International Journal of Epidemiology, 2014, 1–9
doi: 10.1093/ije/dyu229
Cohort Profile

Cohort Profile

Cohort Profile: LifeLines, a three-generation cohort study and biobank

Salome Scholtens,[1]* Nynke Smidt,[1,2] Morris A Swertz,[3,4] Stephan JL Bakker,[5] Aafje Dotinga,[1] Judith M Vonk,[1,2] Freerk van Dijk,[3,4] Sander KR van Zon,[6] Cisca Wijmenga,[4] Bruce HR Wolffenbuttel[7] and Ronald P Stolk[1,2]

[1]LifeLines Cohort Study, Groningen, The Netherlands, [2]Department of Epidemiology, [3]Genomics Coordination Center, [4]Department of Genetics, [5]Department of Internal Medicine, [6]Department of Health Sciences, Community and Occupational Medicine and [7]Department of Endocrinology, University Medical Center Groningen, Groningen, The Netherlands

*Corresponding author. LifeLines Cohort Study and Biobank, PO Box 30.001, 9713 BZ Groningen, The Netherlands. E-mail: s.scholtens@umcg.nl

Accepted 13 November 2014

Abstract

The LifeLines Cohort Study is a large population-based cohort study and biobank that was established as a resource for research on complex interactions between environmental, phenotypic and genomic factors in the development of chronic diseases and healthy ageing. Between 2006 and 2013, inhabitants of the northern part of The Netherlands and their families were invited to participate, thereby contributing to a three-generation design. Participants visited one of the LifeLines research sites for a physical examination, including lung function, ECG and cognition tests, and completed extensive questionnaires. Baseline data were collected for 167 729 participants, aged from 6 months to 93 years. Follow-up visits are scheduled every 5 years, and in between participants receive follow-up questionnaires. Linkage is being established with medical registries and environmental data. LifeLines contains information on biochemistry, medical history, psychosocial characteristics, lifestyle and more. Genomic data are available including genome-wide genetic data of 15 638 participants. Fasting blood and 24-h urine samples are processed on the day of collection and stored at -80 °C in a fully automated storage facility. The aim of LifeLines is to be a resource for the national and international scientific community. Requests for data and biomaterials can be submitted to the LifeLines Research Office [LLscience@umcg.nl].

Key Messages

- LifeLines is one of the major biobanks which make up the Dutch national biobank infrastructure (Biobanking and BioMolecular resources Research Infrastructure, BBMRI-NL). It makes its data and biomaterials available for researchers worldwide.
- The LifeLines data collection, the highly automated LifeLines laboratory and storage facility at −80 °C (over 8 million aliquots) are based on high standards with ISO9001 certification.
- Data are released to researchers within a remote system (LifeLines workspace) running on a high-performance computer cluster to ensure data quality and security.
- Up to now, 215 requests for data, biomaterials and for additional data or biomaterial collection among the LifeLines participants have been granted. Currently about 60 researchers worldwide are working with LifeLines data and biomaterials.

Why was the cohort set up?

The LifeLines Cohort Study, a large observational population-based cohort study in the north of The Netherlands, was established in 2006 as a resource for international researchers.[1] The overall aim of the LifeLines Cohort Study is to gain insight into the aetiology of healthy ageing. Further understanding of why, for instance, some people with overweight develop diabetes, some develop cardiovascular disorders and others remain free of disease will enable us to move forward to more personalized prevention and personalized medicine, ultimately resulting in adding additional years of healthy life.

Healthy ageing is a broad concept and does not only mean being free of age-related chronic diseases, but also includes physical and cognitive capability and social and psychological well-being.[2] Since these aspects of healthy ageing do not stand on their own, but are all connected and result from complex interactions during the life course between a wide range of environmental exposures, phenotypic characteristics and genomic factors, the study of healthy ageing challenges existing study designs.[3–5] The LifeLines Cohort Study, encompassing a prospective design with long-term follow-up, a large number of participants and an extensive biomaterial and data collection including multiple exposure variables and endpoints, is designed to address the important questions on healthy ageing. In combination with a three-generation design, the LifeLines Cohort Study allows for research on the disentanglement of genetic, lifestyle and environmental contributions to the development of chronic diseases, the study of between-generation similarities and the identification of preclinical stages of ageing at an early age.[1]

The LifeLines Cohort Study is a national endeavour in coordination with the University Medical Center in Groningen, The Netherlands. The population in the north of The Netherlands is highly suitable for this type of study because of its homogeneous composition and low migration rates relative to other parts of The Netherlands. By the end of 2006 the first participants were included, and in December 2013 the recruitment period was closed after reaching the target number of 165 000 participants. At that time, the total number of participants included was 167 729, who will be followed for at least 30 years.

