Universitat Autònoma de Barcelona

MULTIMORBIDITY PATTERNS IN AN OLDER POPULATION FROM NORTH EUROPE

Albert Roso Llorach

Universitat Autònoma de Barcelona

Escola de Doctorat

TESI DOCTORAL

MULTIMORBIDITY PATTERNS IN AN OLDER POPULATION FROM NORTH EUROPE

Doctorand: Albert Roso Llorach

Directora de la tesi: Dra. Concepció Violán Fors

Tutora de la tesi: Dra. Maria Teresa Puig Reixach

Programa de Doctorat en Metodologia de la Recerca Biomèdica i Salut Pública

Departament de Pediatria, Obstetrícia i Ginecologia i Medicina Preventiva i Salut Pública de la Universitat Autònoma de Barcelona

Universitat Autònoma de Barcelona, 2022

Disseny de la coberta: Yasmin Garcia Muñoz

A la meua família i especialment a la Ivette,

AGRAÏMENTS

Mentre escric aquestes línies me n'adono com de prop està el final d'aquest viatge, sembla que va ser ahir quan tot va començar i llavors es veia tan lluny el final... Doncs bé, aquest viatge estrany i molt llunyà està arribant a la seva fi i m'agradaria donar les gràcies a totes les persones, que en major o menor grau, directa o indirectament, han col·laborat o participat d'aquest.

Primer de tot, voldria agrair a l'IDIAP Jordi Gol per donar-me l'oportunitat de poder dur a terme aquesta tesi doctoral. Vull donar les gràcies a la direcció, en Josep Basora, l'Anna Berenguera i la Sandra Illán, per totes les facilitats i suport que he rebut. Tampoc no em vull oblidar d'en Boni Bolíbar per apostar per aquesta tesi.

A continuació, donar les gràcies també a la Universitat Autònoma de Barcelona i al programa de doctorat en Metodologia de la Recerca Biomèdica i Salut Pública per permetre'm realitzar aquesta tesi doctoral. En especial, agrair a la meva tutora, la Teresa Puig, per tota la seva predisposició i ajuda durant aquest difícil camí.

Com deia, aquest camí va començar fa molts anys, en aquells anys on encara era una persona que començava en el món on m'he desenvolupat personal i professionalment. Per això, també voldria agrair a la gent que em va ajudar a començar, com són la gent de Tortosa, en Carlos, Ramón, Marylene, Roger, on vaig aprendre a ser científic, i a la gent de Lleida, la Montse i el Carles, que em van ensenyar el que era l'ofici d'estadístic. Aquesta tesi, també, és en part gràcies a vosaltres.

Seguidament, la vida em va fer tornar a Barcelona, a l'IDIAPJGol, on m'he sentit com a casa durant els darrers 10 anys de la meva vida. Gràcies a tota la gent que heu passat per l'IDIAP i m'heu ajudat en aquest camí. En especial vull agrair al meu company Tomàs, per compartir penúries estadístiques, de doctorat i personals al nostre despatx. A la Maria i l'Edurne, a veure si algun dia quadrem un dinar! Al Ramón, a l'Ana, al Carles, a la Silvia, a la Carmen, a la Mònica, a la Rosa i en general a tota la gent de la UEM, deu n'hi do en sou molts, no m'estranya que us quedeu tota la zona de la UTR... a la gent d'Innovació, el David, el Marc i la Diana, a la gent de Quali, del RWEpi, de la UGP i del SIDIAP, també a la gent de les USR, als nous, ..., la llista es fa llarga!! No em voldria deixar ningú! A la gent de l'Àrea de Recursos Corporatius, la Laura, la Laia, la Stela i companyia, per posar-nos les coses fàcils i també compartir penes i alegries. També recordar els que han marxat en aquests anys, l'Aina, l'Helena, la Ma Angeles, la Pau (segur que ens veiem pel barri!) i voldria també tenir un record especial per la Ceci, que va ser de les persones que em va animar a seguir aquest difícil camí, et trobem molt a faltar... aquesta tesi també va per tu...

A continuació, voldria agrair a tot l'equip de multimorbiditat, del qual he format part des de que vaig entrar a l'IDIAPJGol i és on m'he pogut desenvolupar com a estadístic i investigador. En aquest sentit, gràcies al Jordi Real, per introduir-me al món apassionant dels clústers. També moltes gràcies al Quintí per ensenyar-me el que eren els patrons de multimorbiditat. Al Chema Valderas per la seva visió crítica i aportacions. A la Kika, moltes gràcies per sempre estar disposada a ajudar-nos i per la teva forma de ser. A la Mariona, moltes gràcies, tens gran part de culpa de que aquesta tesi s'hagi portat a terme. A la Amalia i la Noemí, per ajudar-nos amb la part de polifarmàcia, ànims Noemí que ara et toca a tu! També agrair a la Marina per la teva energia i passió, gràcies per seguir-me ajudant en aquest camí. Moltes gràcies a l'equip de la UPC, en especial a la Marga Cabrera per descobrir-nos aquestes tècniques que ens han ajudat tant. Gràcies Sergio per tota la feinada que vas fer i per totes les teves aportacions, gran part del mèrit d'aquesta tesi és teu! Per últim voldria acabar, amb les noves incorporacions, Lucía i Carlos, moltes gràcies pel vostre suport, us passo el relleu, sé que el futur és vostre!

Seguidamente, quería darte las gracias a ti Amaia, muchas gracias por acogerme y conectar con nosotros. Fue una gran experiencia estar en Aging Research Center, ¡lástima que la pandemia no nos dejó repetir! Ya sabes que es una gozada trabajar contigo, espero que esto sea uno más de los proyectos que nos quedan por hacer, ¡gracias! In addition, I would also thank you Davide, for all your help, your wisdom, and your always on point comments. You are one of the most brilliant people in this field, it was a pleasure to collaborate with you. Moreover, thanks to Professor Laura Fratiglioni for allowing me to collaborate with ARC and working with the SNAC-K data. Grazie mille!

Ara voldria dedicar unes línies a la meva directora, la Conxa Violán, sense ella res d'això hagués estat possible. Des de l'inici quan vaig arribar a l'IDIAP, la Conxa va ser la persona que em va acollir i em va guiar en aquest camí i m'ha ensenyat a ser l'investigador i el professional que sóc avui, no puc mostrar-li res més que no sigui total agraïment per tot. Moltes gràcies Conxa, saps que ens ha costat però ho hem aconseguit. Segur que això és un punt i seguit, i coneixent-te sé que tenim molts més reptes per endavant! Gràcies per estar-hi sempre!

Per últim, he reservat les últimes línies per a la gent que no forma part de la meva vida acadèmica o professional. Gràcies a tots els amics que durant aquest temps m'ha donat suport d'alguna manera o altra en aquest viatge, a partir d'ara em podreu anomenar doctor, hehe, però la veritat que això no canvia res, gràcies pel suport!

Als meus germans Eduard i Francesc per haver-me d'aguantar des de l'inici del temps, estic molt orgullós de vosaltres, vull dedicar-vos també aquesta tesi. Parlant d'aguantar, voldria agrair als meus pares, Eduard i Vicenta, per tot l'esforç que van fer per garantir-me poder arribar on sóc ara. Sé que el camí ha estat difícil i que hi ha hagut coses bones i dolentes, però vull que sabeu que us ho dec tot, vos estimo.

Per últim, voldria agrair-te Ivette tots aquest anys. Ets la meva crítica més ferotge però alhora el meu suport més gran. Sense tu no ho hauria pogut aconseguir ni seria la persona que sóc ara, tu em fas ser millor en tot, gràcies per ser-hi. Sé que ens queden moltes aventures per viure i vull viure-les amb tu. T'estimo molt...

RESUM

1.- Antecedents

La multimorbiditat, coneguda com la presència de dues o més malalties cròniques en una persona, és una creixent condició de salut relacionada amb l'envelliment. És ben sabut que les persones que pateixen múltiples malalties cròniques tendeixen a agrupar-se en grups homogenis, i les malalties cròniques solen tenir una llarga durada i, en general, una progressió lenta. Per tant, els mètodes per estimar patrons de multimorbiditat han de ser prou flexibles per identificar aquests patrons i la seva evolució al llarg del temps.

2.- Objectiu de l'estudi

L'objectiu d'aquesta tesi és identificar patrons longitudinals de multimorbiditat en una cohort poblacional sueca de gent gran. Els objectius específics són 1) estimar patrons de multimorbiditat i les seves característiques sociodemogràfiques, d'estil de vida, clíniques i funcionals. 2) Traçar l'evolució dels patrons i detectar les trajectòries clíniques i la mortalitat al llarg del temps i 3) estimar l'evolució longitudinal de la gent gran i el seu temps de permanència a mesura que es mouen entre els patrons.

3.- Mètodes

S'han realitzat tres estudis per donar resposta als objectius de la tesi. Les dades provenen de l'Estudi Nacional Suec sobre Envelliment i Cura a Kungsholmen (SNAC-K), un estudi poblacional que inclou 3.363 individus de la comunitat i institucionalitzats de ≥60 anys. Per a l'estudi 1 i 2) els participants multimorbids van ser agrupats per l'algorisme de clúster *fuzzy c-means*. Per a l'estudi 3, la mostra global es va estratificar en grups d'edat considerant tres dècades. Es van aplicar models ocults de Markov per modelar l'evolució temporal tant dels patrons de multimorbiditat com de les transicions dels individus durant un seguiment de 12 anys.

4.- Resultats

En el primer estudi, els individus multimorbids es van classificar en sis clústers mitjançant l'algorisme de clusterització *fuzzy c-means*. Aquests clústers van mostrar perfils sociodemogràfics, d'estil de vida, clínics i funcionals significativament diferents. En el segon estudi, es van identificar sis clústers d'individus utilitzant *fuzzy c-means*. Durant 12 anys, els canvis en la composició del clúster, les transicions dels participants d'un clúster a un altre i la mortalitat dels participants van mostrar un quadre clínic dinàmic però ben definit. En el tercer estudi, es van identificar quatre patrons longitudinals de multimorbiditat per a cada dècada utilitzant models ocults de Markov. A mesura que augmenta l'edat, l'estabilitat clínica disminueix i el temps de permanència dins d'un mateix patró de multimorbiditat és més curt.

5.- Conclusions

Els patrons de multimorbiditat van mostrar significativament diferències sociodemogràfiques, d'estil de vida i funcionals. Les trajectòries clíniques indicaven un gran dinamisme i complexitat, però identificable al llarg del temps. Diferents clústers es van associar de forma diferenciada amb la mortalitat. El dinamisme entre els patrons de multimorbiditat es va reflectir en els diferents temps de permanència entre patrons. Els mètodes de *fuzzy c-means* i de models ocults de Markov van capturar la naturalesa longitudinal dels patrons de multimorbiditat. Els resultats obtinguts poden ajudar a comprendre millor la complexitat de la multimorbiditat, i a millorar les intervencions preventives en salut.

RESUMEN

1.- Antecedentes

La multimorbididad, conocida como la presencia de dos o más enfermedades crónicas en una persona, es una creciente condición de salud relacionada con el envejecimiento. Es bien sabido que las personas que padecen múltiples enfermedades crónicas tienden a agruparse en grupos homogéneos, y las enfermedades crónicas suelen tener una larga duración y, en general, una progresión lenta. Por lo tanto, los métodos para estimar patrones de multimorbididad deben ser lo suficientemente flexibles para identificar estos patrones y su evolución a lo largo del tiempo.

2.- Objetivo del estudio

El objetivo de esta tesis es identificar patrones longitudinales de multimorbididad en una cohorte poblacional sueca de personas mayores. Los objetivos específicos son 1) estimar patrones de multimorbididad y sus características sociodemográficas, de estilo de vida, clínicas y funcionales. 2) Trazar la evolución de los patrones y detectar las trayectorias clínicas y la mortalidad a lo largo del tiempo y 3) estimar la evolución longitudinal de las personas mayores y su tiempo de permanencia a medida que se mueven entre los patrones.

3.- Métodos

Se han realizado tres estudios para dar respuesta a los objetivos de la tesis. Los datos provienen del Estudio Nacional Sueco sobre Envejecimiento y Cuidado en Kungsholmen (SNAC-K), un estudio poblacional que incluye a 3.363 individuos de la comunidad e institucionalizados de ≥60 años. Para el estudio 1 y 2) los participantes multimorbidos fueron agrupados por el algoritmo de clúster fuzzy c-means. Para el estudio 3, la muestra global se estratificó en grupos de edad considerando tres décadas. Se aplicaron modelos ocultos de Markov para modelar la evolución temporal tanto de los patrones de multimorbididad como de las transiciones de los individuos durante un seguimiento de 12 años.

4.- Resultados

En el primer estudio, los individuos multimorbidos se clasificaron en seis clústeres mediante el algoritmo de clusterización fuzzy c-means. Estos clústeres mostraron perfiles sociodemográficos, de estilo de vida, clínicos y funcionales significativamente diferentes. En el segundo estudio, se identificaron seis clústeres de individuos utilizando fuzzy c-means. Durante 12 años, los cambios en la composición del clúster, las transiciones de los participantes de un clúster a otro y la mortalidad de los participantes mostraron un cuadro clínico dinámico, pero bien definido. En el tercer estudio, se identificaron cuatro patrones longitudinales de multimorbididad para cada década utilizando modelos ocultos de Markov. A medida que aumenta la edad, la estabilidad clínica disminuye y el tiempo de permanencia dentro de un mismo patrón de multimorbididad es más corto.

5.- Conclusiones

Los patrones de multimorbididad mostraron significativamente diferencias sociodemográficas, de estilo de vida y funcionales. Las trayectorias clínicas indicaban un gran dinamismo y complejidad, pero identificable a lo largo del tiempo. Diferentes clústeres se asociaron de forma diferenciada con la mortalidad. El dinamismo entre los patrones de multimorbididad se reflejó en los diferentes tiempos de permanencia entre patrones. Los métodos de fuzzy c-means y de modelos ocultos de Markov capturaron la naturaleza longitudinal de los patrones de multimorbididad. Los resultados obtenidos pueden ayudar a comprender mejor la complejidad de la multimorbididad, y a mejorar las intervenciones preventivas en salud.

ABSTRACT

1.- Background

Multimorbidity, known as the presence of two or more chronic diseases in one person, is a growing health condition related to aging. It is well known that people suffering multiple chronic diseases tend to cluster into homogenous groups, and chronic diseases tend to have a long duration and, generally, a slow progression. Therefore, the methods applied to estimate multimorbidity patterns should be flexible enough to identify those patterns and their evolution over time.

2.- Objective

The aim of this thesis was to identify longitudinal multimorbidity patterns in a Swedish population-based cohort of older adults. The specific aims were 1) to estimate multimorbidity patterns and their sociodemographic, lifestyle, clinical and functional characteristics; 2) to trace the patterns' evolution and detect clinical trajectories and mortality over time; and 3) to estimate the longitudinal evolution of older individuals and their permanence time as they move among patterns.

3.- Methods

We conducted three studies to meet the aims of this thesis, using data from the Swedish National Study on Aging and Care in Kungsholmen (SNAC-K), a population-based study including 3363 community-dwelling and institutionalized individuals aged 60 years and older. For Study 1 and Study 2, we used the fuzzy c-means cluster algorithm to cluster multimorbid participants. For Study 3, we stratified the overall sample into three ten-year age groups and applied Hidden Markov Models to track the temporal evolution of multimorbidity patterns and individuals' transitions over 12 years of follow-up.

4.- Results

In Study 1, the clusters showed significantly different sociodemographic, lifestyle, clinical and functional profiles. In Study 2, changes in cluster composition, participants' transitions from one cluster to another and participant mortality over 12 years generated a dynamic but well-defined clinical picture. In Study 3, we identified four longitudinal multimorbidity patterns for each decade, observing that with increasing age, clinical stability, and the permanence time within a single multimorbidity pattern both decreased.

5.- Conclusions

Multimorbidity patterns showed significant sociodemographic, lifestyle and functional differences. Clinical trajectories showed great dynamism and complexity but can be tracked over time. Different clusters were differentially associated with mortality. The dynamism among

multimorbidity patterns was reflected by the varying permanence times across patterns. Fuzzy c-means and Hidden Markov Models captured the longitudinal nature of multimorbidity patterns. Our results may help to clarify the concept of multimorbidity and improve preventive health interventions.

TABLE OF CONTENTS

FIGURE LIST

TABLE LIST

1. Introduction

1.1. Global burden of disease

In middle- and high-income countries, life expectancy has increased dramatically over the course of the 20th and 21st centuries (1), due to improvements in health resources and medical sciences, combined with decreases in preventable mortality (2). While people are certainly living longer on average, this does not necessarily reflect better health, as an increase in life expectancy anticipates an increase in morbidity (3–5). The acquisition of multiple chronic illnesses or long-term conditions is termed multimorbidity; it occurs in people of all ages but is more frequent in those aged 65 years and older (6). The estimated prevalence of multimorbidity in the general population ranges from 13% to 72%, depending on the setting and age group studied (7), and has increased in recent decades (8–11).

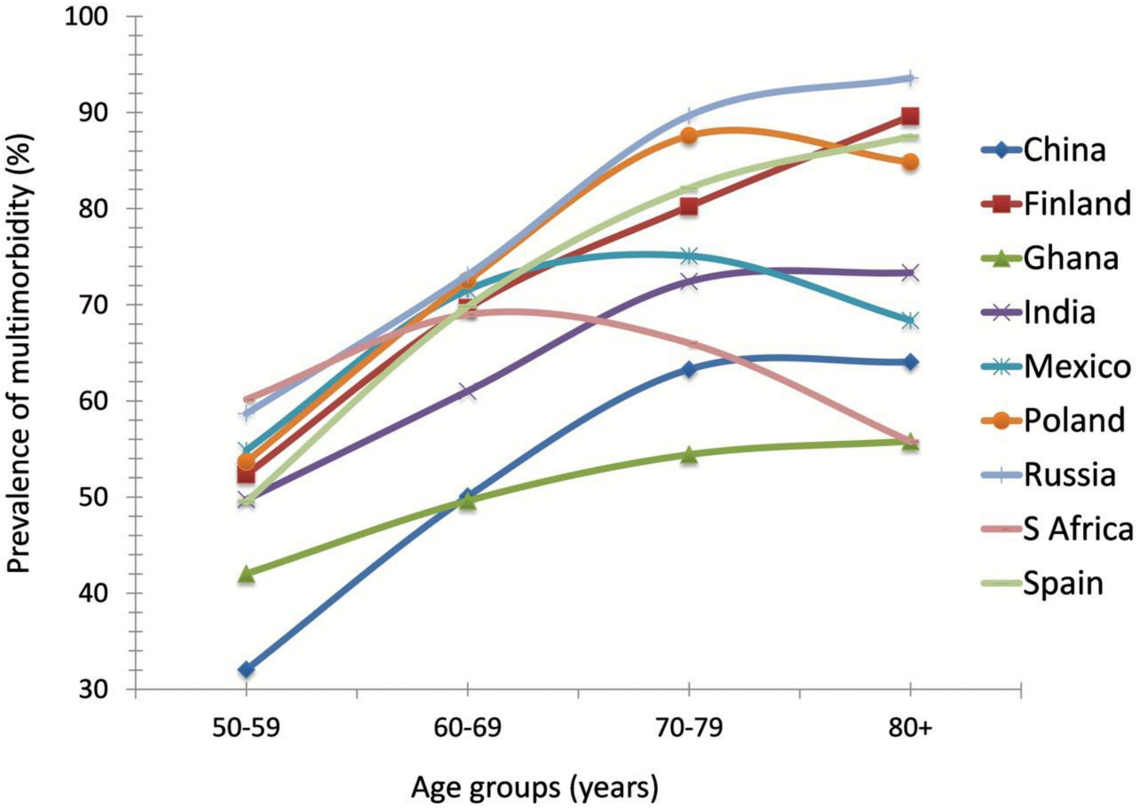


Figure 1.1: Multimorbidity prevalence across age groups by country. Garin et al. Journals of Gerontology: Medical Sciences, 2016.

Multimorbidity adversely affects risk of death, health-related quality of life, functional ability and mental well-being (12,13). Multimorbidity poses major challenges to the delivery of health care worldwide, as health systems are often focused on the management of single diseases and lack appropriate coordination and continuity of care across different sectors (14,15).

1.2. Multimorbidity

Despite some terminological inconsistencies, the literature generally supports the definition of multimorbidity as the coexistence of two or more chronic diseases in one person (16). In 2018, the UK Academy of Medical Sciences defined multimorbidity as the coexistence of two or more chronic health conditions, which can include long-term physical non-communicable diseases, mental health conditions of long duration or long-term infectious diseases (6).

Multimorbidity differs conceptually from comorbidity, which can be defined as the presence of additional diseases in relation to an index disease in an individual (17). This concept revolves around the idea that a principal disease (index disease) largely dictates the patient’s course of treatment for other biologically related diseases (comorbid diseases) (18).

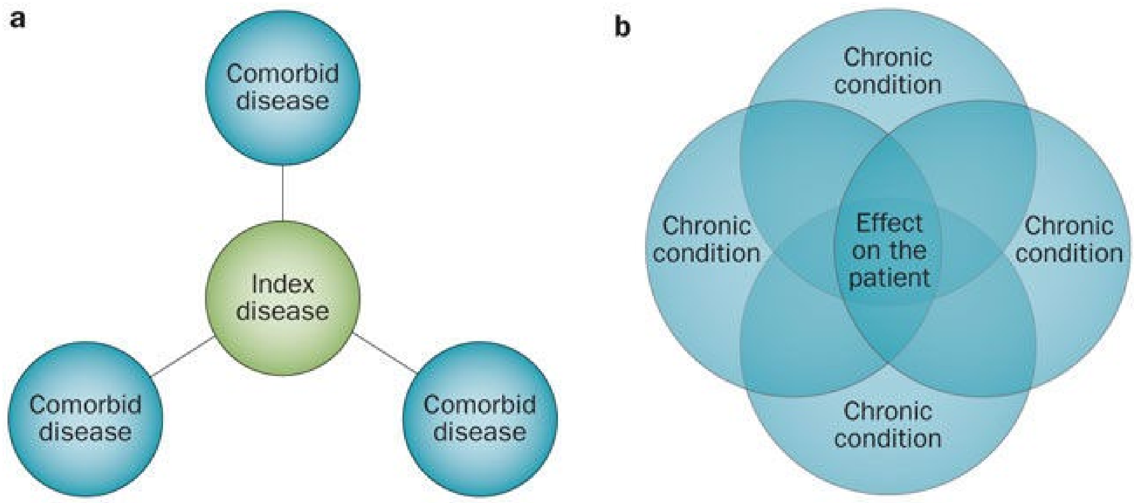


Figure 1.2: Conceptual Diagram of Comorbidity and Multimorbidity. Cynthia M. Boyd, Martin Fortin. Public Health Reviews, 2010.

The difficulty in establishing what qualifies as a chronic disease has led to a lack of coherency regarding the definition and measurement of multimorbidity between different studies and cohorts (19). This has resulted in heterogeneous estimates of multimorbidity prevalence and burden (20–22). Recent studies have presented operational definitions that vary in terms of:

1) the number and types of conditions included;
2) the cut-off number of conditions for defining when multimorbidity is present;
3) whether conditions are simply counted or are weighted in relation to predefined outcomes; and
4) the data sources and data collection methods used (23–26).

The Swedish National Study on Aging and Care in Kungsholmen (SNAC-K) targeted this issue by categorizing the sixty most prevalent chronic diseases in multimorbid patients to operationalize the classification of chronic diseases (27). This methodology was based on a consensus definition: an international multidisciplinary team classified all four-digit level codes from the International Classification of Diseases, 10th revision (ICD-10) as chronic or non-chronic, before grouping the chronic codes into broader categories according to clinical criteria.

COPD, EMPHYSEMA, CHRONIC BRONCHITIS	
Included ICD-10 codes and labels	
J41	Simple and mucopurulent chronic bronchitis
J42	Unspecified chronic bronchitis
J43	Emphysema
J44	Other chronic obstructive pulmonary disease
J47	Bronchiectasis
DEMENTIA	
Included ICD-10 codes and labels	
F00	Dementia in Alzheimer disease
F01	Vascular dementia
F02	Dementia in other diseases classified elsewhere
F03	Unspecified dementia
F051	Delirium superimposed on dementia
G30	Alzheimer disease
G31	Other degenerative diseases of nervous system, not elsewhere classified
DIABETES	
Included ICD-10 codes and labels	
E10	Insulin-dependent diabetes mellitus
E11	Non-insulin-dependent diabetes mellitus
E13	Other specified diabetes mellitus
E14	Unspecified diabetes mellitus
E891	Postprocedural hypoinsulinemia

HYPERTENSION	
Included ICD-10 codes and labels	
I10	Essential (primary) hypertension
I11	Hypertensive heart disease
I12	Hypertensive renal disease
I13	Hypertensive heart and renal disease
I15	Secondary hypertension

Table 1.1: Descriptors of ICD-10 codes included and excluded in each chronic disease category. Calderón-Larrañaga et al. J Gerontol A Biol Sci Med Sci, 2017.

