Additional file 1: Search strategies

Search strategy 1

((Health status monitoring [Title/Abstract] OR Surveillance [Title/Abstract]) AND Linked data [Title/Abstract]))

Search strategy 2

((Health status monitoring [Title/Abstract]) OR Surveillance [Title/Abstract]) AND Machine learning approach [Title/Abstract]))

Additional file 2: Definitions of data sources

Here we describe the definitions of different types of data sources, artificial intelligence techniques applied, health outcome, determinants and intervention indicators.

Different types of data sources:

1. ***Health surveys/Population health surveys*** collect information of risk factors, health behaviors and non-health care determinants of health[1]. Health surveys are used to measure the prevalence of risk factors and healthy behavior, monitor the effects of interventions, measure community attitudes to health policy initiatives, as well as assess trends in health and disease outcomes. Health surveys could involve health interview surveys or health examination surveys. These health surveys will include those surveys which are performed either at national or sub-national levels.
2. ***Disease-specific or population-based registries:*** A registry is a collection of information about individuals, usually focused around a specific diagnosis or condition[2]. Many registries collect information about people who have a specific disease or condition, while others seek participants of varying health status who may be willing to participate in research about a particular disease. Individuals provide information about themselves to these registries on a voluntary basis. Registries could be disease-specific registries, screening registries, immunization registries, etc.
3. ***National cohorts*** are performed to investigate the causes of development of major chronic diseases, i.e. cardiovascular diseases, cancer, diabetes, neurodegenerative/-psychiatric diseases, musculoskeletal diseases, respiratory and infectious diseases, and their pre-clinical stages or functional health impairments at the national level.
4. ***Clinical trials data*** include data on safety and efficacy of interventions. This data may be available through national or international trial registries. For example, ClinicalTrials.gov, Cochrane Library, WHO International Clinical Trials Registry Platform (ICTRP), European Union Clinical Trials Database, etc.
5. ***Administrative data sources*** were initially developed for administrative use, not for public health surveillance and have a larger coverage of population. For example, birth certificates, death certificates, census, biobank data (i.e., it is a biorepository that accepts, processes, stores and distributes bio specimens [i.e., blood, urine, spinal fluid, etc.] and associated data for use in research and clinical care[3]), GIS (Geographical Information System/GPS/Geodata), socioeconomic data, and retirement/pension data, etc.
6. ***Electronic health/patients/medical records (EHRs)*** include a summary of administrative data, clinical data of patients and determinants of health indicators (i.e., various types of exposures). In scientific literature, electronic health records (EHRs) are often refer to patients' record or electronic medical records. These data sources are flexible to link with different types of data sources.

   "EHRs are described as a repository of patient data in digital form[4]and include the following information: active and past diagnosis; past medical history; physical examinations; laboratory test orders and results; current prescriptions; radiological images and reports; hospitalization information; consultant reports; details of emergency care; immunizations; pathology reports; social history; lifestyle; allergies; genetic information; health screening study results; physicians, nurse, social worker, physical therapy notes at admission and discharge[5]."

EHRs may include the following data sources:

1. *Hospital Record (HR)/In hospital medical record* includes information about a patient generated during a period of hospitalization with written accounts of consultants' opinions as well as nurses' observations and treatments[6].
2. *Hospital Discharge Record is* a clinical report prepared by a physician or other health professional that summarizes the patient's chief complaint, the diagnostic findings, the therapy administered and the patients' response to it, and recommendations on discharge[7].
3. *General Practitioner (GPs)/Primary care* include information on diagnoses and symptoms, laboratory test results, referrals to specialists and drug or healthcare product prescribed[8].
4. *Specialist care* (i.e., cardiologists, neurologists, gynecologists, etc.) includes a highly skilled in a specific and restricted medical field[9].
5. *Emergency care (i.e., emergency room and outpatient emergency)* include information on evaluation and initial treatment of medical conditions caused by trauma or sudden illness[10].
6. *Health insurance claim (i.e., healthcare reimbursement)* is a detailed invoice that a health care provider (such as doctor, clinic, or hospital) sends to the health insurer to reimburse the expenses spent on health services (i.e., drugs, diagnostic/laboratory tests, etc.) received by a patient[1].
7. *Drug prescription* include information on prescription date, type of drug, strength, dosage regimen, quantity, and route of administration[11].
8. *Genomic/DNA data* sources (i.e., information about functions of specific genes and to assess the association of gene mutations in certain diseases [2] such as for breast cancer BRCA 1/2).

7. *X-data sources)*: These type of data sources provide precise information on determinants of health and can include data on various exposures such as biological parameters, social behavior, life style, physical environment, nutrition, etc. These sources are considered as part of big data (i.e., voluminous amount of structured, semi structured and unstructured data that has the potential to be mined for information [12]).

