












OPEN

Clustering and trajectories of key noncommunicable disease risk factors in Norway: the NCDNOR project

Knut Eirik Dalene^{1,7}, Simon Lergenmuller^{2,7}, Erik R. Sund^{3,4}, Laila A. Hopstock⁵, Trude Eid Robsahm⁶, Yngvar Nilssen², Wenche Nystad¹, Inger Kristin Larsen² & Inger Ariansen¹

Noncommunicable diseases (NCDs) are a leading cause of premature death globally and have common preventable risk factors. In Norway, the NCDNOR-project aims at establishing new knowledge in the prevention of NCDs by combining information from national registries with data from population-based health studies. In the present study, we aimed to harmonize data on key NCD risk factors from the health studies, describe clustering of risk factors using intersection diagrams and latent class analysis, and identify long-term risk factor trajectories using latent class mixed models. The harmonized study sample consisted of 808,732 individuals (1,197,158 participations). Two-thirds were exposed to ≥ 1 NCD risk factor (daily smoking, physical inactivity, obesity, hypertension, hypercholesterolaemia or hypertriglyceridaemia). In individuals exposed to ≥ 2 risk factors (24%), we identified five distinct clusters, all characterized by fewer years of education and lower income compared to individuals exposed to < 2 risk factors. We identified distinct long-term trajectories of smoking intensity, leisure-time physical activity, body mass index, blood pressure, and blood lipids. Individuals in the trajectories tended to differ across sex, education, and body mass index. This provides important insights into the mechanisms by which NCD risk factors can occur and may help the development of interventions aimed at preventing NCDs.

Noncommunicable diseases (NCDs) accounted for ~41 million (74%) of all deaths in 2019, and >17 million died from NCDs before age 70.¹ NCDs represent an individual burden,^{2,3} and substantial global economic costs.⁴ In response, the World Health Organization (WHO) and United Nations aim to reduce premature NCD mortality by 33% within 2030.^{5,6} Key modifiable NCD risk factors include the behavioural factors tobacco use, physical inactivity, an unhealthy diet, and harmful use of alcohol, and the biological factors obesity, hypertension, hyperlipidaemia and hyperglycaemia.⁷ While studies have investigated behavioural and biological risk factors separately,^{8–11} few have explored the clustering of these factors combined.^{12–14} Previous research has identified risk factor patterns in different populations, using various methods,^{8,9,15,16} but large population-based studies in adults from westernized countries using appropriate model-based techniques are lacking. Furthermore, studies have indicated that risk factor patterns may vary over time,^{17–21} yet to our knowledge, few studies investigated trajectories of three or more risk factors in prospectively collected data.^{18,19} As longevity increases, it is likely that NCD prevalence will increase, but this may not affect all population groups equally. Hence, to counteract further widening of social inequality in health,²² we need to identify vulnerable groups of individuals who cluster with respect to NCD risk factors.

“A life-course approach to prevent noncommunicable diseases in an ageing population – NCDNOR”, is a research project aiming at establishing new knowledge in the prevention of NCDs by combining NCD endpoints

¹Department of Chronic Diseases, Norwegian Institute of Public Health, PO Box 222, 0213 Oslo, Skøyen, Norway. ²Department of Registration, Cancer Registry of Norway, PO Box 5313, 0304 Oslo, Majorstuen, Norway. ³Department of Public Health and Nursing, Norwegian University of Science and Technology, Oslo, Norway. ⁴Levanger Hospital, Nord-Trøndelag Hospital Trust, Levanger, Norway. ⁵Department of Health and Care Sciences, UiT The Arctic University of Norway, Oslo, Norway. ⁶Department of Research, Cancer Registry of Norway, Oslo, Norway. ⁷These authors contributed equally: Knut Eirik Dalene and Simon Lergenmuller. ✉email: keda@fhi.no

across somatic disciplines, examining effects of socioeconomic circumstances, health behaviours, biological markers, and mental health throughout the life-course. The project utilizes unique data from administrative and health registries covering the entire Norwegian population and includes information on both behavioural and biological NCD risk factors from Norwegian population-based health studies covering ~ 800,000 individuals nationwide and ~ 1.2 million person-observations. In the present study, we aim to (i) describe the health studies included in NCDNOR and the harmonization of data on key NCD risk factors to create one study sample, (ii) describe the clustering of key NCD risk factors, and (iii), for individuals with repeated measurements, identify trajectories of key NCD risk factors.

Methods

Study sample. The NCDNOR project includes health study data from women and men who participated in at least one of the following studies: the Norwegian Counties Study^{23,24}; the Age 40 Program Oslo²⁵; the Age 40 Program²⁶; the Trøndelag Health Studies (HUNT1–HUNT4/HUNT4ST)²⁷; the Tromsø Study (Tromsø4–Tromsø7)²⁸; and/or the Cohort of Norway (CONOR) studies²⁹ (Fig. 1). The different studies have been described elsewhere,^{23–29} and thus only briefly described in the following. More details are given in Supplementary Table S1.

The Norwegian Counties Study included participants from three Norwegian counties, with three waves of data collection 1974–1988 (Fig. 1). For these first three waves (participation 84–88%),³⁰ all residents aged 35–49 years and random samples of residents aged 20–34 were invited. The second and third wave also invited previous participants. In 2006–2008, a fourth wave was conducted among a sample of previous participants (participation 59%).²⁴ We included 93,568 individuals from the Norwegian Counties Study (Fig. 2), of which > 50,000 individuals participated at least three times.

The Age 40 Program Oslo was a cardiovascular disease (CVD) risk factor screening program conducted by the Oslo municipality health council in collaboration with the National Health Screening Service in Norway.³¹ All residents who turned 40 years between 1981 and 1988 were invited the year they turned 40 (participation 55%).²⁵ A restructured program continued until 1999. Due to poor quality of the data, we did not include data collected after 1988.²⁵ We included 23,886 individuals from the Age 40 Program Oslo (Fig. 2).

The Age 40 Program was similar to the Age 40 Program Oslo and conducted by the National Health Screening Service in the other Norwegian counties in 1985–1999. All residents aged 40–42 years were invited.²⁶ In addition, some counties invited residents aged < 40, 43–44 and 65–67 years.³² Some individuals moved between counties, resulting in ~ 10,000 individuals participating several times. The overall participation was ~ 69% (81% in 1985, 52% in 1999).²⁶ We included 416,944 individuals from the Age 40 Program (Fig. 2).

The HUNT Study has been conducted in four waves (1984–2019) (Fig. 1).²⁷ All residents ≥ 20 years of age in the former Nord-Trøndelag county were invited (participation 89% (HUNT1), 54% (HUNT4)). In 2019, HUNT also invited all residents aged ≥ 18 years in the former Sør-Trøndelag county (HUNT4ST, participation 43%). In total, ~ 250,000 individuals have participated, yielding ~ 380,000 person-observations. We included 225,916 individuals from the HUNT Study (Fig. 2).

The Tromsø Study has been conducted seven times (1974–2016).²⁸ NCDNOR includes individuals from Tromsø4–Tromsø7 (Fig. 1). In Tromsø4 and Tromsø7 all residents of Tromsø municipality aged ≥ 25 and ≥ 40 years were invited, respectively. In Tromsø5 and Tromsø6, previous participants and age-group specific random samples were invited (participation 65–79%).³³ We included 36,648 individuals from Tromsø4–Tromsø7 (Fig. 2), of which ~ 20,000 individuals participated at least two times.

