BMJ Open

Study protocol for an epidemiological study 'Multimorbidity – identifying the most burdensome patterns, risk factors and potentials to reduce future burden (MOLTO)' based on the Finnish health examination surveys and the ongoing register-based follow-up

Tuija Jääskeläinen ,[1] Päivikki Koponen,[1] Annamari Lundqvist ,[1] Jaana Suvisaari,[1] Jutta Järvelin,[2] Seppo Koskinen[1]

To cite: Jääskeläinen T, Koponen P, Lundqvist A, *et al*. Study protocol for an epidemiological study 'Multimorbidity – identifying the most burdensome patterns, risk factors and potentials to reduce future burden (MOLTO)' based on the Finnish health examination surveys and the ongoing register-based follow-up. *BMJ Open* 2022;**12**:e056073. doi:10.1136/bmjopen-2021-056073

► Prepublication history for this paper is available online. To view these files, please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2021-056073).

Received 04 August 2021
Accepted 22 February 2022

For numbered affiliations see end of article.

Correspondence to
Dr Tuija Jääskeläinen; tuija.jaaskelainen@thl.fi

ABSTRACT

Introduction Multimorbidity, defined as the co-occurrence of two or more long-term medical conditions, is an increasing public health concern worldwide causing enormous burden to individuals, healthcare systems and societies. The most effective way of decreasing the burden caused by multimorbidity is to find tools for its successful prevention but gaps in research evidence limit capacities to develop prevention strategies. The aim of the MOLTO study (Multimorbidity - identifying the most burdensome patterns, risk factors and potentials to reduce future burden) is to provide novel evidence required for cost-effective prevention of multimorbidity by defining the multimorbidity patterns causing the greatest burden at the population level, by examining their risk and protective factors and by estimating the potentials to reduce the future burden.

Methods and analysis The MOLTO study is based on the data from the Finnish population-based cross-sectional (FINRISK 2002–2012, FinHealth 2017 the Migrant Health and Well-being Study 2010–2012) and longitudinal (Health 2000/2011) health examination surveys with individual-level link to administrative health registers, allowing register-based follow-up for the study participants. Both cross-sectional and longitudinal study designs will be used. Multimorbidity patterns will be defined using latent class analysis. The burden caused by multimorbidity as well as risk and protective factors for multimorbidity will be analysed by survival analysis methods such as Cox proportional hazards and Poisson regression models.

Ethics and dissemination The survey data have been collected following the legislation at the time of the survey. The ethics committee of the Hospital District of Helsinki and Uusimaa has approved the data collection and register linkages for each survey. The results will be published as peer-reviewed scientific publications.

Strengths and limitations of this study

- ⇒ The study is based on the data from the large, population-based health examination surveys with standardised, reliable methodology and possibility to follow-up the participants via health registers.
- ⇒ A possibility to analyse a large variety of measures of burden caused by multimorbidity as well as potential risk and protective factors for multimorbidity.
- ⇒ The definition of multimorbidity will be based on the multimorbidity patterns, which cause the greatest burden at the population level to produce the results with high scientific and public health impact.
- ⇒ Despite the large population-based data sets, low prevalence of certain chronic diseases may limit the identification of the multimorbidity patterns, especially when studying the differences between the population subgroups.

INTRODUCTION

Multimorbidity, usually defined as the co-occurrence of two or more long-term medical conditions, is an increasing phenomenon wordwide.[1–3] The prevalence of multimorbidity increases strongly with the increase in age being nearly 100% in older persons.[4–6] On the other hand, multimorbidity is comparatively common also among young and middle-aged adults especially in socioeconomically deprived populations.[7] Multimorbidity is a major challenge for healthcare systems causing enormous costs to societies.[1 8] From an individual perspective, multimorbidity is associated with disability and functional decline as well as reduced quality of life and life expectancy.[8] Further, multimorbidity is

associated with higher overall vulnerability to diseases and decreased resistance to acute health threats.[2]

From a methodological perspective, major challenge concerning multimorbidity research is that no clear consensus exists for measuring multimorbidity: number and types of chronic conditions included as well as research settings, designs, data sources and methodology vary widely between the different studies.[9 10] This impairs the comparability of the estimates on multimorbidity prevalence and burden across the countries and studies. A systematic review of over 500 studies on multimorbidity published in 2021 highlighted that reporting should be improved by stating clearly which conditions were included in multimorbidity measurement, their clinical code sets and why these conditions were chosen.[9]

Knowledge of the most common clusters of diseases (ie, multimorbidity patterns) and which of them are the most burdensome from the individual and societal perspectives both in general populations and in certain population groups is still limited.[1 4] This restricts possibilities to identify those individuals with a single disease who are at the greatest risk to develop another one as well as development and evaluation of intervention strategies designed specifically to prevent the relevant chronic conditions simultaneously.[1] While some multimorbidity patterns are well known, such as cardiorespiratory or metabolic patterns, others need further identification, especially those including both, somatic and psychiatric diseases.[11–14]

Well-established sociodemographic risk factors for multimorbidity include older age and lower socioeconomic status.[1 6–8 14] The results concerning sex are somewhat contradictory, although, according to systematic reviews[6 8] women tend to have higher risk for multimorbidity compared with men. Regarding lifestyle-related factors, the evidence is still contradictory, possibly due to methodological differences between the studies, and mainly based on cross-sectional data.[1] Smoking, physical inactivity and obesity have been identified as potential risk factors for multimorbidity in several studies,[1 4] but the associations between alcohol consumption or nutrition and multimorbidity have been studied less.[1] Some studies have focused on individual dietary factors, like fruit and vegetable intake, but the results are mixed.[15–17] The association between quality of diet and multimorbidity is still poorly known. Furthermore, despite the fact that lifestyle-related risk factors are typically clustering,[18] the research focusing on the combined effects of different lifestyle-related factors is still limited and based on cross-sectional design,[19] specific disease clusters[16 20] or older populations.[17] Furthermore, the interactions between the lifestyle-related risk factors and genetic risk of multimorbidity are poorly known.

Systematic reviews have summarised the main consequences of multimorbidity, including reduced self-rated health and quality of life, functional decline and disability, increased risk of mortality and increased use of health services and costs.[1 4 6 21] Different disease combinations may affect differentially these outcomes. There is also evidence that the higher the number, and the more severe the diseases are, the greater the negative impact.[21] Some studies have also indicated that certain clusters of diseases may have greater negative impact on these outcomes than could be predicted based on the sum of the individual conditions.[22]

The most effective way of decreasing the burden caused by multimorbidity is to find tools for its successful prevention.[3] Despite the rapidly growing number of research publications on multimorbidity,[23] there are still gaps in our knowledge causing challenges to identify the risk groups for multimorbidity, to develop the cost-effective prevention strategies as well as to develop healthcare services to the special needs of multimorbid patients. In addition to the lack of agreed definition of multimorbidity and other heterogeneity in measurements described above, several other research gaps have been identified.[1 3 10 23] The present study, 'Multimorbidity identifying the most burdensome patterns, risk factors and potentials to reduce future burden' (the MOLTO study) aims to fill the following research gaps:

- The majority of the previous studies have been conducted in an observational, cross-sectional design. The MOLTO study will be conducted in population-based, longitudinal design to clarify the temporal aspects of the associations and to provide nationally representative results.
- Most previous studies have focused on older populations. In addition to older populations, the MOLTO study will provide more information on the causes and consequences of multimorbidity among younger age groups.
- In previous research, the definition of multimorbidity is highly heterogeneous, and in many studies, based on counting of diseases. The MOLTO study will deepen the analysis by revealing the multimorbidity patterns, which cause the greatest burden in the general population and its subgroups. The burden will be defined in several perspectives including mortality, functional decline and work disability, quality of life as well as healthcare utilisation and costs. The coexistence of both physical and mental diseases will be taken into account. Identifying multimorbidity patterns, leading to the most burdensome consequences, is essential to improve health services and to develop care guidelines.
- The comprehensive picture of the risk and protective factors for multimorbidity is still unclear. The MOLTO study aims to provide more evidence by identifying the genetic, sociodemographic, lifestyle-related and biological pathways to multimorbidity. Furthermore, the combined effects and the interactions between the risk factors will be studied in detail. The results will help to develop new tools for early identification of persons at risk of multimorbidity as well as determination of the factors and their combinations, which should be prioritised to improve cost-effective

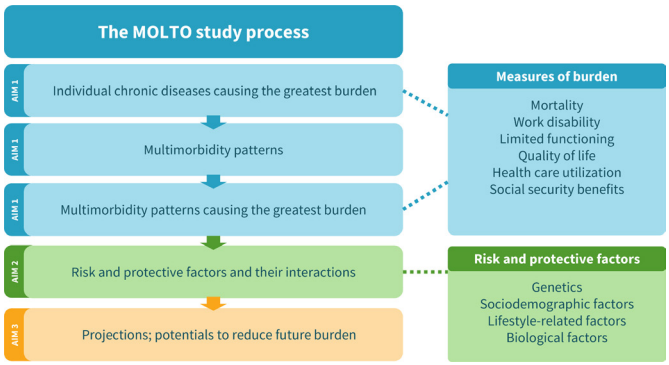

Figure 1 Study process for the MOLTO study (Multimorbidity - Identifying the most burdensome patterns, risk factors and potentials to reduce future burden).

prevention of multimorbidity. Furthermore, the results can be used to identify vulnerable population groups to target public health interventions and health promotion actions to population subgroups, where most multimorbidity cases could potentially be prevented.

- The MOLTO study will produce reliable projections of the future multimorbidity burden and the related needs for social and healthcare services to enable the development of evidence-based public health programmes to reduce the future burden of multimorbidity.

The overall aim of the study is to provide up-to-date, reliable and novel evidence required for cost-effective prevention of multimorbidity. We will focus on the multimorbidity patterns, which cause the greatest burden in the adult population to produce results with high scientific and public health impact. The study process is summarised in figure 1. Specific aims and research questions are as follows:

Aim 1: to identify the multimorbidity patterns that cause the greatest burden among Finnish adults

Q1: Which chronic diseases cause the greatest burden in Finland?

Q2: What kind of multimorbidity patterns will be identified based on these diseases? Which of these multimorbidity patterns cause the greatest burden?

Q3: Is the burden caused by these multimorbidity patterns greater than could be expected on the basis of the burden caused by the individual diseases?

Q4: Does the burden caused by these multimorbidity patterns differ between the population groups?

Aim 2: to examine the main sociodemographic, lifestyle and biological risk and protective factors of the most burdensome multimorbidity patterns

Q5: Which sociodemographic, lifestyle and biological factors are associated with the risk of multimorbidity?

Q6: How do these risk factors cluster and interact?

Q8: Does the genetic risk of multimorbidity patterns defined by polygenic risk scores modify the association between other risk factors and multimorbidity?

Aim 3: to examine how to best reduce the future burden of multimorbidity

Q9: How will multimorbidity patterns evolve by 2050?

Q10: How influencing the risk factors of multimorbidity could change these projections and reduce future burden caused by multimorbidity?

METHODS AND ANALYSIS

Health examination surveys

The MOLTO study is conducted between 1 January 2020 and 31 December 2026. It is based on the pooled data from the following Finnish population-based health examination surveys (HESs) coordinated by Finnish Institute for Health and Welfare (THL). The sampling designs of the HESs were developed with the main aim that the individuals selected are representative of the target population. The samples have been drawn from nationwide population registers, covering all residents in Finland.[24]

1. The longitudinal Health 2000/2011 Survey.[25 26] A nationally representative two-stage stratified cluster sample of the Finnish adult population examined in 2000–2001 and re-examined in 2011–2012. The target population consisted of individuals aged 18 or over and living in mainland Finland. The sample frame was regionally stratified according to five university hospital regions (strata): Helsinki, Turku, Tampere, Kuopio and Oulu. In the first stage of sampling, 80 health centre districts, 16 from each university hospital regions, were sampled as a cluster. The 15 towns with the largest populations were selected with probability 1 and the sample size for each health centre district was proportional to its population. Other 65 health centre districts were selected using systematic sampling with probabilities proportional to size (PPS-SYS design). Systematic random sampling was used to draw the sample from each health centre district using the Finnish Population Information System. This two-stage stratified cluster sample represents the adult population living in mainland Finland.
2. The cross-sectional FINRISK studies conducted in 2002, 2007 and 2012.[27] The target population of the FINRISK studies consisted of individuals aged 25–74 years of selected areas in Finland. Each year, an independent random sample of Finnish adults from selected areas (North Karelia and Northern Savo from Eastern Finland, cities of Turku and Loimaa from Southwestern Finland, Helsinki and Vantaa from Southern Finland, Northern Ostrobothnia and Kainuu in Northern Finland) was drawn from the Finnish Population Information System. These different geographical areas in Finland cover nearly half of the population.
3. The cross-sectional FinHealth 2017 Study.[28] A nationally representative two-stage stratified cluster sample of the Finnish adult population examined in 2017. The target population consisted of individuals aged 18 or over and living in mainland Finland. The sampling design was based on the Health 2000 sampling design. In

Table 1 Characteristics of the health examination survey data

	H2000	H2011	FR2002	FR2007	FR2012	FH2017	Maamu
Study years	2000–2001	2011–2012	2002	2007	2012	2017	2010–2012
Age range (years)	30+	30+	25–74	25–74	25–74	30+	30–64
Sample (n)	8028	7964	13 437	11 953	9905	9288	3000*
Participation rate (%), any stage	93	73	71	67	65	71	70/51/63†
Participation rate (%), health examination	84	59	66	53	59	60	47/38/52‡

*Russian origin n=1000, Somali origin n=1000, Kurdish origin n=1000.
†70% for Russian origin subjects, 51% for Somali origin subjects, 63% for Kurdish origin subjects.
‡47% for Russian origin subjects, 38% for Somali origin subjects, 52% for Kurdish origin subjects.
FH, FinHealth; FR, FINRISK; H2000, Health 2000; H2011, Health 2011; Maamu, Migrant Health and Wellbeing Study.

FinHealth 2017 Study, 50 health centre districts out of the 80 health centre districts of the Health 2000 were selected.