Several ongoing studies are using these categories (28–31), but we are still far from a universal standard classification of chronic diseases. Despite a lack of consensus on its operationalization, multimorbidity affects more than half of the older population (32), and 60% of older adults suffer from six or more chronic diseases (33). The main determinants of multimorbidity are older age, female gender and low socioeconomic status (22,34).

1.3. Determinants

1.3.1. Age

Advanced age is strongly associated with multimorbidity, owing to a combination of biological factors. As age increases, the body generally experiences a physical decline, perhaps most noticeably a reduction in muscle mass and strength and an increase in body fat and frailty (3,35). Invisible to the eye are numerous changes to the organ systems, including reduced elasticity of the heart and increased vascular stiffness, reduction of renal mass and impaired renal response, reduced intestinal absorption and impaired digestive response, and altered hormone levels (3,35). These changes lead to a weaker, less capable body that is more likely to acquire multiple chronic illnesses (3).

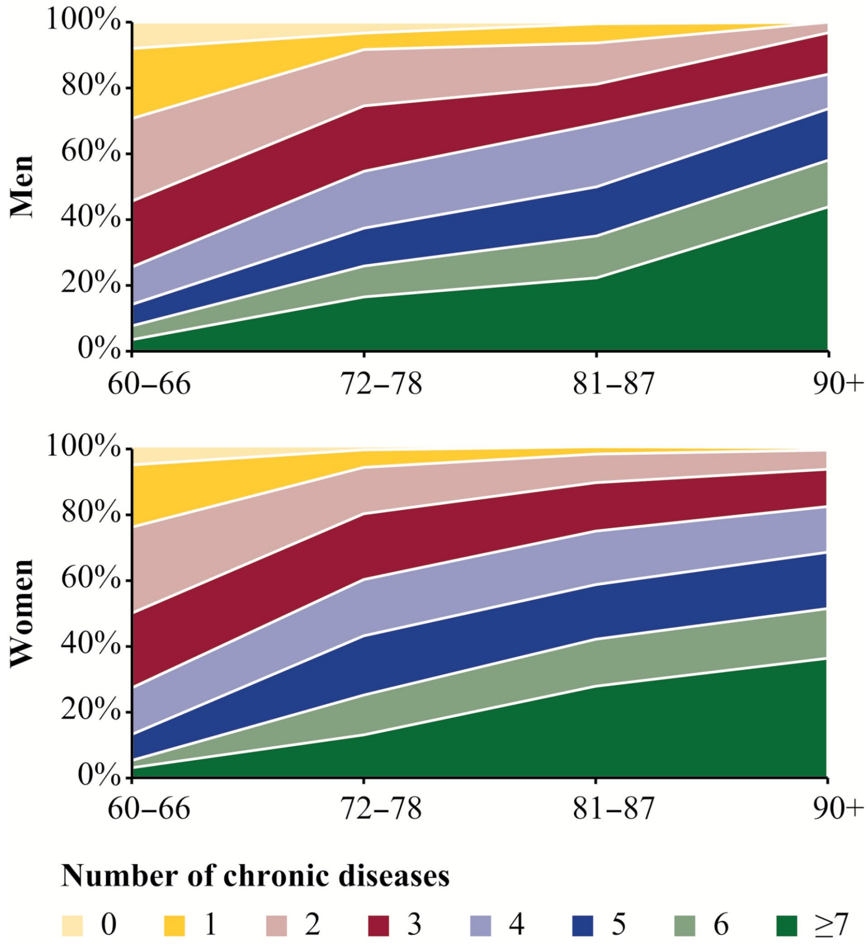


Figure 1.3: Percent distribution of number of chronic disease categories by sex and age group. Calderón-Larrañaga et al. J Gerontol A Biol Sci Med Sci, 2017.

1.3.2. Sex

While the literature on multimorbidity has extensively studied the association with sex, it has produced conflicting findings. Although most studies associate female sex with a higher prevalence of multimorbidity, many show no such association (6,22). Given that women live longer than men on average (36), and so have more years of life to develop chronic diseases, it remains unclear whether the determinant of increased multimorbidity is biological sex itself or rather a combination of other gender-specific and societal factors such as sexism, gender-based violence, poverty or the fact that women are more likely to seek healthcare (6,22,37). One recent longitudinal study showed that, besides prevalence, the incidence of multimorbidity was higher in females than males over time (38).

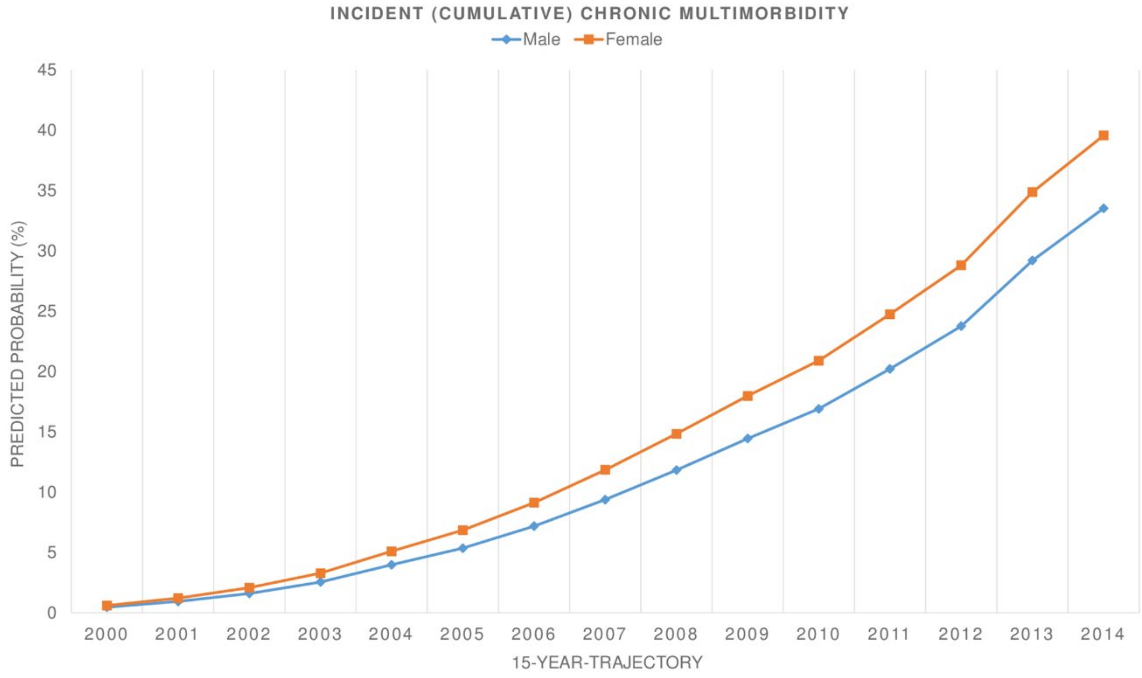


Figure 1.4: Incident (cumulative) multimorbidity (%) during 15-year trajectory stratified by sex. Vos R, Boesten J, van den Akker M. PLOS ONE, 2022

1.3.3. Socioeconomic status

In high-income countries, lower socioeconomic status is associated with higher prevalence of multimorbidity (6,22,39), owing to environmental factors such as living conditions, consumption of high-calorie foods and tobacco use (6). One study conducted in Scotland (UK) showed that people with the lowest socioeconomic status developed multimorbidity 10 to 15 years earlier than those with the highest socioeconomic status (39). On the other hand, higher prevalence of multimorbidity is associated with higher socioeconomic status in low- and middle-income countries (6). This may be because wealthier individuals in low- and middle-income countries have greater access to lifestyle factors that contribute to multimorbidity, such as high-calorie foods, tobacco and alcohol; as well as greater access to healthcare, which ultimately leads to higher levels of disease diagnosis (6).

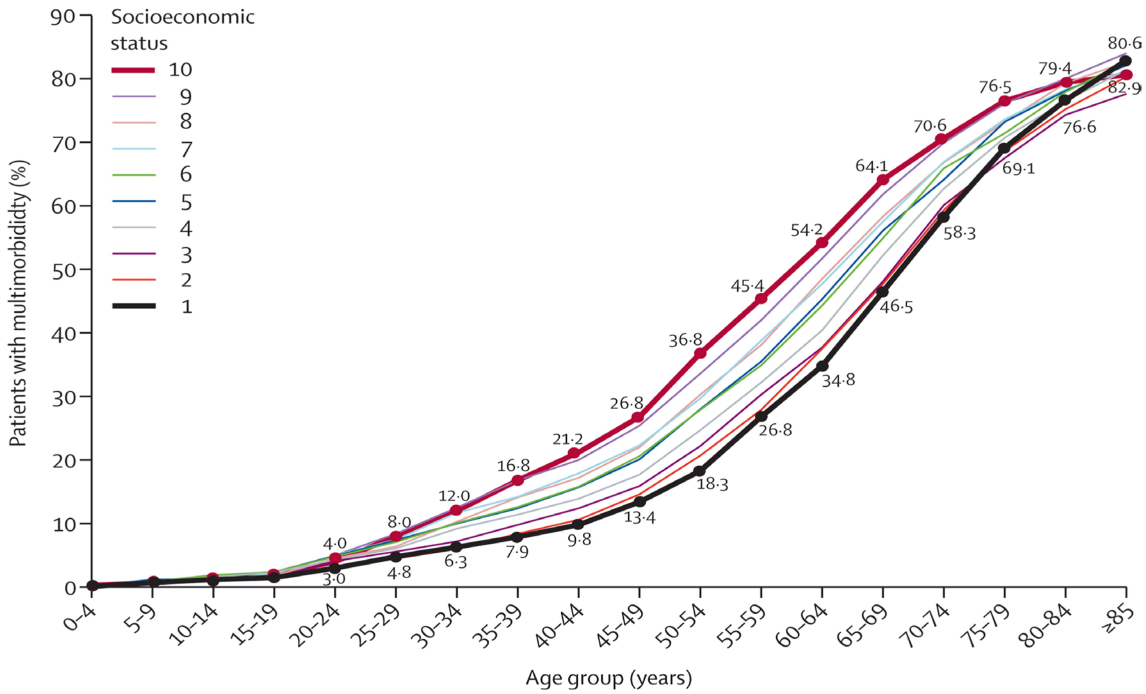


Figure 1.5: Prevalence of multimorbidity by age and socioeconomic status. On socioeconomic status scale, 1=most affluent and 10=most deprived. Barnett et al, Lancet 2012.

To date, public policies addressing socioeconomic disparities have been largely insufficient. Given that health inequalities are expected to increase in the near future, substantial support is needed in low-income regions and low-income strata of society to break the chain of inequality. Adopting a long-life multidimensional approach to population health is key (40).

Individual and parental educational levels can serve as a proxy for socioeconomic status. A recent study showed that both these proxy variables influence the risk of multimorbidity, highlighting the need to address inequality at all stages of life (41).

1.3.4. Lifestyle behaviors and environmental exposures

Increased multimorbidity is associated with a wide variety of lifestyle behaviors and other environmental exposures. Tobacco use and alcohol consumption negatively impact the body, and are among the health behavior determinants most strongly associated with increased multimorbidity prevalence (6).

According to findings from longitudinal studies, walking speed and handgrip strength are inversely associated with the onset of multimorbidity and new diseases in general (42). Research also supports the hypothesis that better physical fitness slows the accumulation of chronic diseases, with various studies observing a link between lack of physical activity and multimorbidity (43–46). However, the findings of non-longitudinal studies may be biased by reverse causality, as people with multimorbidity could be less physically active because of low fitness (47).

Overuse of medical services, especially in high-income countries, can lead to overdiagnosis of illnesses in individuals who are in good general health or who do not have severe symptoms that would negatively impact their quality of life or lifespan (48). Area of residence (urban versus rural) may also be a relevant factor (49), but more research is needed to achieve consensus on this topic.

1.4. Impact

1.4.1. Polypharmacy

The impact of multimorbidity on individuals in high-income countries encompasses a wide array of issues. Multimorbidity is strongly associated with lower quality of life, a decline in physical functionality, disability and higher risk of mortality (6,18,22,39,49). Multimorbid individuals require greater care, in both the healthcare setting and at home (6,18). As a result, they may have a considerable treatment burden, defined as the time and effort required to coordinate care, attend appointments and access treatments, and the negative impact this has on their lives (6).

Polypharmacy (use of multiple medications) is strongly associated with multimorbidity (18,22,50) and decreased quality of life (51). As with multimorbidity, there is no universally accepted definition for polypharmacy, although the most common definition in the literature is the concurrent use of five or more drugs (52,53). The World Health Organization (WHO) acknowledged these facts in a recent report, which stated “Polypharmacy is the concurrent use of multiple medications. Although there is no standard definition, polypharmacy is often defined as the routine use of five or more medications. This includes over-the-counter, prescription and/or traditional and complementary medicines used by a patient” (54).

Regarding the interrelation of multimorbidity and polypharmacy, a recent study using a large database of older adults found that 93.1% of the population satisfied the criteria for multimorbidity and 50% for polypharmacy, and almost 50% had both conditions. Almost all people with polypharmacy had multimorbidity, and 53% of the multimorbid people had polypharmacy (55).

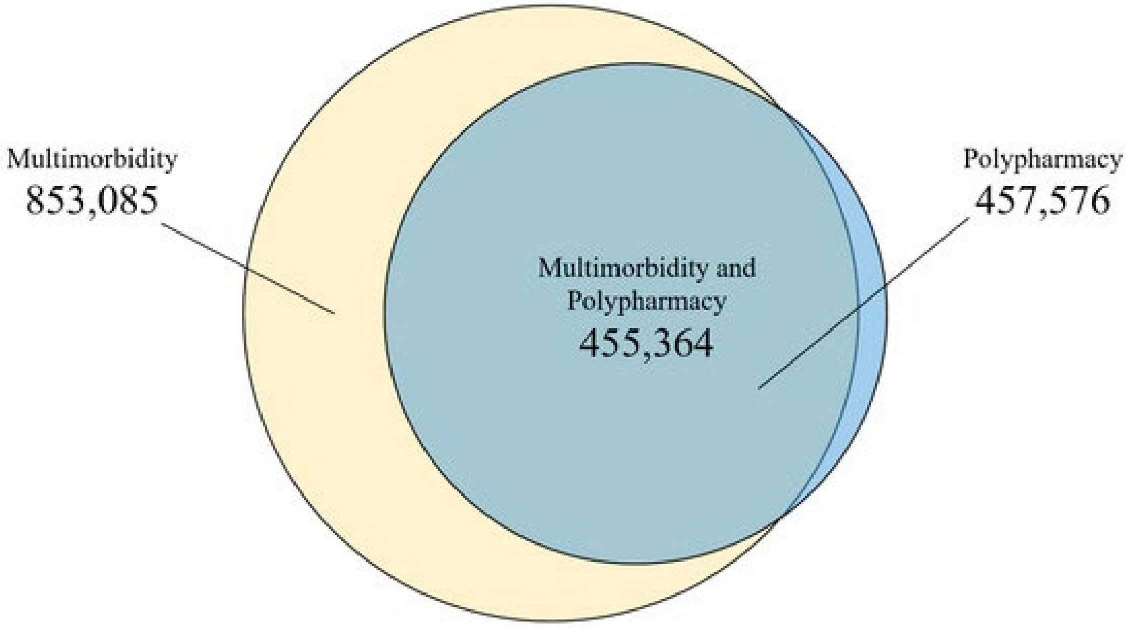


Figure 1.6: Multimorbid and polymedicated individuals in the study aged 65–99 years. Stafford et al. Int J Environ Res Public Health, 2021

1.4.2. Financial burden

In countries such as the USA, where individuals must pay out-of-pocket for some or all of each medical service bill (6), people with multimorbidity can face a heavy financial burden. In general, healthcare systems around the world are designed to treat individual conditions separately rather than multiple conditions jointly; this creates a system-patient disconnect that ultimately results in lackluster care for multimorbid individuals (39). The financial burden of multimorbidity has a negative impact on affected individuals. To effectively address the issue, policymakers and healthcare providers must be aware of this negative impact and promote continuity of care (56).

Multimorbidity also poses a considerable challenge for healthcare systems. In view of the single-disease paradigm described above, multimorbid individuals attend primary care centers and are admitted to hospital with far greater frequency than their non-multimorbid counterparts, thus placing a heavy burden on healthcare service resources (6,18,39). In countries with universal, state-sponsored healthcare, the government bears the associated financial strain. In addition, the cost of treating multimorbidity appears to be considerably greater than the sum of its parts (the cost of treating each condition separately)(6).

1.5. Frailty

Frailty is an emerging concept in geriatric medicine. The term frailty refers to the predisposition of biologically older people to develop adverse outcomes and experience rapid changes in health status (57). Some authors define frailty as a clinical state of increased vulnerability to dependency and/or mortality in the presence of a stressor (58). Studies show that frailty and even pre-frailty are significantly associated with mortality in middle-aged and older adults (59).

Studies have researched frailty with two main models/measures, the first of which is based on an in-depth evaluation of the frailty phenotype (e.g., physical examination, performance measures and questionnaires). The frailty phenotype developed by Fried et al. (60) was based on the following items:

1) Slow walking speed
2) Decreased grip strength
3) Weight loss
4) Physical inactivity
5) Exhaustion

The second type are cumulative deficit models, or frailty indexes, which are constructed with variables such as counts of diseases, laboratory measures, and social and functional impairments, to generate a frailty score (61), with higher values indicating greater degrees of frailty. The frailty index model by Rockwood was based on the ratio between the number of deficits present divided by the number of deficits considered (62). The deficit inclusion criteria were as follows:

1) Biological association with health status
2) Accumulation with age
3) Saturation not occurring at an early age

In addition, Clegg et al. proposed and validated an electronic frailty index (eFI) based on 36 deficits that can be identified in primary care electronic health records (57) . Researches in Catalonia created another index based on the eFI (eFRAGICAP index) using electronic health records; they concluded that their tool had good discriminative capacity to identify frail subjects compared to other frailty scales and predictive outcomes (63).

The ageing population is characterized by multimorbidity and frailty; both are complex syndromes of aging (64). Seven out of every 10 frail individuals are multimorbid, while fewer than two out of every 10 multimorbid individuals are frail. In older people, both measures are associated with risk of disability, hospitalization and mortality, as well as escalating health-related costs (65).

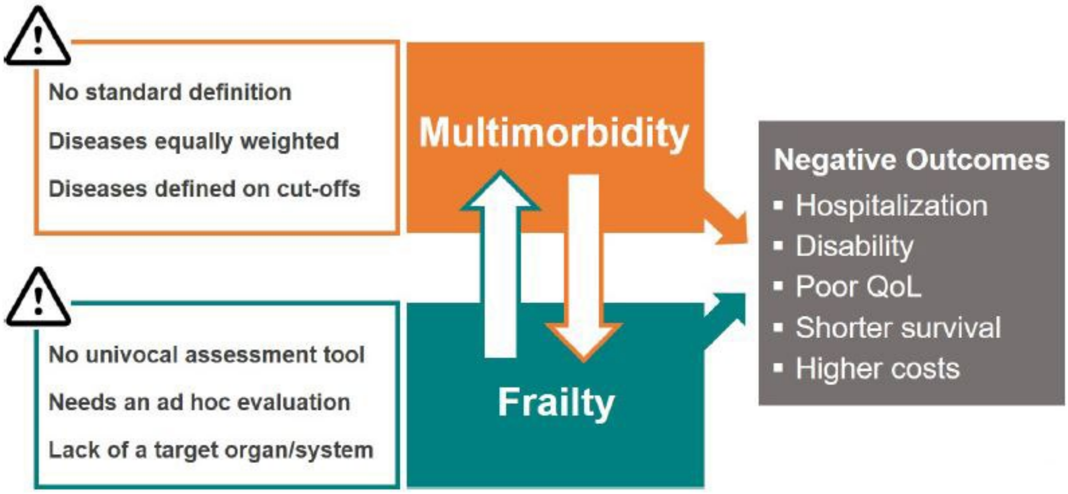


Figure 1.7: Multimorbidity and frailty: two constructs with close relationship, similar consequences and equal challenges. Vetrano DL et al. J Gerontol A Biol Sci Med Sci, 2018.

Some studies have found an association between multimorbidity and frailty (64,66,67). Chronic diseases contribute to the development of frailty (57,63,68,69), while frailty-related health deterioration may lead to the development of comorbidities and thus multimorbidity (68). Research has confirmed the existence of this bidirectional association (64), suggesting some overlap between the two concepts. A recent study carried out in Catalonia analyzed the dynamics of both conditions as people age and calculated the associated risk of death, nursing home admission and need for home care (70). The authors observed that the nature of multimorbidity and frailty varies with the age of the individual, as does the impact of these variables on health status. People become frailer as they age, and their frailty is increasingly characterized by disability and other symptoms rather than by diseases. Mortality is strongly associated with the number of comorbidities, whereas frailty-related deficits are associated with needing specialized care. However, more studies are needed to assess this relationship quantitatively and understand how it evolves over time.

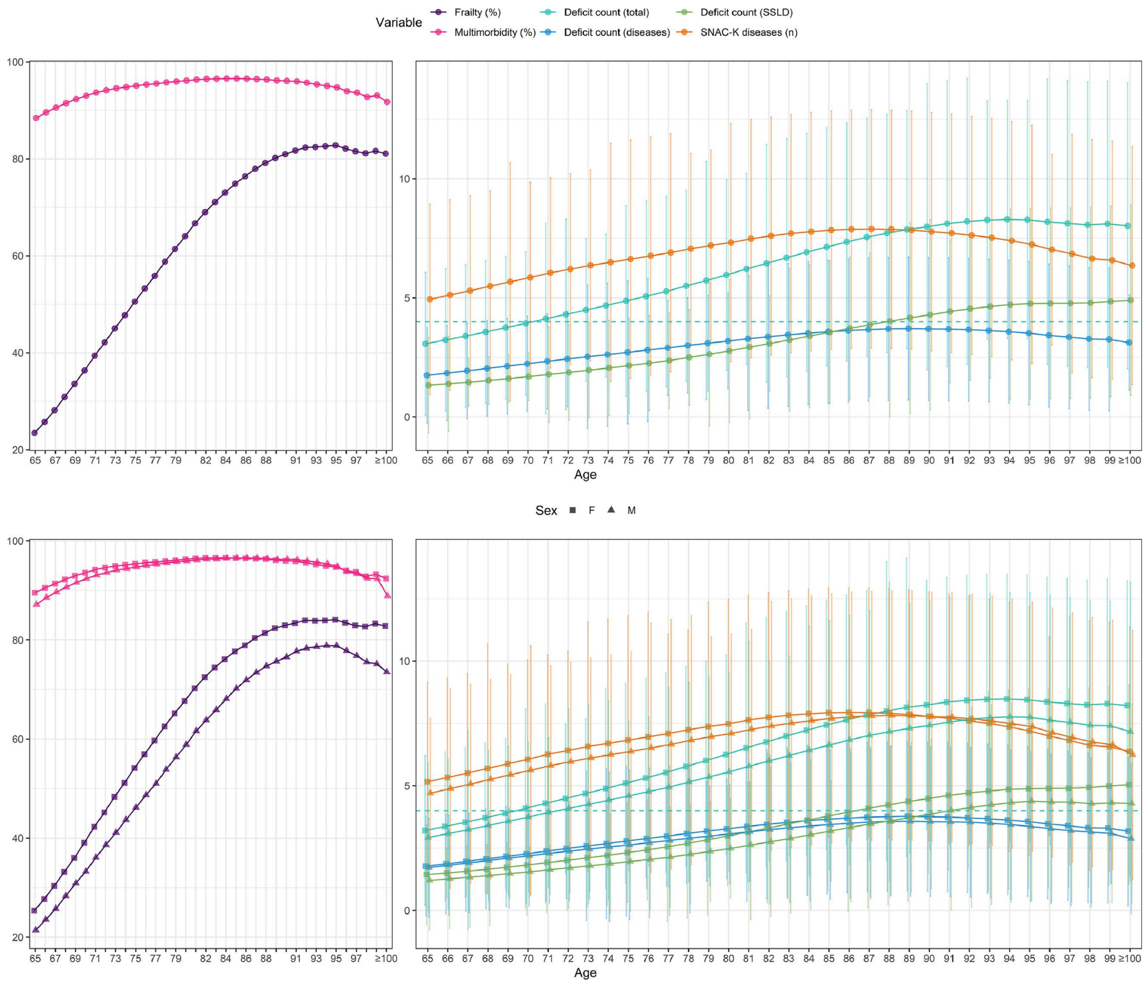


Figure 1.8: Dynamics of frailty and multimorbidity with age. Carrasco-Ribelles et al. eClinicalMedicine, 2022.

1.6. Multimorbidity patterns

A pattern is a combination of variables that show a set of characteristics in a group. Multimorbidity patterns share a set of diseases, which can be commonly defined based on all the diseases diagnosed in a patient (acute and chronic), only considering the chronic conditions or incorporating the functionality and alterations of the psychosocial sphere. The majority of multimorbidity studies have defined patterns based on chronic diseases because of their relevance to health, as well as their durability and progression over time (16,22,34).