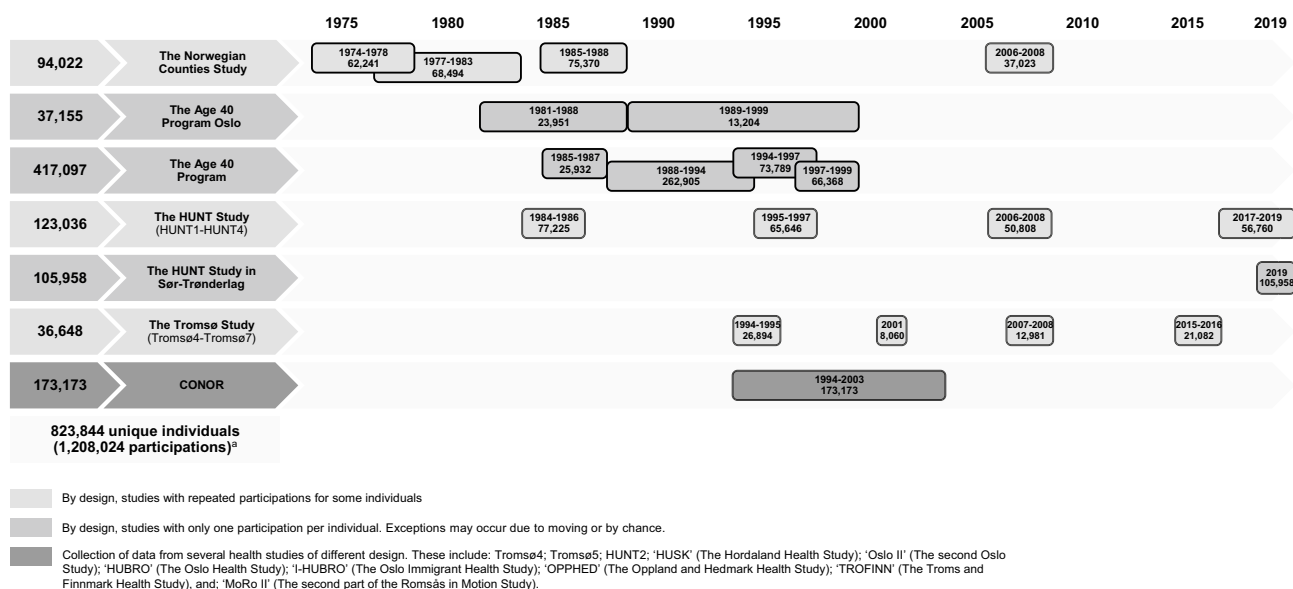


Figure 1. Timeline for the health studies included in this study. Abbreviations: HUNT, The Trøndelag Health Study; CONOR, Cohort of Norway. ^aCONOR includes data from HUNT2 and Tromsø4–Tromsø5. These participations were only counted once.

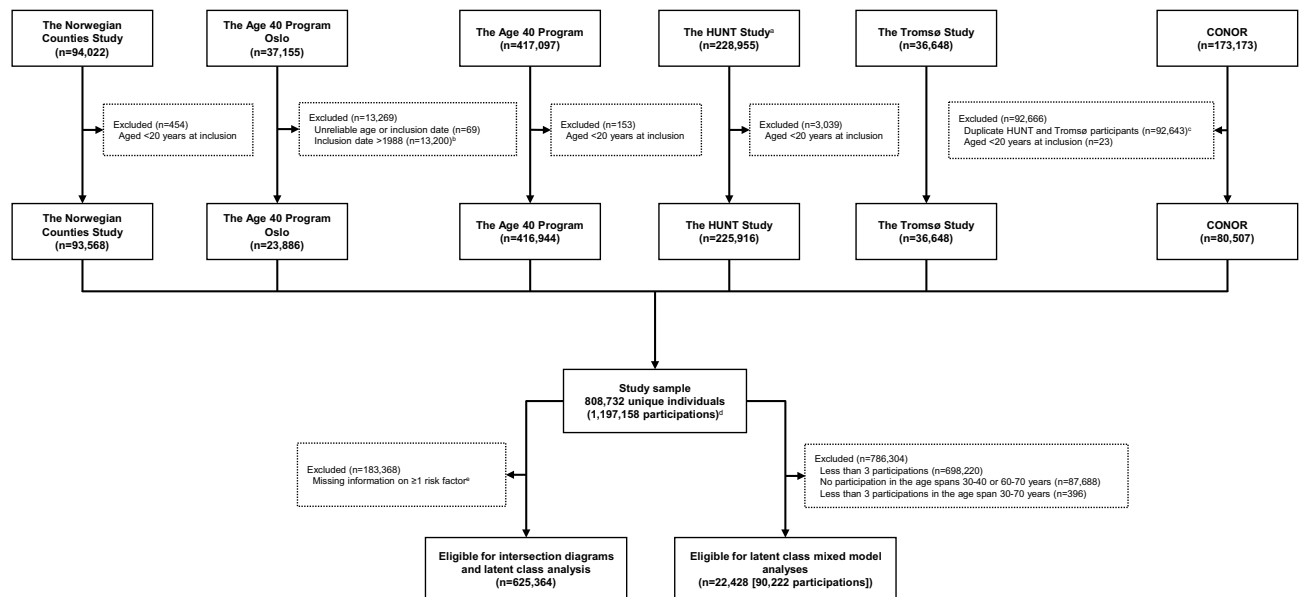


Figure 2. Selection of individuals from the different health studies into the study samples. Abbreviations: HUNT, The Trøndelag Health Study; CONOR, Cohort of Norway. ^aIncluding HUNT Sør-Trøndelag, ^bThe quality of the data collected after 1988 was reported to be poor. ^cCONOR includes data from HUNT2 and Tromsø4-Tromsø5. We excluded those from CONOR. ^d808,732 does not correspond to the sum of the numbers at the bottom of each health study, as individuals may have participated in several different studies. In addition, 181 individuals participated to several health studies < 1 year apart. These participations are not included in 1,197,158. ^eMissing information consists of 77.6% non-measured, non-response (20.7%) and manual cleaning decisions (1.7%).

CONOR results from a collaboration between epidemiological research centres in Norway, and includes harmonized data from 10 regional health studies (including HUNT2, Tromsø4 and Tromsø5, Fig. 1).²⁹ Some studies invited everyone above a specific age, others invited selected age groups.³⁴ After excluding the duplicate HUNT and Tromsø participants, we included 80,507 individuals from the remaining seven CONOR studies (participation 30–66%) (Fig. 2).

NCDNOR also includes health study data from the Norwegian Youth Health Studies,³² and the Young-HUNT studies²⁷ not included in the present study.

Harmonization of data on NCD risk factors. The protocols of the health studies included in NCDNOR were similar,³¹ facilitating harmonization and pooling of data. All studies included questionnaires covering behavioural factors. Except for HUNT4ST and the 4th Norwegian Counties Study, studies included anthropometric and biomarker measurements conducted by trained personnel. A full description of the data can be found elsewhere.^{33,35,36} Based on data-availability, and the global strategies for the prevention of NCDs,⁷ we consider the following key NCD risk factors: *daily smoking*, *physical inactivity*, *harmful use of alcohol* (only available for a subsample), *obesity*, *hypertension*, *hypercholesterolaemia* and *hypertriglyceridaemia*. A brief description of data harmonization and risk factor definitions follows (additional details in Supplementary Tables S2–S4).

Behavioural risk factors. Questions about smoking varied between studies, but all questionnaires allowed categorization of participants into smoking status (never, former, or current daily smokers [risk factor *daily smoking*]), and calculation of *smoking intensity* (cigarettes/day [currently or previously]) and smoking duration (years) (Supplementary Tables S2a & S2b).

Leisure-time physical activity (*LTPA*) was assessed either by the Saltin-Grimby Physical Activity Level Scale (SGPALS), the CONOR instrument,³⁷ the HUNT instrument,³⁸ or a combination (Supplementary Tables S3a–d).³⁹ These were harmonized into the SGPALS categories (Supplementary Tables S3e–f): (1) Reading, TV-watching, or other sedentary activities (risk factor *physical inactivity*), (2) Riding a bicycle or walking (including to/from work) or doing other activities of light-moderate intensity ≥ 4 h/week, (3) Heavy gardening, sports, exercise ≥ 4 h/week, and (4) Regular hard physical exercise, participation in competitive sports, etc. several times/week.

Information on frequency of alcohol consumption, alcohol units usually consumed over 14 days, frequency of heavy episodic drinking, and alcohol problems were available for a subset of studies (Supplementary Table S4). We define *harmful use of alcohol* as > 5 units of beer, wine or spirits on at least one occasion in the past 30 days (corresponding to ≥ 59 g of pure alcohol) or self-reported alcohol problems.^{40,41} We also calculated *mean alcohol consumption* (grams of pure alcohol/14 days) by multiplying units of beer (0.33L), wine (0.15L) and spirits (0.04L) by 11.8, 14.0 and 12.6, respectively⁴¹.

Biological risk factors. With two exceptions, height and weight were measured by trained staff (self-reported in HUNT4ST and the 4th Norwegian Counties Study). We removed extreme values ($< 120 / > 220$ cm, $< 30 / > 260$ kg) and calculated BMI (kg/m^2). BMI ≥ 30 defined *obesity* (values $< 10 / > 80$ were set to missing). For the 4th Norwegian Counties Study, only weight was self-reported, and BMI calculated using height measured in the 1st, 2nd or 3rd Norwegian Counties Study²⁴.

In the 1st and 2nd Norwegian Counties Study, HUNT1, and the Age 40 Program Oslo, blood pressure was measured twice using calibrated mercury sphygmomanometers (≥ 1 -min interval) after ≥ 4 min seated rest.⁴² Thereafter, three automatic oscillometric blood pressure measurements were taken (1-min intervals) after ≥ 2 min seated rest. We harmonized measurements using the second sphygmomanometer measurement and the rounded arithmetic mean of oscillometric measurements two and three.^{43,44} If only two oscillometric measurements were available, we used the last measurement.⁴³ We excluded extreme blood pressure values (diastolic $< 30 / > 160$ mmHg, systolic < 60 mmHg). In the present study, we defined *hypertension* as a *systolic blood pressure* ≥ 160 mmHg and/or *diastolic blood pressure* ≥ 100 mmHg (Grade 2 or 3 hypertension),⁴⁴ as outlined in the national guidelines for the prevention of cardiovascular disease when considering blood pressure as an individual risk factor⁴⁵.