4. The cross-sectional Migrant Health and Wellbeing Study (Maamu).[29] A large-scale population survey on the health and well-being of adults of Russian, Somali and Kurdish origin. The three groups were selected to represent different kinds of large foreign-born groups in Finland. The sample was randomly selected from the Finnish Population Information System, including individuals from selected large Finnish cities (Helsinki, Espoo, Vantaa, Turku, Tampere and Vaasa). The cities were selected from the metropolitan area and other parts of the country, with a higher proportion of foreign-born persons than in most of the other areas. Maamu data will be used in the substudies where migrant background will be taken into account.

Sample sizes, participation rates and age ranges of the HESs are presented in table 1. All HESs included a comprehensive health examination with blood sampling, blood pressure measurements and anthropometric measurements (tables 2 and 3). All other HESs except FINRISK surveys included also functional capacity tests. Furthermore, all HESs included self-administered questionnaires and/or interviews to gather information on sociodemographic factors, lifestyle, quality of life, healthcare utilisation, functional capacity and work ability as well as diagnosed chronic diseases and medication. Measurements in these HESs were conducted using standardised[30] and mainly comparable methods allowing us to combine different data sets and use pooled data.

National registers

Register-based data will be used both as aggregated population level data to define the chronic diseases, which cause the greatest burden among Finnish adults and individually linked to HES data using personal identity code. The following Finnish administrative register-based data will be used:

1. Causes of death register: dates, primary and contributory causes of deaths (Statistics Finland).[31]
2. The Register of Completed Education and Degrees[32] and socioeconomic status,[33] occupation[34] (Statistics Finland).
3. Care Register for Health Care: dates and diagnoses of hospitalisations and outpatient visits within public healthcare, including primary care since 2011 (THL).[35]
4. Finnish Cancer Registry: date and diagnoses of cancers.[36]
5. Registers of the Social Insurance Institution (Kela): entitlement to specially reimbursed medications due to specific chronic conditions, purchase of medicines, sickness absence, disability allowance, rehabilitation.[37]
6. Registers of then Finnish Centre for Pensions: earnings-related pensions.[38]
7. Population Information System maintained by Digital and population data services agency of Finland: Spatial information.[39]

Measurements

Chronic diseases

Prevalent chronic diseases at baseline and incident diseases during the follow-up period will be defined based on the national register data (Causes of Death,[31] Care Register for Health Care,[35] Cancer Registry[36] and medications[37]), using health examination measurements and self-reported information on diagnosed chronic diseases as a complement when feasible. The classification of the chronic diseases will be based on the 10th revision of the International Classification of the Diseases (ICD-10) complemented with the International Classification of Primary Health Care coding and the Anatomical Therapeutic Chemical Classification System codes for medication. Regarding mental health, Composite International Diagnostic Interview,[40] the Beck Depression Inventory[41] and the Hopkins Symptom Checklist-25[42] were used (see table 2).

Multimorbidity patterns

In the MOLTO study, multimorbidity will be defined as the patterns of two or more chronic conditions as defined by WHO.[2] We will focus on the chronic diseases, which cause the greatest burden in Finland. We have estimated the burden of diseases based on the latest aggregated register-based data. First, the proportion of register-based outcomes caused by each chronic diseases among Finnish adult population was defined (eg, what proportion of deaths is caused by malignant neoplasms). The chronic

Table 2 Health examination survey data on mental health, quality of life, physical functioning, work ability and healthcare utilisation

	H2000	H2011	FR2002	FR2007	FR2012	FH2017	Maamu
Mental health							
M-CIDI* (diagnosis for mental disorders)	+	+					
BDI† (depressive symptoms)	+	+			+	+	
HSCL-25‡ (depressive and anxiety symptoms)		+					+
Questions about being low-spirited or depressed during the last 12 months			+	+	+	+	
Quality of life (QOL)							
15D§	+	+					
EQ-5D ¶	+	+					
EuroHIS-8 **		+				+	+††
Question about perceived QOL during past month ‡‡	+			+	+		
Functional capacity performance tests							
Grip strength (dominating hand)	+	+				+	+
Chair stand 1, 5, 10 times §§	+	+				+	+
Walking test 4 m/6,1 m ¶¶	+	+					
Joint function tests for age group 55+ ***	+	+				+	
Standing balance	+	+					+
Functional capacity questions, for example							
Ability to walk 500 m without resting	+	+	+	+	+	+	+
Ability to walk 2 km without resting	+	+					
Ability to run a short distance (100 m)	+	+	+	+	+	+	+
Ability to run a long distance (500 m)	+	+	+	+	+		
Ability to climb stairs for one flight	+	+	+	+	+	+	
Ability to climb stairs for several flights	+	+					+
Walking difficulties due to knee pain	+	+	+	+	+	+	
Walking difficulties due to hip pain	+	+	+			+	
Work ability questions, for example,							
Work ability estimate†††	+	+		+	+	+	+
Work ability score ‡‡‡	+	+		+	+		
Days being absent from work or being unable to do daily chores (in last 12 months)	+	+	+	+	+	+	
Healthcare utilisation questions, for example,							
Number of visits to a doctor during the last 12 months	+	+	+	+	+	+	+
Number of visits to a nurse during the last 12 months	+	+	+	+	+	+	
Participation in health check (eg, in occupational healthcare)	+	+	+	+	+	+	
Using of health services due to mental health problems in the past 12 months	+	+				+	+
Physiotherapy during the last 12 months	+	+				+	+

*Computerised version of the Composite International Diagnostic Interview allowing the estimation of diagnoses for mental disorders during the past 12 months, according to the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV).[40]
†The Beck Depression Inventory.[41]
‡The Hopkins Symptom Checklist-25.[42]
§Finnish health-related quality of life instrument.[52 53]
¶Generic measure of health status developed by the EuroQol Group.[54]
**EuroHIS-QoL 8-item index.[55]
††Maamu included also some questions from WHOQOL-BREF -measurement.
‡‡From 0 (worst possible) to 10 (best possible).
§§H2000 only 1 and 5 times and only for age group 55+, Maamu only 10 times.
¶¶H2000 only 6,1 m and only for age group 55+.
***H2011 and FH2017 only shoulder and squatting.
†††Work ability estimate.[56]
‡‡‡Question is part of the Work Ability Index.[57]
FH, FinHealth; FR, FINRISK; H2000, Health 2000; H2011, Health 2011; Maamu, Migrant Health and Wellbeing Study.

