitation letter was also sent to the old sibling.

The identification procedure was very time consuming because only approximately 10% of the GPs were willing to give names and addresses of potential old participants.

Therefore, other methods of identifying potential participants, including advertising and personal contacts, were used. However, the number of old participants recruited through personal contact and as result of advertising was low.

The number of persons contacted, including those contacted through advertisements, was approximately 10,500, leading to the identification of 84 GEHA families. Of these families, a total of 77 families gave positive response and were included in the study.

Younger control persons were mainly spouses of a family member of an old participant, and interview and blood sample collection took place at the time of the older person's visit. In addition, as a result of advertisement a number of volunteers from "Sheltering Housing" or "Age Concern NI" Sectors were included.

## Newcastle, United Kingdom

The recruitment area covers the regions Newcastle upon Tyne, North Tyneside, Gateshead, Sunderland, and Northumberland. The population is about 1.2 million inhabitants.

The identification of families with at least two nonagenarian siblings was problematic and slow. Within the UK there is no civil register from which surviving nonagenarian siblings can be identified. The recruitment strategy therefore required contacting all nonagenarians within the target area and asking whether they had a living nonagenarian sibling. If they

answered yes, the consent of the proband to participate and to contact his/her sibling(s) was sought.

Two different approaches were used to locate nonagenarians. First, an indirect method was attempted; this method involved contact with a general practitioner in order to gain permission to subsequently contact potential participants from the individuals registered with the GP. This method was found to be very unproductive because of a very low participation rate from the GPs.

Therefore, a second method that required directly contacting nonagenarians was employed. After obtaining the necessary permissions lists of nonagenarians were requested and obtained from the Contractor Service Authority (CSA). Information about the study was sent to the GPs in the study area. All nonagenarians in the target area were contacted by mail and were asked whether they had a living nonagenarian sibling(s), and whether they wanted to participate in the study. When a positive reply was received, the potential participant was contacted by telephone to find out the exact details of his/her sibling.

The number of persons contacted (including those who were contacted through advertisements) to identify GEHA families was 8898, leading to identification of 109 GEHA families. Of these families, a total of 106 families gave positive response and were included in the study.

Younger control subjects were initially recruited from family and friends of staff working within the Institute for Ageing and Health at Newcastle University. When this pool had been exhausted, an advertisement was placed on the University web site inviting people to take part in the study. Each potential volunteer was contacted by telephone and was provided with a verbal explanation of the study and an information booklet prior to an appointment being made for a face-to-face visit from a research nurse for completion of the questionnaire and collection of blood samples.

## Leiden, the Netherlands

The recruitment area covers all of the Netherlands which has about 16.4 million inhabitants.

Information on name, date and place of birth, and place of residence for all living inhabitants aged 90 years and over was requested from the local population register in each municipality in the Netherlands (467 municipalities in 2005). In addition, information about the parents of these individuals (i.e., their names and places and dates of birth) was obtained. Living nonagenarian sibling pairs were identified by searching all files on the same parental names and dates of birth.

Information about the GEHA project was sent to the identified living nonagenarian sibling pairs, and they were asked whether they would like to be provided with additional information about the study. If so, a nurse visited the sibling pair to inform them in detail about the GEHA project and the interview. For those who wished to participate in the project, a second visit was planned for the interview, the informed consent and blood collection.

A significant fraction of the younger control persons recruited for the study comprised the spouses of children of old participants. In addition, young controls were found among acquaintances and colleagues of the families of the participants or the recruitment team.

## Kiel, Germany

The recruitment area covers Schleswig-Holstein and the city of Hamburg in the northern part of Germany. The area has about 4.7 million inhabitants.

The names and addresses of persons aged 90 years or older were obtained from the registry offices in this region. Because no information about potential living siblings was available *a priori*, an introduction letter together with a reply card was sent to all individuals aged 90 years or older asking whether they were willing to participate in the study. On the reply card it was possible to indicate a telephone number and address for contact and whether the respondent had any living siblings also aged 90 years or more. Only potential participants who returned the reply card were contacted by phone. During the telephone call, it was determined whether the person and his/her old sibling were eligible for participation; if so, appointments for home visits were made.

In addition to the mailings, advertisements in newspapers and magazines and on radio and television were used to attract potential participants. This method was responsible for 6% of the recruited sib pairs.