We grouped these types of data sources as "X-data sources". Some of these data sources are enlisted below and more could be possible:

1. *m-Health (mobile-Health)* is the use of mobile phones, wireless health apps and wearable devices and measure a set of biological parameters which could be used for disease surveillance and health care services[13]. Data from these apps can also be used for disease surveillance, treatment support, epidemic outbreak tracking and chronic disease management[13].
2. *Social media* is an electronic communication through which users create online communities to share information, ideas, personal messages, and other content (such as videos) and the related data may reflect user's social behavior towards different aspects[14].
3. *Mobility mode data* (i.e., commuting for work by walking, using bicycle, public transport such as metro, train, bus, etc.) provide information about the mobility mode of general population whether walking or using different means of transport.
4. *Build physical environment* include data regarding green spaces, environmental exposure in terms of cleanliness, sound pollution, air quality, etc.
5. *Nutrition:* direct producers of seasonal/local fruits and vegetables, etc.
6. *Housing infrastructure* data source provide information about housing space and location.

Artificial intelligence (AI) techniques

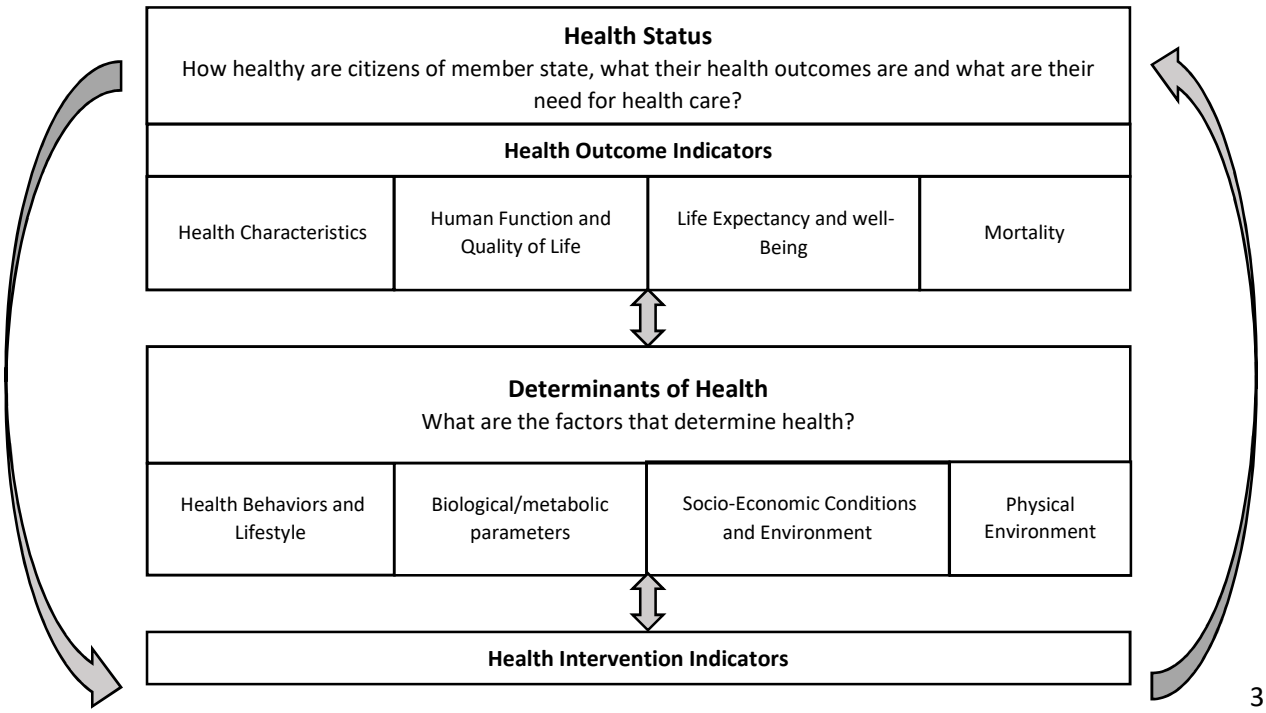
The artificial intelligence techniques following techniques: machine learning, natural language processing, markov decision process, support vector machine, data mining, regression, etc., to analyze, estimate and predict the health indicators either from linked data or using an individual data set.

Artificial Intelligence					
Machine learning	Natural language processing	Markov decision process	Support vector machine	Data mining	Others

Health outcome and intervention related indicators and determinants of health

We have selected the following non-communicable diseases based on burden of disease with higher incidence of mortality and morbidity across EU member states[15]: cardiovascular diseases, cancer, chronic obstructive pulmonary diseases, diabetes, neurodegenerative disease, mental health, accidents/trauma, maternal and perinatal health and any other disease.
Health outcomes indicators which are estimated from linked data and/or by using AI techniques to an individual data set, will be identified.