Table 3 Health examination survey data on lifestyle and biological risk factors

	H2000	H2011	FR2002	FR2007	FR2012	FH2017	Maamu
Smoking status							
Cigarettes, cigars, pipefuls	+	+	+	+	+	+	+
Smokeless tobacco (snus)			+	+	+	+	
Alcohol consumption							
AUDIT/AUDIT-C*		+		+		+	+
Questions concerning alcohol consumptions (frequency, amount)	+	+	+	+	+	+	
Dietary habits							
Food frequency questionnaire†	+	+		+	+	+	
Questions concerning dietary habits, for example							
Use of vegetables (frequency)	+	+	+	+		+	+
Use of fruits and berries (frequency)		+	+	+		+	+
Fat spread on bread	+	+	+	+	+	+	+
Cooking fat	+	+	+	+	+	+	+
Glass of milk per day (frequency and/or fat per cent)	+		+	+		+	
Slices of bread per day (dark/mixed/white)	+	+	+	+			+
Physical activity/sedentary behaviour							
Leisure-time physical activity	+	+	+	+	+	+	+
Commuting physical activity	+	+	+	+	+	+	+
Time of sitting per day	+	+		+	+	+	
Sleep and sleeping questions							
Hours of sleep per day	+	+		+	+	+	+
Getting enough sleep (self-estimated)		+		+	+	+	
Anthropometric measures‡							
Weight and height (body mass index calculated)	+	+	+	+	+	+	+
Waist circumference	+	+	+	+	+	+	+
Hip circumference	+		+	+	+	+	
Body composition determined by bioimpedance analysis	+	+		+	+	+	
Blood pressure measurement‡	+	+	+	+	+	+	+
Laboratory analyses (serum, plasma or blood)							
Glucose, mmol/l (fasting)¶	+	+	+	+		+	+
HbA1c, % and/or mmol/mol	+	+	+		+	+	+
Total cholesterol, mmol/l	+	+	+	+	+	+	+
LDL-cholesterol§, mmol/l	+	+	+	+	+	+	+
HDL-cholesterol, mmol/l	+	+	+	+	+	+	+
Triglycerides, mmol/l	+	+	+	+	+	+	+
25-hydroxyvitamin D, nmol/l	+	+			+	+	+
CRP, mg/l	+	+	+	+	+	+	+

*The Alcohol Use Disorders Identification Test.[58 59]
†Food frequency questionnaire.[43 44]
‡EHES manual.[30]
§Direct measurement and/or calculated with the Friedewald's formula.
¶Participants were asked to fast four hours before health examination
CRP, C reactive protein; EHES, European Health Examination Survey; FH, FinHealth; FR, FINRISK; H2000, Health 2000; H2011, Health 2011; HbA1c, haemoglobin A1c; HDL, high density lipoprotein; LDL, low density lipoprotein; Maamu, Migrant Health and Wellbeing Study.

diseases (excluding, eg, infectious and parasitic diseases, conditions related to childbirth and pregnancy, injury as well as very rare chronic diseases) were categorised according to ICD-10 chapters and blocks. The following register-based outcomes were included: mortality, the use of health services, disability pensions, sickness and disability allowances and entitlements to specially reimbursed medication (for detailed information, see tables 4 and 5).

Second, this information was summarised by choosing those chronic diseases (ie, disease blocks), which were among the 10 most common causes according to at least one of the register-based outcomes mentioned above. This summarised information is presented in tables 4 and 5, showing that diseases of circulatory system as well as mental and behavioural disorders were particularly burdensome from several different perspectives. Concerning individual diseases (ie, disease blocks), 'diabetes mellitus' (E10-E14) and 'other degenerative diseases of the nervous system' (G30-G32; including G30 Alzheimer disease) appeared to be particularly burdensome. Diseases of the musculoskeletal system and connective tissue caused remarkable burden especially in terms of use of healthcare services, disability pensions and sickness allowance. Malignant neoplasms were significant causes of deaths and use of healthcare services.

Information on the chronic diseases causing the greatest burden in Finland, presented in tables 4 and 5, will be the basis when defining multimorbidity patterns in the MOLTO—study.

The burden caused by multimorbidity patterns

When defining the most burdensome multimorbidity patterns, the burden caused by multimorbidity will be estimated using the following outcome variables: mortality, work disability, limited functioning, quality of life, healthcare utilisation and costs. Information on mortality will be obtained from Statistics Finland.[31] Functional capacity, work ability and quality of life will be estimated based on HES data (table 2). When estimating work ability, also information concerning disability pensions will be used.[38] Healthcare utilisation will be determined based on register-based information from the Care Register for Health Care[35] as well as self-reported information from HESs (table 2).

Potential risk and protective factors of multimorbidity patterns

All HESs include comprehensive information concerning potential sociodemographic, lifestyle and biological risk and protective factors of multimorbidity (table 3). Information on sociodemographic factors will be complemented with register-based information on education, socioeconomic status and occupation[32–34] as well as information on degree of urbanisation.[39] Lifestyle variables (table 3) are determined by self-administered questionnaires or interviews. Data on diet have been collected by a validated self-administered food frequency questionnaire,[43 44] allowing us to assess both individual dietary indicators (eg, use of vegetables) and the overall diet defined by a dietary score. Biological factors, that is, anthropometrics and blood pressure, have been measured with standardised methods by trained nurses. Blood collection, sample processing and management were performed by trained laboratory personnel. The blood samples were processed and frozen immediately after sampling. The laboratory analyses were performed at the biochemistry laboratory at THL, which has taken part in External Quality Assessment Schemes organised by Labquality, Helsinki, Finland. Health 2000/2011, FINRISK and FinHealth cohorts have also been genotyped with genome-wide genotyping arrays (Illumina Inc and Thermo Fisher Scientific) and imputed with population-specific reference panel to contain over 12 million genomic variants.

Patient and public involvement

Patient or the public were not involved in the design or conduct of the study. The MOLTO study is based on the Finnish HESs carried out among general population.

Statistical analyses

Both cross-sectional and longitudinal study designs will be used. Longitudinal designs will be based on register-based follow-up (all HESs) and repeated survey measurements (Health 2000/2011).

Aim 1

Multimorbidity patterns will be determined using latent class analysis[45] based on chronic diseases, which cause the greatest burden in Finland (tables 4 and 5). The burden caused by multimorbidity will be analysed by survival analysis methods[46] such as Cox proportional hazards and Poisson regression models.

Aim 2

The risk and protective factors of multimorbidity will be analysed using Cox proportional hazard models. As the multimorbidity pattern can be determined at any time point using the event times obtained from the register data for each individual, we will also apply multistate models to analyse the transitions between the different multimorbidity categories in real time.[47] We will examine the combined effects of the risk factors on selected multimorbidity patterns by testing both additive (the relative excess risk due to interaction) and multiplicative interactions in the same Cox proportional hazards models.[48] The clustering of the risk factors and the accumulation of multiple risk factors in the same individual will be examined using cluster and latent class analyses. To evaluate the relative importance of the risk factors at the population level, we will assess population attributable factors for the risk factors of multimorbidity incidence.[49]

Aim 3

For projections novel techniques using data-driven Bayesian hierarchical and/or multistate models which have been developed in the Finnish research project 'Projections of the burden of disease and disability in

Table 4 The proportions (%) of register-based outcomes caused by chronic diseases according to ICD-10 chapters and blocks in men