Disease patterns	Disease combinations	Common diseases
• Cardiovascular and metabolic diseases • Mental health problems • Musculoskeletal disorders	• Depression comorbid with 8 other conditions (e.g. hypertension, arthritis, diabetes)	• Diabetes • Heart disease • Cancer • Hypertension
	• Hypertension comorbid with 6 other conditions (e.g. osteoarthritis, diabetes, cancer)	• Depression • COPD • stroke
	• Diabetes comorbid with 6 other conditions (e.g. hypertension, coronary artery disease)	• Arthritis/ osteoarthritis • Osteoporosis • Asthma
	• Arthritis comorbid with hypertension, CVD, dyslipidemia, diabetes, and mental health problems	• Gastrointestinal problems • Heart failure • Dementia • Hearing problems
	• Asthma comorbid with arthritis, CVD, and diabetes	• Vision problems • Urinary problems
	• Osteoarthritis comorbid with CVD and/or metabolic conditions	• Thyroid diseases

Table 1.2: Summary of disease patterns, disease combinations, and common diseases in multimorbidity. Xu et al. Ageing Research Reviews, 2017.

To estimate multimorbidity patterns, researchers need methods that identify and separate certain population groups from others and measure non-random associations between diseases in the sub-groups (16,22,34,71). Ng and colleagues identified five analytical methods used to identify multimorbid condition groups:

1) Factor-analysis method: the role of factor analysis is to identify ‘latent’ factors based on the assumption that variables associated with the same factor share a common underlying trait that is responsible for the correlation among them (72).
2) Hierarchical-clustering method: the aim of cluster analysis is to assign entities (such as health conditions) into groups (called clusters) so that entities in the same cluster are more alike to one another than to entities from different clusters. Cluster analysis is also known as ‘unsupervised classification’, where there is no a priori information regarding the underlying group structure (73).

3) Unified-clustering algorithm: a three-step unified-clustering method to identify groups of multimorbid conditions. This method specifically addresses three statistical issues for using cluster analysis to study multimorbidity patterns, namely adjustment for multimorbidity by chance, the uniqueness of clustering results and control for false discovery (74).
4) Multiple correspondence analysis (MCA): this is a nonparametric multivariate method that uses graphical procedures to reveal the association between categorical variables (binary, nominal and ordinal). It attempts to present multivariate categorical data in a low-dimensional space (a counterpart of principal component analysis for categorical data) (75).
5) Network and cluster analyses: these reveal networks of conditions from which to identify sub-networks or groups of connected health conditions (76).

	Factor1	Factor2	Factor3	Factor4
Disorders of lipoid metabolism	**0.25**	−0.16	0.04	−0.06
Osteoporosis	**0.40**	−0.07	−0.17	−0.07
Thyroid disease	**0.25**	0.01	0.02	−0.13
Gastro-oesophageal reflux	**0.44**	0.13	−0.14	0.09
Diverticular disease of colon	**0.31**	0.14	−0.08	0.01
Varicose veins of lower extremities	**0.30**	0.03	0.05	−0.02
Arthropathy	**0.30**	−0.04	0.09	0.00
Cervical pain syndromes	**0.28**	−0.09	−0.02	0.01
Low back pain	**0.34**	−0.06	0.05	0.07
Anxiety, neuroses	**0.37**	0.18	−0.07	−0.13
Dermatitis and eczema	**0.27**	0.02	0.03	0.04
Congestive heart failure	−0.04	**0.39**	**0.37**	0.03
Cardiac arrhythmia	−0.02	**0.34**	**0.36**	−0.14
Iron deficiency, other deficiency anaemia	0.10	**0.35**	0.14	0.00
Cerebrovascular disease	−0.03	**0.36**	0.08	0.08
Dementia and delirium	−0.01	**0.42**	−0.13	0.00
Chronic ulcer of the skin	−0.16	**0.50**	0.10	0.02
Ischemic heart disease (excluding infarction)	0.09	0.18	**0.29**	−0.06
Hypertension	0.11	−0.08	**0.44**	0.03
Diabetes	−0.13	0.04	**0.46**	−0.02
Haematologic disorders, other	0.01	0.12	**0.31**	0.05
Obesity	0.13	−0.25	**0.30**	0.13
Behaviour problems	0.17	0.19	−0.06	**0.39**
Depression	0.06	−0.06	0.05	**0.79**

Figure 1.9: Factor-analysis method. Prados-Torres A et al. PLOS ONE, 2012

1.6.1. Hierarchical clustering versus exploratory factor analysis

The most common methods for examining disease clustering are hierarchical cluster analysis (HCA) and exploratory factor analysis (EFA), which offer very different approaches and solutions (16,22,34,71).

The HCA approach assigns diagnoses to groups or clusters, so that diagnoses in the same cluster are more similar to one another than to diagnoses from different clusters (in relation to a given measure). EFA reduces the observed set of diagnoses to a smaller number of latent factors that account for the correlations between them.

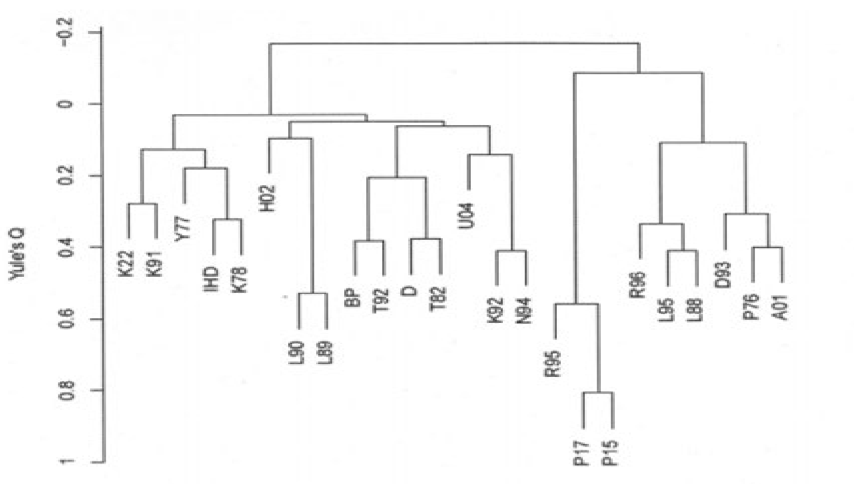


Figure 1.10: Hierarchical cluster analysis, dendrogram for the conditions. Déruaz-Luyet A et al. BMJ Open, 2017

Both HCA and EFA are descriptive methods that identify associations between diagnoses and determine patterns of multimorbidity. HCA clusters tend to contain diagnoses that are similar to each other (in terms of Euclidean distances), but dissimilar to the diagnoses in other clusters; no diagnosis can be included in more than one cluster. In contrast, EFA, like confirmatory factor analysis, is primarily used to test hypothesized relationships between observed measures and latent constructs. In addition, EFA allows for inclusion of any diagnosis in multiple factors, as diagnoses can present significant correlations with more than one factor.

Methodological studies have shown that multimorbidity patterns vary depending on the method of analysis used (HCA versus EFA), and that EFA is useful for describing comorbidity relationships, while HCA could be useful for in-depth study of multimorbidity patterns (77).

1.6.2. Hierarchical versus non-hierarchical cluster analysis

Among cluster analysis methods, there are two main techniques: HCA and non-hierarchical cluster analysis (NHCA) (78).

HCA is often the preferred technique in biomedicine, when the goal is to identify relatively homogeneous groups of cases based on selected characteristics using an algorithm that either agglomerates or divides entities to form clusters. HCA is organized so that one cluster can be entirely contained within another cluster, but no other kind of overlap between clusters is possible. However, the technique is less adequate for robust identification of patterns in data, for several reasons: the hierarchical clusters are susceptible to outliers in the data, the final solution depends on the chosen distance measure, and the algorithms require a large distance matrix and so are inefficient for analyzing large data sets. In addition, HCA methods focus on diseases rather than individuals as the unit of analysis when assessing multimorbidity patterns.

In contrast to HCA, the NHCA approach does not construct groups via iterative division or clustering; instead, it assigns patients to clusters once the number of clusters is specified. The results are less susceptible to outliers in the data, to the influence of choosing a distance measure or to the inclusion of inappropriate or irrelevant variables. Algorithms that do not require a distance matrix can analyze extremely large data sets.

The most frequently used NHCA method is the k-means algorithm, which is composed of the following steps:

1) Place *k* points into the space represented by the patients being clustered. These points represent initial group centroids.
2) Assign each patient to the group with the closest centroid.
3) When all patients have been assigned, recalculate the positions of the k centroids.

Repeat Steps 2 and 3 until the centroids no longer move. This separates the patients into homogenous groups while maximizing heterogeneity across groups.

The K-means method belongs to the family of hard clustering algorithms. Hard clustering forces each individual into a single cluster, whereas soft clustering allows elements to be simultaneously classified into multiple clusters.

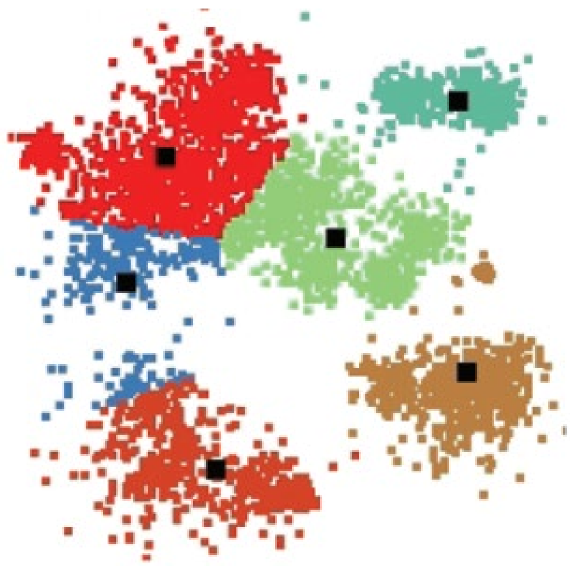


Figure 1.11: K-means clustering

1.6.3. Hard versus soft clustering

Soft techniques present the following advantages over the most commonly used hard clustering algorithms (hierarchical clustering and k-means). First, individuals (and not diseases) are grouped in clusters according to co-occurring diseases. Second, instead of forcing each individual into a specific cluster, these methods assign each individual a probability of membership to each identified cluster. This makes more sense from a biological perspective, as biological mechanisms show that individuals can be associated with multiple diseases and can be classified in different patterns at the same time. Finally, a single disease can characterize more than one cluster, which allows us to build patterns of multimorbidity that take all possible disease combinations into account. In summary, by using soft clustering techniques, we place individuals and not their diseases at the center of our analyses (79).

The fuzzy c-means cluster analysis algorithm is among the most popular methods of the soft clustering algorithms family. It estimates *c* cluster centers (similar to k-means) but with fuzziness, so that individuals may belong to more than one cluster. Compared with hard cluster analysis, fuzzy cluster analysis better accounts for the stochastic nature of some disease associations, the potential noise stemming from the measurement (e.g., disease assessment) and the variance due to between-individual differences.

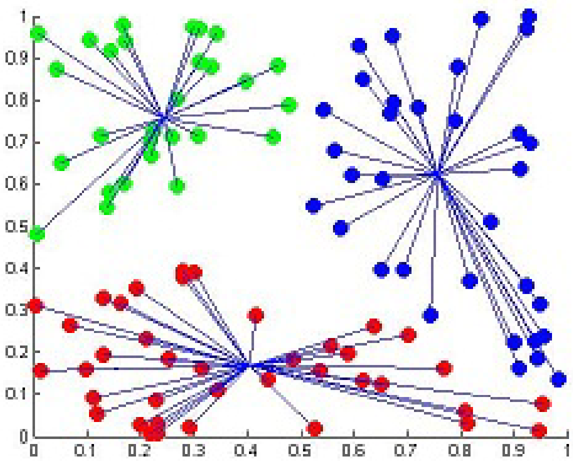


Figure 1.12: Fuzzy c-means clustering

Through this technique, we can obtain the clusters of individuals and a membership matrix that indicates the degree of participation of each subject in each cluster (78).

$$
\begin{array}{l c}
 & \begin{array}{ccc} \text{Cluster 1} & \text{Cluster 2} & \text{Cluster 3} \end{array} \\
\begin{array}{l} \text{Individual 1} \\ \text{Individual 2} \\ \text{Individual 3} \\ \text{Individual 4} \end{array} & \begin{pmatrix} 0.8 & 0.1 & 0.1 \\ 0.02 & 0.9 & 0.08 \\ 0.3 & 0.3 & 0.4 \\ 0.03 & 0.22 & 0.75 \end{pmatrix}
\end{array}
$$

Figure 1.13: Membership matrix

The fuzzy c-means and k-means algorithms are similar in that they both have cluster centers, but the fuzziness in the soft clustering algorithm allows points to belong to more than one cluster.

Algorithm	**How it works**	**Best used…**
K-means	Partitions data into *k* number of mutually exclusive clusters. How well a point fits into a cluster is determined by the distance from that point to the cluster's center.	• when the number of clusters is known. • for fast clustering of large data sets.

Fuzzy c-means	Partition-based clustering where data points may belong to more than one cluster.	• when the number of clusters is known. • for pattern recognition. • when clusters overlap.

Table 1.3: k-means vs fuzzy c-means

1.6.4. Other approaches

There are other approaches that focus on identifying groups of individuals with different patterns of multimorbidity. These methods allow impact analyses using the whole sample simultaneously. The most commonly used techniques include the following:

1) Latent class analysis (LCA): this is a model-based probabilistic clustering approach where the assignment of an individual to a class is probabilistic rather than deterministic. The resultant classes represent probabilistic groups of patients with similar combinations of conditions. As a result, each derived patient cluster has a unique and probabilistic multimorbidity phenotype profile where members do not need to have all included conditions (80–82). Though considered a robust statistical technique for estimating clusters of individuals, the model-based LCA is more computationally demanding than its cluster algorithm counterparts (83).

2) Hierarchical clustering methods for individuals: the same HCA methodology is applied to a set of individuals (84,85). This technique is mainly applied in small data sets because, as mentioned in section 1.5.2, these algorithms require a large distance matrix and so are not efficient for analyzing large data sets.

3) Self-organizing maps (SOMs): this method can be viewed as a nonparametric regression technique that converts multidimensional data spaces into lower dimensional abstractions. A SOM generates a nonlinear representation of the data distribution and allows the user to identify homogenous data groups visually (86). SOMs are mainly used to visualize data dependency among the comorbidities in cluster analyses.

Table 3. Class Proportions and Class-Specific Probabilities from Seven-Latent-Class Model of Chronic Conditions.

	Latent Class						
Class	**1**	**2**	**3**	**4**	**5**	**6**	**7**
Assigned label	**Relatively Healthy**	**Hyper-tension**	**Musculo-skeletal Disorders**	**Headache-Mental Disorders**	**Asthma-Allergy**	**Complex Cardio-metabolic Disorders**	**Complex Respira-tory Disorders**
Class proportion	0.59	0.14	0.10	0.07	0.06	0.03	0.02
Item-response probabilities							
Hypertension	0.05	**0.63**	0.25	0.13	0.05	***0.73***	0.38
Ischemic heart disease	0.00	0.08	0.02	0.03	0.00	*0.30*	0.09
Stroke	0.00	0.05	0.01	0.02	0.00	*0.14*	0.04
Diabetes	0.01	0.23	0.02	0.02	0.01	*0.29*	0.12
Cancer	0.01	0.06	0.06	0.02	0.01	*0.09*	0.07
COPD	0.01	0.06	0.06	0.04	0.01	0.22	***0.69***
Asthma	0.02	0.02	0.01	0.08	0.46	0.16	***0.91***
Allergy	0.14	0.10	0.19	0.33	***0.94***	0.34	0.46
Arthritis	0.05	0.33	**0.77**	0.30	0.08	***0.84***	**0.50**
Osteoporosis	0.01	0.06	0.13	0.02	0.00	*0.19*	0.15
Slipped discs/other back injuries	0.05	0.09	0.37	0.35	0.08	***0.60***	0.30
Mental disorders	0.06	0.06	0.08	*0.42*	0.13	0.30	0.19
Migraine/recurrent headache	0.09	0.05	0.12	***0.65***	0.18	0.37	0.21
Tinnitus	0.07	0.16	0.21	0.22	0.09	*0.34*	0.20
Cataract	0.01	0.12	0.11	0.01	0.00	*0.26*	0.12
Multimorbidity (2+ chronic conditions) (%)	0.00	0.84	1.00	1.00	0.81	1.00	1.00
Number of chronic conditions reported (mean)	0.43	1.91	2.25	2.54	2.25	4.48	5.37

Item-response probabilities > 0.5 in **bold** to facilitate interpretation
Within each item, the class with the highest response probability is in *italic*
COPD = chronic obstructive pulmonary disease

doi:10.1371/journal.pone.0169426.t003

Figure 1.14: Class Proportions and Class-Specific Probabilities from Seven-Latent-Class Model of Chronic Conditions. Larsen et al. PLOS ONE, 2017

1.7. Trajectory of multimorbidity

The evolution of multimorbidity throughout people's lives and the time individuals remain within specific patterns are under-researched aspects of this field.

One systematic review by Ho et al. focused on the definition of multimorbidity patterns and trajectories in 566 multimorbidity studies (16). The primary aim of 19 included studies (3.4%) was to trace the trajectory of multimorbidity by examining the trends of multimorbidity prevalence or multimorbidity development over time. All identified studies were based on longitudinal data, but most performed cross-sectional analyses. Most studies assessed the incidence rather that the evolution of multimorbidity. Others focused on the accumulation of conditions by using dyads and triads out of a selected list of chronic conditions (87). Some studies adopted simple approaches like analysis of variance (ANOVA) and least square means analysis to estimate the association of risk factors at baseline with the evolution of multimorbidity (88).

The most popular method for assessing the trajectory of multimorbidity is logistic regression models, which consider both baseline and/or repeated measurements covariates. Some studies have used simple approaches, adding potential risk factors for developing multimorbidity to the model (10,89,90). Furthermore, models that consider multimorbidity as a binary outcome can be extended by applying a multinomial logistic regression (91).

In contrast, some authors have taken advantage of the longitudinal data, using temporal correlation between individuals to estimate the multimorbidity evolution through multilevel logistic models (92). With this type of approach, researchers can identify the acquisition sequence of multimorbidity and assess the influence of risk factors and determinants on the sequence (93). An additional use for this technique is to examine individual change using multilevel logistic growth curve models (94).

In addition to logistic regression, survival models can be applied to assess the relationship between multimorbidity and mortality, controlling for risk factors and time-dependent covariates (95).

The second most popular approach to measuring multimorbidity trajectories is to apply linear mixed or hierarchical models to estimate the speed of multimorbidity (96), or to estimate the association between baseline variables and the rate of multimorbidity development over time (97). Hierarchical linear models can also be employed to analyze covariate variations in temporal changes of multimorbidity status (98).

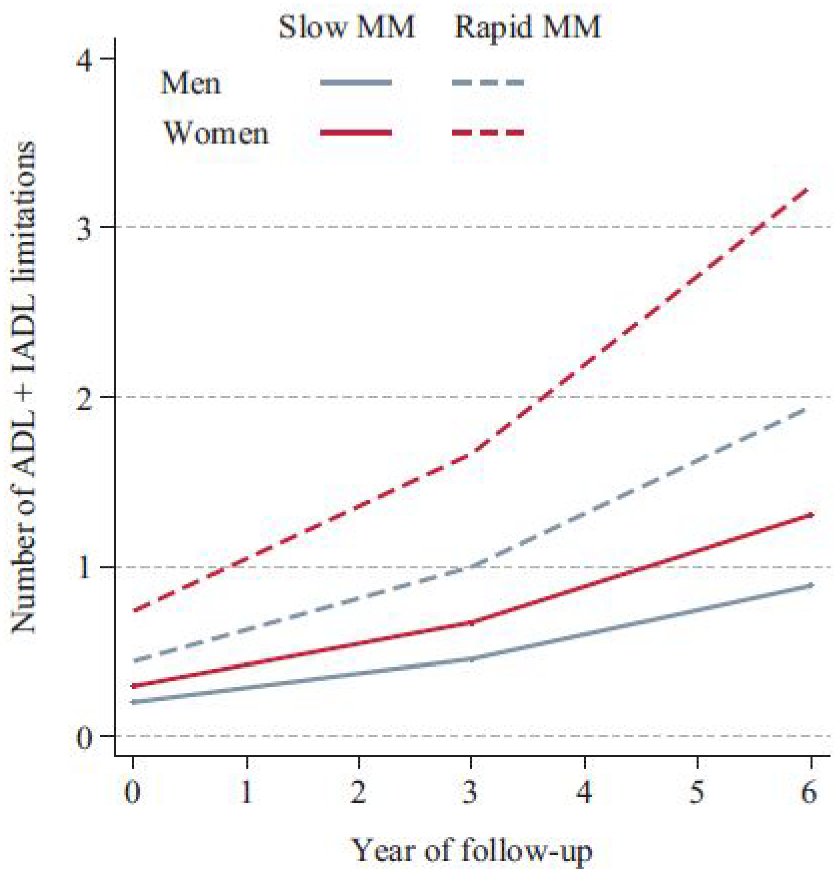


Figure 1.15: Predicted mean number of ADL+IADL limitations associated with rapid versus slow speed of multimorbidity development, stratified by sex. Calderón-Larrañaga et al. J Intern Med, 2018

Several studies have used generalized linear models. Negative binomial or Poisson regression have been applied to assess the trajectories of multimorbidity burden over time by considering multimorbidity outcomes as count data (42,99). With these models, researchers can determine the relationship between risk factors and covariates with both the development of multimorbidity and worsening of multimorbidity. Models such as probit regression have also been applied in a multimorbidity setting (100).

Some studies have used the generalized estimating equations model (GEE) and the multilevel random intercept model with repeated measurements to determine patterns of incident multimorbidity and polypharmacy. The clinical trajectories can be estimated taking into account the correlation of longitudinal data within individuals and the occurrence of repeated events (38).

Other authors have opted to use structural equation modeling (SEM) to study the evolution of multimorbidity determinants like socioeconomic status. This technique investigates the underlying structure of the relationships among all observed and latent variables. The term structural indicates that the parameters are not merely descriptive measures of association, but rather that they reveal a certain 'causal' relation (101).

Simulation studies have been conducted to estimate the evolution of multimorbidity. Dynamic microsimulation models can simulate the characteristics (sociodemographic factors, health behaviors, chronic diseases and geriatric conditions) of individuals over long time periods (102).

Finally, some authors have combined analysis of clustering with regression methods; for example, using latent class growth analysis to identify multimorbidity trajectories over time and using multinomial regression to calculate relative risk ratios. These risk ratios reflect the association between baseline risk factors and multimorbidity trajectory (103).

1.8. Longitudinal multimorbidity patterns

By definition, chronic diseases have a long duration and usually a slow progression. The evolution of diseases affects the composition of the multimorbidity patterns. Several studies have analyzed patterns of multimorbidity across different populations, settings and countries, but most studies have adopted a cross-sectional design or have focused on the progression of comorbidities of index diseases (104,105). In addition to between-study methodological differences, one explanation for heterogenous findings may be related to the dynamic nature of disease clusters, which is not accounted for in cross-sectional studies (22,34).

Multimorbidity patterns evolve over time, and mortality selection plays an important role in shaping the observed population (106). Therefore, multimorbidity patterns must be analyzed longitudinally to determine their evolution and/or stability over time. Guisado-Clavero et al. explored multimorbidity patterns across six years (33).

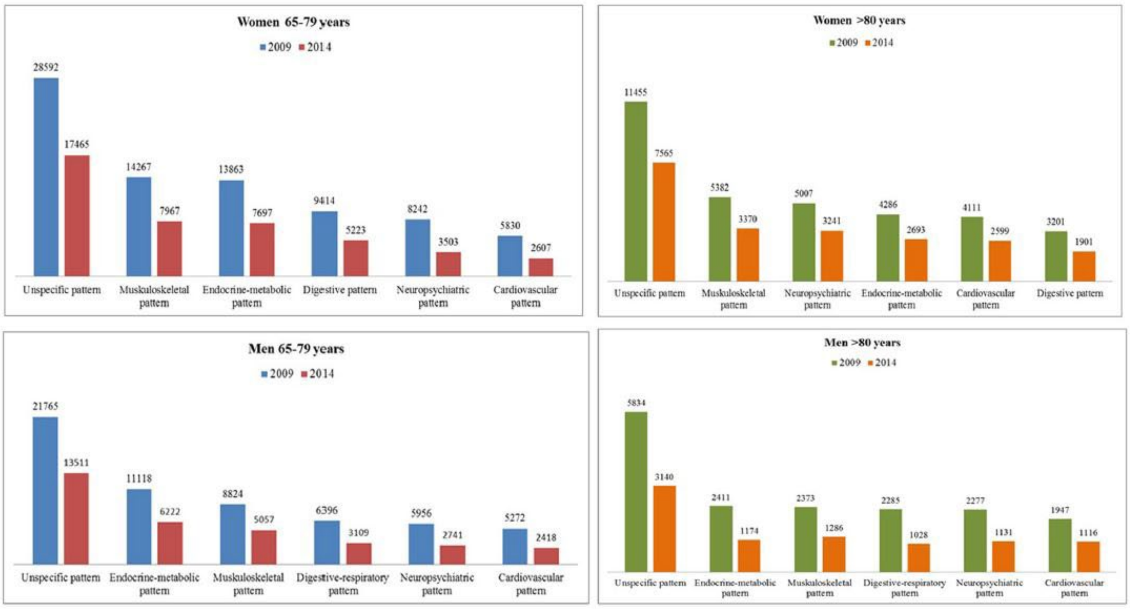


Figure 1.16: Sample corresponding to each pattern and people remaining in that pattern at the end of the study. Guisado-Clavero et al. BMC Geriatrics, 2018

The study authors detected a considerable stability of some patterns over time, and concluded that people may suffer from diseases closely related to more than one multimorbidity pattern. Consequently, it seems more accurate to identify multimorbidity patterns considering that an individual can be classified or distributed across different patterns. This be addressed with advanced statistical methods and machine learning approaches.