We have adopted EuroREACH Framework describing health status monitoring to classify the identified health outcome indicators, non-healthcare determinants of health and health intervention indicators under different categories[11] *(see figure 2).*



Prevention indicator	Promotion indicator	Others

Figure 2: EuroREACH Framework for health status monitoring

Health outcome indicators which are estimated either from linked data or by applying AI techniques to an individual data set, describe the health status of a population in terms of health characteristics (e.g., prevalence of stroke among ≥ 65 years old), human function and quality of life (e.g., quality of life after stroke), life expectancy and well-being (e.g., survival of people with stroke) and mortality (e.g., causes of mortality).
These outcome indicators have the potential to improve health surveillance with more precise information. These health outcome indicators include prevalence, incidence, population attributable risk, population attributable fraction, relative risk, hazard ratio, etc.

Determinants of health which are identified either from linked data or from an individual data set, can provide more variables/parameters to better understand exposures factors related to health behavior and lifestyle (i.e., risk and/or protective behavior, response to health problems, etc.), biological/metabolic parameters (i.e., genetic, body structure and functioning, etc.), socio-economic conditions and environment (i.e., attitudes, social networks, education, employment, living standard, etc.) and physical environment (i.e., water quality, air quality, food safety, etc.). For example, use of public transport to commute for work.
These determinants should be different from already compiled databases of OECD, WHO or Eurostat.

Health intervention indicators which are estimated either from linked data or by applying AI techniques to an individual data set, describe the effect of the interventions applied in terms of prevention (e.g., the use of genetic screening of BRCA1/2 genes among families with a family history of breast cancer) or promotion (i.e., integrated health programs at work place, schools, hospitals, policies and practices on healthy lifestyle, etc.).

References:

1. Madans JH. Health Surveys. In: Smelser NJ, Baltes PB, eds. *International Encyclopedia of the Social & Behavioral Sciences*. Oxford: Pergamon; 2001:6619-6627.
2. Health NIo. What is a registry? https://www.nih.gov/health-information/nih-clinical-research-trials-you/list-registries. 2018.
3. De Souza YG, Greenspan JS. Biobanking Past, Present and Future: Responsibilities and Benefits. *AIDS (London, England).* 2013;27(3):303-312.
4. 20514 IT. Health Informatics-Electronic Health Record-Definition, Scope, and Context: https://www.iso.org/obp/ui/#iso:std:iso:tr:20514:ed-1:v1:en. 2005.
5. Häyrinen K, Saranto K, Nykänen P. Definition, structure, content, use and impacts of electronic health records: A review of the research literature. *International Journal of Medical Informatics.* 2008;77(5):291-304.
6. Katzenellenbogen JM BS, Somerford P, Anderson CS, Semmens JB, Codde JP, Vos T. Disability burden due to stroke in Western Australia: new insights from linked. *Int J Stroke.* 2010;5(4):269-277.
7. White C GM, Johnson B, Corbin T. Social inequalities in adult male mortality by the National Statistics. *Health Stat Q.* 2007;36:6-23.
8. Manuel DG SS. Health-related quality of life and health-adjusted life expectancy of people with. *Diabetes Care.* 2004;27(2):407-414.
9. Tuppin P, Rudant J, Constantinou P, et al. Value of a national administrative database to guide public decisions: From the systeme national d'information interregimes de l'Assurance Maladie (SNIIRAM) to the systeme national des donnees de sante (SNDS) in France. (0398-7620 (Print)).
10. Dunnell K BJ, Wood R, Babb P. Measuring aspects of women's life and work for the study of variations in health. *Am J Ind Med.* 1999;36(1):25-33.
11. EuroREACH. EuroREACH Framework: http://hdn.euhs-i.eu/performance/frameworks/euroreach-framework. 2013.
12. Big data: https://searchdatamanagement.techtarget.com/definition/big-data.
13. Lloyd K, McGregor J, John A, et al. A national population-based e-cohort of people with psychosis (PsyCymru) linking prospectively ascertained phenotypically rich and genetic data to routinely collected records: Overview, recruitment and linkage. *Schizophrenia Research.* 2015;166(1):131-136.
14. Atramont A, Bonnet-Zamponi D, Bourdel-Marchasson I, Tangre I, Fagot-Campagna A, Tuppin P. Health status and drug use 1 year before and 1 year after skilled nursing home admission during the first quarter of 2013 in France: a study based on the French National Health Insurance Information System. *European Journal of Clinical Pharmacology.* 2018;74(1):109-118.
15. WHO. Global status report on noncommunicable diseases: http://www.who.int/nmh/publications/ncd_report_full_en.pdf. 2010.