Serum *total cholesterol*- and *triglycerides* were analysed from non-fasting, venous blood samples using the non-enzymatic Liebermann-Burchard method (Technicon Autoanalyzer) until 1979.⁴⁶ Thereafter, analyses were conducted using enzymatic methods. Results obtained before 1979 were recalculated using a correction factor (enzymatic method = $0.92 \times$ non-enzymatic method + 0.03).⁴⁶ Total cholesterol < 2 mmol/L or > 20 mmol/L and triglycerides < 0.01 mmol/L or > 30 mmol/L were set to missing.^{47,48} We then defined the risk factors *hypercholesterolaemia* (≥ 7 mmol/L) and *hypertriglyceridaemia* (≥ 4 mmol/L), which were guided by national guidelines for the prevention of cardiovascular disease and the 2019 European Society of Cardiology and European Atherosclerosis Society Guidelines for the management of dyslipidaemias^{45,48}.

Other variables. The health study data included in NCDNOR can be linked to national mandatory health and administrative registries that cover the entire population.⁴⁹ We used data from Statistics Norway on geography, education and income.⁵⁰ We categorized highest attained educational into levels corresponding to the current Norwegian standard⁵¹: 1) Primary or lower secondary school (Primary [≤ 10 years of education]); 2) Upper secondary education (Secondary [11–14 years]); 3) Higher education (Tertiary [> 14 years, typically a university/college degree]), and highest attained income using year and sex-specific quintiles.

Statistical analyses. We ensured the correct ordering of the participations within individuals with repeated measurements by excluding participations < 1 year apart ($n = 181$, Fig. 2 (due to data protection risk minimisation, date uncertainties could range 45–365 days)). We defined the time of first participation as ‘study entry’.

We described the occurrence and clustering of NCD risk factors at study entry using intersection diagrams overall, by sex and study entry period (1974–1989, 1990–2004 and 2005–2019).⁵² We further described the clustering of risk factors using latent class analysis (LCA) to identify distinct patterns of risk factors among individuals with at least two such risk factors (Supplementary Methods)^{53,54}.

To identify trajectories of risk factors among individuals with repeated participations, we used latent class mixed models (LCMMs) (Supplementary Methods 2).^{54–56} We used the continuous/categorical version of the NCD risk factors (smoking intensity [among smokers], LTPA, BMI, blood pressure, blood lipids). For blood pressure and blood lipids, we used multivariate LCMMs, allowing simultaneous modelling of trajectories.⁵⁶ We estimated trajectories for those with at least three participations in the time window 30–70 years of age, whereof one participation had to occur within the age span 30–40 years (trajectory entry), and another within the age span 60–70 years (Supplementary Methods).

For LCA and LCMM, the optimal number of classes was identified by fitting models with 1–6 classes and evaluating each model in terms of: quality of the model fit (Akaike Information Criterion, Bayesian Information Criterion), classification power (entropy, posterior probability), and relevance.^{57–59} Once the optimal number of classes was identified, we allocated individuals to the class for which they had the largest posterior probability (Supplementary Methods).

Depending on the model, and the restrictions used to identify classes, we had various degrees of missing data (non-measured, non-response, or manual cleaning decisions). The proportion of missing was considerably higher for alcohol (Supplementary Table S5), we therefore provide results including alcohol as supplements only. We conducted several sensitivity analyses (Supplementary Methods).

Ethics. Access to data in NCDNOR is based on informed consent from the participants, approval from The Regional Committees for Medical and Health Research Ethics South-East (nr 28,561/2019/1203), and The Data Privacy Impact Assessment at the Norwegian Institute of Public Health⁴⁹. The legal basis for processing of personal data is Article 6 (1) (e) of the GDPR and the exemptions pursuant to Article 9 (2) (j) of the GDPR, with a supplementary legal basis in Sects. 8 and 9 of the Personal Data Act. All methods were performed in accordance with relevant guidelines and regulations including the Declarations of Helsinki.

Results

Combined, the health studies sent out 1,905,277 invitations, resulting in 1,208,024 participations (63%) from 823,844 unique individuals (Fig. 1) covering all Norwegian counties (Supplementary Fig. S1). After exclusions, the final study sample consisted of 808,732 individuals (52% women) aged ≥ 20 years (born 1882–1999), who participated in at least one health study between 1974 and 2019, for a total of 1,197,158 participations (Fig. 2).

Repeated measurements were available for 211,705 (26.2%) participants, with 110,512 (13.7%) participating at least three times and 52,222 (6.5%) participating at least four times (Table 1).

At study entry, median age was 41.4 (inter-quartile range [IQR]: 3.3) years, median birth year was 1951 (IQR: 11.5), and the median participation year was 1993 (IQR: 11.2) (Table 2). Cohort characteristics stratified by participation number for those that participated at least three times are shown in Supplementary Table S6. For these, the median age at study entry was 39.4 years (IQR: 11.7), and median birth year 1940 (IQR: 16.7). The median time between participations was 5.0 (IQR: 6.0) years between participation 1 and 2, and 9.4 (IQR: 6.0) years between participation 2 and 3.

At study entry, the proportion of missing information ranged from 2.6% (obesity) to 21.0% (hypercholesterolaemia) (Supplementary Fig. S2) but was considerably higher for harmful use of alcohol (73.4%). Further details on missing information in Supplementary Table S5.

Clustering of NCD risk factors. *Intersection diagrams.* At study entry, 625,364 (77.3%) participants had available information on daily smoking, physical inactivity, obesity, hypertension, hypercholesterolaemia and hypertriglyceridaemia (Fig. 3). Of these, 62.3% had at least one risk factor, 24.0% had at least two risk factors and 6.2% had at least three risk factors. Daily smoking (38.3%), physical inactivity (23.2%), and hypercholesterolaemia (12.2%) were the most prevalent risk factors. The three most common combinations of two risk factors were daily smoking and physical inactivity (7.3%), daily smoking and hypercholesterolaemia (2.9%) and physical inactivity and obesity (1.4%). The most common combination of three risk factors was daily smoking, physical inactivity, and hypercholesterolaemia (1.2%) (Fig. 3). Patterns were similar in women and men, although higher proportions of men had hypercholesterolaemia, hypertension, hypertriglyceridaemia, and multiple risk factors (Supplementary Fig. S3).

The proportion of individuals with at least one risk factor was higher for those included in the periods 1974–1989 (64.8%) and 1990–2004 (63.8%) than for those included later (2005–2019, 56.4%) (Supplementary Fig. S4). From 1974–1989 to 2005–2019, we observed a decrease in daily smoking (from 43.8% to 17.0%) and hypercholesterolaemia (from 16.5% to 9.0%) and an increase in obesity (from 7.1% to 22.8%). Individuals included in 1974–1989 and 1990–2004 were also younger at study entry than those included in 2005–2019 (median [IQR] age 40.6 [2.4], 44.3 [2.4], and 52.2 [21.2] years, respectively). Consequently, the proportion exposed to at least one risk factor was similar or higher among e.g., 40–49-year-olds in 1974–1983 compared to 70–79-year-olds in 2004–2013 (Supplementary Fig. S5).

In the sub-sample that also had data on harmful use of alcohol ($n = 78,941$), 66.4% had at least one risk factor. The prevalence of harmful use of alcohol was 21.5%, making it the third most prevalent risk factor after daily smoking (26.2%) and physical inactivity (25.8%) (Supplementary Fig. S6).