A total of 37,748 persons received introduction letters. Among the 2832 individuals who returned the reply cards, a total of 146 families were identified. Of these families, old siblings of 100 families were visited and included in the study. The remaining 46 families were excluded because of refusal, dementia, death of a sibling or other reasons.

The younger controls were primarily sought among spouses of children of the old participants. When no one was available or volunteered, the young controls were recruited by the same means as the long-lived individuals, i.e., suitable candidates were contacted by mail after obtaining their names and addresses from the respective registry offices.

## Louvain, Belgium

The recruitment area covers the French speaking part of Belgium and has about 4.3 million inhabitants.

From the information in the national population register, an address list was constructed comprising groups of two persons aged 90 years or more with the same family name who were born in the same village. One of the two persons aged 90 years or more was first contacted by telephone to determine whether the two were actually siblings and whether the other was still alive. Whether both siblings were in good mental and physical condition and able to understand the purpose of the study and to agree with it was also determined.

If the response was positive, a letter describing the GEHA project was sent to all identified siblings and to all individuals aged 90 years or more when no information on relationship was available. After approximately 2 weeks, possible participants were contacted again whether they wished to participate.

The number of persons contacted (including those from advertisements) to identify GEHA families was 3001, leading to identification of 372 GEHA families. Of these families, a total of 105 families gave positive responses and were included in the study.

The preferred method of recruitment of younger control persons was to find a spouse of a child of the old participants, but this was rarely possible. Therefore, younger control persons were primarily recruited through advertisements at weekly conferences of the “Université des Aînés” (University for elderly-retired persons).

## Warsaw, Poland

The recruitment area covers four regions of Poland: Mazowieckie, Wielkopolskie, Lodzkie, and Kujawsko-Pomorskie. The total population of these four regions is about 13.2 million.

From the national electronic census register (PESEL) the Ministry of Interior and Administration in Poland provided a nationwide list of potential sib pairs containing names, family names, places and dates of birth, and the actual addresses of a total of 15192 individuals. Sib pairs were identified according to the first and maiden name of the mother and the first name and family name of the father. Within the recruitment area a total of 1084 families were identified and contacted, and agreement to participate was obtained from 145 families.

A few younger control persons were found among the spouses of GEHA-families' offspring. However, the majority of controls were found by contacting families, acquaintances and colleagues who knew people aged 50–75 years born in the same area as the recruited sib pair but unrelated to them and with no relatives among the recruited sib pairs.

## Montpellier, France

The recruitment took place in three regions, Languedoc-Roussillon, Rhône-Alpes, and to some extent Ile de France. The population of the first two regions is about 8.5 million.

Two different ways of identifying old participants were used. In the first part of the study period, general practitioners, nursing homes and municipalities were contacted and were asked whether they knew any persons aged 90 years or more. If they did, details were obtained, a family member was contacted by telephone, the project was explained, and the caller was offered to send an informational letter. The contact was also asked whether he or she had any living siblings aged 90 years or more. If they did, the family member was contacted again after a week, and if both old siblings wished to participate, an appointment was made for a home visit. It was estimated that 1 in 20 contacted nonagenarians had a living nonagenarian brother or sister. This procedure, which was responsible for almost 40% of the total recruitment, was very cumbersome and time consuming because of its dependence on indirect contact with families with nonagenarians. Therefore, in the second part of the study, a list of persons receiving old-age pensions was obtained from the Caisse nationale d'assurance vieillesse (CNAV). The list was sorted by surname and place of birth, and nonagenarians with the same surname and birth place were identified and contacted directly.

Advertisement by means of radio, TV, newspapers and magazines had very little effect on the recruitment success rate; in this region, less than 5% of the old participants were recruited as result of such advertising.

Altogether, a total of 18300 persons were contacted (including those from advertisements) to identify a total of 909 GEHA families. Of these families, 607 families did not enter the study because of no interest (68%), poor health (26%), death of one sibling (2%), or other reason (4%).

The younger control persons were identified through electoral lists. Letters were sent to a random sample of persons asking them to participate. In case of a positive response, a nurse arranged for an interview and blood sampling. A total of 1200 younger persons were contacted, yielding 351 younger controls of which 302 were linked to a GEHA family.