	Causes of deaths*		Primary health care†		Specialised medical care†					
	Primary	Contributory	Outpatient doctor visits	Hospitalisation (days)	Outpatient visits	Hospitalisation (days)	Disability pensions‡	Sickness allowance§	Disability allowance¶	Prescription medicines**
Year	2018	2018	2018	2018	2018	2018	2019	2018	2018	2018
Age range (years)	30+	30+	30+	30+	30+	30+	30+	25–67	16+	30+
Total (n)	26 756	28 522	1 602 189	1 279 916	3 155 758	1 597 397	79 987	111 879	94 208	1 022 755
C00-C97 Malignant neoplasms (%)	**25.0**	**4.6**	**1.8**	**9.7**	**15.1**	**7.9**	**2.1**	**3.1**	**2.6**	**5.2**
C15-C26 Digestive organs (%)	8.7	0.7	0.3	2.7	2.5	2.4	0.5	NA	NA	NA
C30-C39 Respiratory and intrathoracic organs (%)	5.7	0.5	0.1	1.7	1.2	0.8	0.2	NA	NA	NA
C60-C63 f Male genital organs (%)	3.4	1.7	0.6	1.8	5.0	0.8	0.2	NA	NA	2.6
C81-C96 Lymphoid, haematopoietic and related tissue (%)	2.3	0.7	0.2	1.0	2.1	1.7	0.4	NA	NA	1.1
E00-E90 Endocrine, nutritional and metabolic diseases (%)	**1.5**	**10.2**	**8.5**	**2.0**	**2.3**	**1.3**	**1.8**	**1.0**	**2.9**	
E10-E14 Diabetes mellitus (%)	1.1	8.2	5.3	1.3	1.3	0.6	1.2	0.5	2.3	18.4
F00-F99 Mental and behavioural disorders (%)	**3.9**	**8.9**	**6.3**	**9.2**	**11.6**	**24.0**	**46.4**	**18.6**	**35.2**	
F00-F09 Organic, including symptomatic, mental disorders (%)	2.7	1.6	0.7	4.2	0.2	1.8	0.8	NA	9.3	NA
F10-F19 Mental and behavioural disorders due to psychoactive substance use (%)	0.8	5.7	1.2	3.8	2.0	1.6	1.8	0.3	1.4	NA
F20-F29 Schizophrenia, schizotypal and delusional disorders (%)	0.1	0.7	0.5	0.5	3.5	15.9	14.9	1.5	9.3	4.1††
F30-F39 Mood (affective) disorders (%)	<0.1	0.7	1.7	0.4	4.2	3.9	18.0	10.6	2.3	NA
F40-F48 Neurotic, stress-related and somatoform disorders (%)	0	0.1	1.2	0.2	1.1	0.6	2.8	5.6	NA	NA
G00-G99 Diseases of the nervous system (%)	**13.1**	**7.0**	**3.0**	**11.3**	**5.6**	**4.0**	**8.7**	**5.2**	**21.2**	
G20-G26 Extrapyramidal and movement disorders (%)	1.9	0.8	0.3	1.9	0.7	0.4	1.1	NA	2.7	1.0
G30-G32 Other degenerative diseases of the nervous system (%)	10.1	4.6	0.7	5.7	0.2	0.5	2.0	NA	10.0	NA
G40-G47 Episodic and paroxysmal disorders (%)	0.2	1.0	1.2	1.6	3.2	1.4	1.1	NA	1.0	2.8
H00-H59 Diseases of the eye and adnexa (%)	**0**	**<0.1**	**2.1**	**0.1**	**6.0**	**0.3**	**1.1**	**2.0**	**3.2**	
H40-H42 Glaucoma (%)	0	0	<0.1	<0.1	1.1	<0.1	0.1	NA	NA	3.7
I00-I99 Diseases of the circulatory system (%)	**35.9**	**42.7**	**14.5**	**20.4**	**8.0**	**19.0**	**7.4**	**6.9**	**15.8**	
I10-I15 Hypertensive diseases (%)	3.3	9.9	5.7	0.9	0.6	0.3	0.1	0.4	0.4	20.2
I20-I25 Ischaemic heart diseases (%)	20.2	9.9	2.2	2.0	1.2	3.9	1.4	1.9	2.0	11.3
I30-I52 Other forms of heart disease (%)	3.6	15.1	4.1	4.5	3.6	6.0	1.3	1.8	1.7	7.3‡‡

Continued

Table 4 Continued

	Causes of deaths*		Primary health care†		Specialised medical care†					
	Primary	Contributory	Outpatient doctor visits	Hospitalisation (days)	Outpatient visits	Hospitalisation (days)	Disability pensions‡	Sickness allowance§	Disability allowance¶	Prescription medicines**
I60-I69 Cerebrovascular diseases (%)	6.6	4.6	0.8	9.7	0.8	5.7	4.2	1.5	10.5	NA
I70-I79 Diseases of arteries, arterioles and capillaries (%)	1.7	2.3	0.5	2.5	1.0	2.4	0.4	0.5	0.8	NA
J00-J99 Diseases of the respiratory system (%)	**4.9**	**10.6**	**7.3**	**10.8**	**3.4**	**7.5**	**1.3**	**5.4**	**1.7**	
J40-J47 Chronic lower respiratory diseases (%)	3.2	5.5	2.1	1.7	0.9	0.9	1.1	0.6	1.4	9.6
K00-K99 Diseases of the digestive system (%)	**4.8**	**4.4**	**3.1**	**3.3**	**5.5**	**7.0**	**0.7**	**6.6**	**0.6**	
K50-K52 Noninfective enteris and colitis (%)	<0.1	0.2	0.3	0.1	1.4	0.3	0.3	NA	NA	2.2
K70-K77 Diseases of liver (%)	2.9	2.6	0.2	0.8	0.3	0.7	0.2	NA	NA	NA
M00-M99 Diseases of the musculoskeletal system and connective tissue (%)	**0.3**	**0.8**	**15.7**	**3.5**	**7.5**	**4.4**	**20.3**	**31.7**	**4.9**	
M00-M25 Arthropaties (%)	0.3	0.4	5.1	1.3	3.2	2.3	6.5	5.9	2.0	NA
M40-M54 Dorsopathies (%)	<0.1	0.2	5.0	1.2	2.4	1.3	11.1	13.2	2.1	NA
M60-M79 Soft tissue disorders (%)	<0.1	0.1	5.2	0.4	1.3	0.3	2.0	5.2	NA	NA
N00-N99 Diseases of the genitourinary system (%)	**0.4**	**4.3**	**3.6**	**3.7**	**4.7**	**3.1**	**0.4**	**1.1**	**0.9**	
N17-N19 Renal failure (%)	0.2	3.3	0.4	0.6	1.8	0.7	0.2	NA	NA	NA

The proportions of outcomes in order of size (from largest to smallest) in each column.
1.–5. (very dark red), 6.–10. (dark red), 11.–15. (red), 16.–20.(light red), 21.—(very light red), ICD-10 chapters (grey), NA Not available in the present study.
The proportions are calculated from the total number (n) based on the aggregated register-based data covering the Finnish adult population.
*Causes of deaths register, Statistics Finland.[31]
†Care register for Health Care, Finnish Institute for Health and Welfare; main diagnosis.[35]
‡Disability pension granted after 1995; primary diagnosis; Finnish Centre for Pensions (ETK).[38]
§Social Insurance Institution (Kela).[60]
¶Social Insurance Institution (Kela).[61]
**Social Insurance Institution (Kela); not including limited special reimbursement.[62]
††Including ICD-10 codes: F01, F03, F06.0-F06.3, F20-F25, F28, F29, F30.1, F30.2, F31, F32.3, F33.3, F84, G10, G20, G30.0, G30.1, G30.8, G30.9, G31.0, G35, G40.9.
‡‡Including ICD-10 codes: I11.0, I13, I47–I49 I50, I97.1, P29.0.