The current paper outlines the study design, the study population and an overview of the data and biomaterial collected.

Who is in the cohort?

Inclusion

Since all inhabitants in The Netherlands are registered with a general practitioner (GP), eligible participants were invited to participate in the LifeLines Cohort Study through their GP. A large number of GPs within the northern three provinces of The Netherlands (Friesland, Groningen and Drenthe) were involved and invited all their patients between the ages of 25 and 50 years, unless the participating GP considered the patient not eligible based on the following criteria:

- severe psychiatric or physical illness;
- limited life expectancy (<5 years);
- insufficient knowledge of the Dutch language to complete a Dutch questionnaire.

Subsequently, individuals who were interested to participate received detailed information by mail about the LifeLines Cohort Study, and an informed consent form. After the signed informed consent was received by the LifeLines organization, the participants received a baseline questionnaire and an invitation to a comprehensive health assessment at the LifeLines research site. During the visit, participants were asked to indicate whether their family members, such as partners, parents, parents-in-law and

children would also be willing to participate in the study. If so, permission was asked to send them an invitation to participate. Children could only participate if one of their parents was a participant. In addition, inhabitants of the northern provinces could also register themselves via the LifeLines website.

Participant and non-responder characteristics

Overall, the majority of the invited GPs agreed to participate ($n = 562/812$; 73%). From those GPs who agreed to participate, 333 307 persons were approached (Figure 1). The main reason for GPs not participating in the LifeLines Cohort Study was the anticipated additional workload, due to the questions of patients (with regard to feedback on their health that they could receive from LifeLines). All participants had the option to receive feedback on the results of physical examinations performed during the visits to the LifeLines research site, including anthropometric outcomes, blood pressure, ECG, pulmonary function, cholesterol levels and other biochemistry. The GP of the participant was always informed about the results of the clinical tests.

In total, 49% (81 652/167 729) of the included participants were invited through their GP, 38% (64 489/167 729) were recruited via participating family members and 13% (21 588/167 729) self-registered via the LifeLines website. Almost 25% of the invited persons agreed to participate and were included in the study (Figure 1). The cohort includes children from 6 months of age and elderly up to 93 years of age and somewhat more women than men (Table 1).

Over two-thirds of the participants had at least one family member participating in LifeLines (67%, $n = 112\,596$). The partner of 34% ($n = 56\,944$) of the participants was included and a child of 27% ($n = 45\,169$) of the participants was included. In total, 51% of the participants ($n = 84\,888$) were part of a two-generation relationship that was embedded in the study, and 12% ($n = 20\,362$) of a three-generation relationship.

Ethical considerations

All participants signed an informed consent form before they received an invitation for the physical examination. The LifeLines Cohort Study is conducted according to the principles of the Declaration of Helsinki and in accordance with research code UMCG. The LifeLines study is approved by the medical ethical committee of the University Medical Center Groningen, The Netherlands.

How often have they been followed up?

Every 1.5 year a follow-up questionnaire is administered to all participants, and they are invited for a renewed physical examination at the LifeLines research site on average every 5 years. After completion of the inclusion in December 2013, the cohort have been invited for a comprehensive follow-up physical examination and biobanking from January 2014 onwards. Linkage is being established between the LifeLines database and records of general practitioners, records on drugs dispensed by pharmacies, and other health and national registries, using a third-party pseudonymization system. The LifeLines protocol is open to amendments by any researcher. Within the standing infrastructure, additional questionnaires, measurements and biomaterial collections can be implemented by