## Kiev, Ukraine

The recruitment area mainly covers the Kiev region and regions within 300 km of Kiev. However, a small fraction (9%) of the participants resided at distances up to 800 km from Kiev in regions in the West Ukraine (the Carpathians) and the South Ukraine (the Crimea).

The source for identifying possible old participants was registers of retired people from the local social security offices in the included regions. Initial contact with the nonagenarians was either by telephone or with the help of employees of the local social services. From this contact, the existence of living siblings was determined and possible further participation was clarified.

The number of persons contacted (including those from advertisements) to identify GEHA families was 13204, leading to identification of 316 GEHA families. Of these families, a total of 121 families gave a positive response.

Younger controls were selected from the same areas where the nonagenarian siblings were residing. In general, neighbours belonging to the same ethnic group as the nonagenarian siblings were chosen.

## Athens, Greece

The recruitment area covers the whole country, which has a population of about 11.2 million.

In Greece, the identification of families with old siblings involved several different methods. The main source for identification of potential old participants was lists of persons receiving old-age pensions from the Agricultural Organization of Insurance. From this list, potential participants were identified from their surnames and the names of their parents. However, because only their names and regions were listed, additional information including addresses and telephone numbers had to be obtained from municipalities, post offices, and telephone companies. The first contact to potential participants was a phone call in which an explanation of the project was given together with an offer to have some biochemical tests made. If the person had a living old sibling and if agreed to participate, a home visit was arranged.

Advertisements by means of radio, TV, newspapers and magazines were also used to recruit families with old siblings, but this had very little effect on the recruitment success rate, and less than 5% of the old participants were recruited in this manner.

In total, 459 old persons were contacted, and the number of families with at least 2 old siblings identified was 295.

The younger controls were primarily sought among spouses of children or other relatives of the old participants; when no one was available or volunteered, the young controls were sought among persons living in the same neighbourhood as sibs or in a Senior Centre Open Protection in the same area.

## Bologna, Italy

The recruitment area covers mainly the Emilia Romagna region (located in Northern Italy with about 4.1 million inhabitants) and some cities in the regions of Toscana, Veneto, Marche, Liguria, Lombardia, Abruzzo, and Piemonte. About 80% of the recruited families came from the Emilia Romagna region.

The population Register Office of Bologna (Emilia Romagna), Livorno (Tuscany) and Varese Ligure (Liguria) provided a list of sib pairs with individuals aged 90 years or more that contained information on name, date and place of birth, and residence, thus allowing researchers to contact the subjects in the list without gathering additional information. For the other geographical areas, because of the exhaustion of the sib pairs lists provided by the official Population Registers, sib pairs were enrolled on the basis of an advertisement-based recruitment strategy. Articles on the GEHA project asking for sib pairs with individuals aged 90 years or more with specific requirements were published in local newspapers and popular magazines. Moreover, the principal investigator participated in televised scientific programs in which he asked sib pairs with individuals aged 90 years or more from Northern Italy to contact the recruiting centre to be enrolled in the project.

In total, 804 families were contacted, of which at least two old siblings agreed to participate in 258 families.

The younger controls were primarily recruited among spouses of children of the old participants. If this was not possible, a young control was randomly recruited in the population or sought among subjects already known to the recruitment team.

## Rome, Italy

The recruitment area covers the city of Rome and has about 2.5 million inhabitants.

A list of all residents aged 86 years or more that included the names of each individual's parents was received from the municipality of Rome. From this list, a total of 1246 families with 2 or more sibs were identified and linked to the Rome telephone registry, resulting in a database with addresses and telephone numbers of 863 families with at least one sibling aged 90 years or more. All these families were contacted by mail. Approximately 2 weeks after the invitation letter, a follow-up was made by telephone; of 623 families contacted, 106 families entered the study.

The younger control persons were primarily sought among spouses of the children of the oldest participants. In total, 76% of the younger control persons were spouses of a child of a participant, while the remaining younger controls were neighbours.

## Calabria, Italy

The recruitment area covers the region of Calabria in southern Italy and has about 2.0 million inhabitants.

Families with old siblings were identified on the basis of demographic information obtained from local municipalities. All 409 municipalities of Calabria were initially contacted asking for lists of persons born in 1915 or earlier who resided in the municipalities. Of 409 municipalities, 176 (43%) sent the relevant list. From a population of 12551 nonagenarians included in the lists, potential siblings were identified by visiting local municipalities and performing genealogical reconstruction of families. At the same time, the age and family relations of the selected persons aged 90 years or more were confirmed, and birth certificates were obtained. Eligible persons were contacted by phone, and the aim of the GEHA project and procedure for participation was explained. If the person agreed to participate, an appointment for a home visit was made.