Additional file 3: Abbreviation of European Countries with contact detail of survey respondents

S/No	Abbreviations	Country	First name	Last name	Institute	Email
1	AT	Austria	Stefan	Mathis-Edenhofer	Public Health Institute	stefan.mathis-edenhofer@goeg.at
2	BE	Belgium	Herman	Van Oyen	Sciensano	herman.vanoyen@sciensano.be
3	BG	Bulgaria	Raina	Nikolova	Public Health Institute	r.nikolova@ncpha.government.bg
4	HR	Croatia	Ivan	Pristas	CIPH (HZJZ)	ivan.pristas@hzjz.hr
5	CY	Cyprus	Vasilios	Scoutellas	Public Health Institute	VScoutellas@mphs.moh.gov.cy
6	CZ	Czech Republic	Sarka	Dankova	UZIS	sarka.dankova@uzis.cz
7	CZ	Czech Republic	Jiri	Jarkovsky	Institute of Health Information and Statistics	Jiri.Jarkovsky@uzis.cz
8	DK	Denmark	Mette Bjerrum	Koch	Danish Health Data Authority	mebk@sundhedsdata.dk
9	DK	Denmark	Maja Bæksgaard	Jørgensen	Statens Institut for Folkesundhed	mbha@sdu.dk
10	EE	Estonia	Eleri	Lapp	MoSA	eleri.lapp@sm.ee
11	FI	Finland	Mika	Gissler	THL	mika.gissler@thl.fi
12	DE	Germany	Angelika	Schaffrath Rosario	RKI	SchaffrathRosarioA@rki.de
13	EL	Greece	Spyridon	Goulas	Governmental Health Insurance Organization	sgoulas@eopyy.gov.gr
14	FR	France	Anne	Gallay	Public Health Institute	Anne.GALLAY@santepubliquefrance.fr
15	FR	France	Jennifer	Zeitlin	INSERM U 1153 – Research Institute	jennifer.zeitlin@inserm.fr
16	IE	Ireland	Sheona	Gilsenan	DOH	sheona_gilsenan@health.gov.ie
17	IT	Italy	Luigi	Palmieri	ISS	luigi.palmieri@iss.it
18	IT	Italy	Brigid	Unim	ISS	brigid.unim@iss.it
19	LV	Latvia	Janis	Misins	CDPC	janis.misins@spkc.gov.lv
20	LT	Lithuania	Rita	Gaidelyte	HI	rita.gaidelyte@hi.lt
21	LU	Luxembourg	Anne-Charlotte	Lorcy	Health-Directorate	anne-charlotte.lorcy@ms.etat.lu
22	MT	Malta	Neville	Calleja	MFH	neville.calleja@gov.mt
23	NL	Netherlands	Peter	Achterberg	RIVM	peter.achterberg@rivm.nl
24	NO	Norway	Hakon	Haaheim	National Health Directorate	hakon.haaheim@helsedir.no
25	PL	Poland	Jakub	Adamski	MoH	j.adamski@mz.gov.pl
26	PL	Poland	Piotr	Nowosielski	MoH	p.nowosielski@mz.gov.pl
27	PT	Portugal	Paulo	Nogueira	DGS	paulo.nogueira@dgs.min-saude.pt
29	PT	Portugal	Martins	José	Public Health Institute	josemartins@dgs.min-saude.pt
30	PT	Portugal	Carlos	Dias	Public Health Institute	carlos.dias@insa.min-saude.pt
31	RO	Romania	Silviu	Radulescu	Public Health Institute	silviu.radulescu@insp.gov.ro
32	RS	Serbia	Maja	Krstic	Batut	maja_krstic@batut.org.rs

33	SK	Slovakia	Jan	Cap	National Health Information Centre	Jan.Cap@nczisk.sk
34	SI	Slovenia	Metka	Zaletel	NIJZ	metka.zaletel@nijz.si
35	ES	Spain	Beatriz	Perez-Gomez	ISCIII	bperez@isciii.es
36	SE	Sweden	Hanna	Lobosco	MoH	hanna.lobosco@folkhalsomyndigheten.se
	SE	Sweden	Rosita	Wigand	Public Health Agency of Sweden	rosita.wigand@folkhalsomyndigheten.se
	SE	Sweden	Jenny	Borlin	Public Health Agency of Sweden	jenny.borlin@folkhalsomyndigheten.se
37	UK-ENG	England	Robert	Aldrige	Welsh Government	r.aldridge@ucl.ac.uk
38	UK-SC	Scotland	Ian	Grant	National Statistics Department	ian.grant@nhs.net
39	UK-WL	Wales	Ronan	Lyons	University of Swansa	r.a.lyons@swansea.ac.uk

Additional file 4: Data sources used for linkage across European countries in 2019