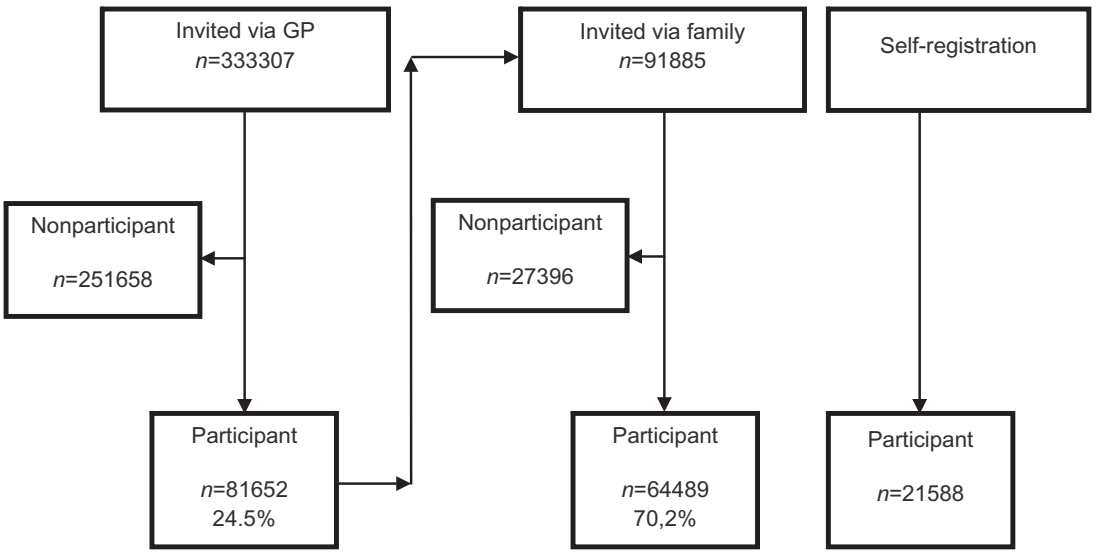

Figure 1. Flow chart of inclusion route and response to invitation to participate.

Table 1. Demographic characteristics of the population of the LifeLines Cohort Study

Characteristics	Total (167729)		GP (81652)		Family (64489)		Self-registration (21588)	
Age (years, mean SD)	41.5	16.1	41.2	7.4	39.5	22.6	48.7	14.9
Age categories								
0–<18 years	14801	8.8%	–	–	14801	23.0%	–	–
18–<65 years	140222	83.6%	81652	100%	39954	62.0%	18616	86.2%
65–93 years	12706	7.6%	–	–	9734	15.0%	2972	13.8%
Sex								
Female	97055	57.9%	47271	57.9%	35978	55.8%	13806	64.0%
Male	70674	42.1%	34381	42.1%	28511	44.2%	7782	36.0%

submitting a research proposal to LifeLines. Depending on the research question of the applicant, the complete LifeLines population or a selection of the population will be approached for additional data and/or biomaterial collection.

What has been measured?

Physical examination

Participants aged 8 years and older were invited to one of the 12 local research sites in the north of The Netherlands for physical examination. The baseline assessment consisted of two visits. During the first visit (duration 60 min) physical examinations were performed by a trained research nurse (Table 2) and containers for collection of a 24-h urine sample were handed out accompanied by oral and written instruction on how to collect this sample. Approximately 2 weeks after the first visit, a second visit (duration 10 min) was arranged to collect a fasting blood sample and hand in the collected 24-h urine.

Questionnaires

The baseline questionnaire consisted of two parts containing questions on, among other topics, demographics, health status, lifestyle and psychosocial aspects (Table 3). The first part was sent to the participant after LifeLines received the informed consent form. The second part was handed out during the first visit. For participants of 65 years and older, the questionnaire was slightly adjusted to the circumstances of elderly people. Children were included in LifeLines from the age of 6 months. All parents of participating children received a questionnaire containing questions on pregnancy, childbirth and health of the child in the first 6 months. Additional questionnaires were administered with questions on lifestyle, behaviour and environmental exposure, specifically suited to the age of the child.