In total, 425 families were contacted, of which 200 agreed to participate. The remaining 225 families did not enter the project — 72 because of the death of at least one member in the period between the acquisition of the list and the contact for the visit (3 months on average),

81 because at least one member refused, 21 were untraceable, and 51 were excluded for dementia.

Younger control persons were primarily found among the spouses of the children of the old participants. When it was impossible to find a spouse, a younger control person was found by contacting a general practitioner and by looking for a person aged between 50 and 75 years of age born in the same area as the recruited sib pair but unrelated to them and with no relatives among the recruited sib pairs.

## Sardinia, Italy

The recruitment area covers the whole island of Sardinia, which has about 1.6 million inhabitants.

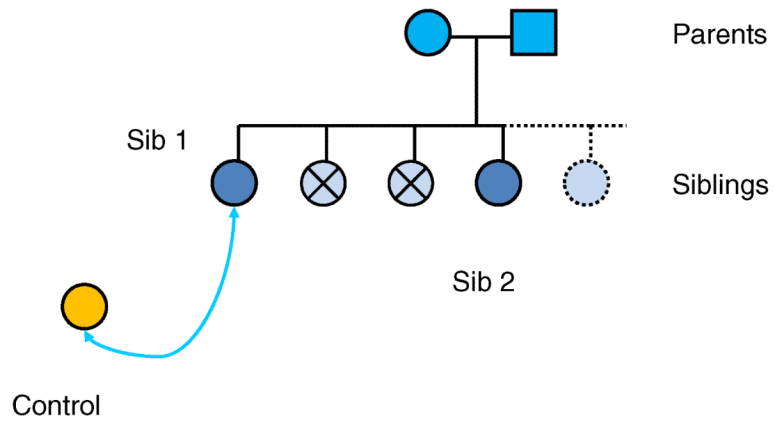
Demographic information was obtained from local municipalities. A letter was sent to each of the 377 municipalities asking for information about any families with nonagenarian siblings, and some municipalities replied with a list of persons. The first contact with potential participants was a telephone call in which a description of the project was given; additionally the participants were offered some biochemical tests. If the person had a living old sibling and agreed to participate, a home visit was arranged. In total, 159 families were contacted, and 124 gave a positive response.

Younger controls were selected in the same areas where the nonagenarian siblings were residing. These controls were primarily sought among the relatives of the old participants.