Latent class analysis. Among individuals with at least two risk factors ($n = 150,355$), we identified five risk factor clusters. Based on the most prevalent risk factors in each cluster, these were labelled: dyslipidaemia ($n = 14,028$), inactive smokers ($n = 45,720$), smokers with hypercholesterolemia ($n = 30,954$), obesity ($n = 31,692$)

	Number of individuals in the health studies with at least k participations ^a						Total no. individuals with at least k participations
	The Norwegian Counties Study	The Age 40 Program Oslo	The Age 40 Program	The HUNT Study	The Tromsø Study	CONOR	
Part. 1, No ^b	93,568	23,884	416,935	225,914	36,648	80,507	808,732
Age ^c	41.0 (8.4)	40.3 (0.6)	41.4 (1.6)	45.5 (29.5)	45.2 (18.4)	45.5 (19.3)	
Part. 2, No	77,853		9,628	68,922	19,909	90	211,705
Age ^c	46.9 (10.8)		44.8 (3.3)	51.6 (22.3)	59.7 (14.7)	47.0 (14.1)	
Part. 3, No	51,620		2,197	38,949	9,647		110,512
Age ^c	51.4 (9.5)		54.7 (16.1)	60.4 (16.5)	66.5 (14.3)		
Part. 4, No	18,768		17	19,290	2,812		52,222
Age ^c	70.5 (6.1)		–	69.8 (13.2)	74.8 (9.9)		
Part. 5, No							10,502
Part. 6, No							2,280
Part. 7–9, No							915
Total No. Participations	241,809	23,884	428,777	353,075	69,016	80,597	1,197,158

Table 1. Number of participations by health studies and overall ($n = 808,732$). *Part* participation, *No.* number, *HUNT* The Trøndelag Health Study, *CONOR* Cohort of Norway. ^aThe total number of individuals in each health study differs from the number of individuals in each study shown in Fig. 2, because participations < 1 year apart from each other (between studies) were excluded ($n = 181$). The total number of individuals in the final study sample remains unchanged. ^bThe number of individuals with at least one participation is also the total number of individuals in each health study. Individuals can participate in several different health studies, therefore, the sum of the numbers in the first row ($n = 877,456$) does not correspond to the total number of individuals with at least one participation in the final study sample ($n = 808,732$). ^cMedian (Inter Quartile Range [IQR]), not shown for the Age 40 Program Participation 4 due to low n ($n = 17$).

Characteristic	Women	Men	Overall
Participants, No. (%)	420,840 (52.0)	387,892 (48.0)	808,732 (100.0)
Age, median (IQR)	41.4 (3.1)	41.5 (3.4)	41.4 (3.3)
Birth year, median (IQR)	1951 (11.5)	1951 (11.5)	1951 (11.5)
Participation year, median (IQR)	1993 (11.2)	1993 (11.4)	1993 (11.2)
Highest attained education ^a	417,109	384,566	801,675
Primary	28.0	24.9	26.5
Secondary	44.3	46.6	45.4
Tertiary	27.7	28.5	28.1
Highest attained income quintile ^a	398,050	378,748	776,798
Quintile 1	7.0	3.6	5.4
Quintile 2	7.9	7.2	7.6
Quintile 3	16.1	16.7	16.4
Quintile 4	26.4	24.4	25.5
Quintile 5	42.4	48.0	45.2
Smoking status (daily)	393,910	364,052	757,962
Never	44.3	36.1	40.4
Former	21.5	27.1	24.2
Current	34.2	36.8	35.5
Smoking intensity (cigarettes/day) ^b	388,957	353,458	742,415
Median (IQR)	5.0 (10.0)	10.0 (15.0)	6.0 (12.5)
Smoking duration (years) ^b	390,408	360,569	750,977
Median (IQR)	5.0 (20.0)	11.0 (22.0)	8.0 (20.0)
Leisure-time physical activity	405,499	372,748	778,247
Category 1	23.9	24.1	24.0
Category 2	57.2	45.2	51.5
Category 3	14.5	23.4	18.7
Category 4	4.5	7.3	5.8
Harmful use of alcohol	110,467	104,581	215,048
No	89.2	76.6	83.1
Yes	10.8	23.4	16.9
Mean alcohol consumption (grams/14 days)	151,803	131,383	283,186
Median (IQR)	28.8 (71.4)	64.5 (100.2)	47.6 (84.1)
Body mass index (kg/m ²)	409,848	377,462	787,310
Median (IQR)	23.9 (5.0)	25.4 (4.2)	24.7 (4.8)
Underweight	1.8	0.4	1.1
Normal weight	59.6	44.7	52.5
Overweight	27.4	44.4	35.6
Obesity	11.1	10.5	10.8
Blood pressure	368,420	345,653	714,073
Systolic (mmHg), median (IQR)	123.0 (20.0)	133.0 (19.0)	128.0 (21.0)
Diastolic (mmHg), median (IQR)	76.0 (14.0)	80.0 (14.0)	78.0 (14.0)
Blood lipids	329,591	309,130	638,721
Total cholesterol (mmol/L), median (IQR)	5.4 (1.4)	5.8 (1.5)	5.6 (1.4)
Triglycerides (mmol/L), median (IQR)	1.1 (0.8)	1.7 (1.3)	1.4 (1.1)

Table 2. Participant characteristics at study entry (n = 808,732). Data are presented as number of participants or column percentages, unless otherwise indicated. Because of rounding, percentages may not sum up to 100% No. number, IQR inter-quartile range. ^aAt study entry. ^bFor current and former smokers.

and hypertension (n = 27,961) (Fig. 4). Compared to individuals with no (n = 236,023) and one (n = 238,986) risk factor, individuals in the five risk factor clusters had fewer years of education, lower income, smoked more, and were less physically active (Supplementary Table S7). The cluster labelled hypertension had the lowest proportions of individuals with tertiary education and high income.

NCD risk factor trajectories. Among the 110,512 individuals that participated at least three times, a total of 22,428 (20.3%) individuals participated at least three times between 30 and 70 years of age (90,222 participations in total) (Fig. 5).

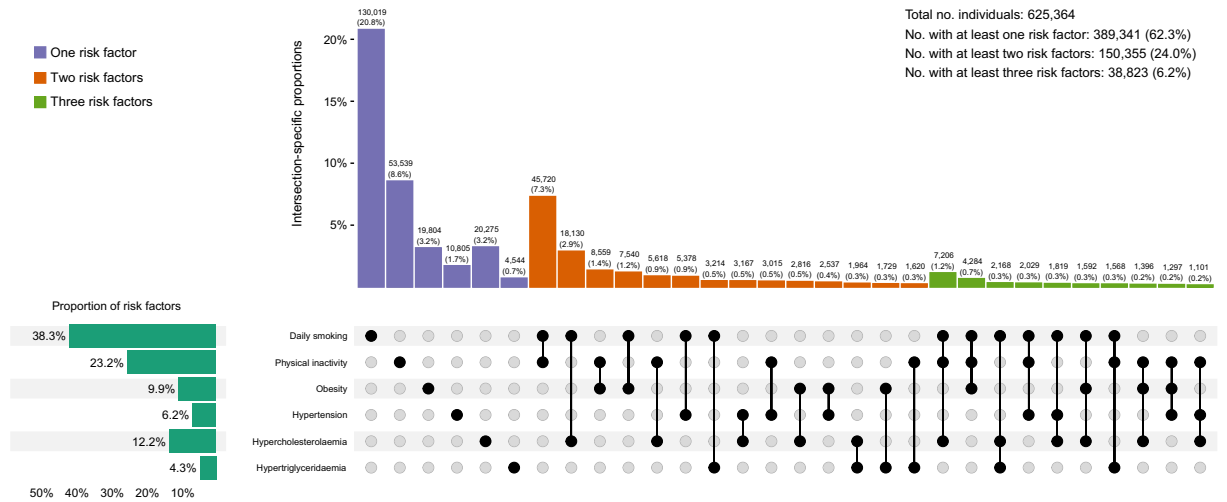


Figure 3. Intersection diagram showing the occurrence and clustering of noncommunicable disease risk factors at study entry (n = 625,364). This shows combinations of noncommunicable disease risk factors in the study sample at study entry. All single risk factor-intersections are shown, as well as the 24 most common combinations of at least two risk factors.