Biomaterial collection and biobanking

During the second visit, blood was drawn and the collected 24-h urine was taken in for direct measurements and for long-term biobanking (Table 4). The collected blood samples included: three 10-ml tubes and one 4-ml tube, which were anticoagulated with EDTA; one 3.5-ml tube which was anticoagulated with heparin; one 6-ml tube which was anticoagulated with citrate; one 2-ml tube which was anticoagulated with sodium fluoride; and two tubes for separation of serum, of which one 10-ml tube was prepared with gel separation and one 6-ml tube was prepared with clot activator. From the 24-h urine collection, three 10-ml vials were drawn for analyses and biobanking. After timed and protocolled coagulation of the tubes intended for separation of serum, the blood and urine samples were placed at 4°C and transported from the LifeLines research site to the LifeLines laboratory in Groningen, under tightly controlled and continuously monitored conditions. From the LifeLines laboratory, part of the samples were directly transferred to the central laboratory of the University Medical Center Groningen, to perform routine clinical chemistry assays on fresh samples. The other part of the samples remained at the LifeLines laboratory, where urine, serum and plasma samples were prepared and DNA was isolated. These serum, plasma, DNA and urine samples were stored in two-dimensional barcoded aliquots at −80 °C, to allow for future measurements.

Genomics

Genome-wide genotype data based on the Illumina CytoSNP-12v2 array are currently available for 15 638 samples, and all independent and Caucasian-ancestry

Table 2. Content of the baseline visit at the LifeLines research site. Details of the content can be found in the LifeLines data catalogues at [www.lifelines.net]

Measurement	Measurement description and tool	Age
Anthropometry	Height without shoes: SECA222 stadiometer Weight without shoes and heavy clothing: SECA 761 scale Waist and hip circumference: SECA 200 measuring tape	From 8 yrs
Blood pressure	10 measurements during 10 min with Dynamap, PRO 100V2	From 8 yrs
Pulmonary function	1 measurements with Welch Allyn version 1.6.0.489	From 8 yrs
Electrocardiogram	1 measurement with Welch Allyn DT100, 12 leads	From 13 yrs
Skin autofluorescence	1 measurement at the lower arm with Advanced Glycation End products (AGE)-reader[6]	From 18 yrs
Neuropsychiatric health	MINI (MINI International Neuropsychiatric Interview)[7] Dutch version 5.0.0, interview by research nurse	From 18 yrs
Cognition	MMSE (minimal mental state examination), interview by research nurse[8] Ruff Figural Fluency Test (RFFT)[9]	From 65 yrs From 18 yrs
Biomaterial collection		
Blood (fasting)		From 8 yrs
Urine	Spot morning sample 24-h urine collection Timed overnight urine collection	From 18 yrs From 18 yrs 8–18 yrs

Table 3. Overview of the content of the baseline questionnaires for adults (18 years and older). Details of the content can be found in the LifeLines data catalogues at [www.LifeLines.net]

Topic	Dimensions (examples)
GENERAL INFORMATION	
Demographics	Marital status, household composition, nationality
Family composition	Date of birth, date of death father and mother, children
Work	Employment, function, income, absences
Education	Highest level of education
HEALTH	
Health status	History and prevalence of diseases, disabilities and symptoms, ECRHS II[10]
Medication use	Current prescription, doses
Healthcare use	Contact with healthcare professionals, receiving informal care
Questions for females	Number of pregnancies, age at menopause and age at menarche
Birth and development	Birthweight, birth defects, breastfeeding
Body weight	Highest and lowest body weight in past 5 years, unwanted weight loss
LIFESTYLE AND ENVIRONMENT	
Physical activity	SQUASH[11]
Nutrition	Food frequency questionnaire (FFQ)[12], alcohol use
Smoking	Current and past active and passive smoking behaviour
Daytime	Sleeping, activities, volunteer work
Physical environment	Type of floor covering at home, domestic animals
PSYCHOSOCIAL PARAMETERS	
Quality of life, well-being	The Short Form (36) Health Survey (SF-36),[13] Positive and Negative Affect Schedule (PANAS)[14]
Health perception, somatization	Symptom Checklist (SCL-90)[15]
Personality	Neuroticism-Extroversion-Openness Personality Inventory,[16], Anxiety Sensitivity Index[17]
Stress	List of Threatening Experiences (LTE),[18] Longterm Difficulties Inventory (LDI)[18]
Social support, independence	Social production function (SPF-IL),[19] Groningen Frailty Index (GFI),[20] INTERMED[21]