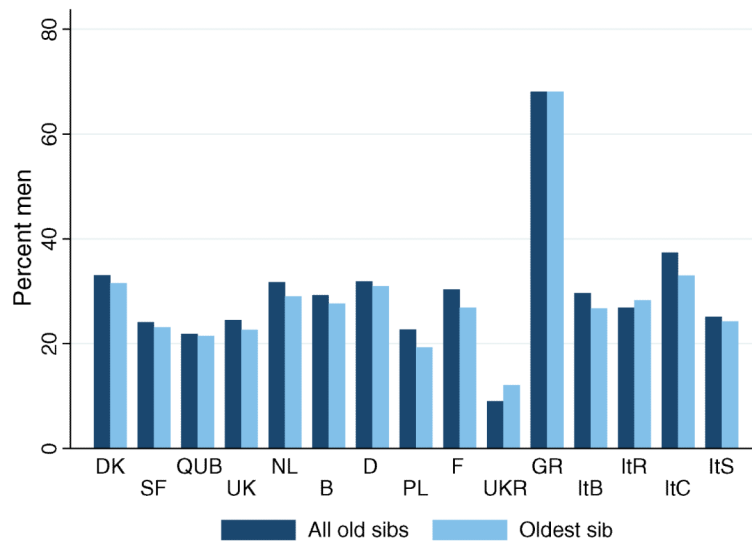
## References

- Beekman M, Blauw GJ, Houwing-Duistermaat JJ, Brandt BW, Westendorp RGJ, Slagboom PE. Chromosome 4q25, microsomal transfer protein gene, and human longevity: novel data and a meta-analysis of association studies. *J. Gerontol. A Biol. Sci. Med. Sci.* 2006; 61 A:355–362. [PubMed: 16611701]
- Christensen K, Johnson TE, Vaupel JW. The quest for genetic determinants of human longevity: challenges and insights. *Nat. Rev. Genet.* 2007; 7:436–448. [PubMed: 16708071]
- Christensen K, Doblhammer G, Rau R, Vaupel JW. Ageing populations: the challenges ahead. *Lancet.* 2009; 374:1196–1208. [PubMed: 19801098]
- Dato S, Passarino G, Rose G, Altomare K, Belizzi D, Mari V, Feraco E, Franceschi C, De Benedictis G. Association of the mitochondrial DNA haplogroup J with longevity is population specific. *Eur. J. Hum. Genet.* 2004; 12:1080–1082. [PubMed: 15470367]
- De Benedictis G, Tan Q, Jeune B. Recent advances in human gene-longevity association studies. *Mech. Ageing Dev.* 2001; 122:909–920. [PubMed: 11348658]
- Eurostat. Regions-Nomenclature of territorial units for statistics-NUTS-2003/EU25 2004 edition. Office des publications officielles des Communautés européennes; Luxembourg: 2004.
- Flachsbart F, Caliebe A, Kleindorp R, Blanche H, von Eller-Eberstein H, Nikolaus S, Schreiber S, Nebel A. Association of FOXO3A variation with human longevity confirmed in German centenarians. *Proc. Natl. Acad. Sci. U.S.A.* 2009; 106:2700–2705. [PubMed: 19196970]
- Franceschi C, Bezzukov V, Blanché H, Bolund L, Christensen K, De Benedictis G, Deiana L, Gonos E, Hervoonen A, Yang H, Jeune B, Kirkwood TBL, Kristensen P, Leon A, Pelicci PG, Peltonen L, Poulain M, Rea IM, Remacle J, Robine JM, Schreiber S, Sikora E, Slagboom PE, Spazzafumo L, Stazi MA, Toussaint O, Vaupel JW. Genetics of Healthy Aging in Europe. The EU-Integrated Project GEHA (Genetics of Healthy Aging). *Ann. N. Y. Acad. Sci.* 2007; 1100:21–45. [PubMed: 17460163]
- Herskind AM, McGue M, Holm NV, Sørensen TI, Harvald B, Vaupel JW. The heritability of human longevity: a population based study of 2874 Danish twin pairs born 1870–1900. *Hum. Genet.* 1996; 97:319–323. [PubMed: 8786073]

- Hjemborg, J.v.B.; Iachine, I.; Skytthe, A.; Vaupel, JW.; McGue, M.; Koskenvuo, M.; Kaprio, J.; Pedersen, NL.; Christensen, K. Genetic influence on human lifespan and longevity. *Hum. Genet.* 2006; 119:312–321. [PubMed: 16463022]
- Bøyum A. Isolation of lymphocytes, granulocytes and macrophages. *Scand. J. Immunol. Suppl.* 1976; 5:9–15.
- Katz S, Ford AB, Moskowitz RW, Jackson BA, Jaffe MW, Cleveland MA. Studies of illness in the aged. The Index of ADL: a standardized measure of biological and psychosocial function. *J. Amer. Med. Assoc.* 1963; 185:914–919.
- Lauritsen, JM., editor. EpiData Data Entry, Data Management and Basic Statistical Analysis System. EpiData Association; Odense, Denmark: 2000-2008. <http://www.epidata.dk>
- Molloy DW, Alemayehu E, Roberts R. Reliability of a Standardized Mini-Mental State Examination compared with the traditional Mini-Mental State Examination. *Am. J. Psychiatry.* 1991; 148:102–105. [PubMed: 1984692]
- Schächter F, Cohen D, Kirkwood T. Prospects for the genetics of human longevity. *Hum. Genet.* 1993; 91:519–526. [PubMed: 8340104]
- Willcox BJ, Donlon TA, He Q, Chen R, Grove JS, Yano K, Masaki KH, Willcox DC, Rodriguez B, Curb JD. FOXO3A genotype is strongly associated with human longevity. *Proc. Natl. Acad. Sci. U.S.A.* 2008; 105:13987–13992. [PubMed: 18765803]



**Fig. 1.** Organization of a trio consisting of a GEHA family with 2 living siblings and a younger control matched for gender with the oldest participating sibling.