Table 5 The proportion (%) of register-based outcomes caused by chronic diseases according to ICD-10 chapters and blocks in women

	Causes of deaths*		Primary health care†		Specialised medical care†					
	Primary	Contributory	Outpatient doctor visits	Hospitalisation (days)	Outpatient visits	Hospitalisation (days)	Disability pensions‡	Sickness allowance§	Disability allowance¶	Prescription medicines**
Year	2018	2018	2018	2018	2018	2018	2019	2018	2018	2018
Age group	30+	30+	30+	30+	30+	30+	30+	25–67	16+	30+
Total (n)	27 127	27 524	2 464 262	1 739 410	3 979 539	1 571 446	82 142	164 372	137 000	1 120 386
C00-C97 Malignant neoplasms (%)	**21.5**	**2.8**	**1.3**	**7.1**	**13.4**	**7.6**	**3.0**	**4.0**	**2.0**	**6.1**
C15-C26 Digestive organs (%)	7.1	0.5	0.1	2.2	1.6	1.9	0.4	NA	NA	NA
C30-C39 Respiratory and intrathoracic organs (%)	3.0	0.3	<0.1	0.8	0.6	0.5	0.2	NA	NA	NA
C50-C50 Breast (%)	3.1	0.7	0.5	1.0	5.7	1.0	1.1	2.2	NA	3.7
C51-C58 Female genital organs (%)	2.7	0.2	0.1	0.9	1.4	0.9	0.3	NA	NA	0.4
C81-C96 Lymphoid, haematopoietic and related tissue (%)	1.9	0.4	0.1	0.7	1.5	1.5	0.4	NA	NA	0.9
E00-E90 Endocrine, nutritional and metabolic diseases (%)	**1.2**	**9.3**	**6.7**	**2.0**	**2.6**	**1.6**	**1.3**	**1.4**	**2.6**	
E00-E07 Disorders of thyroid gland (%)	<0.1	0.5	1.4	0.2	0.8	0.2	<0.1	0.6	NA	6.4
E10-E14 Diabetes mellitus (%)	0.8	7.3	3.0	0.9	0.8	0.4	1.0	0.2	1.9	14.0
F00-F99 Mental and behavioural disorders (%)	**5.9**	**5.2**	**6.1**	**8.3**	**13.3**	**20.0**	**48.6**	**27.0**	**28.4**	
F00-F09 Organic, including symptomatic, mental disorders (%)	5.3	2.4	0.6	5.1	0.2	1.1	0.5	NA	11.5	NA
F20-F29 Schizophrenia, schizotypal and delusional disorders (%)	0.2	0.8	0.3	0.9	2.9	11.2	10.2	0.8	6.1	4.4††
F30-F39 Mood (affective) disorders (%)	<0.1	0.6	2.1	0.7	6.6	5.7	28.0	15.1	3.4	NA
F40-F48 Neurotic, stress-related and somatoform disorders (%)	0	0.1	1.7	0.4	2.0	1.0	3.6	10.2	NA	NA
G00-G99 Diseases of the nervous system (%)	**23.4**	**9.3**	**3.0**	**11.7**	**4.6**	**4.0**	**8.9**	**5.4**	**25.2**	
G20-G26 Extrapyramidal and movement disorders (%)	1.6	0.6	0.3	1.5	0.7	0.4	1.0	NA	1.8	0.1
G30-G32 Other degenerative diseases of the nervous system (%)	20.6	7.2	0.7	7.7	0.2	0.8	0.9	NA	16.2	NA
G35-G37 Demyelinating diseases of the central nervous system (%)	0.2	0.1	0.1	0.2	0.5	0.3	3.0	NA	2.0	0.1
G40-G47 Episodic and paroxysmal disorders (%)	0.2	1.1	1.3	1.2	2.0	1.2	1.0	NA	0.8	2.6
H00-H59 Diseases of the eye and adnexa (%)	**0**	**0.1**	**2.2**	**0.2**	**6.9**	**0.4**	**0.9**	**1.8**	**3.4**	
H40-H42 Glaucoma (%)	0	<0.1	0.2	<0.1	1.3	0.1	0.1	NA	NA	5.4

Continued

Table 5 Continued

	Causes of deaths*		Primary health care†		Specialised medical care†					
	Primary	Contributory	Outpatient doctor visits	Hospitalisation (days)	Outpatient visits	Hospitalisation (days)	Disability pensions‡	Sickness allowance§	Disability allowance¶	Prescription medicines**
I00-I99 Diseases of the circulatory system (%)	**34.2**	**48.9**	**11.3**	**20.9**	**5.2**	**15.5**	**4.0**	**3.1**	**14.3**	
I10-I15 Hypertensive diseases (%)	5.7	13.1	5.3	1.8	0.4	0.4	<0.1	0.3	1.0	20.0
I20-I25 Ischaemic heart diseases (%)	14.9	8.8	0.9	2.3	0.4	2.2	0.4	0.4	2.5	6.1
I30-I52 Other forms of heart disease (%)	3.3	18.9	3.1	6.0	2.3	5.7	0.5	0.7	2.8	6.1‡‡
I60-I69 Cerebrovascular diseases (%)	8.3	5.0	0.4	8.6	0.5	4.9	2.8	0.6	7.1	NA
I70-I79 Diseases of arteries, arterioles and capillaries (%)	1.5	2.0	0.3	1.4	0.5	1.4	0.1	0.3	0.5	NA
J00-J99 Diseases of the respiratory system (%)	**3.4**	**7.8**	**8.9**	**7.7**	**2.9**	**6.1**	**0.9**	**6.6**	**1.5**	
J40-J47 Chronic lower respiratory diseases (%)	1.7	3.5	2.0	1.0	0.9	1.0	0.7	0.8	0.6	13.0
M00-M99 Diseases of the musculoskeletal system and connective tissue (%)	**0.5**	**2.2**	**18.4**	**6.1**	**9.7**	**6.5**	**24.5**	**30.6**	**13.5**	
M00-M25 Arthropaties (%)	0.3	1.1	6.1	2.3	4.8	3.8	10.6	6.1	7.8	NA
M40-M54 Dorsopathies (%)	<0.1	0.2	5.3	2.1	2.5	1.7	10.5	11.8	3.2	NA
M60-M79 Soft tissue disorders (%)	<0.1	0.1	6.3	0.6	1.4	0.3	2.1	4.8	NA	NA
N00-N99 Diseases of the genitourinary system (%)	**0.4**	**4.5**	**3.9**	**4.6**	**5.4**	**3.5**	**0.3**	**2.9**	**0.5**	
N17-N19 Renal failure (%)	0.2	3.3	0.2	0.4	0.8	0.5	0.1	NA	NA	NA

The proportions of outcomes in order of size (from largest to smallest) in each column.
1.–5.(very dark red), 6.–10. (dark red), 11.–15. (red), 16.–20.(light red), 21.—(very light red), ICD-10 chapters (grey), NA Not available in the present study.
The proportions are calculated from the total number (n) based on the aggregated register-based data covering the Finnish adult population.
*Causes of deaths register, Statistics Finland.[31]
†Care register for Health Care, Finnish Institute for Health and Welfare; main diagnosis.[35]
‡Disability pension granted after 1995; primary diagnosis; Finnish Centre for Pensions (ETK).[38]
§Social Insurance Institution (Kela).[60]
¶Social Insurance Institution (Kela).[61]
**Social Insurance Institution (Kela); not including limited special reimbursement.[62]
††Including ICD-10 codes: F01, F03, F06.0-F06.3, F20-F25, F28, F29, F30.1, F30.2, F31, F32.3, F33.3, F84, G10, G20, G30.0, G30.1, G30.8, G30.9, G31.0, G35, G40.9.
‡‡Including ICD-10 codes: I11.0, I13, I47–I49 I50, I97.1, P29.0.