Table 4. Overview of available haematological and biochemical measures

Group	Measure
BLOOD	
Haematology	Haemoglobin
	Haematocrit
	Leukocytes and differentiation
	Thrombocytes
	Red cell dimensions
Diabetes	Glucose
	HbA1c
Lipids	Total cholesterol
	HDL-cholesterol
	LDL-cholesterol
	Triglycerides
	Apolipoprotein A1
	Apolipoprotein B100
Electrolytes	Sodium
	Potassium
	Calcium
	Phosphorus
Renal function	Creatinine
	Urea
	Uric acid
Liver and inflammation	Aspartate aminotransferase
	Alanine aminotransferase
	Alkaline phosphatase
	Gamma glutamyl transferase
	Albumin
	High sensitivity C-reactive protein
Thyroid function	Thyroid stimulating hormone
	Free T4
	Free T3
URINE	
Spot morning sample	Albumin
	Creatinine
24-h collection sample	Albumin
	Creatinine

samples ($n = 13\,436$) have also been imputed using the Genome of The Netherlands (GoNL) release 5[22] and the 1000 Genomes phase1 v3[23] reference panels. Quality controls of the data are based on SNP filtering on minor allele frequency (MAF) above 0.001, Hardy-Weinberg equilibrium (HWE) *P*-value >1e-4, call rate of 0.95 using Plink,[24] and principal component analysis (PCA) to check for population outliers—resulting in 268 407 SNPs and 13 436 samples kept for genome-wide association analysis. Before imputation, the genotypes were pre-phased using SHAPEIT2[25] and aligned to the reference panels using Genotype Harmonizer [www.molgenis.org/systemsgenetics] in order to resolve strand issues. The samples were imputed using Minimac[26] version 2012.10.3, yielding 19 562 004 and 28 681 763 SNPs, respectively, for GoNL and 1000 Genomes. The GoNL panel showed more accurate genotypes when imputing European samples compared with the 1000 Genomes Phase1 v3 reference panel.[27] The MOLGENIS compute[28] imputation pipeline was used to generate and monitor our job scripts on the distributed file system. In the near future, additional genomics sets will be added such as additional GWAS, methylation chips, metabolomics and RNAseq.

What has it found? Key findings and publications

Up to 1 September 2014, a total of 215 requests for data and biomaterials and for additional data or biomaterials collection among the LifeLines participants have been approved by the LifeLines scientific board. Requests were submitted by researchers from a wide variety of disciplines, including internal medicine, genetics, public health, environmental epidemiology, nutrition, sociology, demography and economics.

In 2010 the first data became available from participants with whole genome genetic data. From 2012 onwards a number of data releases have followed, with expanding subject numbers and data items. The first complete data release was provided in 2013, including 94 516 participants and more than 180 000 person-years of follow-up. All the baseline data will be available by the end of 2014.

Besides requests for data or biomaterials, 16 proposals for add-on studies have been granted and implemented by LifeLines, leading to additional data collection in subgroups of LifeLines. Examples are 'LifeLines deep', questionnaires among informal caregivers and studies to validate different instruments for cognitive functioning. In 'LifeLines deep', additional biomaterials were collected such as exhaled air and faeces, and extensive gut-related phenotypes are recorded in a subgroup of 1500 participants of the LifeLines population (Tichelaar *et al.*, submitted for publication). The questionnaires on experiences with informal care were sent to approximately 1000 informal caregivers in the LifeLines population. LifeLines is also part of the Genome of The Netherlands Consortium and thereby has contributed to detailed characterization of the genetic variation of the Dutch population and improvement of the imputation quality [www.nlgenome.nl].[22]

Up to September 2014, 18 papers have been published with LifeLines as the main data source, and LifeLines has participated in 28 GWAS papers. A complete list of publications with LifeLines data or biomaterials can be found at [www.LifeLines.net].

Although LifeLines is a relatively young cohort, the papers published so far already show the unique

opportunities for science. The large study population, the wide variety of exposure and disease outcomes and availability of genetic data enabled researchers to study multimorbidity and gene-environment interactions. For example, in pulmonary disease, LifeLines showed interactions between genetic predisposition and occupational exposures in the development of both large and small airways obstruction.[29,30] Also, LifeLines contributed to further insight in the interactions between genetic variants and environmental exposures or between different environmental factors on, among others, lung function level,[31] diabetes[32] and the metabolic syndrome.[33]

What are the main strengths and weaknesses?