1.8.1. Statistical and machine learning modelling approaches

Today, it is generally assumed that researchers should analyze multimorbidity patterns longitudinally, assessing their evolution and/or stability over time. New statistical techniques have been applied to find homogeneous groups of people who suffer from similar multimorbidity patterns, while allowing for the temporal evolution of patterns.

New advanced statistical techniques have been developed to meet the challenge of modelling the complex data in large longitudinal data sets. In parallel, machine learning techniques have been gaining popularity in this type of analysis. Machine learning is a sub-field of the computer science field of artificial intelligence robotics, pattern recognition software, etc.). Although machine learning evolved separately from statistics, somewhere along the way it started relying heavily on statistical principles, and some techniques can be considered to belong to both fields. Generally, statistics draws population inferences from a sample, and machine learning finds generalizable predictive patterns. One of the main advantages of machine learning is that it does not assume any data model/structure, which makes it more flexible that statistical modelling in some situations.

To develop these algorithms, three modalities of learning can be applied:

1) Supervised learning measures a series of characteristics in a set of observations and a response variable in the same set of observations. As such, the algorithm combines questions and answers and can obtain predictions.
2) A second, unsupervised learning modality is based on statistical techniques that analyze a series of characteristics measured in a set of observations. It cannot make predictions because the variable answer is not available; rather, its function is to group and/or observe relationships between variables.
3) In contrast to supervised or unsupervised learning, reinforcement learning consists of training machine learning models to make a sequence of decisions. To date, multimorbidity studies have made greater use of supervised and unsupervised learning.

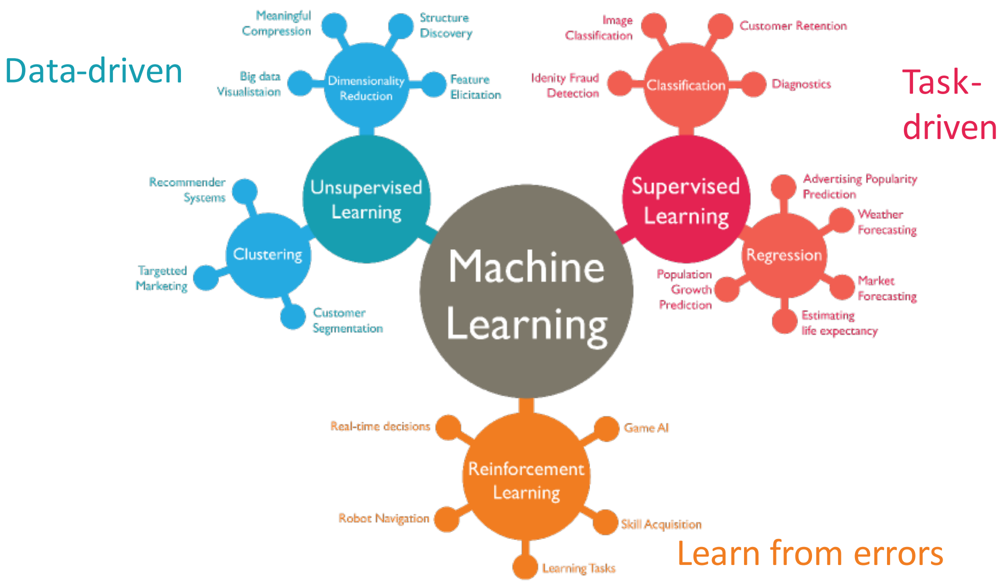


Figure 1.17: Machine Learning techniques

1.8.2. Longitudinal trajectory

Despite the growth and popularity of new statistical and ML techniques, few published studies have adopted longitudinal approaches to date, although this trend is changing (107).

Two previous studies analyzed disease progression and multimorbidity pattern trajectories using latent class growth models in the UK (108) and the USA (109). In terms of the analytical approach, the latent class growth models employed were based on the distribution of the repeated measures of binary diagnosis outcomes to identify longitudinal trajectories.

Lappenschaar et al. used multilevel temporal Bayesian networks (MBN), which are aimed at analyzing relationships between diseases (i.e., networks), in a large cohort in the Netherlands (110). In an MBN, the disease variables are also represented as nodes in a network, but the associations have a direction, and probabilistic associations are represented by arrows. Temporal arrows always point from the past to the future, and a causal interpretation can be assumed. Another study investigating multimorbidity trajectory networks within large databases took place in Denmark (111).

One example of the new developed methodologies is the algorithm proposed by Faruqui and colleagues, an unsupervised multi-level temporal Bayesian network designed to represent the relationship among emergence of multiple chronic conditions and patient-level risk factors over time (112). The authors also performed the comparison with several methods, and concluded that longest path algorithm from the Bayesian network identified the most probable sequence from/to a specific disease.

Another example of the new developed methodologies is the work by Giannoula et al., which focused on the identification of complex time-dependent disease associations using dynamic time warping (113). The proposed clustering algorithm, illustrated in Fig. 1.18 belongs to the class of unsupervised machine learning methods. The study authors represented the disease-history vectors of patients of a Catalan health data set as time sequences of ordered disease diagnoses. They identified statistically significant pairwise disease associations and assessed the temporal directionality of these associations. Subsequently, they applied an unsupervised clustering algorithm, based on dynamic time warping, to the common disease trajectories to group them according to shared temporal patterns. More recently, the authors further applied the algorithm to identify disease trajectories by integrating data from electronic health records with genetic and phenotypic information (114).

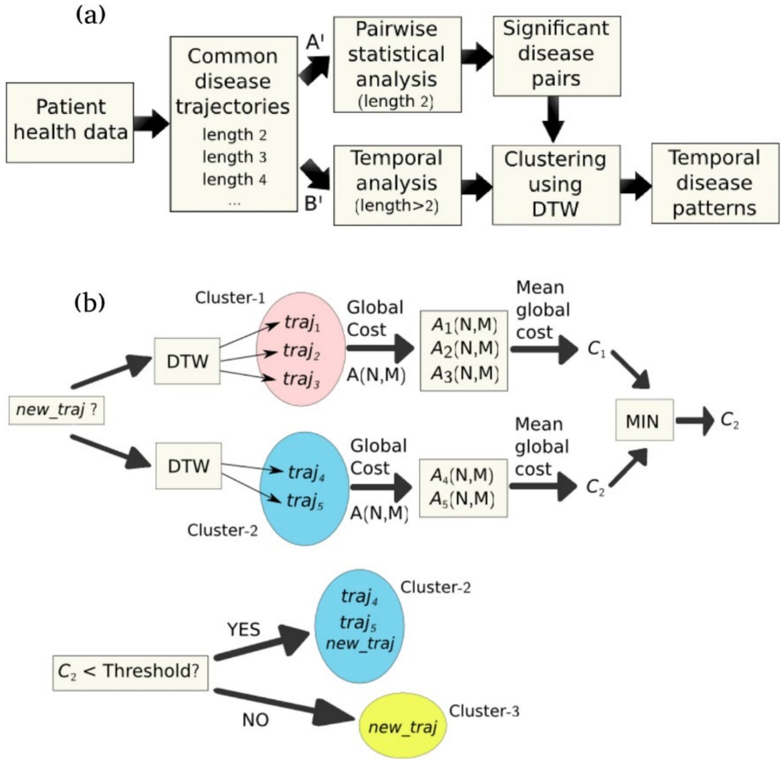


Figure 1.18: Flow-charts of the proposed methodology. (a) A flow-chart of the proposed methodology for the extraction of time-dependent disease associations and (b) the unsupervised clustering method of the common disease trajectories using the dynamic time warping algorithm. Giannoula et al. Sci Rep, 2018

1.8.3. Longitudinal transitions

In the longitudinal study of multimorbidity, it is crucial to track longitudinal shifts or transitions across periods of time. One of the more straightforward ways to assess the transition between diseases and patterns is Alluvial plots or Sankey diagrams. Xu et al. constructed a Sankey diagram to characterize the dynamic changes of different combinations of three conditions over time (115).

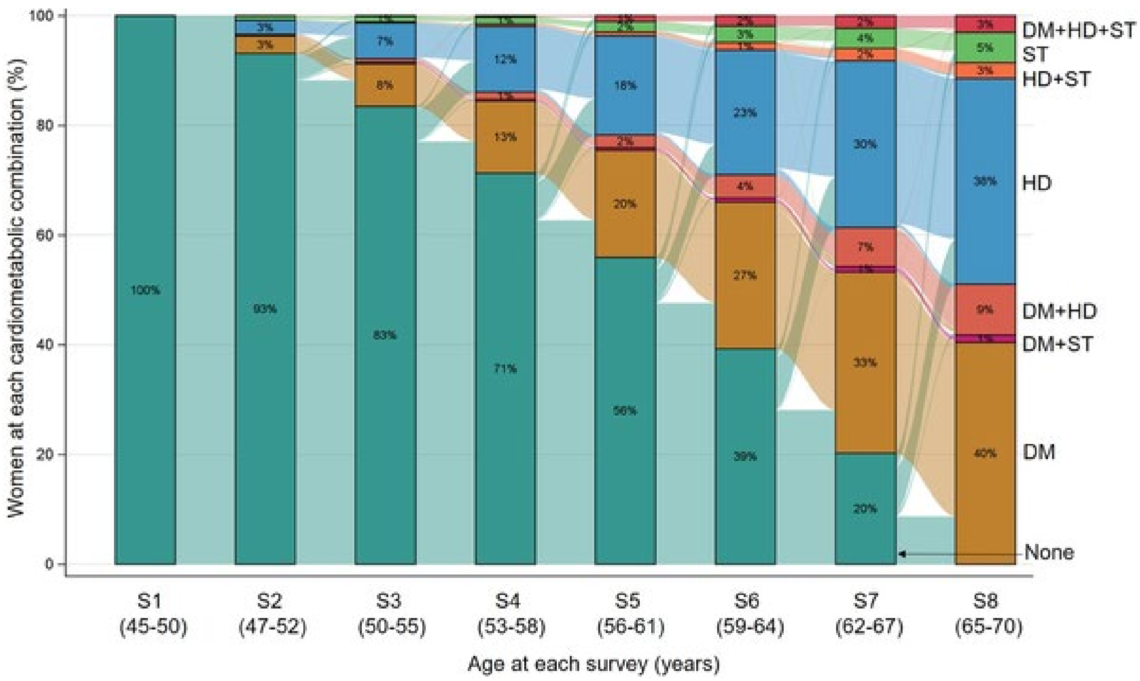


Figure 1.19: Sankey diagram showing the longitudinal progression and transitions among different combinations of diseases. Xu et al. PLoS Med, 2018

However, this descriptive approach, although informative, cannot characterize the whole random process. Some studies have taken a step further by using multistate models to analyze longitudinal multimorbidity data. Multistate models enable the analysis of longitudinal data in which individuals may experience more than one health event. Multistate models are defined by states and transitions between them (116–118). States can be transient, where individuals can enter and exit, or absorbing, where individuals never exit once they enter (e.g., death). This type of model can be analyzed using survival analysis methods.

Some studies have used multistate models to define an interconnected progressive chronic disease system for older adults (119). In this type of modelling, there are different clinical states that an individual can occupy at a given time point. An individual starts from one of the single disease states and moves towards the absorbing state, usually death, either directly or through different intermediate multi-disease states. Freisling et al. assumed a multistate modelling for transitions to cancer, cardiovascular disease (CVD), type 2 diabetes and subsequently to multimorbidity using cox proportional hazards (120).

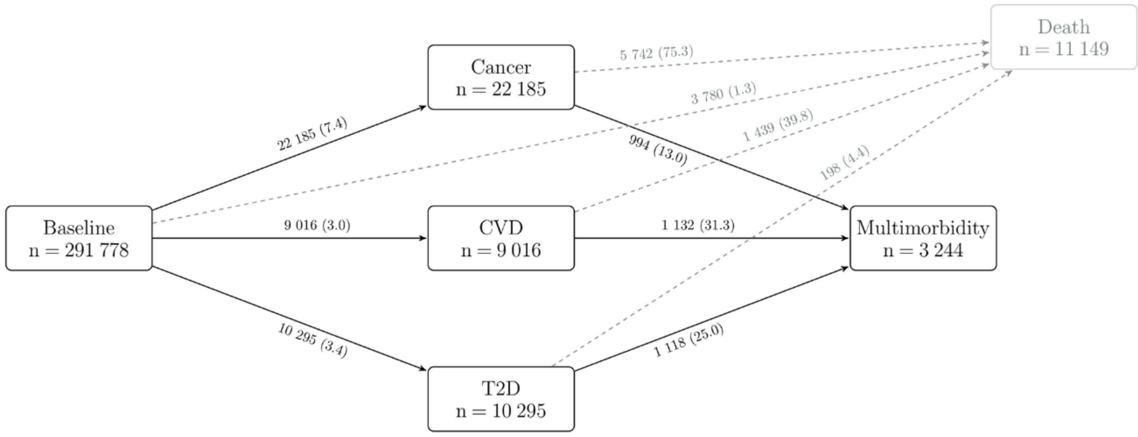


Figure 1.20: Transitions from baseline to cancer, CVD, T2D, and subsequent cancer-cardiometabolic multimorbidity. Freisling et al. BMC Medicine, 2020.

Multistate Markov models can estimate the transition hazard (the instantaneous risk of transitioning from one state into another), as well as transition probabilities and the mean sojourn time in a given state. The main draw of Markov models is their simplicity. A multistate model is considered Markov if it assumes that the probability of transitioning to a new state depends only on the current value of the model. In general, a random process can be described as a Markov model if it determines future probabilities solely based on its current values. This means that the past, current and future states of the system are independent of one another. For this reason, Markov processes are sometimes described as 'memoryless'.

One example of applying Markov chain models in the multimorbidity setting is found in the study of Alaeddini et al., who modelled disease transitions using Markov chain models, placed in a latent regression Markov mixture model to incorporate subject-specific covariates (e.g., age, sex, race/ethnicity). The study authors used a Markov clustering algorithm to identify patterns of disease progression (121).

1.9. Hidden Markov Models

There is a growing trend of applying dynamic machine learning methodologies to identify multimorbidity patterns. One method that has influenced the study of multimorbidity patterns is Hidden Markov Models (HMMs). These models overcome several of the limitations of previously employed methods; for example, they can account for the variability in chronic disease interactions over time (122).

HMMs integrate a dynamic Bayesian network that works with the temporal sequence of the observed patient's data (123,124). In HMM, the observations are random variables conditioned by a hidden state or cluster. For instance, if we consider that each patient belongs to one multimorbidity pattern each year, it is not possible to observe the cluster directly (Figure 1.21).

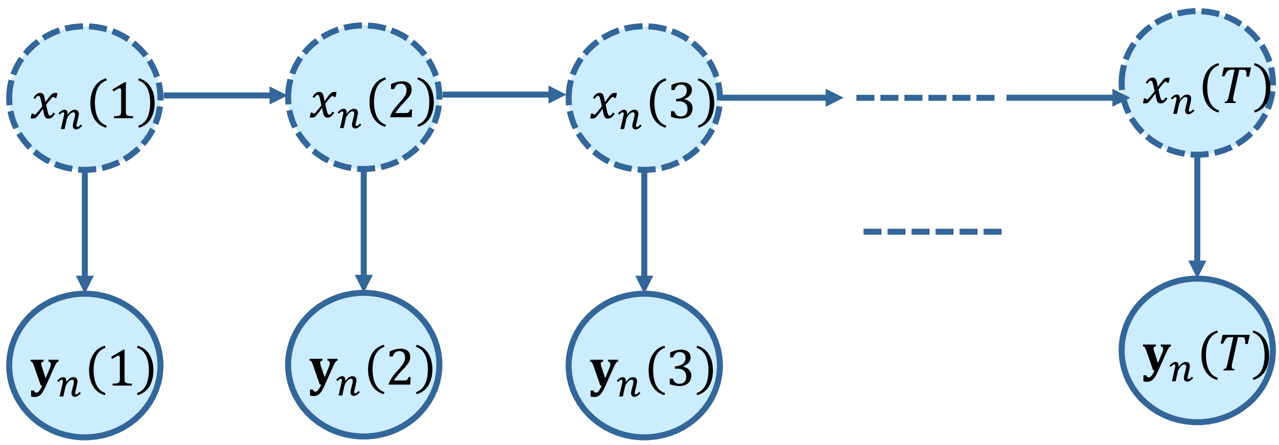


Figure 1.21: Hidden states or clusters. UPC Signal processing & communications.

The main characteristics of HMM are the following:

1) Each subject in each year belongs to a cluster or state.
2) The hidden variable (x) indicates the cluster to which a subject belongs in each year.
3) The information available for each subject in each year is the observations data (y).
4) The temporal evolution of a subject is modeled using hidden variables and observable variables.

To apply this model, we must assume some properties of the stochastic process. The two main assumptions of HMM are: the future is independent of the past given the present, and the observations are independent of the future and past given the present.

HMM considers the individual's characteristics and their evolution over time. In contrast to other methods, HMM estimates use all longitudinal information. Transition to other clusters depends on the evolution of the chronic diseases burden that an individual is accumulating longitudinally. The model can estimate:

- Most likely pathway for a subject having its data.

- Probability of a certain pathway for a specific subject.

By refocusing the analysis on individuals through HMM, we can obtain a better characterization of the population groups with multimorbidity.

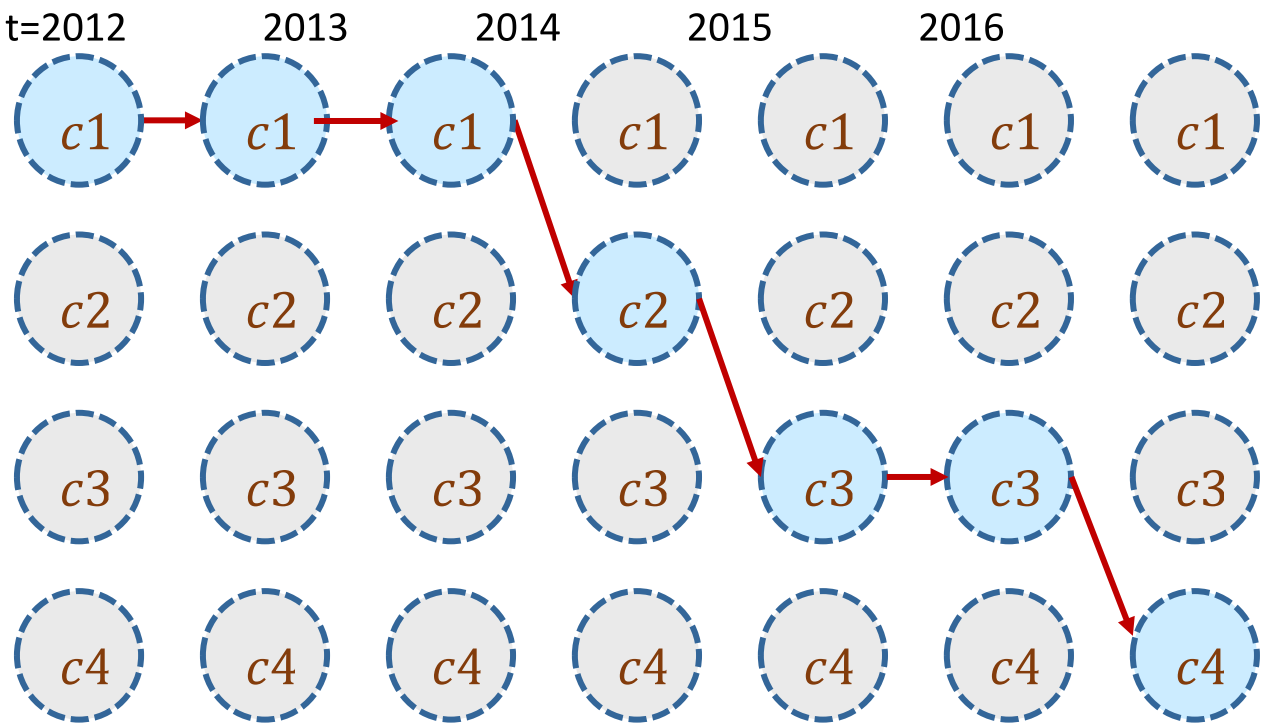


Figure 1.22: Pathway for a specific subject over time. UPC Signal processing & communications.

The longitudinal multimorbidity patterns obtained with HMM methods provide a comprehensive picture of the evolution of multimorbidity over a patient's lifetime. The model can predict the likely multimorbidity pattern of a person over the next few years. Based on this information, health professionals and decision makers can implement preventive interventions to alter many trajectories and even shift causes of mortality.

Previous studies have applied dynamic Bayesian networks for health analysis. As mentioned in section 1.8.2, a Dutch analysis applied this type of methodology to a large primary care data set (110). Other examples relate to the decomposition of shared latent factors using Bayesian multimorbidity dependency maps and healthcare predictive risk modelling (125,126).

Despite the potential of HMM, only one previous register-based study has used this technique for the longitudinal study of multimorbidity and polypharmacy (122,127). It demonstrated the feasibility of characterizing multimorbidity patterns over time. Multimorbidity trajectories were generally stable, although the study authors observed changes in specific multimorbidity patterns. Ultimately, they showed that HMM is useful for modelling transitions across multimorbidity patterns and mortality risk.

2. Justification

New concepts have emerged in geriatric epidemiology, public health and primary care research to approach health complexity in older people. For instance, the term comorbidity refers to the existence of additional conditions beyond an index disorder, while multimorbidity refers to the coexistence of two or more chronic diseases in one person. It is necessary to identify associations between diseases and the risk factors for these associations, and to analyze the trajectories of multimorbidity patterns and their longitudinal evolution over time, to improve the organization of health services based on the groupings of disease that a person presents and their sociodemographic characteristics.

Despite a lack of consensus on its operationalization, multimorbidity affects more than half of the older population, and 60% of older adults suffer from six or more chronic diseases. The main determinants of multimorbidity are older age, female gender, low socioeconomic status, unhealthy lifestyle behaviors and environmental exposures. The impact of multimorbidity includes a wide array of issues, from polypharmacy to heavy financial burden on individuals and the health system. In addition, the interrelationship of multimorbidity and frailty further complicates the study of multimorbidity in older people. Therefore, it is important to analyze multimorbidity in the context of sociodemographic, lifestyle, clinical and functional characteristics.

Researchers have applied several statistical techniques to find homogeneous groups of people suffering from similar multimorbidity patterns. Factor analysis can define multimorbidity patterns based on the mutual relation among diseases, while hard clustering techniques (e.g., k-means cluster analysis), identify non-overlapping groups of people with common diseases where individuals are assigned to one group. In contrast, soft clustering techniques (e.g., fuzzy c-means) do not force individuals into one specific cluster, but rather assign each individual a probability of membership to all identified clusters, which makes more sense from a biological perspective. It seems more useful to identify multimorbidity patterns where individuals can be classified or distributed across different patterns.

By definition, chronic diseases have a long duration and usually a slow progression. Therefore, it is fundamental to analyze multimorbidity patterns longitudinally, assessing their evolution and/or stability over time. While several studies have used longitudinal data, most have adopted a cross-sectional design or focused on the trends of multimorbidity prevalence or the development of multimorbidity over time. This highlights the need to apply methodologies that identify multimorbidity patterns considering all longitudinal information. Modelling this complex data requires advanced statistical methods and machine learning approaches.

In the longitudinal study of multimorbidity, it is crucial to track longitudinal shifts or transitions across periods of time. This explains the growing trend of applying dynamic machine learning methodologies to identify longitudinal multimorbidity patterns.

For all the reasons outlined above, research in this field should make use of flexible statistical and machine learning techniques such as fuzzy c-means and Hidden Markov Models, which can assign people to more than one pattern and track their longitudinal shifts from one pattern to another over long periods of time.

3. Hypotheses

Main hypothesis:

Statistical and machine learning techniques adapted to the longitudinal nature of multimorbidity patterns can identify such patterns and their characteristics, and detect their evolution and underlying dynamics.