**Fig. 2.** Proportion of men among all old siblings and proportion of men who are the oldest siblings in GEHA families. DK = Odense, SF = Tampere, QUB = Belfast, UK = Newcastle, NL = Leiden, B = Louvain, D = Kiel, PL = Warsaw, F = Montpellier, UKR = Kiev, GR = Athens, ItB = Bologna, ItR = Rome, ItC = Calabria, ItS = Sardinia.

**Table 1**

GEHA partners involved in recruitment.

Recruitment centre	GEHA partner	Country	Recruitment region
Odense	University of Southern Denmark	Denmark	Whole country
Tampere	Tampere School of Public Health	Finland	Whole country
Belfast	Queens University of Belfast	United Kingdom	Northern Ireland
Newcastle	University of Newcastle	United Kingdom	Newcastle upon Tyne, North Tyneside, Gateshead, Sunderland, and Northumberland
Leiden	Leiden University Medical Centre	The Netherlands	Whole country
Louvain	Catholic University of Louvain	Belgium	French speaking part (Wallonia and Bruxelles)
Kiel	Institute of Clinical Molecular Biology	Germany	Schleswig-Holstein, Hamburg
Warsaw	Nencki Institute of Experimental Biology	Poland	Mazowieckie, Wielkopolskie, Lodzkie, Kujawsko-Pomorskie
Montpellier	University of Montpellier UM1	France	Languedoc Roussillon, Rhône Alpes, and to some extent in Ile de France
Kiev	Institute of Gerontology	Ukraine	Kiev region and regions within 800 km from Kiev
Athens	National Hellenic Research Foundation	Greece	Whole country
Bologna	CIG - University of Bologna	Italy	Emilia Romagna, Toscana, Veneto, Marche, Liguria, Lombardia, Abruzzo, and Piemonte
Rome	Istituto Superiore di Sanità	Italy	Rome
Calabria	University of Calabria	Italy	Calabria
Sardinia	University of Sassari	Italy	Sardinia

**Table 2**

Approval of the GEHA project and periods of recruitment by partners.

Recruitment centre	Approval by local ethical committee	Recruitment started	Recruitment ended	Comments
Odense	October 11, 2004	August 2004 (pilot study February 2004)	October 2007	Approval by ethical committee was obtained as an extension/amendment to an ongoing study at SDU (the Longitudinal Study of Aging Danish Twins)
Tampere	May 17, 2004	June 2004	January 2008	
Belfast	February 9, 2005	September 2005	May 2009	Obtaining approval from Ethical Committee was a lengthy process because of changes in UK-wide and local procedures for the evaluation of research projects from October 2004
Newcastle	February 9, 2005	May 2005	May 2008	Newcastle led a joint ethical application with Belfast. UK regulatory procedures were very significantly tightened just prior to GEHA application. Thus, substantial amendments to the consent procedures were required to comply with UK regulations. Some reworking of protocols was also required. Additional local approval had to be obtained from each local committee and Primary Care organisations within the study area
Leiden		May 2004	July 2007	Approval by ethical committee was obtained as an extension/amendment to an ongoing study at LUMC (the Leiden Longevity Study)
Louvain	January 12, 2005	February 2005	March 2007	
Kiel	December 9, 2004	January 2005	August 2008	
Warsaw	October 5, 2004	January 2005	August 2008	
Montpellier	December 14, 2004	November 2004	February 2007	
Kiev	April 1, 2006	April 2006	August 2008	
Athens	September 20, 2004	February 2005	May 2008	Approval was obtained from the Hellenic Data Protection Authority on 17 February 2005
Bologna	First approval July 20, 2004	November 2004	May 2008	Second and third approvals were obtained from the ethical committee after modifications to the questionnaire and the informed consent form. Final approval January 24, 2007
Rome	December 2, 2004	April 2005	August 2008	
Calabria	November 1, 2004	November 2004	April 2007	
Sardinia		June 2005	October 2008	



**Table 3**

Recruitment effort for recruitment centres *without* prior knowledge of familial status. Number of persons contacted.

Recruitment centre	Number of persons contacted	Number of eligible families	Number of families with positive response	Percent with positive response
Belfast	10500	84	77	91.6
Newcastle	8898	187	109	58.3
Louvain	3001	372	105	28.2
Kiel	37748	146	100	68.5
Montpellier	18,300	909	304	33.4
Kiev	13204	316	121	38.3
Total	91651	2014	816	40.5

**Table 4**

Recruitment effort for recruitment centres *with* prior knowledge of familial status. Number of families contacted.