Finland—health policy prospects (PoDDy- HePo)'[50] will be used.

The effects of non-participation will be handled using multiple imputation, inverse probability weights[51] or other suitable methods in all analyses. Complex sampling designs will be taken into account in the analyses.

Ethics and dissemination

The HES data have been collected following the legislation at the time of the survey. The following ethics committees of the Hospital District of Helsinki and Uusimaa have approved the data collection and register linkages for each survey:

- Ethical committee for research in epidemiology and public health: Health 2000 (407/E3/2000) and FINRISK 2002 (558/E3/2001).
- Coordinating ethics committee: FINRISK 2007 (299/EO/06), FINRISK 2012 (162/13/03/00/2011), Health 2011 (45/13/03/00/11), Migrant Health and Wellbeing study (325/13/03/00/2009), FinHealth 2017 (37/13/03/00/2016).

The participants were fully informed, and they participated in the surveys voluntarily. The participants also provided written informed consent for the use of their data and register linkage. Permissions for record linkage have been obtained from data controllers. Both survey and register data include sensitive personal information. To ensure data confidentiality, only a very limited number of persons working with raw data (the data managers) have access to personal identification information, which is used to link survey data to register information. Researchers will work on pseudonymised data sets. When processing the personal data, EU General Data Protection Regulation will be followed.

Dissemination is targeted at the scientific community, health authorities and policymakers as well as the media and general public. Dissemination platforms will include 6–8 peer-reviewed scientific publications, conference and workshop presentations, website (under THL website) and social media such as Facebook and Twitter accounts of THL.

Author affiliations
[1]Department of Public Health and Welfare, Finnish Institute for Health and Welfare, Helsinki, Finland
[2]Department of Knowledge Brokers, Finnish Institute for Health and Welfare, Helsinki, Finland

Contributors TJ, PK, AL and SK were mainly responsible for the study concept and design. JS provided expertise on mental health and JJ on register-based data. TJ drafted the first version of the manuscript. All authors critically commented the manuscript.

Funding This work was supported by the Social Insurance Institution, Finland (grant number 65/26/2019).

Competing interests JJ is Senior Medical Officer of the Care Register for Health Care.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Consent obtained directly from patient(s).

Provenance and peer review Not commissioned; externally peer reviewed.

ORCID iDs
Tuija Jääskeläinen http://orcid.org/0000-0001-6263-2298
Annamari Lundqvist http://orcid.org/0000-0002-0262-8585

REFERENCES

1 The Academy of Medical Sciences. Multimorbidity: a priority for global health research, 2018. Available: https://acmedsci.ac.uk/file-download/82222577
2 World Health Organization. *Multimorbidity: technical series on safer primary care*. Geneva, 2016.
3 The Lancet. Making more of multimorbidity: an emerging priority. *Lancet* 2018;391:1637.
4 Xu X, Mishra GD, Jones M. Evidence on multimorbidity from definition to intervention: An overview of systematic reviews. *Ageing Res Rev* 2017;37:53–68.
5 Fortin M, Stewart M, Poitras M-E, *et al*. A systematic review of prevalence studies on multimorbidity: toward a more uniform methodology. *Ann Fam Med* 2012;10:142–51.
6 Marengoni A, Angleman S, Melis R, *et al*. Aging with multimorbidity: a systematic review of the literature. *Ageing Res Rev* 2011;10:430–9.
7 Barnett K, Mercer SW, Norbury M, *et al*. Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study. *Lancet* 2012;:;380:37–43. 7.
8 Violan C, Foguet-Boreu Q, Flores-Mateo G, *et al*. Prevalence, determinants and patterns of multimorbidity in primary care: a systematic review of observational studies. *PLoS One* 2014;9:e102149.
9 Ho IS-S, Azcoaga-Lorenzo A, Akbari A, *et al*. Examining variation in the measurement of multimorbidity in research: a systematic review of 566 studies. *Lancet Public Health* 2021;6:e587–97.
10 Stirland LE, González-Saavedra L, Mullin DS, *et al*. Measuring multimorbidity beyond counting diseases: systematic review of community and population studies and guide to index choice. *BMJ* 2020;368:m160.
11 Sinnige J, Braspenning J, Schellevis F, *et al*. The prevalence of disease clusters in older adults with multiple chronic diseases--a systematic literature review. *PLoS One* 2013;:;8:e79641. 11.
12 Prados-Torres A, Calderón-Larrañaga A, Hancco-Saavedra J, *et al*. Multimorbidity patterns: a systematic review. *J Clin Epidemiol* 2014;67:254–66.
13 Garin N, Koyanagi A, Chatterji S, *et al*. Global Multimorbidity Patterns: A Cross-Sectional, Population-Based, Multi-Country Study. *J Gerontol A Biol Sci Med Sci* 2016;71:205–14.
14 Nguyen H, Wu Y-T, Dregan A, *et al*. Multimorbidity patterns, all-cause mortality and healthy aging in older English adults: Results from the English Longitudinal Study of Aging. *Geriatr Gerontol Int* 2020;20:1126–32.
15 Sakakibara BM, Obembe AO, Eng JJ. The prevalence of cardiometabolic multimorbidity and its association with physical activity, diet, and stress in Canada: evidence from a population-based cross-sectional study. *BMC Public Health* 2019;19:1361.
16 Singh-Manoux A, Fayosse A, Sabia S, *et al*. Clinical, socioeconomic, and behavioural factors at age 50 years and risk of cardiometabolic multimorbidity and mortality: A cohort study. *PLoS Med* 2018;15:e1002571.
17 Dhalwani NN, Zaccardi F, O'Donovan G, *et al*. Association Between Lifestyle Factors and the Incidence of Multimorbidity in an Older English Population. *J Gerontol A Biol Sci Med Sci* 2017;72:528–34.
18 Noble N, Paul C, Turon H, *et al*. Which modifiable health risk behaviours are related? A systematic review of the clustering of Smoking, Nutrition, Alcohol and Physical activity ('SNAP') health risk factors. *Prev Med* 2015;81:16–41.
19 Fortin M, Haggerty J, Almirall J, *et al*. Lifestyle factors and multimorbidity: a cross sectional study. *BMC Public Health* 2014;14:686.
20 Freisling H, Viallon V, Lennon H, *et al*. Lifestyle factors and risk of multimorbidity of cancer and cardiometabolic diseases: a multinational cohort study. *BMC Med* 2020;18:5.