Specific hypothesis

H1) Multimorbidity patterns differ according to sociodemographic, lifestyle, clinical and functional characteristics.

H2) Multimorbidity patterns change over time. Clinical trajectories and mortality depend on the longitudinal multimorbidity pattern.

H3) An individual's longitudinal shifts from one pattern to another over time depend on the characteristics and multimorbidity evolution of that individual.

4. Aims

Overall aim:

The aim of this thesis was to implement a statistical and machine learning technique adapted to the longitudinal nature of multimorbidity patterns in a Swedish population-based cohort of older adults followed up for 12 years.

Specific aims:

A1) To identify clusters of older people based on their multimorbidity patterns, and to analyze differences among clusters according to sociodemographic, lifestyle, clinical and functional characteristics.

A2) To identify multimorbidity patterns, trace their evolution and detect clinical trajectories and mortality over time.

A3) To estimate the longitudinal evolution of older individuals as they move among patterns, using statistical and machine learning methods to detect the dynamics underlying such patterns.

5. Material and methods

To respond to the hypotheses of this doctoral thesis, we published three articles in international indexed journals (one article for each specific objective). This section will describe the methodology of each study.

5.1. Study population

This thesis is based on data from the population-based Swedish National Study on Aging and Care in Kungsholmen (SNAC-K) (128), which is an ongoing longitudinal population-based study of individuals aged 60 years and older residing at home or in an institution in the Kungsholmen area of Stockholm, Sweden. SNAC-K is one of the four subprojects included in the Swedish National Study on Aging and Care (SNAC). The ultimate goal of SNAC-K is to understand the aging process, and to identify possible preventive strategies for improving health and care in older adults (Figure 5.1).

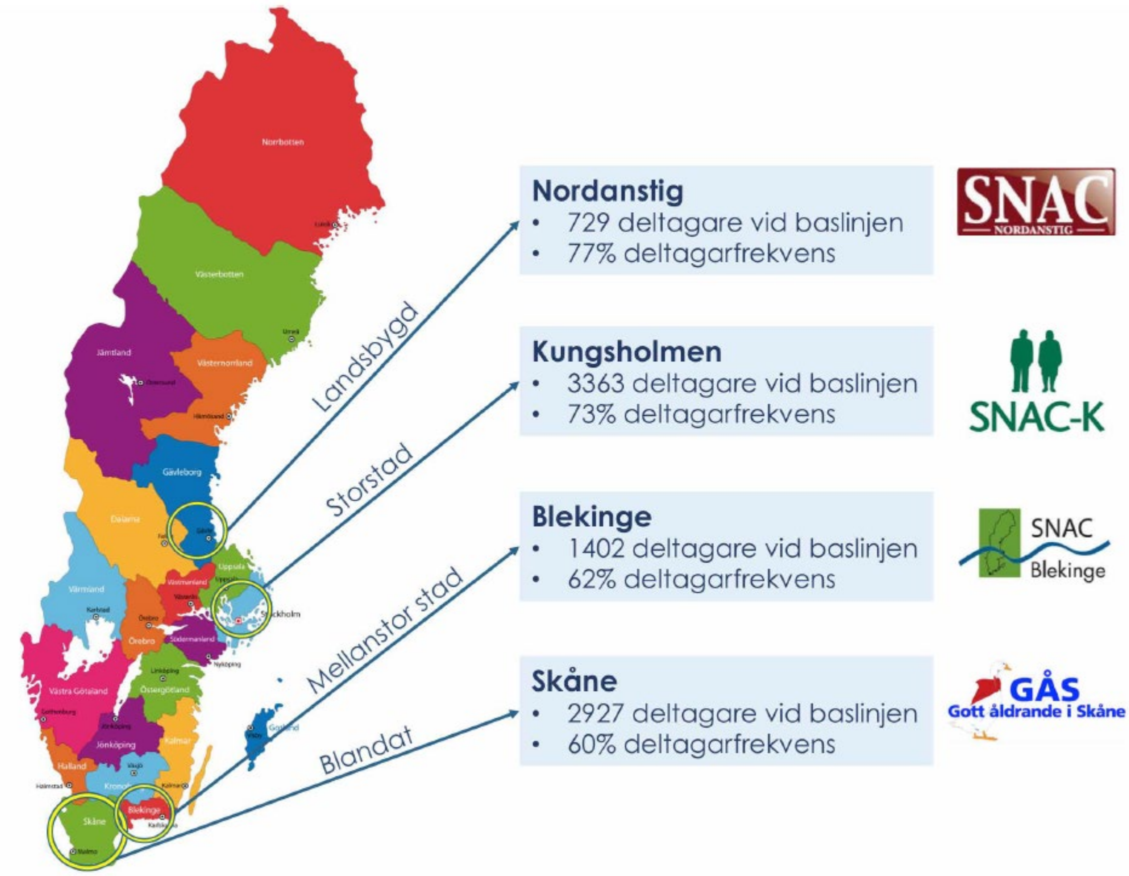


Figure 5.1: SNAC study description

The investigators invited a random sample of 11 age cohorts (60 years, 66 years, 72 years, 78 years, 81 years, 84 years, 87 years, 90 years, 93 years, 96 years and 99 years and older) born between 1892 and 1939 (the youngest and oldest age cohorts were oversampled) to participate in the study. Main causes of ineligibility were deafness, language issues, move to other area or no contact information. Eligible people who agreed to participate were evaluated for the first time between 2001 and 2004, and were subsequently followed up every six years (for those aged under 78 years) or every three years (for those aged 78 years and older). At baseline, 3363 people were examined (participation rate 73%) (Figure 5.2). The main reasons for non-participation were proxy refusal, participant refusal and withdrawal.

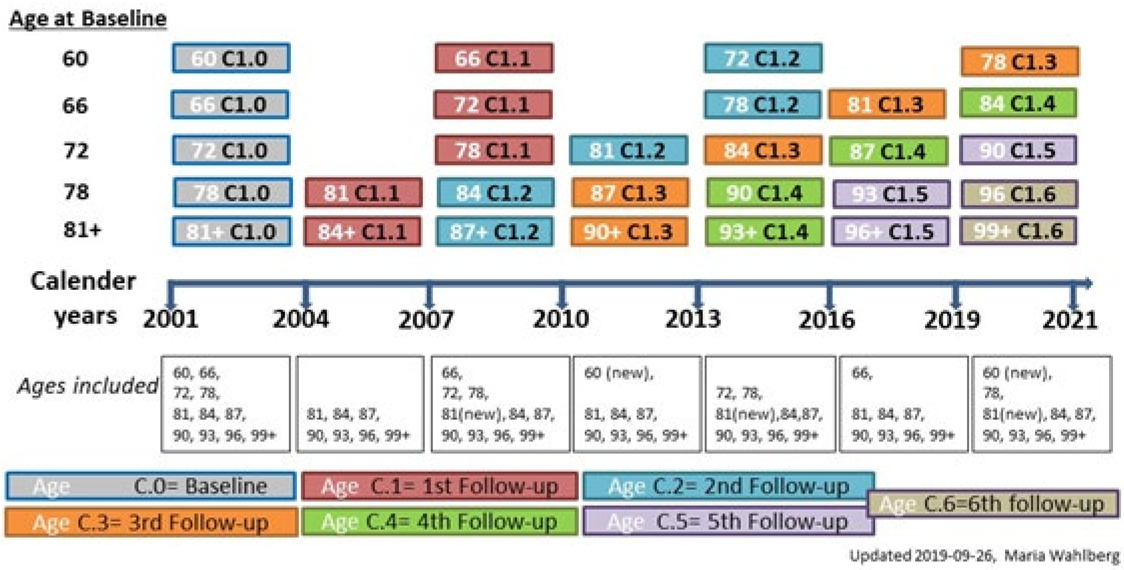


Figure 5.2: SNAC-K study waves

Table 5.1 presents the main sociodemographic characteristics of the SNAC-K cohort at baseline (129).

Item	n (%)
No. of participants	3363
Age	
60–66 years	1034 (38.8)
72–78 years	939 (27.9)
81–87 years	634 (18.9)
90 + years	486 (14.5)

Female sex	2182 (64.9)
Education	
Elementary/ High school	590 (17.5)
University	2741 (81.5)
Living in a nursing home	191 (5.7)

Table 5.1: Sociodemographic characteristics of the SNAC-K population

5.2. Study design and selection criteria

For A1, we used a cross-sectional study design. Of 3363 participants, we excluded 432 because they did not fulfill the inclusion criteria of having multimorbidity (i.e. two or more chronic diseases) at baseline. This resulted in a sample size of 2931 people. As expected, the people we excluded were younger, more educated and less likely to be female than those we included ($p < 0.001$).

For A2, we adopted a longitudinal design, following up the 2931 multimorbid participants to six years (1716 participants) and 12 years (1016 participants). Mortality and dropout were the main causes of loss to follow-up.

We also used a longitudinal study design to meet A3, following up all 3363 SNAC-K participants. We stratified the sample into three age groups: sexagenarians (age cohorts of 60 years and 66 years), septuagenarians (age cohorts of 72 years and 78 years) and octogenarians and beyond (all remaining age cohorts).

5.3. Data collection

The investigators collected data on participants' current status and past history through interviews, clinical examinations and specific tests. The health professionals involved (nurses, psychologists and physicians) received ad hoc training aimed at standardizing procedures. At baseline and at each follow-up visit, participants were examined for an average of six hours. The examination included a biographic assessment and measurement of physical functioning by a nurse (two hours); clinical examination by a physician for the geriatric, neurological and psychological assessment (two hours); and cognitive evaluation by a psychologist (two hours).

5.4. Study variables

Of all variables collected for the SNAC-K study, we included the following variables in our analyses:

- Clinical parameters, lab tests, medication and inpatient and outpatient care data used to identify specific conditions
- Diagnoses according to ICD-10, classified into 60 chronic disease categories in accordance with a clinically driven methodology (27)
- Drug codes, according to the Anatomical Therapeutic Chemical (ATC) classification.
- Educational attainment (elementary, high school, university or higher)
- Main occupation (manual, non-manual; based on the longest job held during the person's lifetime)
- Civil status (unmarried, married, divorced, widowed)
- Smoking status (never smoker, former smoker, current smoker)
- Alcohol consumption (never/occasional, light/moderate, heavy)
- Intensity of physical activity, categorized into three groups as per the recommendations of WHO and the American College of Sports Medicine (ACSM): inadequate (no more than two or three times per month of light and/or moderate/intense exercise), health-enhancing (light exercise several times per week or every day) and fitness-enhancing (moderate/intense exercise several times per week or every day) (130,131)
- Life satisfaction, measured using the self-reported index developed by Neugarten et al. (LSI-A), which captures five components: zest versus apathy, resolution and fortitude, congruence between desired and achieved goals, positive self-concept and mood (132). The LSI-A consists of twelve positive and eight negative items; in SNAC-K, the negative items were reversed and the final scores transformed to a 0–100 scale with higher values indicating greater life satisfaction (133).
- Social network index: a combination of indicators of self-reported social connections and social support, according to the procedure adopted in the National Social Life, Health, and Aging Project (NSHAP Study) (134). For the SNAC-K study, these indicators were categorized into tertiles (poor, moderate, rich) (96).
- Self-rated health, assessed by asking participants, "In general, how would you describe your health?" and categorized as very good/excellent and good/poor
- Level of disability, defined as the number of basic activities of daily living (ADL; bathing, dressing, toileting, continence, transferring, eating) and instrumental activities of daily

living (IADL; grocery shopping, meal preparation, housekeeping, laundry, managing money, using the telephone, taking medications, using public transportation) a participant was unable to perform independently. People living in institutions were assumed to depend on others for grocery shopping, meal preparation, housekeeping and laundry.

- Balance, defined as the time (in seconds) a participant could stand on one leg (up to 60 seconds)
- Grip strength, measured with a dynamometer and converted to kilograms. Participants were seated with their arm resting on a table and their elbow flexed at 90 degrees during measurement.
- Walking speed, assessed by asking participants to walk six meters, or 2.44 meters if the participant reported walking slowly. If the participant was unable to walk or attempted unsuccessfully to walk, a value of 0 was recorded.
- Cognitive status, assessed by physicians with the Mini-Mental State Examination (MMSE), which ranges from 30 to 0 (from best to worst possible score)
- Serum albumin (g/L), creatinine (µmol/L), and C-reactive protein (CRP) (mmol/L) levels, measured in the laboratory of Karolinska Institutet according to standard procedures

5.5. Vital status and loss to follow-up

The SNAC-K investigators obtained information about vital status from death certificates provided by Statistics Sweden, the Swedish governmental statistics agency, and assessed survival status throughout the follow-up period. Participants were considered lost to follow up if they or a proxy declined to participate, could not be contacted, had moved out of the study area or cancelled an assessment.

5.6. Potential bias

This thesis may be affected by several sources of bias:

1) Firstly, although the SNAC-K study used random sampling to create the list of potential participants, selection bias may have arisen from the fact that frail older people and healthy young people are less likely to agree to participate. The investigators oversampled the youngest and oldest people to minimize this effect.

2) Another type of selection bias of study participants arises from longitudinal attrition, when individuals die or decide to leave the study. This can affect estimation in the later waves of follow-up because survival bias can arise. The SNAC-K investigators took this potential bias into account when deciding which variables to include.

3) Self-reported variables may be subject to information bias, although the comprehensive data collection and standardized procedures in the SNAC-K study may have helped to minimize this effect.

5.7. Statistical analyses

In the three studies included in the present thesis, we reported participants' characteristics as absolute numbers and proportion (%), or mean ± standard deviation (SD) with 95% confidence intervals (95% CIs), as appropriate. We carried out all analyses using Stata version 17 and earlier and R version 4.1.2 and earlier. The significance level was set at $\alpha = 0.05$. Specific analytical strategies were adopted in each of the three studies (Table 5.2).

Study	Outcome	Exposures	Potential confounders	Analytical approach
S1	Multimorbidity clusters	Multimorbidity clusters	—	Fuzzy c-means
	Sociodemographic, lifestyle, clinical and functional variables			ANOVA, chi-squared tests

S2	Multimorbidity clusters Mortality	Multimorbidity clusters	Age, sex and education	Fuzzy c-means Logistic regression
S3	Multimorbidity clusters Clinical and functional characteristics	Multimorbidity clusters	Age, sex and education	Fuzzy c-means Hidden Markov Models Linear mixed models

Table 5.2: Analytical approach of the three studies included in the thesis.

5.7.1. Analysis for Study 1

First, we excluded diseases with a prevalence of 2% or less at baseline to reduce statistical noise and thus prevent spurious findings in the models. The initial data set was composed of the information of each patient at each wave divided into the selected disease groups, so that the original data set was defined as $\boldsymbol{X} \coloneqq \{\boldsymbol{x}_1, \boldsymbol{x}_2, \dots, \boldsymbol{x}_N\}$, denoting by $\boldsymbol{x}_n \in \mathbb{R}^D$ for $n = 1, \dots, N$ the vector representing patient n out of the N total participants. We initially characterized each patient by a vector of binary variables that indicated the presence/absence of a disease group at each time. Since all selected features were categorical rather than quantitative variables, we preprocessed the data set by applying a multiple correspondence analysis (MCA), a data analysis technique for nominal categorical data that can detect and represent underlying structures in the data set. By using this method, researchers can represent in a multidimensional space a set of dichotomous or categorical variables (disease groups) that would be difficult to observe in contingency tables; in this way we formed groups of patients with the same characteristics (75).

MCA also enables direct representation of patients as points (coordinates) in geometric space, transforming the original binary data to continuous data (Figure 5.3). Our MCA was based on the indicator matrix. We inspected the optimal number of dimensions and percentages of inertia using a scree plot. We applied the Karlis-Saporta-Spinaki rule to select the extracted dimensions, according to the eigenvalues of the MCA and the number of features and individuals in the data set (135). To reduce the dimensionality, we used the MCA method included in the PCAmix algorithm, as described by Chavent et al. (136). This new data set was defined as $\boldsymbol{\mathcal{Y}} := \{\boldsymbol{y}_1, \boldsymbol{y}_2, \dots, \boldsymbol{y}_N\}$, with $\boldsymbol{y}_n \in \mathbb{R}^d$ for $n = 1, \dots, N$ denoting the new vector representing patient n.

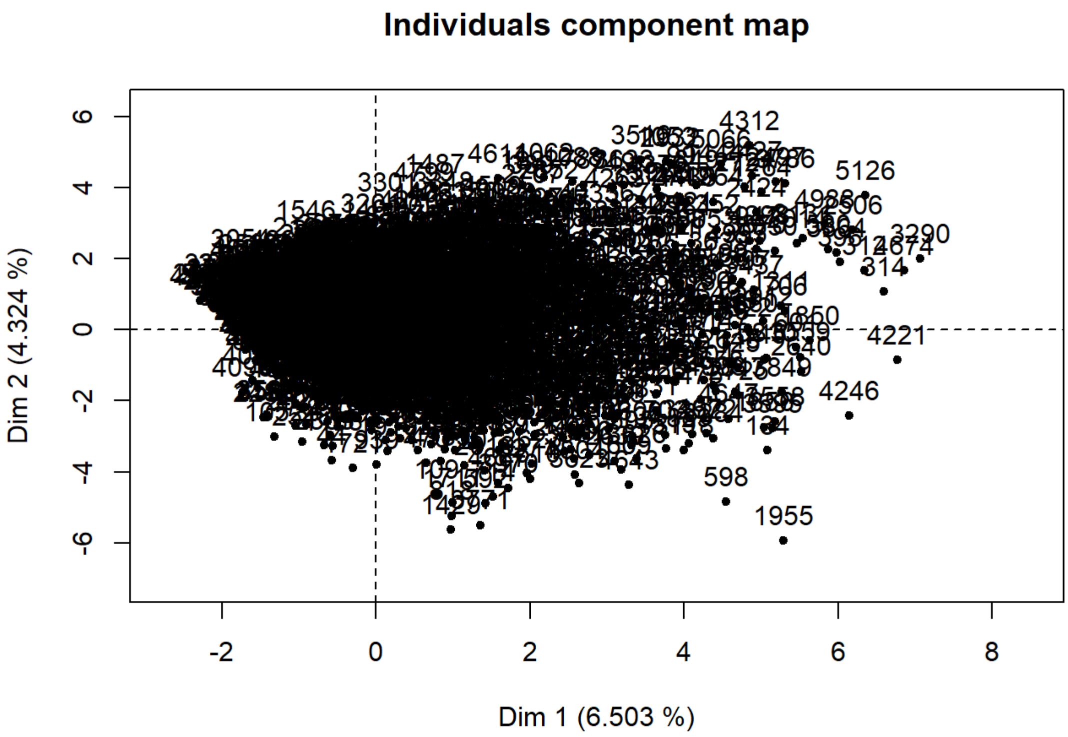


Figure 5.3: PCAmix first 2 dimensions

After computing the transformed data set $\boldsymbol{\mathcal{Y}}$, we identified multimorbidity patterns using the fuzzy c-means cluster analysis algorithm, which belongs to the family of soft clustering algorithms. The algorithm estimates c cluster centers (similar to k-means) but with fuzziness, so that individuals may belong to more than one pattern.

Originally introduced by Bezdek (137), the fuzzy c-means algorithm yields an unsupervised form of grouping in which individuals can belong to more than one cluster. To do this, the model associated individuals with an appropriate set of K membership values, where K denotes the number of clusters. The parameters that determine the clustering process are a set of K centroids $\mathbf{V} = \{\boldsymbol{v}_1, \dots, \boldsymbol{v}_K\}$ where $v_k \in \mathbb{R}^d$ for $k = 1, \dots, K$ and a set of membership factors $\mathbf{U} = \{u_{jn}; j = 1, \dots, K; n = 1, \dots, N\}$ with $0 \leq u_{jn} \leq 1$. Factor u_{jn} indicates the degree to which individual n^{th} belongs to cluster j^{th}. Both centroids $\mathbf{V}$ and membership factors $\mathbf{U}$ are obtained by iteratively minimizing the objective function $J_m(\mathbf{U}, \mathbf{V}, \boldsymbol{\mathcal{Y}})$, which is the weighted sum of squared errors within clusters:

$$J_m(\mathbf{U}, \mathbf{V}, \boldsymbol{\mathcal{Y}}) = \sum_{n=1}^{N} \sum_{j=1}^{K} \left(u_{jn}\right)^m \left\|\boldsymbol{y}_n - \boldsymbol{v}_j\right\|^2; \quad 1 < m < \infty \quad (1)$$

The fuzziness weighting parameter m is selected to adjust the blending of the different clusters; it can be any real number greater than 1. High m values produce a fuzzy cluster set, so that individuals tend to be equally distributed across clusters, whereas lower m values generate a non-overlapped set of clusters, similar to hard clustering.

Since clustering algorithms are unsupervised techniques, the model fitting is traditionally computed through cost functions that depend on both the data set and the clustering parameters, and that are denoted as validation indices. We computed different well-known validation indices to obtain the optimal number of clusters K and the optimal value of the fuzziness parameter m. Methods included were the Fukuyama index, Xie-Beni index, Partition coefficient index, Partition entropy index and Calinski-Harabasz index (138). The decision rules for each index were as follows:

1. Optimum Fukuyama index has to be minimum.
2. Optimum Xie-Beni index has to be minimum.
3. Optimum Partition coefficient index has to be maximum.
4. Optimum Partition entropy index has to be minimum.
5. Calinski-Harabasz index has to be maximum.

Different degrees of fuzzification $m = 1.1, 1.2, 1.4, 1.5, 2, 4$ and number of clusters $K = 2, .., 20$ were considered to estimate the optimal number of clusters (Figure 5.4).

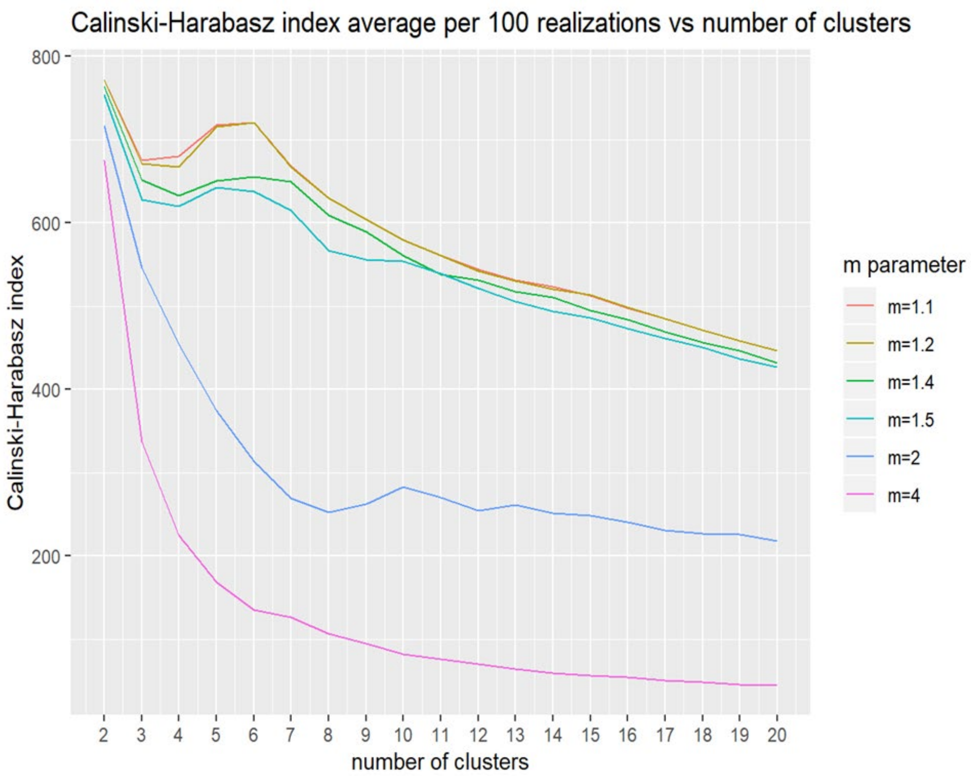


Figure 5.4: Calinski-Harabasz validation index

Given the stochastic nature of the clusters, we ran 100 independent clustering repetitions to obtain the average final solution. To evaluate the consistency and utility of the final clusters, we evaluated the clinical relevance of the findings in the context of previous literature and discussed the findings within the research team (two primary care physicians, two geriatricians, three epidemiologists and two statisticians).

For cross-validation of the model, we randomly sorted participants into two independent data sets and compared their validation indices. Indices were computed and averaged over 100 repetitions.

To examine the disease patterns characterizing each cluster, we used the observed/expected $(O/E)_{dj}$ ratio and the exclusivity ratio EX_{dj}, deciding whether each disease d was overrepresented in any given cluster j.