Recruitment centre	Number of families contacted	Number of families with positive response	Percent with positive response
Odense	1229	476	38.7
Tampere	325	234	72.0
Leiden	377	229	60.7
Warsaw	1084	145	13.4
Athens	459	295	64.3
Bologna	804	258	32.1
Rome	623	106	17.0
Calabria	425	200	47.1
Sardinia	159	124	78.0
Total	5485	2067	37.7

**Table 5**

Number of participants by recruitment centre. Numbers under the heading “Interviews” are based on data from the phenotypic database, while numbers under the heading “Samples” are based on the samples received in Helsinki.

Recruitment centre	<u>Interviews with participants</u>		<u>Participants donating samples</u>	
	Old siblings	Young controls	Old siblings	Young controls
Odense	1004	482	1003	482
Tampere	453	177	450	176
Belfast	160	68	156	68
Newcastle	204	105	204	105
Leiden	428	178	428	178
Louvain	212	105	212	105
Kiel	204	100	204	100
Warsaw	295	138	295	138
Montpellier	662	351	657	351
Kiev	167	75	167	75
Athens	273	136	273	136
Bologna	549	254	543	254
Rome	216	87	216	87
Calabria	412	200	411	200
Sardinia	179	92	162	82
Total	5418	2548	5381	2537

**Table 6**

Number of trios by recruitment centre. Complete trios have blood samples collected from at least two old siblings and the younger control. Incomplete trios have blood samples from fewer than two old siblings and/or from the younger control. *Based on data from samples received in Helsinki.*

Recruitment centre	Sib-ship composition of trio					Total
	2 old siblings	3 old siblings	4 old siblings	5 old siblings	Incomplete	
Odense	396	46	4	0	446	16 462
Tampere	129	27	1	1	158	55 213
Belfast	51	4	1	0	56	27 83
Newcastle	95	1	0	0	96	10 106
Leiden	151	15	2	0	168	35 203
Louvain	84	1	0	0	85	20 105
Kiel	94	5	0	0	99	1 100
Warsaw	134	4	0	0	138	7 145
Montpellier	246	28	3	1	278	26 304
Kiev	47	2	0	0	49	34 83
Athens	95	3	0	0	98	33 131
Bologna	192	24	6	1	223	12 235
Rome	79	2	0	0	81	25 106
Calabria	192	7	1	0	200	0 200
Sardinia	72	2	0	0	74	9 83
Total	2057	171	18	3	2249	310 2559

**Table 7**

Descriptive characteristics of participants. Note that old siblings aged 89 years at interview were included in the study provided they lived past their 90th birthdays.

	Minimum age	Median age	Maximum age	Number of participants
<i>All old siblings</i>				
Men	89.22	92.68	103.98	1678
Women	89.01	93.05	107.13	3740
<i>Oldest sibling</i>				
Men	90.30	94.30	103.98	766
Women	90.00	94.67	107.13	1868
<i>Younger controls</i>				
Men	50.01	63.94	75.63	856
Women	50.00	61.75	75.92	1692

**Table 8**

Gender composition of enrolled GEHA-families with interviewed participants. F = Female, M = Male. FF indicates 2 sisters; FM indicates one sister and one brother.

Sib-ship composition	Gender	Number of families	Percent	Percent of all families
– 2 old siblings	FF	1176	50.8	46.39
	FM	848	36.6	33.45
	MM	291	12.6	11.48
	Total	2315	100.0	91.32
– 3 old siblings	FFF	67	34.4	2.64
	FFM	76	39.0	3.00
	FMM	48	24.6	1.89
	MMM	4	2.0	0.16
	Total	195	100.0	7.68
– 4 old siblings	FFFF	5	23.8	0.20
	FFFM	6	28.5	0.24
	FFMM	8	38.1	0.32
	FMMM	1	4.8	0.04
	MMMM	1	4.8	0.04
	Total	21	100.0	0.84
– 5 old siblings	FFFFF	1	25.0	0.04
	FFFFM	1	25.0	0.04
	FFFMM	2	50.0	0.08
	Total	4	100.0	0.16
All families		2535		100.00