S/No	Data sources used for linkage		European countries		
			Advanced N = 24		In progress N = 2
1	*Health-related administrative data sources (i.e., Electronic Health Records)* Ʊ	Primary care visits, emergency care, referral records, hospital discharge, prescribed medications, health insurance claims, diagnostics procedures, laboratory tests, biobank	22	AT, BE, CY, CZ, DE, DK, EE, ES, FI, FR, HR, IT, LT, MT, NL, NO, PT, SI, SK, SRB, SW, UK[ENG, SC, WL]	LV
2	*Non-health related administrative data sources* ǂ	Birth and mortality database, education level, income tax, GIS, occupation, housing conditions, criminal statistics, land and housing, socioeconomic, census (demographic), house of handicap persons, environmental, road and transport, air pollution, UV light exposure	22	BE, CY, CZ, DE, DK, EE, ES, FI, FR, HR, IT, LT, MT, NL, NO, PL, PT, SI, SK, SRB, SW, UK[ENG, SC, WL]	IE, LV
3	*Disease-specific registries*	Cancer, diabetes, cardiovascular, congenital malformation, tuberculosis, HIV/AIDS, inflammatory bowel disease, renal, reproductive health, dementia, organ transplantation, traffic accidents/trauma or injury, hospital registry of domestic and leisure accidents	22	BE, BG, CY, CZ, DE, DK, EE, ES, FI, FR, HR, IE, LV, MT, NL, NO, PL, PT, SK, SRB, SW, UK [ENG, SC, WL]	LV
4	*National health surveys**	National health examination and interview surveys	15	BE, CZ, DK, DE, EE, ES, FI, FR, IT, NL, NO, PT, SI, SW, UK [ENG, SC, WL]	PL
5	*Population-based epidemiological cohort/National cohorts*	DANCOS, IDEFICS, CONSTANCE, ELFE, Growing up in Scotland, HealthWise Wales cohort, Millennium cohort, Caerphilly cohort study	7	DK, EE, FI, FR, NO, PL, UK [ENG, SC, WL]	
6	*Clinical trials data*	FINGER, PRISMATIC	3	DK, FI, UK [ENG, WL]	

Ʊ Latvia is developing data linkage techniques to link EHRs with other data sources.

ǂ *In Ireland, income database is linked with EHRs of prescribing medicine at small level. Latvia is developing data linkage techniques to link birth and mortality databases either with EHRs or with disease-specific registries.*

** Poland is planning to link this national health survey data with other health data sources in near future. In Ireland, this is done for specific surveys such as housing and health conditions at small scale.*

Additional file 5: Examples of different combinations of data linkages across European countries in 2019

S/No	European countries	Different combinations of data linkages (N = ~ 85)
1	Austria	Hospital discharge with outpatient visit (primary care visit)
2	Belgium	Hospital discharge with health insurance claim
		Educational attainment with mortality database
		Census with mortality database
		Health interview survey with mortality (cause-specific mortality, StatBEL)/use of care (IMA)/prescribed medication (INAMI)
		Disease-specific registries with mortality database
3	Bulgaria	Registry of rare diseases with oncology registry
4	Croatia	Primary health care visits with hospital discharge/health insurance claim/mortality (cause-specific mortality)
		Cancer registry with geospatial registry
5	Cyprus	Hospital discharge with mortality database
		Cancer, diabetes, HIV/AIDS registries with mortality database
6	Czech Republic	Health insurance claims with mortality database
		Registry of hospitalizations is linked with mortality database
		Disease specific registries: Cancer, cardiovascular surgery and intervention, reproductive health, TBC registry, registry of injuries with mortality database/health insurance claims/registry of hospitalizations
7	Denmark	National patient health register is linked with education, income, housing, transfer payments, socioeconomic status, criminal statistics, etc.
8	Estonia	Hospital stay, primary and special ambulatory care linked with health insurance claims
		Health insurance claims and prescriptions are linked with causes-specific mortality
		Birth register linked with causes-specific mortality
		Cancer, tuberculosis and myocardial infarction linked with causes-specific mortality
		Chernobyl Cleanup workers (cancer, causes of death) families and children data with birth register (on irregular basis)
	(In progress)	Genomic database linked with EHRs
	(In progress)	Estonia health insurance database linked with prescription and diagnostics procedures
9	Finland	KANTA and KANSA health register linked with
		Finis birth cohort 87 and 97 linked with
		National HES from 1972-2017 linked with
		National HIS since 1978 onwards linked with
10	France	Hospital discharge linked with health insurance claims and mortality database (national health database: SNDS)
		Population-based epidemiology cohorts (CONSTANCES & ELFE) linked with national health database
		Cancer, congenital malformation, cardiovascular, inflammatory bowel disease and traffic accidents registries linked with national health database
		National health surveys (Esteban examination/interview) linked with national health database
	(In progress)	UV light and air pollution exposure linked with national health database
	(In progress)	House of handicap person s' health and social assistance linked with national health database
11	Germany	National health examination survey in adults linked with mortality database
		National health examination survey in adults linked with health insurance claims
		Cancer registry operated by the public health institute and included in health reporting
		National health surveys use national and sub-national data for weighting
		National health examination surveys use inter-metropolitan socioeconomic data for improvement of field work (in progress)
		Use of socioeconomic data at the metropolitan level for small area estimation (in progress)
		Use of real-time emergency room data for surveillance of infectious diseases (in progress in a local project)
		Linkage of data from national health surveys, health insurance data, cancer registry and other data sources for national burden-of-disease calculation (in progress)
12	Greece	No
13	Ireland (in progress)	Cancer registry linked with Hospital admission linked and mortality database
		Census data linked with mortality database (one off)
		Prescribed medication data Medical eligibility and claims data linked with income level (one off)
14	Italy	Hospital discharge linked with mortality database and national health examination survey
15	Latvia (in progress)	Hospital discharge, primary health care, emergency care records linked with birth and mortality database
		Patient register with specific diseases linked with mortality database
1.6	Lithuania	Compulsory health insurance information system (inpatient, outpatient specialized, primary care, emergency care) linked with causes-specific mortality database
17	Luxembourg	No