The LifeLines Cohort Study is designed for research on complex interactions in the development of chronic disease and is well suited for novel research questions and innovative multidisciplinary studies. The three-generation design of the LifeLines Cohort Study goes a step further than that of traditional cohort studies, where individual participants and first-degree relatives are included. Inclusion of three or more generations in one study opens up unique possibilities for the separation of non-genetic and genetic familial transmission as well as the assessment of (epi)genetic influences and imprinting.[1] Furthermore, the design is important for the investigation of social characteristics like socioeconomic mobility and partner preferences. The open protocol and adjustable content of follow-up measures ensure flexibility in data collection in order to adapt to new social, environmental and scientific developments. Apart from traditional questionnaires we expect to implement new technologies for biobanking research like apps, wearable devices, and linkage to other sources like consumer registries.

The response to the invitation to participate was comparable to that in other large-scale population cohort studies.[34–36] A study on the representativeness of LifeLines for the source population and the population of The Netherlands is currently in progress and will be submitted by the end of 2014. Preliminary results indicate that the LifeLines population is representative of the general population of the north of The Netherlands when the difference in age-distribution has been taken into account, owing to the invitation strategy of three generations.

The large biomaterial collection challenges the existing methodology and capacity for sample handling and storage. Conditional transport of blood and urine samples to the highly automated LifeLines laboratory guarantees complete traceability. Samples are analysed and stored on the day of collection. The storage facility of the LifeLines Cohort Study is a completely automated sample storage with a capacity of 8+ million aliquots at −80°C, which will do fully automated overnight sample retrieval

The LifeLines Cohort Study has invested in a dedicated high-quality IT system for data collection, storage and data release, including a participant invitation program, optical reading of paper questionnaires, online questionnaires and linkage with laboratory results and external data. After the data have gone through a data verification procedure, the research data are pseudonymized and separated from the personal data. The LifeLines research data are stored in a sophisticated three-layer database architecture in another physical location for maximal privacy protection, with layers for: (i) staging data from different sources, such as questionnaires, experiments and registries; (ii) harmonizing all data into one model; and (iii) producing datasets with study-specific pseudonyms per research project. Each layer adds pseudonymization of identifiers to prevent unwanted identification and cross-study integration. All 'big'genetic data are available on a High Perfomance Computing environment. Researchers can request data via the MOLGENIS[37] catalogue which is available as open source from [www.molgenis.org] for use by other biobanks. The LifeLines IT solutions have been developed in international collaboration with the Target project, Center for Information Technology for high-performance computing and BBMRI-NL, CTMM TraIT, BioSHARE and BioMedBridges, for data management and integration.

To the entire process of data and biomaterial collection, storage and release, LifeLines applies high-quality standards. Privacy, security and traceability are ensured in all aspects, to comply with current and future requirements from both the scientific and the public community. As one of the visible results, LifeLines received the ISO9001 quality certificate.

Since LifeLines data and biomaterial can be retrieved by researchers globally and LifeLines employees are not involved in the design or analyses of individual studies, consistency of study results and proper reference of the LifeLines design remain a challenge. The LifeLines Cohort Study establishes and facilitates sharing of data constructs and derivates for consistency of study results. In addition, all publications are reviewed by the LifeLines Research Office for proper description of LifeLines and consistency with previous publications.

As a centre of expertise in biobanking, the LifeLines Cohort Study takes part in a number of strategic alliances, national and international biobank networks and research collaborations. Through these networks the LifeLines Cohort Study established itself as one of the world leaders in the science of biobanking and pooling of data from multiple biobank studies, as well as the advancement of public-private partnerships. The LifeLines Cohort Study forms part of the

Dutch national infrastructure on biobanks (BBMRI-NL) together with the Parelsnoer Institute (hospital biobanks) and several others. Internationally, LifeLines participates in several GWAS consortia, as well as with international initiatives like BioSHaRE, BBMRI-ERIC, BBMRI-LPC and Maelstrom Research (part of P3G).

Can I get hold of the data? Where can I find out more?

The LifeLines Cohort Study is a resource for the international scientific community for research on healthy ageing. The online LifeLines data catalogue at [www.lifelines.net] provides detailed information on all collected variables. It allows researchers worldwide to browse through the data and select the data that they need for answering their research question. Researchers from public institutes can apply for the data and biomaterial through a proposal that is submitted to the LifeLines Research Office [LLscience@umcg.nl]. All proposals are reviewed on scientific quality and methodology by the LifeLines scientific board, which includes scientists from different academic institutes in The Netherlands. Data and biomaterials are provided on a fee-for-service basis and may be used for scientific research only. Data can be accessed via a remote desktop (LifeLines workspace) running on a high-performance computer cluster, which ensures data quality and security. Datasets of individual researchers are archived for at least 10 years and may be accessed by reviewers and evaluation committees. The LifeLines website provides the details of the application process, the data collection and an overview of publications with LifeLines data.