We calculated the $(O/E)_{dj}$ ratio by dividing disease prevalence in the cluster O_{dj} by disease prevalence in the overall population E_d. For the fuzzy c-means algorithm, we denoted membership of an individual n in a cluster j by a membership degree factor u_{nj}. We computed the observed disease prevalence O_{dj} in a cluster j as the ratio between the sum of the membership degree factors corresponding to all individuals with the disease d and the sum of all the membership degree factors corresponding to the cluster j. Assuming that there are n_d individuals with the disease d and that they are grouped in the set I_d, we computed the observed prevalence as

$$O_{dj} = \frac{\sum_{n \in I_d} u_{nj}}{\sum_{n=1}^{N} u_{nj}}$$

and the expected prevalence as

$$E_d = \frac{n_d}{N}$$

Therefore, the Observed/Expected ratio was

$$(O/E)_{dj} = O_{dj} / E_d = \frac{\sum_{n \in I_d} u_{nj}}{\sum_{n=1}^{N} u_{nj}} / \frac{n_d}{N}$$

Exclusivity ratio EX_{dj}, defined as the proportion of individuals with the disease d included in the cluster j over the total number of individuals with the disease n_d, was computed as

$$EX_{dj} = \frac{\sum_{n \in I_d} u_{nj}}{n_d}$$

We considered a disease to be associated with a given cluster when the O/E-ratio was 2 or greater, or the exclusivity was 25% or greater (33,139). In this way, we named multimorbidity patterns after the predominant diseases.

Lastly, we compared the clusters according to the distribution of sociodemographic, lifestyle, clinical and functional variables using analysis of variance (ANOVA) and chi-squared tests.

5.7.2. Analysis for Study 2

We applied the same clustering methodology of Study 1 to identify baseline clusters, 6-year clusters and 12-year clusters. We then evaluated the most likely clinical trajectories of the participants as they moved between clusters over time. Each individual was assigned to the cluster with the highest membership score at each time point. Due to the dynamism of the phenomenon, the names of the clusters changed over time, reflecting the evolving combinations of diseases that characterize them at each time point. We calculated shifts between clusters by cross-tabulating individuals between each wave (baseline to six-year follow-up and six-year to 12-year follow-up) after forcing the individuals into the cluster where they were more likely to belong. We computed frequencies (percentages) of participants who changed from one cluster to another to assess the overlap between waves. Mortality and dropout status were considered as fixed clusters at six-year follow up and at 12-years follow-up.

To estimate the association between clusters and mortality, we fitted logistic regression models adjusted by age, sex and education, using the unspecific cluster as the reference group. We adjusted odds ratios (ORs) and 95% CIs for age, sex and education, and we adjusted all comparisons for multiplicity. When the explanatory variable was normally distributed, we used the Tukey method; otherwise, we used the Benjamini and Hochberg method.

5.7.3. Analysis for Study 3

We used the analyses performed in the first two studies for Study 3, but also considered the 3-year and 9-year follow-up data for the participants aged over 80 years, including in the analysis all diseases that achieved a median prevalence of 2% across all follow-up waves.

For the longitudinal analysis, the observed data was assumed to be a time series of discrete time, for instance, the n^{th} patient was represented by the observed time sequence $\boldsymbol{y}_n(t), t = 1, \ldots, T$. Therefore, to model the temporal evolution of patients through the different clusters or patterns, the sequential individual observations were assumed to follow a dynamic random process represented by a Hidden Markov Model (HMM), so that each cluster was associated with a hidden state or multimorbidity pattern, $x_n(t)$. This means that each patient followed a longitudinal trajectory over $T = 12$ years $t_n := \{x_n(1), \ldots, x_n(T)\}$, through the clusters (122). For example, the n^{th} patient could belong to cluster 1 at baseline, change to cluster 2 at six years, and evolve into cluster 3 at 12 years. In this case, their longitudinal trajectory would be $t_n = \{1, 2, 3\}$ (Figure 5.5).

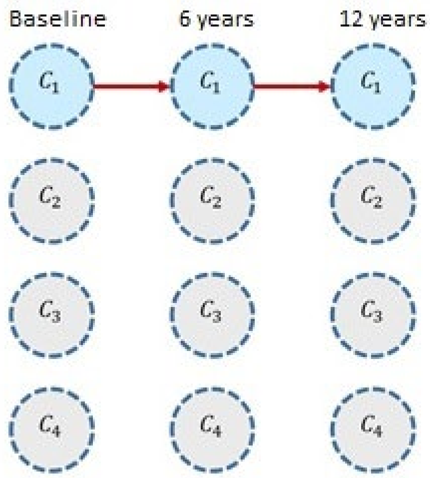


Figure 5.5: Individual trajectory in HMM

We adjusted the observed time sequences $\boldsymbol{y}_n(t), t = 1, \dots, T, n = 1, \dots, N$ to an HMM. In this process, the longitudinal trajectory vector $t_n := \{x_n(1), \dots, x_n(T)\}$ associated with the n^{th} patient plays the role of a latent variable, as there is no direct access to it, but it can be estimated once all the parameters of the model have been identified (Figure 5.6). Each observed vector $\boldsymbol{y}_n(t)$ is conditioned on the state of the corresponding latent variable $x_n(t)$.

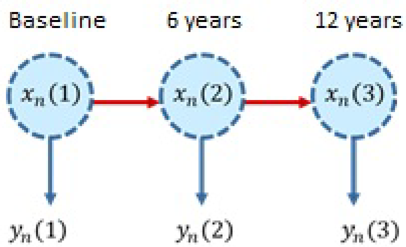


Figure 5.6: Latent variables in HMM

To develop the HMM, we considered all features of all individuals at each study wave. We estimated the HMM in two stages: first, we pre-processed the data set by applying an MCA to the categorical features to reduce the number of features on the new data set; and second, we applied an FCM on the new data set to identify an initial set of clusters. Additionally, we could include participants who died or dropped out in the model by including absorbing states. An absorbing state is a state that, once entered, cannot be left.

In the second stage, we estimated the following parameters of a first order HMM:

1) Initial state probability π_j, releated to j^{th} cluster
2) Transition probabilities, defined as $p_{ij} = \Pr\{x_n(t) = i|x_n(t-1) = j\}$, where p_{ij} is the probability that any patient jumps from the j^{th} group to the i^{th} group in a defined time
3) Parameters of the observed variables distribution $\mathcal{N}(\mathbf{m}_i; \mathbf{C}_i)$, where $\boldsymbol{m}_i$ is the mean vector and $\boldsymbol{C}_i$ is the covariance matrix associated with the hidden state $x_n(t) = i$

We fitted the set of HMM parameters into the observation data set by applying the Baum-Welch (BW) algorithm (123,124). The BW algorithm is well documented in the literature. It is a procedure that iteratively alternates between the expectation step (E-Step) and the maximization step (M-Step). It must be initialized by choosing starting values, for example, by using the centroids from FCM and randomly initializated transition probabilities. Once the algorithms have converged, the final set of model parameters are estimated. Therefore, the longitudinal trajectories $\{t_n; n = 1, \dots, N\}$ followed by the individuals can be inferred (Figure 5.7).

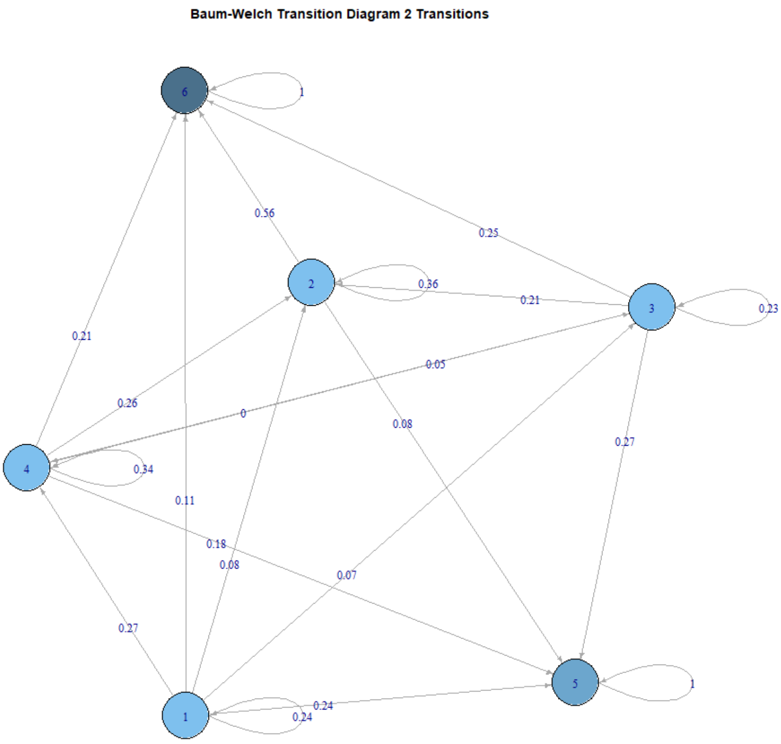


Figure 5.7: Markov Chain diagram

The best cluster trajectory is computed by maximizing the probability of the observed sequence conditioned to the set of final parameters. This problem is efficiently solved by applying the well-known Viterbi algorithm and repeating N times, one for each patient (140,141). To validate the model, we compared BW and Viterbi transition probabilities, finding a good agreement between theoretical and observed values.

The time unit considered for each transition across clusters/states was the time between follow-up waves: six years for sexagenarians and septuagenarians and three years for octogenarians and beyond. The time T_i spent in a specific cluster/state i before moving to other cluster/state j was assumed to follow a geometric probability distribution:

$$\Pr\{T_i = m\} = {p_{ii}}^{m-1}(1 - p_{ii}), m \geq 1$$

Subsequently, we computed the expected average time spent or mean sojourn (permanence) time as follows:

$$E(T_i) = \frac{1}{(1 - p_{ii})}$$

To optimize the performance of the selected mathematical model, we initialized the iterative process involved in the application of the BW algorithm using a range of 100 different values of the parameters to be learned. We selected the best model using a procedure that is equivalent to applying the Bayesion information criterion to choose the best set of HMM parameters.

For the HMM clusters characterization, we denoted membership u_{nj} of an individual n in a cluster j as a binary variable. We considered a disease to be associated with a given cluster when the O/E-ratio was 2 or greater and the exclusivity was relaxed to a 20% threshold.

Finally, we used linear mixed models to estimate the longitudinal trends of clinical and functional characteristics (number of chronic diseases, number of drugs, walking speed and MMSE) associated with the multimorbidity patterns, assuming a random intercept and including an interaction between the patterns and follow-up time both as linear and quadratic. We also adjuted the models by age, sex and education.

5.8. Statement of ethics

All studies were approved by the Regional Ethics Review Board in Stockholm, Sweden. Participants in the study completed and signed a written informed consent form as stipulated by the ethics board. For participants with prevalent or incident cognitive impairment, next of kin provided consent.

6. Results

The PhD thesis is based on 3 scientific articles. All of them have been published in scientific peer-reviewed journals with impact factor:

1. Marengoni A, **Roso-Llorach A***, Vetrano DL, Fernández-Bertolín S, Guisado-Clavero M, Violán C, Calderón-Larrañaga A. Patterns of Multimorbidity in a Population-Based Cohort of Older People: Sociodemographic, Lifestyle, Clinical, and Functional Differences. J Gerontol A Biol Sci Med Sci. 2020 Mar 9;75(4):798-805. doi: 10.1093/gerona/glz137. PMID: 31125398.

2. Vetrano DL, **Roso-Llorach A***, Fernández S, Guisado-Clavero M, Violán C, Onder G, Fratiglioni L, Calderón-Larrañaga A, Marengoni A. Twelve-year clinical trajectories of multimorbidity in a population of older adults. Nat Commun. 2020 Jun 26;11(1):3223. doi: 10.1038/s41467-020-16780-x. PMID: 32591506; PMCID: PMC7320143.

3. **Roso-Llorach A**, Vetrano DL, Trevisan C, Fernández S, Guisado-Clavero M, Carrasco-Ribelles LA, Fratiglioni L, Violán C, Calderón-Larrañaga A. 12-year evolution of multimorbidity patterns among older adults based on Hidden Markov Models. Aging (Albany NY). 2022 Nov 23;14. doi: 10.18632/aging.204395. Epub ahead of print. PMID: 36435509.

* First shared authorship.

6.1. Results by study

6.1.1. Study 1

In this first study, individuals were classified into six clusters using fuzzy c-means clustering algorithm. Around half of the SNAC-K cohort of multimorbid older adults were grouped into five clinically meaningful clusters, named psychiatric and respiratory diseases (PSY-RESP), heart diseases (HEART), respiratory and musculoskeletal diseases (RESP-MSK), cognitive and sensory impairments (CNS-IMP), and eye diseases and cancer (EYE-CANCER). These clusters showed significantly different sociodemographic, lifestyle, clinical, and functional profiles.

The PSY-RESP cluster was associated with higher values of alcoholism and neuroticism. The HEART cluster grouped people with the highest number of co-occurring chronic diseases and drug usage and the highest levels of serum creatinine and CRP. Individuals in the EYE-CANCER cluster exhibited the lowest muscle strength. The CNS-IMP cluster grouped people of very old ages who lived in nursing homes and had the lowest physical functional status.

The other half of the study population was grouped in a UNSPECIFIC cluster, that included the youngest people with the lowest mean number of chronic diseases, the best functional and cognitive and status, and the highest life satisfaction.

6.1.2. Study 2

In this second study, six clusters of individuals with multimorbidity were identified using fuzzy c-means clustering algorithm. There was a high heterogeneity in the multimorbidity clustering at baseline. Only half of the participants could be grouped into a well-characterized cluster: psychiatric and respiratory diseases, heart diseases, respiratory and musculoskeletal diseases, cognitive and sensory impairment, and eye diseases and cancer.

The other half of the participants were sorted into an unspecific cluster and were characterized by having a younger age, lower numbers of co-occurring diseases and drugs, good functional and cognitive abilities, and a high percentage of cardiovascular risk factors.

Over 12 years, changes in cluster composition, participants' transitions from one cluster to another, and participant mortality generated a dynamic but well-defined clinical picture.

The first remarkable trajectory involved the group of people part of the unspecific cluster at baseline. The number of participants grouped in this cluster halved at the 6- and 12-year follow-ups as the majority transitioned towards the specific multimorbidity clusters identified at follow-ups. Given the young age and less complex clinical picture of these individuals, they may be considered an at-risk population for developing more complex multimorbidity and as such potentially susceptible to preventive intervention.

The second relevant trajectory was the high mortality of individuals in clusters characterized by cardiovascular and neuropsychiatric diseases, which, despite representing only 25%, 28%, and 29% of the participants at baseline, 6 years, and 12 years, respectively, accounted for 51% and 57% of deaths during the first and second follow-up periods, respectively.

6.1.3. Study 3

In this third study, four longitudinal multimorbidity patterns were identified for each decade among older adults from the SNAC-K cohort using Hidden Markov Models. The time they spent in each pattern as well as the probability of transitioning across different patterns were estimated throughout a twelve-year follow-up period.

The findings highlight the dynamism and heterogeneity underlying multimorbidity. The dynamism among multimorbidity patterns was reflected by the varying sojourn times across patterns, which differed by age group, and the specific patterns people showed.

Individuals in all decades showed the shortest permanence time in an unspecific pattern lacking any overrepresented diseases (range: 4.6-10.9 years), but the pattern with the longest permanence time varied by age. Sexagenarians remained longest in the Psychiatric-endocrine and sensorial pattern (15.4 years); septuagenarians in the Neuro-vascular and skin-sensorial pattern (11.0 years); and octogenarians and beyond in the Neuro-sensorial pattern (8.9 years).

Transition probabilities varied across age groups. In general, sexagenarians showed the highest levels of stability, as the probabilities of staying in the same patterns were higher than in the other age groups.

An increasing trend was observed for the number of chronic conditions and drugs across age groups, with subjects in the Unspecific patterns consistently showing the lowest values. Conversely, a decreasing trend was observed for walking speed and MMSE in all age groups. While subjects in the Unspecific patterns, except for octogenarians, showed the slowest changes over time, those in the patterns characterized by cardiovascular and/or neurological diseases showed the worse baseline values and fastest declines.

6.2. Published studies

6.2.1. Study 1

Marengoni A, **Roso-Llorach A**, Vetrano DL, Fernández-Bertolín S, Guisado-Clavero M, Violán C, Calderón-Larrañaga A. Patterns of Multimorbidity in a Population-Based Cohort of Older People: Sociodemographic, Lifestyle, Clinical, and Functional Differences. J Gerontol A Biol Sci Med Sci. 2020 Mar 9;75(4):798-805. doi: 10.1093/gerona/glz137. PMID: 31125398.

Journals of Gerontology: Medical Sciences
cite as: *J Gerontol A Biol Sci Med Sci*, 2020, Vol. 75, No. 4, 798–805
doi:10.1093/gerona/glz137
Advance Access publication May 24, 2019

Research Article

Patterns of Multimorbidity in a Population-Based Cohort of Older People: Sociodemographic, Lifestyle, Clinical, and Functional Differences

Alessandra Marengoni, PhD,[1,2,*,†] Albert Roso-Llorach, MSc,[3,4,†] Davide L. Vetrano, MD, PhD,[1,5,6] Sergio Fernández-Bertolín, MSc,[3,4] Marina Guisado-Clavero, MD,[3,4] Concepción Violán, PhD,[3,4] and Amaia Calderón-Larrañaga, PhD[1]

[1]Aging Research Center, Department of Neurobiology, Care Sciences and Society, Karolinska Institutet and Stockholm University, Sweden. [2]Department of Clinical and Experimental Sciences, University of Brescia, Italy. [3]Fundació Institut Universitari per a la recerca a l'Atenció Primària de Salut Jordi Gol i Gurina (IDIAPJGol), Spain. [4]Universitat Autònoma de Barcelona, Bellaterra (Cerdanyola del Vallès), Spain. [5]Department of Geriatrics, Catholic University of Rome, Italy. [6]Centro di Medicina dell'Invecchiamento, Fondazione Policlinico "A. Gemelli," Scientific Institute for Research and Healthcare (IRCCS), Rome, Italy.

*Address correspondence to: Alessandra Marengoni, PhD, Department of Clinical and Experimental Sciences, University of Brescia, Viale Europa 11, 25123 Brescia, Italy. E-mail: alessandra.marengoni@unibs.it
[†]First shared authorship.

Received: February 28, 2019; Editorial Decision Date: May 20, 2019

Decision Editor: Anne Newman, MD, MPH

Abstract

Background: The aim of this study is to identify clusters of older persons based on their multimorbidity patterns and to analyze differences among clusters according to sociodemographic, lifestyle, clinical, and functional characteristics.
Methods: We analyzed data from the Swedish National Study on Aging and Care in Kungsholmen on 2,931 participants aged 60 years and older who had at least two chronic diseases. Participants were clustered by the fuzzy *c*-means cluster algorithm. A disease was considered to be associated with a given cluster when the observed/expected ratio was ≥2 or the exclusivity was ≥25%.
Results: Around half of the participants could be classified into five clinically meaningful clusters: respiratory and musculoskeletal diseases (RESP-MSK) 15.7%, eye diseases and cancer (EYE-CANCER) 10.7%, cognitive and sensory impairment (CNS-IMP) 10.6%, heart diseases (HEART) 9.3%, and psychiatric and respiratory diseases (PSY-RESP) 5.4%. Individuals in the CNS-IMP cluster were the oldest, with the worst function and more likely to live in a nursing home; those in the HEART cluster had the highest number of co-occurring diseases and drugs, and they exhibited the highest mean values of serum creatinine and C-reactive protein. The PSY-RESP cluster was associated with higher levels of alcoholism and neuroticism. The other half of the cohort was grouped in an unspecific cluster, which was characterized by gathering the youngest individuals, with the lowest number of co-occurring diseases, and the best functional and cognitive status.
Conclusions: The identified multimorbidity patterns provide insight for setting targets for secondary and tertiary preventative interventions and for designing care pathways for multimorbid older people.

Keywords: Multimorbidity pattern, Older adults, Swedish National Study on Aging and Care in Kungsholmen (SNAC-K)

Since the beginning of the last century, chronic diseases have progressively replaced infectious diseases in terms of their prevalence and impact on human health. Caring for people with chronic conditions has emerged as one of the major challenges facing health care systems, which remain rooted in episodic and acute care. The worldwide aging phenomenon, along with individuals' longer survival following formerly fatal events, are the main drivers of the increasing burden of chronic diseases. As a consequence, the prevalence of *multimorbidity*, defined as the coexistence of two or more chronic diseases in the same person, is as high as 90% in older adults (1).

Despite the increasing number of studies on the occurrence of multimorbidity across age, gender, and socioeconomic strata, its epidemiology remains poorly understood. In fact, given the wide heterogeneity of people suffering from multimorbidity, no single definition or operationalization seems to serve both research and clinical purposes effectively. For example, the exclusive use of a quantitative approach (ie, the number of co-occurring chronic diseases) fails to capture the clustering of chronic diseases in patterns of multimorbidity (2). Studies attempting to describe multimorbidity patterns have used different methodological approaches to address this issue, such as estimation of observed to expected ratios or odds ratios among the most commonly coexisting dyads or triads of chronic conditions, or cluster and factor analyses to identify systematic groupings among diseases. However, these statistical techniques limit interpretation of results and clinical applicability in, for example, their need for large samples, multiple comparisons, overestimation of effect sizes, and the forcing of diseases into single clusters according to similarity or dissimilarity measures. Moreover, previous studies have focused on identifying patterns of diseases rather than clusters of individuals (3), which has prevented researchers from characterizing such patterns in terms of their clinical and social significance. Although some studies have described multimorbidity patterns in terms of their associated burden of polypharmacy (4) or hospital care (5), other individual-level characteristics such as sociodemographic, lifestyle, clinical, and physical and cognitive functional measures have not yet been explored.

In the present study, we aim to build on previous work by applying a *soft* clustering technique (ie, the fuzzy *c*-means cluster algorithm) to analyze patterns of multimorbidity in a population-based Swedish cohort study of older people. *Soft* techniques (*c*-means) present the following advantages over the *hard* clustering algorithms (in other words, hierarchical clustering, *k*-means) predominantly used in past studies. First, individuals, and not diseases, are grouped in clusters according to their commonly co-occurring diseases. Second, instead of forcing individuals to belong to one specific cluster, participants are assigned a probability of membership in all identified clusters, which makes more sense from a biological perspective. Finally, one disease can characterize more than one cluster, which allows us to build patterns of multimorbidity that take all possible disease combinations into account (6,7). In summary, by using soft clustering techniques, we place individuals and not their diseases at the center of our analyses (8).

The specific objectives of our study are (i) to identify clusters of older people based on their multimorbidity patterns and (ii) to analyze differences among clusters according to sociodemographic, lifestyle, clinical, and functional characteristics.

Methods

Study Population

We used baseline data from the population-based Swedish National Study on Aging and Care in Kungsholmen (SNAC-K) (9). This study consists of community-dwelling and institutionalized older adults aged 60 years and older. Of people born between 1898 and 1943, living in the Kungsholmen district of Stockholm (Sweden), a random sample from 11 age cohorts was invited to participate in the study. Those who accepted were evaluated between 2001 and 2004 for the first time and subsequently followed up every 6 years (those aged <78 years) or every 3 years (those aged ≥78 years). At baseline, 3,363 people were examined (participation rate, 73%). In our study, 432 participants were excluded because they did not fulfill the inclusion criteria of having multimorbidity (ie, two or more chronic diseases) at baseline. As expected, those excluded were younger, more educated, and less likely to be female than those included in the study ($p < .001$).

Study Variables

At each study wave, SNAC-K participants undergo a comprehensive clinical and functional assessment by trained physicians, nurses, and neuropsychologists. Physicians collect information on diagnoses via physical examination, medical history, examination of medical charts, self-reported information, and/or proxy interviews. Clinical parameters, lab tests, medication, and inpatient and outpatient care data are also used to identify specific conditions. All diagnoses are coded according to the International Classification of Diseases 10th revision (ICD-10) and classified into 60 chronic disease categories in accordance with a clinically driven methodology (1). Diseases with a prevalence of <2% were excluded to avoid statistical noise and therefore spurious findings in the models. Drugs are coded in accordance with the Anatomical Therapeutic Chemical (ATC) classification.

Participants' demographics (ie, age, sex, education, occupation, living arrangement, and civil status) and lifestyle factors are collected during nurse interviews. Educational attainment was categorized as elementary, high school, and university or higher; main occupation was categorized as manual or non-manual based on the longest job held during the person's lifetime. Civil status was categorized as unmarried, married, divorced, and widowed; smoking was categorized as never, former, and current; and alcohol consumption was categorized as never/occasional, light/moderate, and heavy. Following the recommendations of the World Health Organization (WHO) and the American College of Sports Medicine (ACSM), participants were categorized in three different groups according to the intensity of their physical activity: inadequate (less than or equal to two to three times per month of light and/or moderate/intense exercise), health-enhancing (light exercise several times per week or every day), and fitness-enhancing (moderate/intense exercise several times per week or every day) (10,11).