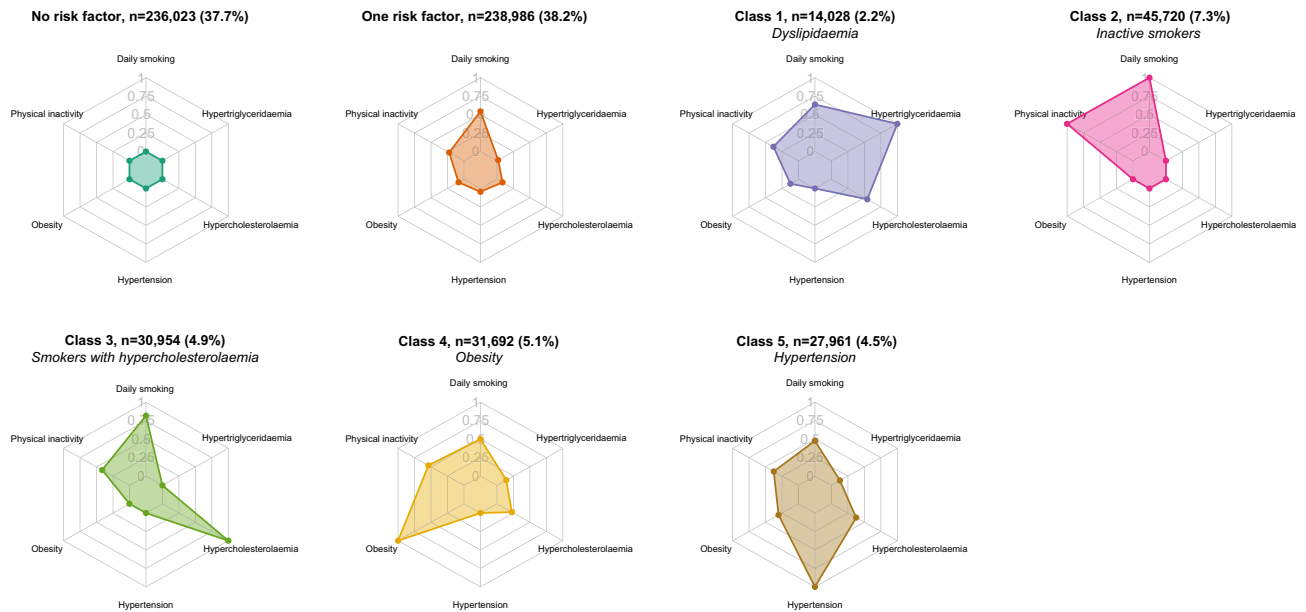


Figure 4. Radar plots showing the distribution of the noncommunicable disease risk factors in each cluster (n = 625,364). For individuals with no or only one risk factor, plots show the proportion of individuals having each risk factor as their single risk factor. For individuals with at least two risk factors, plots show the risk factor distributions of the individuals allocated to each latent class of noncommunicable disease risk factor.

Among smokers, three classes of smoking intensity trajectories were identified: medium-to-low (45.2%), high-to-low (29.0%), and stable high (25.8%) (Fig. 5 & Supplementary Fig. S7). At trajectory entry, individuals in the stable high trajectory had the fewest years of education and the highest prevalence of obesity, whereas individuals in the high-to-low trajectory smoked the highest number of cigarettes/day and had the highest prevalence of physical inactivity, but also the lowest prevalence of obesity (Supplementary Table S8).

We identified three classes of LTPA trajectories: stable moderate (71.9%), moderate-to-vigorous-to-moderate (10.8%), and moderate-to-light-to-moderate (17.3%) (Fig. 5 & Supplementary Fig. S8). At trajectory entry, individuals in the moderate-to-vigorous-to-moderate trajectory had the most years of education and highest income, smoked the least and displayed the lowest prevalence of overweight and obesity (Supplementary Table S9). Two thirds of individuals in this trajectory were men.

We identified three classes of BMI trajectories: normal weight (44.8%), normal weight-to-overweight (44.8%), and overweight-to-obese (10.5%). All trajectories were increasing with age (Fig. 5 & Supplementary Fig. S9). At trajectory entry, individuals in the overweight-to-obese trajectory had fewer years of education, were least physically active and had the highest blood pressure and blood lipid levels (Supplementary Table S10).

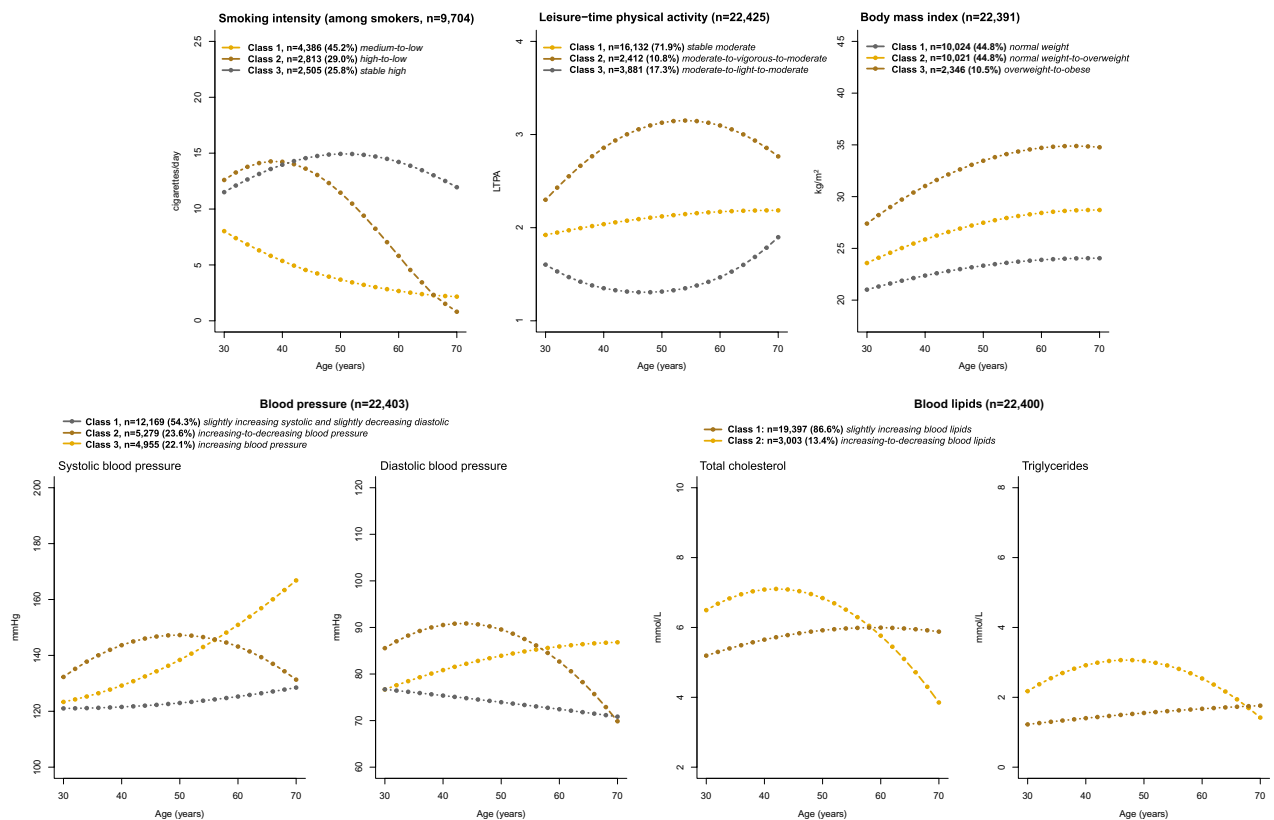


Figure 5. Estimated average trajectories of noncommunicable disease risk factors in each latent class ($n = 22,428$). For each risk factor and for each latent class, the estimated average trajectory of the risk factor considered among individuals allocated to that class.

We identified three classes of trajectories of combined blood pressure: slightly increasing systolic and slightly decreasing diastolic (54.3%), increasing-to-decreasing blood pressure (23.6%), and increasing blood pressure (22.1%) (Fig. 5 & Supplementary Fig. S10). Individuals in the increasing-to-decreasing blood pressure trajectory had higher blood pressure at trajectory entry compared to the other two, and were more likely to be men, have fewer years of education, and to have overweight or obesity. Individuals in the other two trajectories had similar characteristics, although individuals in the slightly increasing systolic and slightly decreasing diastolic trajectory were more likely to be women and have a BMI < 25 compared to the increasing blood pressure trajectory (Supplementary Table S11).

We identified two classes of trajectories of combined blood lipids: slightly increasing blood lipids (86.6%) and increasing-to-decreasing blood lipids (13.4%) (Fig. 5 & Supplementary Fig. S11). Individuals in the increasing-to-decreasing blood lipids trajectory had higher blood lipids and blood pressure at trajectory entry, were more likely to be men, have fewer years of education, smoke more, and to have overweight or obesity than individuals in the other trajectory (Supplementary Table S12).

Discussion

We harmonized data on NCD risk factors collected in population-based health studies conducted over a 45-year period to create a single study sample of 808,732 individuals, of which three-quarters had data on all the six key NCD risk factors *daily smoking*, *physical inactivity*, *obesity*, *hypertension*, *hypercholesterolaemia* and *hypertriglyceridaemia* measured at least once. In this sample, about two-thirds were exposed to at least one risk factor, with about one-fourth exposed to at least two risk factors. Daily smoking, physical inactivity, and hypercholesterolaemia were the most prevalent risk factors. We also observed a reduction in the proportion exposed to risk factors after 2004, mainly driven by a marked reduction in daily smoking, and despite a large increase in the prevalence of obesity. Among individuals exposed to at least two risk factors, we identified five distinct clusters of NCD risk factors (dyslipidaemia, inactive smokers, smokers with hypercholesterolaemia, obesity and hypertension), which were characterized by fewer years of education and lower income compared to individuals exposed to less than two risk factors. Lastly, from 30 to 70 years of age, we identified three distinct trajectories of smoking intensity, LTPA, BMI, blood pressure, and two distinct trajectories of blood lipids. Individuals in the trajectories tended to differ in their characteristics, especially across sex, education, and BMI.