Funding

LifeLines has been funded by a number of public sources, notably the Dutch Government, The Netherlands Organization of Scientific Research NWO, the Northern Netherlands Collaboration of Provinces (SNN), the European fund for regional development, Dutch Ministry of Economic Affairs, Pieken in de Delta, Provinces of Groningen and Drenthe, the Target project, BBMRI-NL, the University of Groningen, and the University Medical Center Groningen, The Netherlands.

Acknowledgements

The authors wish to acknowledge all participants of the LifeLines Cohort Study, the contributing researchers and everybody involved in the set-up and implementation of the study.

Conflict of interest: None declared.

References

1. Stolk RP, Rosmalen JG, Postma DS *et al*. Universal risk factors for multifactorial diseases: LifeLines: A three-generation population-based study. *Eur J Epidemiol* 2008;**23**:67–74.
2. FUTURAGE. The Future of Ageing Research In Europe: A Road Map. 2011. http://futurage.group.shef.ac.uk/road-map.html (1 September 2014, date last accessed)
3. Hewitt RE. Biobanking: the foundation of personalized medicine. *Curr Opin Oncol* 2010;**23**:112–19
4. Ng JWY, Barrett LM, Wong A *et al*. The role of longitudinal cohort studies in epigenetic epidemiology: challenges and opportunities. *Genome Biol* 2012;**13**:246.
5. Manolio TA, Bailey-Wilson JE, Collins FS. Genes, environment and the value of prospective cohort studies. *Nat Rev Genet* 2006;**7**:812–20.
6. Chaudhri S, Fan S, Davenport A. Pitfalls in the measurement of skin autofluorescence to determine tissue advanced glycosylation content in haemodialysis patients. *Nephrology (Carlton)* 2013;**18**:671–75.
7. Sheehan DV, Lecrubier Y, Sheehan KH *et al*. The mini-international neuropsychiatric interview (M.I.N.I.): The development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *J Clin Psychiatry* 1998;**59**(**Suppl 20**): 22–33; quiz 34–57.
8. Wind AW, Schellevis FG, Van Staveren G, Scholten RP, Jonker C, Van Eijk JT. Limitations of the mini-mental state examination in diagnosing dementia in general practice. *Int J Geriatr Psychiatry* 1997;**12**:101–08.
9. Ruff RM, Evans RW, Light RH. Automatic detection vs controlled search: A paper-and-pencil approach. *Percept Mot Skills* 1986;**62**:407–16.
10. Burney PG, Luczynska C, Chinn S, Jarvis D. The European Community respiratory health survey. *Eur Respir J* 1994;**7**:954–60.
11. Wendel-Vos GC, Schuit AJ, Saris WH, Kromhout D. Reproducibility and relative validity of the short questionnaire to assess health-enhancing physical activity. *J Clin Epidemiol* 2003;**56**:1163–69.
12. Siebelink E, Geelen A, de Vries JH. Self-reported energy intake by FFQ compared with actual energy intake to maintain body weight in 516 adults. *Br J Nutr* 2011;**106**:274–81.
13. Hays RD, Morales LS. The RAND-36 measure of health-related quality of life. *Ann Med* 2001;**33**:350–57.
14. Watson D, Clark LA, Tellegen A. Development and validation of brief measures of positive and negative affect: The PANAS scales. *J Pers Soc Psychol* 1988;**54**:1063–70.
15. Derogatis LR, Lipman RS, Covi L. SCL-90: An outpatient psychiatric rating scale – preliminary report. *Psychopharmacol Bull* 1973;**9**:13–28.
16. Costa PT, MacCrae RR. *Revised NEO Personality Inventory (NEO PI-R) and NEO Five-Factor Inventory (NEO FFI): Professional Manual*. Odessa, FL: Psychological Assessment Resources, 1992.
17. Reiss S, Peterson RA, Gursky DM. Anxiety sensitivity, injury sensitivity, and individual differences in fearfulness. *Behav Res Ther* 1988;**26**:341–45.
18. Rosmalen JG, Bos EH, de Jonge P. Validation of the long-term difficulties inventory (LDI) and the list of threatening experiences (LTE) as measures of stress in epidemiological population-based cohort studies. *Psychol Med* 2012;**42**:2599–608.
19. Nieboer A, Lindenberg S, Boomsma A, Van Bruggen AC. Dimensions of well-being and their measurement: The SPF-IL scale. *Soc Indicators Res* 2005;**73**:313–53.