18	Malta	Health insurance claims, prescribed drugs, surgical operations, laboratory information system, radiology information system, patient administration system, outpatients attendance, patient discharge summaries linked with birth and mortality database
		Congenital anomalies, injuries, cancer, dementia, organ transplants registries linked with mortality database
19	The Netherlands	Health examination and interview surveys linked with mortality database
		Health insurance claims with perinatal data
		Cancer registry data with mortality database
20	Norway	Linkage between almost all sources by means of unique personal identification. Both within health and care services, and across other governmental areas. Big data solution in use for accessibility modulation using national health registries linked with land and housing, road and transport, and GIS databases.
21	Poland	Cancer and tuberculosis registry databases linked with mortality, demographic and GIS databases
	(in progress)	National health surveys linked with electronic health records
22	Portugal	Hospital discharge, primary care and medical records linked with hospital registry of domestic and leisure accidents, e-death certification
		Cancer, tuberculosis, HIV and congenital anomalies registries linked with e-death certification and hospital discharge data
23	Romania	No
24	Slovakia	National registry of EHRs (Hospital discharge, general practitioner record, referrals, prescribed medications, laboratory results, diagnostic procedures medical consultations) linked with national disease-specific registries
		National registry of EHRs linked with national registry of health care workers and heath care providers
25	Slovenia	Hospital discharge, drug prescription and perinatal health linked with mortality database
		Hospital discharge, drug prescription and perinatal health linked with census data on education and socioeconomic variables (inequality analysis)
		Hospital discharge, drug prescription and perinatal health linked with European Health Interview Survey
26	Serbia	Hospital discharge linked with cancer registry
		Mortality database linked with cancer registry
27	Spain	National health interview survey linked with mortality database
		Primary care data linked with drugs prescription and laboratory tests
		National Health survey linked with cause-specific mortality data
		All cohort studies can link their data with cause-specific mortality information through an agreement with the National Institute of statistics
28	Sweden	National Patients register linked with causes-specific mortality database
		National Patients register linked with birth database
		National Patients register linked with dental health database
		National Patients register linked with vaccination database
		National Patients register linked with education, income tax, occupation, country of origin and population based register
		National health surveys (ULF/SILC), environmental health survey (MHE) and European health interview survey (EHIS) linked national health care quality registries (each deals with a disease-specific condition)
29	UK-England	UK Cancer Registry is linked with Hospital and Mortality Records
	UK-Scotland	EHRs linked with each other: General and Psychiatric Hospital Stays/ day cases (including intensive care/high dependency stays), Outpatient attendances, Emergency department attendances, Maternity, birth records and Neonatal Care
		EHRs linked with mortality database and census (demographic) database
		Cancer and diabetes registries linked with hospital and mortality records
		Scottish health interview survey linked with hospital and mortality records
	UK-Wales	EHRs linked with each other: Primary care general practice datasets linked with hospital inpatient, emergency department visits, outpatient attendances, child health dataset, congenital anomalies, maternity records, population register and laboratory results in the Secure Anonymized Information Linkage (SAIL) database www.saildatabank.com
		EHRs linked with mortality database, GIS and census (demographic) database
		Cancer, trauma and renal registries linked with all of the above
		Welsh Health Survey and National Survey for Wales (interviews) linked with all of the above
		Healthwise Wales Cohort, Millennium Cohort, Caerphilly Cohort study and UK-Biobank linked to SAIL and all of the above
		EHRs linked Education Attainment records and GIS derived metrics e.g. pollution, housing quality, urban design, alcohol outlets etc.