Harmonization and pooling of data on individual NCD risk factors to create large datasets are not uncommon,^{60,61} and several of the health studies included in NCDNOR have previously been combined,⁶² but none

included all the data sources that we have used. To our knowledge, no study harmonized data on both behavioural and biological risk factors to identify clusters and trajectories of NCD risk factors.

The clustering of NCD risk factors is dependent on which NCD risk factors are considered, how they are measured, the cut-offs used to define risk, and the method used to identify clusters. We chose risk factors based on data-availability and the global strategies for the prevention of NCDs, but also cut-offs believed to represent high NCD risk (higher cut-offs than used e.g. by the WHO)⁷ and reduced risk of misclassification (e.g. due to ‘the white coat syndrome’). Using lower cut-offs causes higher proportions of individuals with clustered risk (e.g. changing cut-offs for hypertension and hypercholesterolemia to grade 1 and 6.2 mmol/L, respectively, the proportion of individuals with clustered risk increases from 24 to 60% [data not shown]).

We found five clusters of risk factors among individuals with at least two of the six risk factors considered. A systematic review on the clustering of behavioural risk factors has, as in the present study, inactive smokers as a common cluster, with prevalence ranging 7–20%.⁸ The only study to date using information on both behavioural and biological risk factors other than obesity, found hypertension, high cholesterol, obesity, and physical inactivity to cluster together.¹² However, they used different definitions for hypertension, hypercholesterolaemia, and obesity than the present study, yielding much higher overall prevalence of these risk factors. While LCA has been used in other studies, they differ greatly in the numbers and types of risk factors considered, making direct comparisons challenging. To our knowledge this is the largest study describing the clustering of NCD risk factors using information on behavioural and biological risk factors combined, and the first to do so in a Nordic country, making our results unique.

Three studies investigated long-term smoking trajectories extending to age 50 or beyond,^{63–65} two of which modelled smoking intensity and had a follow-up time similar to the present study.^{63,64} Although one study was conducted on women, and both were of much smaller scale, they identified trajectories similar to the ones we identified (medium-to-low, high-to-low, stable high).

In a recent systematic review of studies on physical activity trajectories, 3–5 trajectories were found most common.⁶⁶ Compared to our study, most studies were smaller and had shorter follow-up, but similar to our results, several studies showed that stable trajectories were the most prevalent^{66–69}.

We identified three different increasing BMI trajectories, which were similar to those found in other studies conducted in westernized countries of comparatively long follow-up.^{70–73} Other studies less comparable to ours have found as many as nine BMI trajectories⁷⁴.

Several studies have characterized blood pressure trajectories,⁷⁵ but few studies follow individuals over several decades.^{76–78} None modelled systolic and diastolic blood pressure trajectories jointly, and baseline ages were higher⁷⁶ and lower⁷⁷ than in the present study, but all found trajectories comparable to the trajectories we found.

Studies have characterized cholesterol and triglyceride trajectories from childhood to young adulthood⁷⁹ and middle-age,⁸⁰ and from middle-age to older age.²¹ The latter identified five total cholesterol and four triglyceride trajectories, and similar to our study, the majority of individuals had stable or slightly increasing total cholesterol and triglycerides, and ~ 10–15% had decreasing total cholesterol from age ~ 45.

In the studies that were comparable to ours and reported trajectory-specific participant characteristics, differences were consistently found between trajectories with respect to sex,^{63,66–71,76–78,80} education,^{63,66–70,77,78} and BMI.^{61,63,64,66–69,76,78} This aligns with our study, therefore contributing to the knowledge base on risk factor trajectories and on social inequalities in NCD risk.

The study has several strengths. This is the first large-scale population-based study describing both clustering and trajectories of behavioural and biological NCD risk factors combined. Its prospective design reduces recall bias, and the large sample size provides high statistical power. The study includes adults of all ages covering the entire country, making it uniquely suited to study key NCD risk factors in the Norwegian population. In addition, the data were collected over a 45-year period, with repeated measurements making it possible to investigate risk factor trajectories within individuals over a long time-period. Furthermore, NCD risk factor variables were harmonized between health studies using similar measurement protocols and questionnaires, which limits measurement errors due to harmonization. Lastly, compared to other methods, the LCA and LCMM modelling provide fit-statistics and posterior probabilities, allowing us to investigate model performance.

Nevertheless, we also highlight important limitations. Individuals participating in health studies often differ from the general population, which can result in both over and underestimation of exposure.^{81,82} Furthermore, despite high participation rates and the inclusion of studies designed to be at least regionally representative, the health studies were conducted at different time points spanning over five decades. In addition, almost 73% were between 39 and 44 years of age at study entry. This precludes us from generalizing our results to the general adult Norwegian population today. Selection bias may also be exacerbated further by the time window used in LCMMs, restricting the sample to those alive in their 60s⁸².

Furthermore, due to limited data-availability across studies, we did not include other important risk factors, such as unhealthy diet or hyperglycaemia. Therefore, it is likely that the true prevalence of having at least two NCD risk factors is higher than in our study sample. The clusters and shape of the trajectories may be affected by several factors such as changes in national regulations (e.g. the Norwegian Tobacco Control Act changed several times since 1975)⁸³ and medication use (e.g. blood pressure and lipid lowering drugs), data on which were not included in the present study. The study is also subject to measurement errors. Personal education and income were used as proxies for socioeconomic position, which is not ideal, especially for women in earlier cohorts,⁸⁴ blood samples were non-fasting,⁴⁷ and the behavioural risk factors assessed via self-reports, which are subject to well-known biases.⁸⁵ Anthropometrics were also self-reported in HUNT4ST and the 4th Norwegian Counties Study. Lastly, the validity of our results depends on the model specifications chosen. All models were carefully evaluated using several recommended fit-statistics,^{57,58} and performed satisfactorily. While the performance for LCMMs was good, it reflected the irregularity of the measurements. Nonetheless, this can be taken into account in further analyses⁵⁹.

In Norway, the unique national identification number assigned to each resident allows linkage of the individuals included in the health studies to several national mandatory health and administrative registries. Moving forward, this will allow us to investigate how NCD risk factor clusters and trajectories relate to NCD (multi-) morbidity and mortality, and to relate this to life-course socioeconomic circumstances. This will improve the identification of high-risk population groups and help inform future public health policies.

Conclusion

We described the health studies included in the NCDNOR project and harmonized data on key NCD risk factors. Our results suggest that a large proportion of the adult population in Norway have been exposed to multiple risk factors simultaneously. We identified five distinct clusters of NCD risk factors, where individuals exposed to two or more risk factors were characterized by fewer years of education and lower income compared to individuals exposed to fewer risk factors. Among individuals with repeated measurements, we identified 2–3 long-term trajectories for each risk factor that tended to differ the most across sex, education, and BMI. This study provides important insights into the mechanisms by which NCD risk factors can occur and may help the development of interventions aimed at reducing the NCD burden by targeting multiple risk factors simultaneously.

Data availability

The data that support the findings of this study are available from the Norwegian Health Data Authority and Statistics Norway, but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Data are however available from the authors upon reasonable request and with permission of the Norwegian Health Data Authority and Statistics Norway.