20. Peters LL, Boter H, Buskens E, Slaets JP. Measurement properties of the Groningen frailty indicator in home-dwelling and institutionalized elderly people. *J Am Med Dir Assoc* 2012;**13**:546–51.
21. Peters LL, Boter H, Slaets JP, Buskens E. Development and measurement properties of the self assessment version of the INTERMED for the elderly to assess case complexity. *J Psychosom Res* 2013;**74**:518–22.
22. Genome of The Netherlands Consortium. Whole-genome sequence variation, population structure and demographic history of the Dutch population. *Nat Genet* 2014;**46**:818–25.
23. 1000 Genomes Project Consortium, Abecasis GR, Altshuler D *et al*. A map of human genome variation from population-scale sequencing. *Nature* 2010;**467**:1061–73.
24. Purcell S, Neale B, Todd-Brown K *et al*. PLINK: A tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007;**81**:559–75.
25. Delaneau O, Zagury JF, Marchini J. Improved whole-chromosome phasing for disease and population genetic studies. *Nat Methods* 2013;**10**:5–6.
26. Howie B, Fuchsberger C, Stephens M, Marchini J, Abecasis GR. Fast and accurate genotype imputation in genome-wide association studies through pre-phasing. *Nat Genet* 2012;**44**:955–59.
27. Deelen P, Menelaou A, Leeuwen EM van *et al*. Improved imputation quality of low-frequency and rare variants in European samples using the "Genome of The Netherlands". *Eur J Hum Genet* 2014;**22**:221–27.
28. Byelas H, Dijkstra M, Neerincx P *et al*. Scaling bio-analyses from computational clusters to grids. *Proceedings of Fifth International Workshop on Science Gateways (IWSG), 2013, 3–5 June. Zurich, Switzerland,* 2013.
29. de Jong K, Boezen HM, Kromhout H *et al*. Occupational exposure to vapors, gases, dusts, and fumes is associated with small airways obstruction. *Am J Respir Crit Care Med* 2014;**189**:487–90.
30. de Jong K, Boezen HM, Kromhout H *et al*. Pesticides and other occupational exposures are associated with airway obstruction: the LifeLines cohort study. *Occup Environ Med* 2014;**71**:88–96.
31. de Jong K, Boezen HM, Hacken NH *et al*. GST-omega genes interact with environmental tobacco smoke on adult level of lung function. *Respir Res* 2013;**14**:83.
32. Jansen H, Stolk RP, Nolte IM *et al*. Determinants of HbA1c in nondiabetic Dutch adults: genetic loci and clinical and lifestyle parameters, and their interactions in the Lifelines Cohort Study. *J Intern Med* 2013;**273**:283–93.
33. Slagter SN, van Vliet-Ostaptchouk JV, Vonk JM *et al*. Combined effects of smoking and alcohol on metabolic syndrome: the LifeLines cohort study. *PLoS One* 2014;**9**:e96406.
34. Allen N, Sudlow C, Downey P *et al*. UK Biobank: Current status and what it means for epidemiology. *Health Policy Technol* 2012;**1**:12326.
35. van den Brandt PA, Goldbohm RA, van 't Veer P, Volovics A, Hermus RJ, Sturmans F. A large-scale prospective cohort study on diet and cancer in The Netherlands. *J Clin Epidemiol* 1990;**43**:285–95.
36. Chen Z, Lee L, Chen J *et al*. Cohort profile: The Kadoorie Study of Chronic Disease in China (KSCDC). *Int J Epidemiol* 2005;**34**:1243–49.
37. Swertz MA, Dijkstra M, Adamusiak T *et al*. The MOLGENIS toolkit: Rapid prototyping of biosoftware at the push of a button. *BMC Bioinformatics* 2010;**11**(**Suppl 12**): S12-2105-11-S12-S12.