Additional file 6: Description of health outcome indicators estimated using linked data across European countries in 2019

S/No	Categories	Health outcomes indicators (N = 46)	European countries
1	**Cardiovascular (N = 14)**		
	Health characteristics	Incidence of stroke among less than and more than 65 years old	CZ, FR, LT, NL, NO, SW, UK-WL
		Prevalence of stroke among less than and more than 65 years old	CZ, LT, NL, NO, SW, UK-WL
		Incidence of myocardial infarction among less than and more than 65 years old	CZ, FR, LT, NL, NO, SW, UK-WL
		Prevalence of myocardial infarction among less than and more than 65 years old	CZ, LT, NL, NO, SW, UK-WL
	Mortality	Mortality due to stroke within 30-days of hospitalization	FR, IT, LV, MT, NO, SW
		Mortality due to myocardial infarction within 30-days of hospitalization	CZ, FR, NO, SW
		Risk of mortality due to myocardial infarction at municipal level	SW
		Risk of mortality due to myocardial infarction at municipal level (in future)	ES
	Human function and quality of life	Neurorehabilitation and functional outcomes of patients after stroke	FR, SW
		Neurorehabilitation and functional outcomes of patients after stroke (in progress)	NO
	Life expectancy and well-being	30- days survival following stroke	FR, NO, SW, UK-WL
		90-days survival following stroke	SW, UK-WL
		365-days survival following critical care for stroke	
		Hospital utilization in the 365 days following intensive care discharge	
		365-days survival following stroke	FR
	All of above categories	DALYs, YLL, YLD	SW, UK-SC
2	**Neurodegenerative disease (N = 6)**		
	Health characteristics	Prevalence of Multiple Sclerosis	CZ, HR, FR, LT, NO, SW, UK-WL
		Prevalence of Alzheimer	CZ, FR, LT, SW, UK-WL
		Incidence of Alzheimer	FR, LT, SW, UK-WL
		Incidence and prevalence of Alzheimer (in progress)	NO
		Prior event rate ratio to estimate the influence of exposure to antipsychotic medication on acute cardiac events and hip fracture due to dementia	SW, UK-WL
	Mortality	Mortality due to Parkinson	FR, SW
		Mortality due to Dementia	FR
3	**Maternal and perinatal health/child health (N = 6)**		
	Health characteristics	Incidence of low birth weight	CY, CZ, PL, FR, LT, NO, SW
		Incidence of pre-term birth rate	SW
		Incidence of pre-term birth rate (in future)	CZ, FR, NO, UK-WL
		Incidence of gestational diabetes	SW
		Incidence of gestational diabetes (in future)	FR, NO
		Prevalence of congenital anomalies	CZ, PT, LT, NO, SW, UK-WL
		Emergency admissions for potentially preventable	UK-WL

		hospitalizations (PPH) between the age of 1 and 5 years	
	Mortality	Stillbirth	FR, NL, LT, NO
4	**Diabetes (N = 6)**		
	Health characteristics	Incidence of diabetes	CZ, FR, LT, SW, UK-WL
		Prevalence of diabetes	
		Incidence and prevalence of diabetes (in progress)	NO
	Mortality	Mortality due to diabetes and related risk factors	BG, CZ, DE, FR, SW
	Human function and quality of life	Amputation rate (related complications)	BE, FR, MT, LT, SW
	Human function and quality of life	Number of patients with installed insulin pump during diabetes curation	FR, PL, SW
	All of above categories	DALYs, YLL, YLD	SW, UK (SC, WL)
5	**Suicide/Trauma/Injury (N = 7)**		
	Health characteristics	Use of health care services before suicide	LT, SW, UK-WL
		Prevalence of morbid conditions before suicide	
		Standardized prevalence of suicide	CZ, PL, FI, LT, NO, SW
		Incidence of suicide	
		Risk of road accident among users of prescribed medicines	FR
		Incidence of injuries	LT, NO, SW, UK-WL
	Mortality	Death rates in road accidents	EE, LT, NO, SW
6	**Cancer (N = 6)**		
	Health characteristics	Incidence of various types of cancer	CY, CZ, DE, FI, FR, PL, NL, NO, SW, UK-WL
		Prevalence of various types of cancer	
		Incidence rate by stage for colorectal cancer	CZ, MT, NO, SW, UK-WL
	Mortality	Mortality rates due to various types of cancer	CY, CZ, ES, FI, IT, PT, LT, NO, SW, UK-WL
	Life expectancy and well-being	5-years relative survival rates	CY, CZ, DE, EE, FI, IT, PT, NL, NO, SK, SW
	Human function and quality of life	Scale of return to work after cancer and determining factors	SW
		Scale of return to work after cancer and determining factors (in future)	BE
7	**Alcoholic liver disease and hepatic failure (N = 1)**		
	Mortality	Standardized mortality ratios at 60 days and 5-years following unscheduled admissions	SW, UK-WL

Additional file 7: Description of health determinants identified using linked data across European countries in 2019