Received: 25 May 2023; Accepted: 29 August 2023

Published online: 02 September 2023

References

1. The World Health Organization. Premature deaths due to noncommunicable diseases (NCD) as a proportion of all NCD deaths. [https://www.who.int/data/gho/data/indicators/indicator-details/GHO/ncd-deaths-under-age-70-\(percent-of-all-ncd-deaths\)](https://www.who.int/data/gho/data/indicators/indicator-details/GHO/ncd-deaths-under-age-70-(percent-of-all-ncd-deaths)). Accessed 14 June 2022.
2. The World Health Organization. Noncommunicable diseases. <https://www.who.int/en/news-room/fact-sheets/detail/noncommunicable-diseases>. Accessed 14 June 2022.
3. Skou, S. T. *et al.* Multimorbidity. *Nat. Rev. Dis. Primers* **8**(1), 48 (2022).
4. Ghebreyesus, T. A. Acting on NCDs: Counting the cost. *The Lancet* **391**(10134), 1973–1974 (2018).
5. The World Health Organization. *Global Action Plan for the Prevention and Control of Noncommunicable Diseases 2013–2020* (World Health Organization, 2013).
6. The United Nations (UN). *Transforming our world: The 2030 Agenda for Sustainable Development*. Geneva: United Nations General Assembly, 2015.
7. The World Health Organization. Noncommunicable diseases: Risk factors. <https://www.who.int/data/gho/data/themes/topics/topic-details/GHO/ncd-risk-factors>. Accessed 30 Jan 2023.
8. Meader, N. *et al.* A systematic review on the clustering and co-occurrence of multiple risk behaviours. *BMC Public Health* **16**(1), 657 (2016).
9. McAloney, K., Graham, H., Law, C. & Platt, L. A scoping review of statistical approaches to the analysis of multiple health-related behaviours. *Prev Med* **56**(6), 365–371 (2013).
10. Sharma, S., Malarcher, A. M., Giles, W. H. & Myers, G. Racial, ethnic and socioeconomic disparities in the clustering of cardiovascular disease risk factors. *Ethn. Dis.* **14**(1), 43–48 (2004).
11. Kontari, P., Fife-Schaw, C. & Smith, K. Clustering of cardiometabolic risk factors and dementia incidence in older adults: A cross-country comparison in England, the United States, and China. *J Gerontol Ser A* **2**, 58 (2022).
12. Bu, F., Steptoe, A. & Fancourt, D. Relationship between loneliness, social isolation and modifiable risk factors for cardiovascular disease: A latent class analysis. *J. Epidemiol. Community Health* **75**(8), 749–754 (2021).
13. Mkuu, R. S. *et al.* Identifying individuals with multiple non-communicable disease risk factors in Kenya: A latent class analysis. *Public Health* **198**, 180–186 (2021).
14. Pelzom, D., Isaakidis, P., Oo, M. M., Gurung, M. S. & Yangchen, P. Alarming prevalence and clustering of modifiable noncommunicable disease risk factors among adults in Bhutan: A nationwide cross-sectional community survey. *BMC Public Health* **17**(1), 975 (2017).
15. Birch, J. *et al.* Clustering of behavioural risk factors for health in UK adults in 2016: A cross-sectional survey. *J. Public Health* **41**(3), e226–e236 (2018).
16. Uddin, R., Lee, E.-Y., Khan, S. R., Tremblay, M. S. & Khan, A. Clustering of lifestyle risk factors for non-communicable diseases in 304,779 adolescents from 89 countries: A global perspective. *Prev. Med.* **131**, 105955 (2020).
17. Perrier, F. *et al.* Life-course trajectories of physical activity and melanoma risk in a large cohort of Norwegian women. *Clin. Epidemiol.* **14**, 1571–1584 (2022).
18. Artaud, F. *et al.* Trajectories of unhealthy behaviors in midlife and risk of disability at older ages in the Whitehall II Cohort Study. *J. Gerontol. Ser. A* **71**(11), 1500–1506 (2016).
19. Hsu, H.-C., Luh, D.-L., Chang, W.-C. & Pan, L.-Y. Joint trajectories of multiple health-related behaviors among the elderly. *Int. J. Public Health* **58**(1), 109–120 (2013).
20. Li, F., Lin, Q., Li, M., Chen, L. & Li, Y. The association between blood pressure trajectories and risk of cardiovascular diseases among non-hypertensive Chinese population: A population-based cohort study. *Int. J. Environ. Res. Public Health* **18**(6), 2909 (2021).
21. Duncan, M. S., Vasani, R. S. & Xanthakis, V. Trajectories of blood lipid concentrations over the adult life course and risk of cardiovascular disease and all-cause mortality: Observations from the framingham study over 35 years. *J. Am. Heart Assoc.* **8**(11), e011433 (2019).
22. Marmot, M. & Bell, R. Social inequalities in health: A proper concern of epidemiology. *Ann. Epidemiol.* **26**(4), 238–240 (2016).
23. Bjartveit, K. & Tverdal, A. Health consequences of sustained smoking cessation. *Tob. Control* **18**(3), 197–205 (2009).
24. Meyer, H. E. & Vollset, T. A. Moderated Poster Session II: Prevention and health policy. *Eur. J. Cardiovasc. Prev. Rehabil.* **16**, S26–S28 (2009).
25. Holme, I., Tonstad, S. & Hjermmann, I. Risikofaktorer for hjerte- og karsykdom blant 40-åringer i Oslo 1981–99. *Tidsskrift for Den Norske Laegeforening* (2004).
26. Tverdal, A. & Selmer, R. M. 40-åringsundersøkelsene—400 000 menn og kvinner har møtt opp. *Tidsskrift for Den norske legeforening* (2002).

27. Åsvold, B. O. *et al.* Cohort Profile Update: The HUNT Study (International Journal of Epidemiology, 2022).
28. Hopstock, L. A. *et al.* The seventh survey of the Tromsø Study (Tromsø7) 2015–2016: Study design, data collection, attendance, and prevalence of risk factors and disease in a multipurpose population-based health survey. *Scand. J. Public Health* **52**, 14034948221092294 (2022).
29. Naess, O. *et al.* Cohort profile: Cohort of Norway (CONOR). *Int. J. Epidemiol.* **37**(3), 481–485 (2008).
30. Tverdal, A. & Bjartveit, K. Health consequences of reduced daily cigarette consumption. *Tob. Control* **15**, 472–480 (2007).
31. Jenum, A. K., Stensvold, I., Bjartveit, K., Thelle, D. S. & Hjerermann, I. Risikofaktorer for hjerte- og karsykdom i Oslo, Sogn og Fjordane og Finnmark. *Tidsskrift for Den Norske Lægeforening* 1998.
32. Norwegian Institute of Public Health Landsomfattende helseundersøkelser (LHU). <https://www.fhi.no/div/helseundersokelser/landsovfattende-helseundersokelser-lhu/>. Accessed 13 Jan 2023.
33. UiT The Arctic University of Norway. The Tromsø Study. <https://uit.no/research/tromsostudy>. Accessed 21 March 2023.
34. Norwegian Institute of Public Health Cohort of Norway (CONOR). <https://www.fhi.no/studier/cohort-of-norway/>. Accessed 16 Jan 2023.
35. Norwegian University of Science and Technology. HUNT Databank. <https://www.ntnu.edu/hunt/databank>. Accessed 21 March 2023.
36. The Directorate for e-Health. Helsedata. <https://helsedata.no/en/>. Accessed 21 March 2023.
37. Graff-Iversen, S., Anderssen, S. A., Holme, I. M., Jenum, A. K. & Raastad, T. Two short questionnaires on leisure-time physical activity compared with serum lipids, anthropometric measurements and aerobic power in a suburban population from Oslo, Norway. *Eur. J. Epidemiol.* **23**(3), 167–174 (2008).
38. Kurtze, N., Rangul, V., Hustvedt, B. E. & Flanders, W. D. Reliability and validity of self-reported physical activity in the Nord-Trøndelag Health Study: HUNT 1. *Scand. J. Public Health* **36**(1), 52–61 (2008).
39. Saltin, B. & Grimby, G. Physiological analysis of middle-aged and old former athletes. Comparison with still active athletes of the same ages. *Circulation* **38**(6), 1104–15 (1968).
40. Norwegian Institute of Public Health Heavy episodic drinking (Indicator 4). <https://www.fhi.no/en/nc/Indicators-for-NCD/alcohol/heavy-episodic-drinking-indicator-4/?term>. Accessed 01 Aug 2023.
41. Degerud, E. *et al.* Life course socioeconomic position, alcohol drinking patterns in midlife, and cardiovascular mortality: Analysis of Norwegian population-based health surveys. *PLoS Med* **15**(1), e1002476 (2018).
42. Jenum, A. K., Graff-Iversen, S., Selmer, R. & Sjøgaard, A.-J. Risikofaktorer for hjerte-og karsykdom og diabetes gjennom 30 år. *Tidsskrift for Den norske legeforening* 2007.
43. Lund-Larsen, P. G. Blood pressure measured with a sphygmomanometer and with Dinamap under field conditions—a comparison. *Norsk Epidemiologi* **7**(2), 235–241 (1997).
44. Williams, B. *et al.* 2018 ESC/ESH Guidelines for the management of arterial hypertension: The Task Force for the management of arterial hypertension of the European Society of Cardiology (ESC) and the European Society of Hypertension (ESH). *Eur. Heart J.* **39**(33), 3021–3104 (2018).
45. Klemsdal, T. O. *et al.* New guidelines for the prevention of cardiovascular disease. *Tidsskrift for den Norske lægeforening tidsskrift for praktisk medicin, ny raekke* 2017; **137**(16).
46. Foss, O. & Urdal, P. Cholesterol for more than 25 years: Could the results be compared throughout all this time. *Norsk Epidemiologi* **13**(1), 85–88 (2003).
47. Keirns, B. H., Sciarillo, C. M., Koemel, N. A. & Emerson, S. R. Fasting, non-fasting and postprandial triglycerides for screening cardiometabolic risk. *J. Nutr. Sci.* **10**, e75 (2021).
48. Mach, F. *et al.* 2019 ESC/EAS Guidelines for the management of dyslipidaemias: Lipid modification to reduce cardiovascular risk. *Eur. Heart J.* **41**(1), 111–188 (2020).
49. Norsk Epidemiologi. The 27th Norwegian Conference on Epidemiology. *Norsk Epidemiologi* **29**(1) (2021).
50. Statistisk sentralbyrå (Statistics Norway). Official statistics since 1876. <https://www.ssb.no/en> (accessed 17.03.2023).
51. The Norwegian Agency for Quality Assurance in Education (NOKUT). General information about education in Norway. <https://www.nokut.no/en/norwegian-education/general-information-about-education-in-norway/>. Accessed 17 Mar 2023.
52. Lex, A. & Gehlenborg, N. Sets and intersections. *Nat. Methods* **11**(8), 779 (2014).
53. Linzer, D. A. & Lewis, J. B. polCA: An R package for polytomous variable latent class analysis. *J. Stat. Softw.* **42**(10), 1–29 (2011).
54. Nguena Nguefack, H. L. *et al.* Trajectory modelling techniques useful to epidemiological research: A comparative narrative review of approaches. *Clin. Epidemiol.* **12**, 1205–1222 (2020).
55. Proust-Lima, C., Philipps, V. & Lique, B. Estimation of extended mixed models using latent classes and latent processes: The R package lmm. *J. Stat. Softw.* **78**(2), 1–56 (2017).
56. Proust-Lima, C., Philipps, V., Diakite, A., Lique, B. & Proust, M. C. Package 'lmm'. 2022.
57. van de Schoot, R., Sijbrandij, M., Winter, S. D., Depaoli, S. & Vermunt, J. K. The GRoLTS-checklist: Guidelines for reporting on latent trajectory studies. *Struct. Equ. Model.* **24**(3), 451–467 (2017).
58. Morgan, G. B., Hodge, K. J. & Baggett, A. R. Latent profile analysis with nonnormal mixtures: A Monte Carlo examination of model selection using fit indices. *Comput. Stat. Data Anal.* **93**, 146–161 (2016).
59. Lergenmuller, S. *et al.* Lifetime sunburn trajectories and associated risks of cutaneous melanoma and squamous cell carcinoma among a cohort of Norwegian women. *JAMA Dermatol* **158**(12), 1367–1377 (2022).
60. Ezzati, M. *et al.* Worldwide trends in body-mass index, underweight, overweight, and obesity from 1975 to 2016: A pooled analysis of 2416 population-based measurement studies in 128.9 million children, adolescents, and adults; 2017.
61. Zhou, B. *et al.* Worldwide trends in hypertension prevalence and progress in treatment and control from 1990 to 2019: A pooled analysis of 1201 population-representative studies with 104 million participants. *The Lancet* **398**(10304), 957–980 (2021).
62. Dalene, K. E. *et al.* Occupational physical activity and longevity in working men and women in Norway: A prospective cohort study. *The Lancet Public health* **6**(6), e386–e395 (2021).
63. Bahorik, A. L. *et al.* Early to midlife smoking trajectories and cognitive function in middle-aged US adults: The CARDIA study. *J. Gen. Intern. Med.* **37**(5), 1023–1030 (2022).
64. Brook, D. W., Rubenstein, E., Zhang, C. & Brook, J. S. Trajectories of cigarette smoking in adulthood predict insomnia among women in late mid-life. *Sleep Med.* **13**(9), 1130–1137 (2012).
65. Frosch, Z. A. K., Dierker, L. C., Rose, J. S. & Waldinger, R. J. Smoking trajectories, health, and mortality across the adult lifespan. *Addict. Behav.* **34**(8), 701–704 (2009).
66. Lounassalo, I. *et al.* Distinct trajectories of physical activity and related factors during the life course in the general population: A systematic review. *BMC Public Health* **19**(1), 271 (2019).
67. Barnett, T. A., Gauvin, L., Craig, C. L. & Katzmarzyk, P. T. Distinct trajectories of leisure time physical activity and predictors of trajectory class membership: A 22 year cohort study. *Int. J. Behav. Nutr. Phys. Act.* **5**(1), 57 (2008).
68. Laddu, D. R. *et al.* 25-year physical activity trajectories and development of subclinical coronary artery disease as measured by coronary artery calcium: The coronary artery risk development in young adults (CARDIA) study. *Mayo Clin. Proc.* **92**(11), 1660–1670 (2017).
69. Rovio, S. P. *et al.* Longitudinal physical activity trajectories from childhood to adulthood and their determinants: The Young Finns Study. *Scand. J. Med. Sci. Spr.* **28**(3), 1073–1083 (2018).
70. De Rubeis, V. *et al.* Trajectories of body mass index, from adolescence to older adulthood, and pancreatic cancer risk: a population-based case-control study in Ontario, Canada. *Cancer Causes Control* **30**(9), 955–966 (2019).