Life satisfaction was measured using the self-reported index developed by Neugarten and colleagues (12) (LSI-A), which captures five components: zest versus apathy, resolution and fortitude, congruence between desired and achieved goals, positive self-concept, and mood. The LSI-A consists of 12 positive and 8 negative items; in this study, the negative items were reversed and the final scores transformed to a 0–100 scale with higher values indicating greater life satisfaction (13). The social network index combined indicators of self-reported social connections and social support according to the procedure adopted in the National Social Life, Health, and Aging Project (NSHAP Study) (14) and was subsequently categorized into tertiles labeled as poor, moderate, or rich (15). Self-rated health was assessed by asking participants; "In general, how would you describe your health?" and categorized as very good/excellent and good/poor.

Level of disability was measured as the number of basic activities of daily living (bathing, dressing, toileting, continence, transferring, and eating) and instrumental activities of daily living (grocery shopping, meal preparation, housekeeping, laundry, managing money, using the telephone, taking medications, and using public transportation) a person was unable to perform independently. People living in institutions were assumed to depend on others for grocery shopping, meal preparation, housekeeping, and laundry. Balance was measured as the time (in seconds) a participant could stand on one leg (up to 60 seconds). Grip strength was measured with a dynamometer and converted to kilograms. Participants were seated with their arm resting on a table and their elbow flexed at 90 degrees. Walking speed was assessed by asking participants to walk 6 m, or 2.44 m if the participant reported walking slowly. If the participant was unable to walk or attempted unsuccessfully to walk, a value of 0 was recorded. Cognitive status was assessed by physicians with the Mini–Mental State Examination, which ranges from 30 to 0 (from best to worst possible score). Participants' serum albumin (g/L), creatinine (μmol/L), and C-reactive protein (mmol/L) levels were measured at Karolinska Institutet's laboratory following standard procedures.

Statistical Analysis

Multimorbidity patterns were identified using the fuzzy c-means cluster analysis algorithm, which belongs to the family of soft clustering algorithms. The algorithm estimates *c* cluster centers (similar to *k*-means) but with fuzziness, so that individuals may belong to more than one pattern. We used the technique to obtain clusters of individuals as well as a membership matrix that indicated the degree of participation of each subject in each cluster. Through dimensionality reduction (that is, multiple correspondence analysis) we then obtained the input data for the clustering of participants. To determine the number of retained dimensions, the Karlis-Saporta-Spinaki rule was used (16). Different degrees of fuzzification and several validation indices were considered to estimate the optimal number of clusters (7). Given the stochastic nature of the clusters, we ran 100 independent clustering repetitions to obtain the average final solution. The consistency and significance of the final solution was evaluated based on clinical criteria. For cross-validation of the model, we randomly sorted individuals into two independent data sets and compared their validation indices. Indices were computed and averaged over 100 repetitions.

To examine the disease patterns characterizing each cluster, observed/expected ratios were calculated by dividing the prevalence of a given disease within a cluster by its prevalence in the overall population. *Disease exclusivity,* defined as the fraction of participants with the disease included in the cluster over the total number of participants with the disease, was also calculated. A disease was considered to be associated with a given cluster when the observed/expected ratio was ≥2 or the exclusivity was ≥25% (17,18). The clusters were further compared for the distribution of sociodemographic, lifestyle, clinical, and functional variables using analysis of variance and chi-square tests. Statistical analyses were performed using R 3.5.1 and Stata 15.

Results

The study population consisted of 2,931 individuals. The participants' mean age was 76.1 years, and 66.6% were female. Six point five percent were living in a nursing home. The mean number of chronic condition was 4.5, and the mean number of drugs was 4.4. A total of 39 chronic disease categories had a prevalence of ≥2% and were included in the cluster analyses (Table 1).

The following multimorbidity patterns were detected in our population: psychiatric and respiratory diseases (PSY-RESP) 5.4%, heart diseases (HEART) 9.3%, eye diseases and cancer (EYE-CANCER) 10.7%, cognitive and sensory impairments (CNS-IMP) 10.6%, and respiratory and musculoskeletal diseases (RESP-MSK) 15.7%. Around half of the study population (48.4%) could not be classified into any of the abovementioned patterns but constituted a cluster where nonspecific chronic conditions were over-represented, and which was named UNSPECIFIC. A clinical description of the patterns in terms of diseases with the highest observed/expected ratio and exclusivity values is reported in Supplementary Table S1. In addition, Figure 1 depicts all diseases with observed/expected ratios ≥ 2 and/or exclusivity values ≥ 25% in each cluster. The PSY-RESP cluster included individuals with neurotic, stress-related and somatoform disorders, depression, sleep disorders (both insomnia and obstructive sleep apnea), and other unspecified neurological and psychiatric conditions; it also included asthma. The HEART cluster included several cardiac diseases along with cerebrovascular disease, diabetes, migraine, and inflammatory arthropathies. The EYE-CANCER cluster included several eye impairments and solid cancers. The CNS-IMP cluster included dementia, psychiatric and cerebrovascular diseases, and visual and hearing problems. The RESP-MSK cluster included the two most frequent respiratory diseases (ie, asthma and chronic obstructive pulmonary disease) and obstructive sleep apnea.

In Table 2, the five clusters are compared with the UNSPECIFIC one in terms of sociodemographic, lifestyle, clinical, and functional characteristics. The PSY-RESP cluster was associated with higher values of alcoholism and neuroticism. The HEART cluster grouped people with the highest number of co-occurring chronic diseases and drug use and the highest levels of serum creatinine and C-reactive protein. Individuals in the EYE-CANCER cluster exhibited the lowest muscle strength. The CNS-IMP cluster grouped people of very old ages who lived in nursing homes and had the lowest physical functional status. Finally, the part of the population not classified in any specific cluster included the youngest people with the lowest mean number of chronic diseases, the best functional and cognitive and status, and the highest life satisfaction.

Discussion

In the present study, around half of a Swedish cohort of older adults could be classified into five clinically meaningful clusters, named psychiatric and respiratory diseases (PSY-RESP), heart diseases (HEART), respiratory and musculoskeletal diseases (RESP-MSK), cognitive and sensory impairments (CNS-IMP), and eye diseases and cancer (EYE-CANCER). These clusters showed significantly different sociodemographic, lifestyle, clinical, and functional profiles. The other half of the study population was grouped in an unspecific cluster characterized by being younger, having lower numbers of co-occurring diseases and drug use, and good functional and cognitive abilities.

The PSY-RESP cluster included people with asthma along with psychiatric conditions. The co-occurrence of these diseases could be a result of chronic drug treatment with steroids, which can increase neuroticism, depression, and sleep disorders (19). Besides, asthma symptoms have been associated with depression, even in older participants (20). This cluster grouped relatively young people with alcohol abuse problems and low life satisfaction. The association between alcohol use and psychiatric disorders is well known (21), and this study confirms such an association in older persons. Poor quality of life in people affected by psychiatric and respiratory disorders has also been reported previously (22,23).

The HEART cluster illustrates the well-known link between cardio- and cerebrovascular diseases; atrial fibrillation and heart failure are both risk factors for stroke (24), and diabetes is a risk

Table 1. Disease Prevalence at Baseline in the Swedish National Study on Aging and Care in Kungsholmen (*N* = 2,931)

Rank	Chronic Conditions	Prevalence (%)
1	Hypertension	73.29
2	Dyslipidemia	50.12
3	Chronic kidney diseases	37.84
4	Ischemic heart disease	17.50
5	Colitis and related diseases	14.43
6	Osteoarthritis and other degenerative joint diseases	14.23
7	Anemia	13.68
8	Deafness, hearing impairment	13.07
9	Obesity	13.07
10	Heart failure	12.04
11	Thyroid diseases	11.84
12	Atrial fibrillation	11.02
13	Dementia	10.85
14	Depression and mood diseases	10.44
15	Solid neoplasm	10.10
16	Diabetes	9.96
17	Cerebrovascular disease	8.94
18	Osteoporosis	7.68
19	Other musculoskeletal and joint diseases	7.44
20	Dorsopathies	7.30
21	Asthma	6.86
22	Glaucoma	6.38
23	Cataract and other lens diseases	6.24
24	Other eye diseases	5.70
25	Chronic obstructive pulmonary disease, emphysema, chronic bronchitis	5.66
26	Autoimmune diseases	5.12
27	Esophagus, stomach, and duodenum diseases	4.95
28	Blindness, visual impairment	4.91
29	Inflammatory arthropathies	4.57
30	Prostate diseases	4.50
31	Other cardiovascular diseases	3.92
32	Neurotic, stress-related, and somatoform diseases	3.55
33	Other genitourinary diseases	2.87
34	Cardiac valve diseases	2.83
35	Migraine and facial pain syndromes	2.59
36	Other psychiatric and behavioral diseases	2.52
37	Sleep disorders	2.39
38	Other neurological diseases	2.18
39	Bradycardias and conduction diseases	2.12

factor for stroke and coronary heart disease (25). The high prevalence of migraine in this cluster may be explained by either brain vascular pathology or by the drugs prescribed for cardiac diseases, such as nitrates (26,27). The high number of co-occurring chronic conditions and drugs were distinctive of individuals belonging to this cluster. This cluster had the second highest percentage of persons, after the CNS-IMP cluster, with limitations in activities of daily living and instrumental activities of daily living. Individuals in this cluster also showed the highest serum creatinine and C-reactive protein levels. Expression of proinflammatory cytokines increases throughout the human life span, and this increase is correlated with cardiovascular health (28). Chronic low-grade inflammation, in turn, promotes autonomic imbalance, stimulates remodeling, depresses cardiac function, prompts endothelial dysfunction, and leads to a progression of atherosclerosis and impaired renal function (29).

The RESP-MSK cluster included osteoporosis possibly related to the chronic treatment of asthma and chronic obstructive pulmonary disease with steroids (30). Vitamin D deficiency could also underlie both respiratory and skeletal disorders; vitamin D supplements are beneficial both in preventing exacerbations of chronic obstructive pulmonary disease and improving bone density measures (31,32). The presence of upper gastrointestinal system disorders can also be related to the treatment of respiratory diseases and the use of diphosphonates in osteoporosis (33), whereas thyroid and other autoimmune diseases are often correlated (34).

The EYE-CANCER cluster included several eye impairments and solid cancers. A high percentage of participants in this cluster were widowed, which is explained by their higher age. Old age may also explain why they have the lowest grip strength.

The CNS-IMP cluster brings to light the association recently found between sensorial impairment and dementia. Hearing deficits have attracted much interest, motivated by strong evidence that impaired hearing is a risk factor for cognitive decline and dementia (35,36). The relationship between vision loss and dementia has been evaluated in cross-sectional and longitudinal studies; the 3C cohort study suggested that poor vision, in particular near vision loss, may be an indicator of dementia risk at short and middle term (37). Retinal microvasculature pathology has been associated with vascular dementia, especially in persons with diabetes (38). Multiple sensorial impairments have also been found to increase dementia

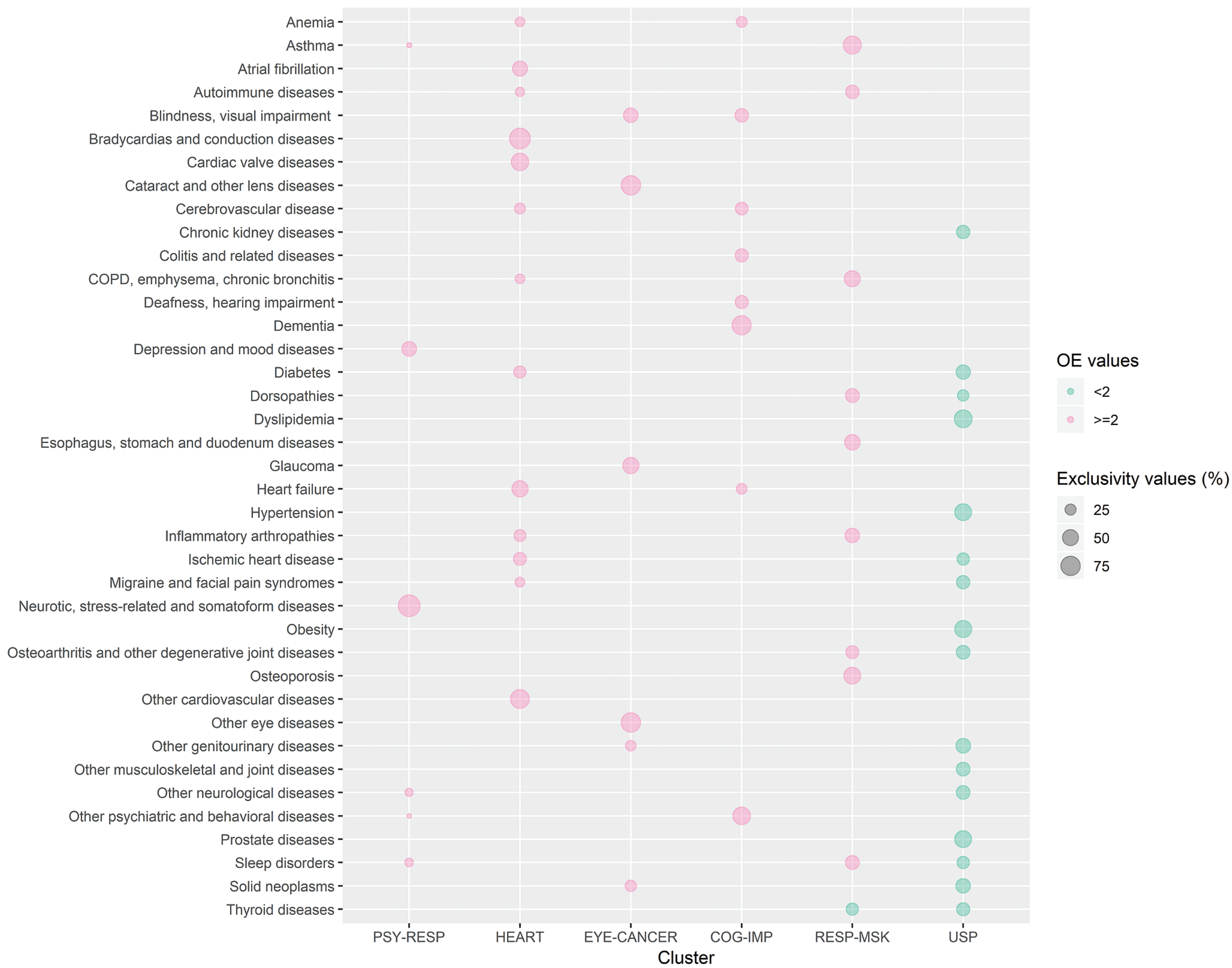


Figure 1. Chronic diseases characterizing clusters of older people identified at baseline in the Swedish National Study on Aging and Care in Kungsholmen (*N* = 2,931). PSY-RESP: psychiatric and respiratory diseases; HEART: heart diseases; EYE-CANCER: eye diseases and cancer; CNS-IMP: cognitive and sensory impairments; RESP-MSK: respiratory and musculoskeletal diseases; O/E ratio: observed/expected ratio.

risk (39). Persons in this cluster were very old and had the worst levels of physical and cognitive function; this justifies why 43% of them were living in a nursing home. Any disease in the cluster could explain the functional impairment, particularly dementia and cerebrovascular diseases (40). This cluster grouped the highest percentage of people who were manual workers with low education. Low educational attainment and a manual occupation during early life have been consistently associated with an increased risk of dementia (41) and poor income in later life. Finally, this cluster was also characterized by having the highest percentage of persons with a poor social network and inadequate physical activity levels.

Clinical and Public Health Implications

Despite the high prevalence of multimorbidity in the older population, knowledge about how chronic diseases co-occur in single individuals is limited. Furthermore, findings from different studies are hardly comparable with the literature in the field because of differing methodological approaches.

In a previous study from an older Swedish population, the coexistence of diseases was evaluated with a cluster analysis approach (2). Cluster analysis groups diseases according to their similarity forcing each disease to be part of one single cluster; this approach can be particularly useful to generate new research hypotheses on the pathophysiological correlations as well as the strength of causal associations between diseases. Conversely, the soft technique employed in the present study has the main advantage to group individuals who, according to their commonly co-occurring diseases, belong to different multimorbidity patterns, enabling one same disease to belong to more than one cluster. In addition, individuals are provided with a probability to belong to each of the identified clusters, and the most probable membership was investigated in our analyses. This methodology can be advantageous to describe the overall health status of a specific population, providing particularly useful information from a clinical point of view.

First, groups of people at high risk of adverse health outcomes can be identified and may benefit from targeted secondary and tertiary preventative interventions. For example, individuals in the HEART cluster may develop disabilities for a number of reasons, such as dyspnea in heart failure, stroke sequelae, and/or peripheral atherosclerosis and neuropathy from diabetes. Yet a correct identification of the exact cause of the functional limitations could be particularly important to this group of individuals to plan correct measures and

Table 2. Sociodemographic, Lifestyle, Clinical, and Functional Differences Among Clusters of Older People at Baseline in Swedish National Study on Aging and Care in Kungsholmen (*N* = 2,931)

	PSY-RESP	HEART	EYE-CANCER	CNS-IMP	RESP-MSK	UNSPECIFIC	ALL
n (%)	159 (5.4)	272 (9.3)	313 (10.7)	309 (10.6)	460 (15.7)	1,418 (48.4)	2,931
Sociodemographic factors							
Female sex (%)	74.2	59.2	72.5	76.9	74.4	61.0	66.6
Age, mean	73.3	82.3	83.2	88.2	73.9	71.8	76.1
Living in a nursing home (%)	5.8	6.0	4.1	43.2	1.8	0.7	6.5
Civil status (%)							
Unmarried	17.0	16.2	16.1	21.1	17.8	16.4	17.1
Married	35.2	35.5	25.1	25.8	38.4	49.5	40.6
Divorced	20.9	8.2	11.2	7.5	16.8	13.0	12.8
Widow	26.9	40.1	47.6	45.6	27.0	21.1	29.5
Education (%)							
Elementary	16.9	25.1	21.8	34.8	19.3	14.4	19.1
High school	47.6	56.1	56.5	47.4	49.5	50.1	50.9
University	35.5	18.8	21.8	17.7	31.1	35.5	30.0
Occupation (%)							
Manual worker	19.1	31.7	33.0	39.8	24.0	21.5	25.6
Non-manual worker	80.9	68.3	67.0	60.2	76.0	78.5	74.4
Life satisfaction score, mean	46.5	49.6	53.3	48.2	57.3	60.1	57.1
Social network (%)							
Poor	39.7	50.1	43.2	67.7	31.8	27.0	34.9
Moderate	35.8	24.4	36.5	21.3	35.4	35.8	33.8
Rich	24.5	25.5	20.3	11.0	32.9	37.2	31.2
Lifestyle factors							
Smoking (%)							
Never	43.6	45.4	59.2	60.1	45.5	46.7	48.7
Former	35.1	44.1	31.2	30.9	39.7	38.3	37.5
Current	21.3	10.6	9.6	9.0	14.8	15.0	13.8
Alcohol consumption (%)							
Never/occasional	39.7	52.7	51.2	77.6	40.4	30.1	40.8
Light/moderate	36.9	37.2	35.0	17.3	42.1	53.1	43.8
Heavy	23.4	10.1	13.8	5.1	17.5	16.9	15.3
Physical activity (%)							
Inadequate	42.4	56.0	44.7	82.9	31.1	22.7	36.9
Health-enhancing	40.2	36.3	45.8	14.4	50.1	52.5	45.2
Fitness-enhancing	17.4	7.7	9.6	2.6	18.8	24.7	17.9
Clinical and functional factors							
Self-rated health (%)							
Very good/excellent	15.4	5.1	23.0	14.8	23.6	43.3	32.5
Good/poor	84.6	94.9	77.0	85.2	76.4	56.7	67.5
Chronic conditions, mean	5.7	7.7	6.0	5.5	4.7	3.2	4.5
Drugs, mean	6.2	7.7	5.0	6.1	5.3	2.8	4.4
Serum albumin (g/L), mean	41.5	40.5	40.0	38.7	40.9	42.0	41.2
Serum creatinine (umol/L), mean	86.1	107.5	95.8	98.6	87.7	87.3	90.9
Serum CRP (mmol/L), mean	6.5	8.7	6.9	8.6	7.3	6.1	6.8
ADL + IADL limitations, mean	1.3	2.1	1.4	7.2	0.7	0.3	1.4
Balance test (s), mean	22.9	9.2	9.6	4.3	22.7	30.1	22.6
Grip strength test (N), mean	22.2	22.7	20.3	23.7	22.2	26.7	24.7
Walking speed (m/s), mean	0.9	0.6	0.7	0.3	0.9	1.1	0.9
MMSE test, mean	27.7	27.6	27.7	16.6	28.6	28.8	27.4

Notes: PSY-RESP = psychiatric and respiratory diseases; HEART = heart diseases; EYE-CANCER = eye diseases and cancer; CNS-IMP = cognitive and sensory impairments; RESP-MSK = respiratory and musculoskeletal diseases; CRP = C-reactive protein; ADL = activities of daily living; IADL = instrumental activities of daily living; MMSE = Mini–Mental State Examination. Missing values (%): age (0.3), civil status (0.5), education (1.1), occupation (2.7), life satisfaction score (39.2), social network (11.6), smoking (3.3), alcohol consumption (3.2), self-rated health (31.4), drugs (0.2), serum albumin (8.6), serum creatinine (8.6), serum CRP (10.0), ADL + IADL limitations (3.8), balance test (10.4), grip strength (25.1), walking speed (4.0), MMSE test (6.8). The distribution of all variables was significantly different across clusters ($p < .001$).

prevent disability. People in the CNS-IMP cluster could be systematically screened for functional impairment; if present, physical rehabilitation could be prescribed that might delay the progression to disability. Although no improvement has been shown in cognitive functions, there is promising evidence that exercise programs may improve the ability to perform activities of daily living for people with dementia (42). People in the PSY-RESP cluster may benefit from specific interventions designed to reduce alcohol abuse and subsequently improve their quality of life. Individuals in the EYE-CANCER cluster, characterized by a low muscle strength, may be screened for sarcopenia, given the known association between low muscle mass and chemotoxicity (43).

Second, care management could be improved for people with specific patterns characterized, for example, by polypharmacy and therefore a high frequency of potentially inappropriate medication and adverse drug reactions (44). Recent guidelines specifically developed for people with multimorbidity underline the potential treatment burden for patients prescribed a high number of drugs (45), such as those in the HEART cluster. In fact, certain therapeutic regimens that are appropriate for diseases affecting people in middle adulthood could be associated with the development of specific patterns of co-occurring diseases in late life. Some examples are chronic steroid treatment for respiratory diseases and the development of skeletal and psychiatric disorders (30), or chronic treatment with anticholinergic drugs for psychiatric diseases and the development of dementia in older age (46).

Finally, the group of people included in the UNSPECIFIC cluster is particularly interesting from both a research and prevention point of view. In fact, people in this pattern, despite suffering from two or more chronic diseases, were relatively younger, suggesting that aging itself is the main driver of disease clustering. This finding strengthens the idea that aging and multimorbidity share pathophysiological mechanisms (28) and that their connection is more evident when we analyze not only the number but also the patterns of diseases. Furthermore, the identification of this group of people is fundamental to plan interventions for the primary prevention of disease accumulation and to distribute health care resources accordingly.

Strengths and Limitations

The main strength of this study is the statistical technique, applied to allow clustering individuals according to their co-occurring diseases. The fuzzy *c*-means cluster algorithm is used for pattern recognition when clusters tend to overlap, which is most often the case in older adults. Other strengths are the high number of very old people in the cohort and the comprehensive list of both mental and physical chronic conditions included in the analyses. Limitations include the cross-sectional design of the study and the average high socioeconomic status of participants in SNAC-K, which limits the external validity of the findings.

Conclusion

In the present study, half of a cohort of older adults could be classified into five clinically meaningful clusters. These clusters showed significantly different sociodemographic, lifestyle, clinical, and functional profiles. This and similar approaches to the epidemiological study of multimorbidity are needed, not only to better understand the complex interactions among co-occurring diseases but also, even more importantly, to improve preventive interventions and optimally address individuals' care needs and the risk of adverse outcomes.

Supplementary Material

Supplementary data are available at *The Journals of Gerontology, Series A: Biological Sciences and Medical Sciences* online.

Funding

This work was supported by the founders of the Swedish National study on Aging and Care (SNAC): the Italian Ministry of Health (PE-2016-02364885); the Ministry of Health and Social Affairs, Sweden; the participating county councils and municipalities; and the Swedish Research Council. Specific grants were obtained from the Swedish Research Council (2016-00981), the Swedish Research Council for Health, Working Life and Welfare (2017-01764).

Acknowledgments

We thank the SNAC-K participants and the SNAC-K group for their collaboration in data collection and management. The study was approved by the Regional Ethics Review Board in Stockholm. Participants in the study completed a written informed consent form as stipulated by the ethics board. For participants with prevalent or incident cognitive impairment, consent was obtained from next of kin.

Author Contributions

A.R.L., D.L.V., and A.C.L. developed the study concept and design. A.R.L. performed the data analysis, and A.M., D.L.V., and A.C.L. contributed to the interpretation of the results. A.M. and A.R.L. drafted the manuscript. All authors provided critical revisions and approved the final version of the manuscript for submission.

Conflict of Interest

None reported.