S/No	Domain/Health condition	Health determinants (N = 34)	Variables can be stratified by	European countries
1	**Physical environment (N = 12)**			
	Emphysema	Air quality	Area of residence	BE, UK-WL
	Injury	Place of injury	Age, sex and area of residence	CY, NO, UK-WL
		Type of injury		
		After injury hospitalized or not		
	Parkinson	Exposure to pesticides (i.e., agricultural activities, in vineyards, metallurgy and solvents, in textile industry)	Area of residence	FR
	Breast cancer mortality	Industrial pollution	Area of residence	ES
	Adiposity	Proximity of fast food outlets from areas of residence	Area of residence	UK (ENG, WL)
	Various chronic health conditions	Alcohol outlet density	Area of residence	NO, SW, UK (SC, WL)
	Mental health	Presence to green-blue spaces		
		Access and visit to green-blue spaces		
		Visit to green-blue spaces		
		Housing quality		
2	**Socioeconomic and environment (N = 10)**			
	Multi-morbidity	Number of single households of older people	Age and living condition	AT, NO, SW
	Breast cancer mortality, injury, diabetes, cardiovascular, mental health	Sociodemographic status	Age, sex and area of residence	EE, ES, FR, NL, NO, PT, SI, SW, UK (SC, WL)
		Socioeconomic status		FR, NO
		Employment status		FR
		Level of education achieved		FR, NO
		Deprivation index		FR
	Pre-term birth	Maternal education to measure social disparities		FR, NO
	Injury	Time and distance between road accident and emergency room		PL
		Standardized absenteeism and attributable indirect costs		
	All types of cancer	Accessibility to linear accelerators for radiotherapy	Area of residence	
3	**Health behavior and life style (N = 6)**			
	Stroke, myocardial infarction, lung cancer, mental health, obesity, other chronic condition	Smoking rate	Age, sex, socioeconomic status and area of residence	BE, CY, FR, IT, MT, NL, NO, UK (SC, WL)
		Alcohol consumption	Age and sex	NO, UK-SC
		Physical activity		UK-SC
		Dietary consumption		UK-SC
		Drug use		NO, UK-SC
	Diabetes	Diabetes risk score	Age, sex and area of residence	CY, NO, SL
4	**Biological/metabolic parameters (N = 3)**			
	Obesity	Self-reported BMI	Age, sex and area of residence (i.e., in county, municipality)	SW, NO, UK-SC
	Diabetes	Blindness	Age, sex and area	CY, FR

		Proteinuria	of residence	
5	**Others (N=3)**			
	Road accidents, neurodegenerative disease	Multi-morbidity	Age and sex	FR
	Chronic health conditions	Disability		FR
		Frailty		FR

Additional file 8: Description of health intervention indicators estimated using linked data across European countries in 2019

S/No	Categories	Domain/Health condition	Health intervention indicators (N = 23)	Member States
1	**Maternal and perinatal health (N = 7)**			
	Prevention	Low birth weight	Prevalence of thyroid gland examination during pregnancy	CZ
	Prevention		Frequency of admission to intensive care unit	CY
	Prevention		Prevalence of maternal smoking and quitting smoking during the pregnancy	FI, FR, NO, SW
	Prevention	Pre-term birth	Percent of births in level III maternity units	FR, NO, SW
	Prevention	Perinatal mortality	Pregnant women with adequate prenatal care (number of visit/timing of initiation)	FR, NO, SW
	Prevention	Prenatal care	Screening programs as preventive check-ups during pregnancy	NO, SI, SW
	Prevention	Neural tube defect	Folic acid supplementation	PT, NO, UK-WL
2	**Cancer (N = 6)**			
	Prevention	Breast, cervical, colorectal and bowl cancer	Screening participation rates, effectiveness and evaluation	BE, EE, CZ, ES, FI, HR, IT, NO, SI, SW, UK-WL
	Prevention	Colorectal cancer	Frequency of surgery	CY, NO, SW
	Prevention	Colorectal cancer	Colonoscopy compliance rate	CZ, SI, NO
	Prevention	Breast cancer	Genetic screening among families (in future)	ES
	Prevention	Breast cancer	Stage distribution of detected cancer	CZ, SI, NO
	Others	All types of cancers	Re-integration in work	BE
3	**Diabetes (N = 4)**			
	Prevention	Diabetes related complications	Foot care	BE, FR, SW
	Prevention		Proportion of diabetics counselled by nurse to avoid complications	SW
	Prevention		Proportion of diabetics counselled by nurse to avoid complications (in future)	CY
	Prevention		Amputation rate	BE, FR, SW
	Prevention	Diabetes	Percentage of diabetics with latest HbA1c above 7.0	FR, MT, SW
4	**Cardiovascular diseases (N =2)**			
	Prevention	Stroke, myocardial infarction	Absolute global CVD risk assessment in primary prevention	IT
	Prevention	Stroke	Aortic aneurysm screening	SW, UK-WL
5	**Neurodegenerative disease (N = 2)**			
	Prevention	Multiple sclerosis	% of patients qualified for pharmacotherapy	PL
	Prevention	Dementia	% of patients using neuroleptic drugs	FR
6	**Trauma/Injury/Suicide (N = 1)**			
	Prevention	Injury/Trauma	Visit to primary care physicians before suicide	SI, LT
7	**Lower/ Upper respiratory infections (N = 1)**			
	Prevention	COPD (Chronic Obstructive Pulmonary Disease)	% of patients with non-invasive ventilations	PL, SW