71. Wang, M. *et al.* Trajectories of body mass index from young adulthood to middle age among Canadian men and women. *Adv. Epidemiol.* **2015**, 121806 (2015).
72. Lavalette, C. *et al.* Body mass index trajectories and prostate cancer risk: Results from the EPICAP study. *Cancer Med.* **9**(17), 6421–6429 (2020).
73. Kelly, S. P. *et al.* Body mass index trajectories across adulthood and smoking in relation to prostate cancer risks: The NIH-AARP Diet and Health Study. *Int. J. Epidemiol.* **48**(2), 464–473 (2018).
74. De Rubeis, V. *et al.* Group-based trajectory modeling of body mass index and body size over the life course: A scoping review. *Obes. Sci. Pract.* **7**(1), 100–128 (2021).
75. Allen, N. B. & Khan, S. S. Blood pressure trajectories across the life course. *Am. J. Hypertens.* **34**(3), 234–241 (2021).
76. Portegies, M. L. *et al.* Mid- to Late-Life trajectories of blood pressure and the risk of stroke: The Rotterdam study. *Hypertension* **67**(6), 1126–1132 (2016).
77. Zhou, H. *et al.* Blood pressure trajectories in early adulthood and myocardial structure and function in later life. *ESC Heart Failure* **9**(2), 1258–1268 (2022).
78. Xu, Y., Möller, J., Wang, R. & Liang, Y. Life-course blood pressure trajectories and cardiovascular diseases: A population-based cohort study in China. *PLoS ONE* **15**(10), e0240804 (2020).
79. Elovainio, M. *et al.* Lipid trajectories as predictors of depressive symptoms: The Young Finns Study. *Health Psychol.* **29**, 237–245 (2010).
80. Yan, Y. *et al.* Associations between life-course lipid trajectories and subclinical atherosclerosis in midlife. *JAMA Netw. Open* **5**(10), e2234862 (2022).
81. Langhammer, A., Krokstad, S., Romundstad, P., Heggland, J. & Holmen, J. The HUNT Study: Participation is associated with survival and depends on socioeconomic status, diseases and symptoms. *BMC Med. Res. Methodol.* **12**, 143 (2012).
82. Hernán, M. A., Hernández-Díaz, S. & Robins, J. M. A structural approach to selection bias. *Epidemiology* **15**(5), 615–625 (2004).
83. Linnansaari, A. *et al.* Towards Tobacco-Free Generation: Implementation of preventive tobacco policies in the Nordic countries. *Scand. J. Public Health* **5**, 14034948221106868 (2022).
84. Galobardes, B., Shaw, M., Lawlor, D. A., Lynch, J. W. & Davey Smith, G. Indicators of socioeconomic position (part 1). *J. Epidemiol. Community Health* **60**(1), 7–12 (2006).
85. Althubaiti, A. Information bias in health research: Definition, pitfalls, and adjustment methods. *J. Multidiscip. Healthc* **9**, 211–217 (2016).

Acknowledgements

The NCDNOR project was established in 2020 and is a collaboration between the Norwegian Directorate of Health, the Norwegian University of Science and Technology, The Trøndelag Health Study, UiT The Arctic University of Norway, the Tromsø Study, the University of Bergen, Haukeland University Hospital, the Norwegian School of Sport Sciences, the Cancer Registry of Norway, and the Norwegian Institute of Public Health. The project is led and coordinated by the Norwegian Institute of Public Health and is funded by the Research Council of Norway (No 319705). We thank everyone that have contributed to establish NCDNOR, and Professor Majid Ezzati for his insightful contributions.

Author contributions

W.N., I.K.H.A., K.E.D., and S.L. conceptualised and designed the study. S.L. and K.E.D. verified, harmonized, pooled, and linked the data. SL analysed the data with help from K.E.D. K.E.D. and S.L. drafted the manuscript. All authors have participated in critical revisions of the manuscript for important intellectual content and have read, provided feedback, and approved the final manuscript. L.A.H. and E.S. contributed to the data collections in the Tromsø and HUNT studies, respectively. S.L. and K.E.D. take responsibility for the integrity of the data and the accuracy of the data analysis.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-023-41660-x>.

Correspondence and requests for materials should be addressed to K.E.D.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023