21 Ryan A, Wallace E, O'Hara P, *et al*. Multimorbidity and functional decline in community-dwelling adults: a systematic review. *Health Qual Life Outcomes* 2015;13:168.
22 Hunger M, Thorand B, Schunk M, *et al*. Multimorbidity and health-related quality of life in the older population: results from the German KORA-age study. *Health Qual Life Outcomes* 2011;9:53.
23 Xu X, Mishra GD, Jones M. Mapping the global research landscape and knowledge gaps on multimorbidity: a bibliometric study. *J Glob Health* 2017;7:010414.
24 Population Information System, 2022. Available: https://dvv.fi/en/population-information-system
25 Heistaro S, ed. *Methodology Report: Health 2000 Survey. Publications of the National Public Health Institute*. Helsinki, Finland: National Public Health Institute, 2008.
26 Lundqvist A, Mäki-Opas T, eds. *Health 2011 – Methods. Publications of the*. Helsinki, Finland: National Institute for Health and Welfare, 2016.
27 Borodulin K, Tolonen H, Jousilahti P, *et al*. Cohort profile: the National FINRISK study. *Int J Epidemiol* 2018;47:696–696i.
28 Borodulin K, Sääksjärvi K. *FinHealth 2017 study – methods. Finnish Institute for health and welfare; report 17/2019*. Helsinki, Finland, 2019.
29 Castaneda A, Rask S, Härkänen T. Enhancing Survey Participation among Foreign-Born Populations. In: *Finnish Yearbook of population research*. , 2018: Vol 53, 89–110.
30 Tolonen H. EHES Manual. In: *Part B. fieldwork procedures*. 2 edn. Helsinki, Finland: National Institute for Health and Welfare, 2016.
31 Statistical Finland. Causes of death, 2021. Available: https://www.stat.fi/til/ksyyt/index_en.html
32 Statistical Finland. Educational structure of population, 2018. Available: https://www.stat.fi/meta/til/vkour_en.html
33 Statistical Finland. Employment, 2021. Available: https://www.stat.fi/meta/til/tyokay_en.html
34 Statistical Finland. Employment, 2020. Available: https://www.stat.fi/til/tyokay/index_en.html
35 Finnish Institute for Health and Welfare. Available: https://thl.fi/en/web/thlfi-en/statistics/information-on-statistics/register-descriptions/care-register-for-health-care [Accessed Access date 03/2021].
36 Finnish cancer registry. Data protection notice, 2021. Available: https://cancerregistry.fi/information/data-protection-notice/
37 Kela. Statistics by topic, 2021. Available: https://www.kela.fi/web/en/statistics-by-topic
38 Finnish center of pension. Register descriptions, 2021. Available: https://www.etk.fi/en/services-for-experts/registers/register-descriptions/
39 Real estate, building, and spatial information, 2022. Available: https://dvv.fi/en/real-estate-building-and-spatial-information
40 Wittchen HU, Lachner G, Wunderlich U, *et al*. Test-retest reliability of the computerized DSM-IV version of the Munich-Composite International Diagnostic Interview (M-CIDI). *Soc Psychiatry Psychiatr Epidemiol* 1998;33:568–78.
41 Beck AT, Beck RW. Screening depressed patients in family practice. A rapid technic. *Postgrad Med* 1972;52:81–5.
42 Derogatis LR, Lipman RS, Rickels K, *et al*. The Hopkins Symptom Checklist (HSCL): a self-report symptom inventory. *Behav Sci* 1974;19:1–15.
43 Männistö S, Virtanen M, Mikkonen T, *et al*. Reproducibility and validity of a food frequency questionnaire in a case-control study on breast cancer. *J Clin Epidemiol* 1996;49:401–9.
44 Paalanen L, Männistö S, Virtanen MJ, *et al*. Validity of a food frequency questionnaire varied by age and body mass index. *J Clin Epidemiol* 2006;59:994–1001.
45 Fop M, Murphy TB. Variable selection methods for model-based clustering. *Stat Surv* 2018;12:18–65.
46 Härkänen T, But A, Haukka J. Non-Parametric Bayesian intensity model: exploring Time-to-Event data on two time scales. *Scandinavian Journal of Statistics* 2017;44:798–814.
47 Iacobelli S, Carstensen B. Multiple time scales in multi-state models. *Stat Med* 2013;32:5315–27.
48 VanderWeele TJ. Sufficient cause interactions and statistical interactions. *Epidemiology* 2009;20:6–13.
49 Laaksonen MA, Knekt P, Rissanen H, *et al*. The relative importance of modifiable potential risk factors of type 2 diabetes: a meta-analysis of two cohorts. *Eur J Epidemiol* 2010;25:115–24.
50 Paalanen L, Härkänen T, Tolonen H. Protocol of a research project 'Projections of the burden of disease and disability in Finland - health policy prospects' using cross-sectional health surveys and register-based follow-up. *BMJ Open* 2019;9:e029338.
51 Härkänen T, Karvanen J, Tolonen H, *et al*. Systematic handling of missing data in complex study designs – experiences from the health 2000 and 2011 surveys. *J Appl Stat* 2016;43:2772–90.
52 Sintonen H. *The 15D measure of health related quality of life: reliability, validity and sensitivity of its health state descriptive system, working paper 41*. Melbourne: National Centre for Health Program Evaluation, 1994.
53 Sintonen H. *The 15D measure of health related quality of life. II feasibility, reliability and validity of its valuation system, working paper 42*. Melbourne: National Centre for Health Program Evaluation, 1995.
54 Rabin R, de Charro F. EQ-5D: a measure of health status from the EuroQol Group. *Ann Med* 2001;33:337–43.
55 Power M. Development of a common instrument for quality of life. In: *Developing common instruments for health surveys*. . Biomedical and Health Research, 2003: Vol 57. 145–63.
56 Aromaa A, Heliövaara M, Impivaara O. *Terveys, toimintakyky JA hoidontarve Suomessa. Mini-Suomi -terveystutkimuksen perustulokset. Kansaneläkelaitoksen julkaisuja AL:32*. Kela: Helsinki ja Turku, 1989.
57 Tuomi K, Ilmarinen J, Jahkola A. *Work ability index*. 2 edn. Helsinki: Finnish Institute of Occupational Health, 1998.
58 Bush K, Kivlahan DR, McDonell MB, *et al*. The AUDIT alcohol consumption questions (AUDIT-C): an effective brief screening test for problem drinking. Ambulatory Care Quality Improvement Project (ACQUIP). Alcohol Use Disorders Identification Test. *Arch Intern Med* 1998;158:1789–95.
59 Dawson DA, Grant BF, Stinson FS, *et al*. Effectiveness of the derived Alcohol Use Disorders Identification Test (AUDIT-C) in screening for alcohol use disorders and risk drinking in the US general population. *Alcohol Clin Exp Res* 2005;29:844–54.
60 Keskinen S, ed. *Kelan sairausvakuutustilasto 2018*. Helsinki, Finland: Kela, 2019.
61 Pösö R, ed. *Kelan vammaisetuustilasto 2018*. Helsinki, Finland: Kela, 2019.
62 Statistical database Kelasto, 2020. Available: https://www.kela.fi/web/en/statistical-database-kelasto?inheritRedirect=true