References

1. Calderón-Larrañaga A, Vetrano DL, Onder G, et al. Assessing and measuring chronic multimorbidity in the older population: a proposal for its operationalization. *J Gerontol A Biol Sci Med Sci.* 2017;72:1417–1423. doi:10.1093/gerona/glw233
2. Marengoni A, Rizzuto D, Wang HX, Winblad B, Fratiglioni L. Patterns of chronic multimorbidity in the elderly population. *J Am Geriatr Soc.* 2009;57:225–230. doi:10.1111/j.1532-5415.2008.02109.x
3. Prados-Torres A, Calderón-Larrañaga A, Hancco-Saavedra J, Poblador-Plou B, van den Akker M. Multimorbidity patterns: a systematic review. *J Clin Epidemiol.* 2014;67:254–266. doi:10.1016/j.jclinepi.2013.09.021
4. Nobili A, Marengoni A, Tettamanti M, et al. Association between clusters of diseases and polypharmacy in hospitalized elderly patients: results from the REPOSI study. *Eur J Intern Med.* 2011;22:597–602. doi:10.1016/j.ejim.2011.08.029
5. Collerton J, Jagger C, Yadegarfar ME, et al. Deconstructing complex multimorbidity in the very old: findings from the Newcastle 85+ Study. *Biomed Res Int.* 2016;2016:8745670. doi:10.1155/2016/8745670
6. Everitt B. *Cluster Analysis*. 5th ed. Chichester, UK: Wiley; 2011.
7. Zhao Q. *Cluster Validity in Clustering Methods* [dissertation]. Joensuu, Finland: University of Eastern Finland; 2012:87. http://cs.joensuu.fi/sipu/pub/qinpei-thesis.pdf. Accessed June 4, 2019.
8. Capobianco E, Lio' P. Comorbidity: a multidimensional approach. *Trends Mol Med.* 2013;19:515–521. doi:10.1016/j.molmed.2013.07.004
9. Lagergren M, Fratiglioni L, Hallberg IR, et al. A longitudinal study integrating population, care and social services data. The Swedish National Study on Aging and Care (SNAC). *Aging Clin Exp Res.* 2004;16:158–168.
10. Rydwik E, Welmer AK, Kareholt I, et al. Adherence to physical exercise recommendations in people over 65–the SNAC-Kungsholmen study. *Eur J Public Health* 2013;23:799–804. doi: 10.1093/eurpub/cks150
11. Nelson ME, Rejeski WJ, Blair SN, et al. Physical activity and public health in older adults: recommendation from the American College of Sports Medicine and the American Heart Association. *Med Sci Sports Exerc.* 2007;39:1435–1445. doi:10.1249/mss.0b013e3180616aa2
12. Neugarten BL, Havighurst RJ, Tobin SS. The measurement of life satisfaction. *J Gerontol.* 1961;16:134–143.
13. Okun MA, Stock WA. The construct validity of subjective well-being measures: an assessment via quantitative research syntheses. *J Community Psychol.* 1987; 15:481–492.

14. Cornwell EY, Waite LJ. Measuring social isolation among older adults using multiple indicators from the NSHAP study. *J Gerontol B Psychol Sci Soc Sci.* 2009; 64(suppl 1):i38–i46. doi:10.1093/geronb/gbp037
15. Calderón-Larrañaga A, Santoni G, Wang HX, et al. Rapidly developing multimorbidity and disability in older adults: does social background matter? *J Intern Med.* 2018;283:489–499. doi:10.1111/joim.12739
16. Sourial N, Wolfson C, Zhu B, et al. Correspondence analysis is a useful tool to uncover the relationships among categorical variables. *J Clin Epidemiol.* 2010;63:638–646. doi:10.1016/j.jclinepi.2009.08.008
17. Guisado-Clavero M, Roso-Llorach A, López-Jimenez T, et al. Multimorbidity patterns in the elderly: a prospective cohort study with cluster analysis. *BMC Geriatr.* 2018;18:16. doi:10.1186/s12877-018-0705-7
18. Violán C, Roso-Llorach A, Foguet-Boreu Q, et al. Multimorbidity patterns with K-means nonhierarchical cluster analysis. *BMC Fam Pract.* 2018;19:108. doi:10.1186/s12875-018-0790-x
19. Lebedeva KA, Caruncho HJ, Kalynchuk LE. Cyclical corticosterone administration sensitizes depression-like behavior in rats. *Neurosci Lett.* 2017;650:45–51. doi:10.1016/j.neulet.2017.04.023
20. de Roos EW, Lahousse L, Verhamme KMC, et al. Asthma and its comorbidities in middle-aged and older adults: the Rotterdam Study. *Respir Med.* 2018;139:6–12. doi:10.1016/j.rmed.2018.04.006
21. Carton L, Pignon B, Baguet A, et al. Influence of comorbid alcohol use disorders on the clinical patterns of major depressive disorder: a general population-based study. *Drug Alcohol Depend.* 2018;187:40–47. doi:10.1016/j.drugalcdep.2018.02.009
22. Boyer L, Baumstarck K, Boucekine M, Blanc J, Lançon C, Auquier P. Measuring quality of life in patients with schizophrenia: an overview. *Expert Rev Pharmacoecon Outcomes Res.* 2013;13:343–349. doi:10.1586/erp.13.15
23. Oga T, Windisch W, Handa T, Hirai T, Chin K. Health-related quality of life measurement in patients with chronic respiratory failure. *Respir Investig.* 2018;56:214–221. doi:10.1016/j.resinv.2018.01.006
24. Haeusler KG, Gröschel K, Köhrmann M, et al. Expert opinion paper on atrial fibrillation detection after ischemic stroke. *Clin Res Cardiol.* 2018;107:871–880. doi:10.1007/s00392-018-1256-9
25. Mitsios JP, Ekinci EI, Mitsios GP, Churilov L, Thijs V. Relationship between glycated hemoglobin and stroke risk: a systematic review and meta-analysis. *J Am Heart Assoc.* 2018;7:e007858. doi:10.1161/JAHA.117.007858
26. Bektas F, Soyuncu S. Nitroglycerin-induced migraine type headache: bilaterally visible temporal arteries. *Int J Emerg Med.* 2010;3:463–464. doi:10.1007/s12245-010-0248-y
27. Mason BN, Russo AF. Vascular contributions to migraine: time to Revisit? *Front Cell Neurosci.* 2018;12:233. doi:10.3389/fncel.2018.00233
28. Ferrucci L, Fabbri E. Inflammageing: chronic inflammation in ageing, cardiovascular disease, and frailty. *Nat Rev Cardiol.* 2018;15:505–522. doi:10.1038/s41569-018-0064-2
29. Smykiewicz P, Segiet A, Keag M, Żera T. Proinflammatory cytokines and ageing of the cardiovascular-renal system. *Mech Ageing Dev.* 2018;175:35–45. doi:10.1016/j.mad.2018.07.006
30. Kanis JA, Stevenson M, McCloskey EV, Davis S, Lloyd-Jones M. Glucocorticoid-induced osteoporosis: a systematic review and cost-utility analysis. *Health Technol Assess.* 2007;11:iii–iv, ix.
31. Zografos GN, Georgiadou D, Thomas D, Kaltsas G, Digalakis M. Drug-induced esophagitis. *Dis Esophagus.* 2009;22:633–637. doi:10.1111/j.1442-2050.2009.00972.x
32. Jolliffe DA, Greenberg L, Hooper RL, et al. Vitamin D to prevent exacerbations of COPD: systematic review and meta-analysis of individual participant data from randomised controlled trials. *Thorax.* 2019;74:337–345. doi:10.1136/thoraxjnl-2018-212092
33. Shahnazari B, Moghimi J, Foroutan M, Mirmohammadkhani M, Ghorbani A. Comparison of the effect of vitamin D on osteoporosis and osteoporotic patients with healthy individuals referred to the Bone Density Measurement Center. *Biomol Concepts.* 2019;10:44–50. doi:10.1515/bmc-2019-0005
34. Luo W, Mao P, Zhang L, Yang Z. Association between systemic lupus erythematosus and thyroid dysfunction: a meta-analysis. *Lupus.* 2018;27:2120–2128. doi:10.1177/0961203318805849
35. Shen Y, Ye B, Chen P, et al. Cognitive decline, dementia, Alzheimer's disease and presbycusis: examination of the possible molecular mechanism. *Front Neurosci.* 2018;12:394. doi:10.3389/fnins.2018.00394
36. Livingston G, Sommerlad A, Orgeta V, et al. Dementia prevention, intervention, and care. *Lancet.* 2017;390:2673–2734. doi:10.1016/S0140-6736(17)31363-6
37. Naël V, Pérès K, Dartigues JF, et al.; Sense-Cog consortium. Vision loss and 12-year risk of dementia in older adults: the 3C cohort study. *Eur J Epidemiol.* 2019;34:141–152. doi:10.1007/s10654-018-00478-y
38. Brenowitz WD, Kaup AR, Lin FR, Yaffe K. Multiple sensory impairment is associated with increased risk of dementia among black and white older adults. *J Gerontol A Biol Sci Med Sci.* 2019;74:890–896. doi:10.1093/gerona/gly264
39. Deal JA, Sharrett AR, Albert M, et al. Retinal signs and risk of incident dementia in the Atherosclerosis Risk in Communities study. *Alzheimers Dement.* 2019;15:477–486. doi:10.1016/j.jalz.2018.10.002
40. Mograbi DC, Morris RG, Fichman HC, et al. The impact of dementia, depression and awareness on activities of daily living in a sample from a middle-income country. *Int J Geriatr Psychiatry.* 2018;33:807–813. doi:10.1002/gps.4765
41. Marengoni A, Fratiglioni L, Bandinelli S, Ferrucci L. Socioeconomic status during lifetime and cognitive impairment no-dementia in late life: the population-based aging in the Chianti Area (InCHIANTI) Study. *J Alzheimers Dis.* 2011;24:559–568. doi:10.3233/JAD-2011-101863
42. Forbes D, Thiessen EJ, Blake CM, Forbes SS, Forbes S. Exercise programs for people with dementia. *Cochrane Database Syst Rev.* 2013; 12: CD006489. doi:10.1002/14651858.CD006489.pub3
43. Wendrich AW, Swartz JE, Bril SI, et al. Low skeletal muscle mass is a predictive factor for chemotherapy dose-limiting toxicity in patients with locally advanced head and neck cancer. *Oral Oncol.* 2017;71:26–33. doi:10.1016/j.oraloncology.2017.05.012
44. Davies EA, O'Mahony MS. Adverse drug reactions in special populations – the elderly. *Br J Clin Pharmacol.* 2015;80:796–807. doi:10.1111/bcp.12596
45. Farmer C, Fenu E, O'Flynn N, Guthrie B. Clinical assessment and management of multimorbidity: summary of NICE guidance. *BMJ.* 2016;354:i4843. doi:10.1136/bmj.i4843
46. Richardson K, Fox C, Maidment I, et al. Anticholinergic drugs and risk of dementia: case–control study. *BMJ.* 2018;361:k1315. doi:10.1136/bmj.k1315

6.2.2. Study 2

Vetrano DL, **Roso-Llorach A**, Fernández S, Guisado-Clavero M, Violán C, Onder G, Fratiglioni L, Calderón-Larrañaga A, Marengoni A. Twelve-year clinical trajectories of multimorbidity in a population of older adults. Nat Commun. 2020 Jun 26;11(1):3223. doi: 10.1038/s41467-020-16780-x. PMID: 32591506; PMCID: PMC7320143.

nature COMMUNICATIONS

ARTICLE

https://doi.org/10.1038/s41467-020-16780-x OPEN

Twelve-year clinical trajectories of multimorbidity in a population of older adults

Davide L. Vetrano[1,2,8]✉, Albert Roso-Llorach[3,4,8], Sergio Fernández[3,4], Marina Guisado-Clavero[3,4], Concepción Violán[3,4], Graziano Onder[5], Laura Fratiglioni[1,6], Amaia Calderón-Larrañaga[1,9] & Alessandra Marengoni[1,7,9]

Multimorbidity—the co-occurrence of multiple diseases—is associated to poor prognosis, but the scarce knowledge of its development over time hampers the effectiveness of clinical interventions. Here we identify multimorbidity clusters, trace their evolution in older adults, and detect the clinical trajectories and mortality of single individuals as they move among clusters over 12 years. By means of a fuzzy c-means cluster algorithm, we group 2931 people ≥60 years in five clinically meaningful multimorbidity clusters (52%). The remaining 48% are part of an unspecific cluster (i.e. none of the diseases are overrepresented), which greatly fuels other clusters at follow-ups. Clusters contribute differentially to the longitudinal development of other clusters and to mortality. We report that multimorbidity clusters and their trajectories may help identifying homogeneous groups of people with similar needs and prognosis, and assisting clinicians and health care systems in the personalization of clinical interventions and preventive strategies.

[1] Aging Research Center, Department of Neurobiology, Care Sciences and Society, Karolinska Institutet and Stockholm University, Tomtebodavägen 18A, SE-171 65 Solna, Stockholm, Sweden. [2] Centro Medicina dell'Invecchiamento, Fondazione Policlinico Universitario "A. Gemelli" IRCCS, and Università Cattolica del Sacro Cuore, L.go Francesco Vito 1, 00100 Rome, Italy. [3] Fundació Institut Universitari per a la recerca a l'Atenció Primària de Salut Jordi Gol i Gurina (IDIAPJGol), Gran Via 587 àtic, Barcelona, Spain. [4] Universitat Autònoma de Barcelona, Campus de la UAB, Plaça Cívica, 08193 Bellaterra (Cerdanyola del Vallès), Spain. [5] Department of Cardiovascular, Endocrine-Metabolic Diseases and Aging, Istituto Superiore di Sanità, Via Giano della Bella 34, 00161 Rome, Italy. [6] Stockholm Gerontology Research Center, Sveavägen 155, 11346 Stockholm, Sweden. [7] Department of Clinical and Experimental Sciences, University of Brescia, Piazza Mercato 15, 25121 Brescia, Italy. [8]These authors contributed equally: Davide L Vetrano, Albert Roso-Llorach. [9]These authors jointly supervised this work: Amaia Calderón-Larrañaga, Alessandra Marengoni. ✉email: davide.vetrano@ki.se

As people age they tend to develop multiple chronic diseases; the term multimorbidity identifies this condition[1]. After 60 years of age, 55–98% of people are affected by two or more chronic diseases, and patients with multimorbidity account for up to 80% of consultations with general practitioners and virtually all consultations with geriatricians[2,3]. Co-occurring diseases interact with each other, increasing the risk of negative events beyond the sum of the risk of each disease[4]. Multimorbidity triggers complex pharmacological regimes, increases the use of health care resources, and reduces the quality and length of life[1,4–6]. It challenges physicians, who are usually trained to consider only a limited number of interactions between diseases and between diseases and drugs, and it puts pressure on health care systems, which struggle to offer older adults with multimorbidity comprehensive assessment, effective treatments, and integrated care paths[6–10]. Moreover, because older people with multimorbidity are usually excluded from randomized clinical trials, there are few clear recommendations about how to provide health care for older adults with multimorbidity. Complexity is thus translated into frustrating uncertainty and powerlessness and affects the quality of care at every level of the health care process[9].

Both clinical experience and epidemiological studies suggest that diseases cluster in the same person according to specific patterns[5,11]. Several clusters of diseases have been identified with some consistency across studies; however, there are a number of discrepancies in study findings[12]. A systematic review by Prados-Torres et al. identified 97 clusters of multimorbidity, and the findings of most of the reviewed studies suggested three clusters of multimorbidity: cardiometabolic, mental health, and musculoskeletal. At the same time, the studies in the review identified many unexplained heterogeneous clusters[12]. In addition to between-study methodological differences, one of the explanations for this finding may lie in the dynamic nature of disease clusters, which is not accounted for in cross-sectional studies. These clusters evolve overtime, and mortality selection plays an important role in shaping the observed population[13]. Capturing such dynamism is the only way to better understand the natural history of multimorbidity and to shed light on previously unexplained findings.

Most previous studies in this field have focused on clusters from the viewpoint of disease analyses rather than the analysis of groups of individuals[12,14]. Focusing on people is in keeping with the principle of patient-centered care and can provide information that facilitates the move toward personalized medicine[15]. A better understanding of older adults' transitions among multimorbidity clusters overtime may help detect homogeneous groups of individuals who may benefit from similar preventive (secondary and tertiary) interventions, treatment, and care. We therefore aimed to identify multimorbidity clusters in a population-based cohort of older adults, trace the evolution of the clusters over 12 years, and follow the clinical trajectories of the individuals as they moved between clusters or to death over time.

We found that multimorbidity clusters change dynamically overtime in older adults, following different clinical trajectories. Different clusters are also associated with different prognosis. Multimorbidity trajectories may help identifying homogeneous groups of people with similar needs and prognosis, and assisting clinicians and health care systems in the personalization of clinical interventions and preventive strategies.

Results

Six clusters of individuals with multimorbidity were identified. The present study is based on data from the Swedish National Study on Aging and Care in Kungsholmen (SNAC-K), an ongoing population-based study started in 2001 and involving 3363 individuals aged ≥60 years from a central area in Stockholm, Sweden. From the whole sample, 432 participants with <2 chronic disease have been excluded (i.e., those without multimorbidity). Those excluded were younger, reported a higher level of education, and were more often male than those included in the study (p for t test < 0.001). At baseline, study participants' mean age was 76.1 ± 11.0 [standard deviation] and 1951 (66.6%) were female. Over the 12 years, 1290 (44%) deaths occurred (28% within the first 6 years and 16% between 6 and 12 years). Moreover, 625 (22%) individuals dropped out (14% within the first 6 years and 8% between 6 and 12 years). At each follow-up, we performed a dimensionality reduction (i.e., multiple correspondence analysis) to obtain the input data for participants' clustering. A fuzzy c-means cluster analysis with optimal a fuzziness parameter at $m = 1.1$ (which outperformed other m values; see "Methods") was employed to identify clusters of individuals based on their underlying patterns of multimorbidity. Using an observed/expected ratio ≥2 (O/E ratio; i.e., the ratio between the prevalence of a given condition in a cluster and its prevalence in the whole sample) and an exclusivity ≥25% (i.e., the proportion of individuals with a given condition in the whole sample that belong to a cluster) for each disease, five clusters of people were identified at baseline: those with *psychiatric and respiratory diseases* (5.4%), *heart diseases* (9.3%), *respiratory and musculoskeletal diseases* (15.7%), *cognitive and sensory impairment* (10.6%), and *eye diseases and cancer* (10.7%). Solutions were evaluated based on their clinical consistency and significance criteria (Supplementary Figs. 1–15). Half of the people (48.7%) were grouped in an additional *unspecific* cluster, as they were affected by prevalent diseases but whose occurrence did not exceed the expected. Similarly, five clusters (plus the unspecific one) were identified at 6 and 12 years. At follow-ups, those diseases characterizing the baseline clusters were regrouped into different multimorbidity clusters. The clinical characteristics of the clusters are reported in Supplementary Table 1.

Individuals had different demographic, clinical and functional profiles across the clusters. Descriptive analyses were carried out to characterize the six clusters of individuals with multimorbidity. At baseline, participants in the *cognitive and sensory diseases*, the *eye diseases and cancer*, and the *heart diseases* clusters were the oldest. Participants in the *heart diseases*, the *eye diseases and cancer*, and the *psychiatric and respiratory diseases* clusters presented the greatest number of chronic diseases (mean number: 7.7 ± 2.4 [standard deviation], 6.0 ± 2.0, and 5.7 ± 2.2, respectively). Participants in the *heart diseases* and *psychiatric and respiratory diseases* clusters and those in the *cognitive and sensory impairment* cluster used the highest number of drugs (mean number: 7.7 ± 3.5, 6.2 ± 3.7, and 6.1 ± 3.4, respectively). Moreover, individuals included in the *heart diseases*, the *eye diseases and cancer*, and the *cognitive and sensory impairment* clusters presented the highest prevalence of disability and slow walking speed. The *cognitive and sensory impairment* and the *psychiatric and respiratory diseases* cluster showed the lowest Mini-Mental State Examination (MMSE) scores. The *unspecific* cluster was characterized by the lowest mean age and the lowest number of chronic diseases and drugs. This group had the lowest prevalence of disability and the highest walking speed, yet it had a high prevalence of hypertension, diabetes, dyslipidemia, and obesity. Such conditions were frequent also among participants in the *heart diseases* and the *eye diseases and cancer* clusters.

At follow-ups, in spite of varied clustering, a similar clinical distribution was observed for the different types of disorders. That is, people in clusters characterized by cardiovascular,

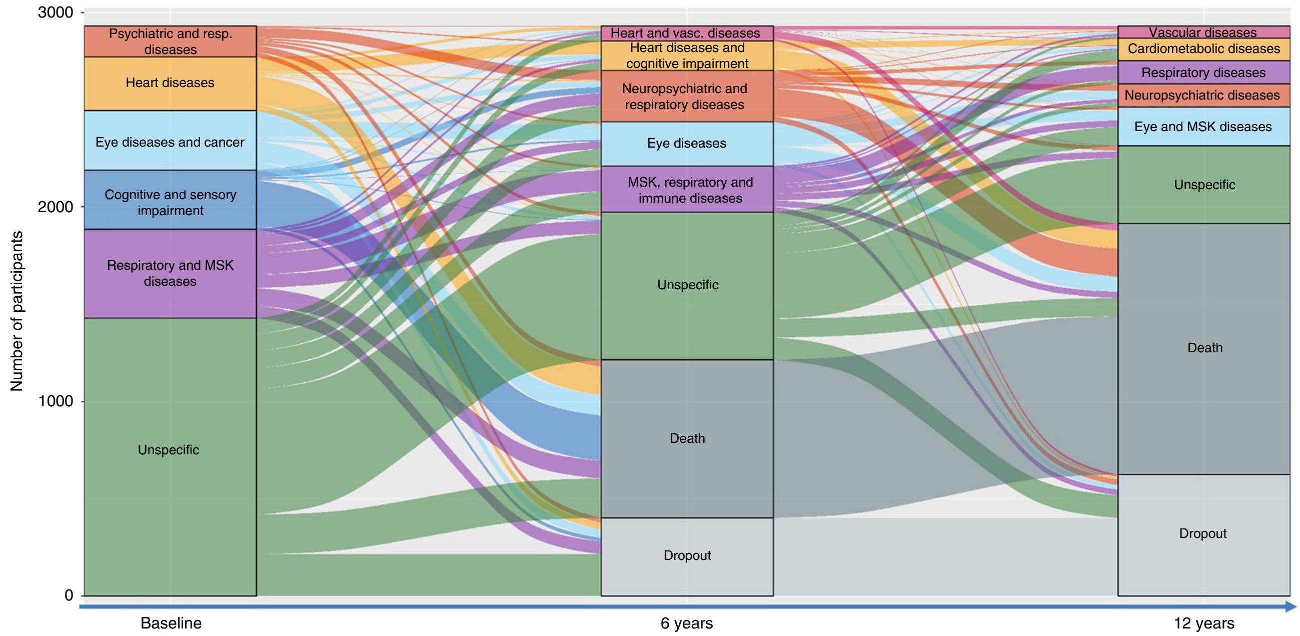


Fig. 1 Evolution of multimorbidity clusters and clinical trajectories of older adults with multimorbidity over 12 years. The height of the boxes and the thickness of the stripes are proportional to the amount of people belonging to the cluster and moving from the cluster, respectively. MSK musculoskeletal. To note, for this analysis participants were assigned to the cluster they were more likely to belong in order to investigate the most likely trajectories.

neuropsychiatric, and respiratory diseases showed the highest number of diseases and drugs and the highest levels of functional impairment.

Patterns of transitions between clusters can be identified over time. Upon assigning the individuals into the cluster they were more likely to belong to, we described their trajectories as they moved between clusters or to death over time. Figure 1 depicts the longitudinal evolution of multimorbidity clusters over 12 years and includes both the change overtime of disease patterns (the diseases that characterize a specific cluster of individuals) and the migration of participants from one cluster to another. The height of the boxes and the thickness of the stripes in the figure are proportional to the amounts of people in the cluster and moving out from the cluster, respectively.

In order to better characterize such transitions, we report in Figs. 2 and 3 the proportion of participants that were part of the 6-year and 12-year follow-ups clusters and that moved from multimorbidity clusters detected at an earlier wave. The percentages of participants moving from baseline and 6-year clusters, to 6-year and 12-year clusters, respectively, are reported in Supplementary Tables 4–7. During both first and second follow-up periods, the main shifts among clusters involved participants in the *unspecific* cluster, who moved primarily to clusters characterized by cardiovascular, eye, respiratory, and musculoskeletal diseases. For example, persons in the *unspecific* group at baseline moved and represented 48.7%, 45.0%, and 38.8% of the 6-year follow-up *heart and vascular diseases*, *musculoskeletal, respiratory and immune diseases*, and *eye diseases* clusters, respectively. Similarly, persons belonging to the *unspecific* group at the 6-year follow-up moved and represented 49.5%, 49.1%, and 20.6% of the 12-year follow up *cardiometabolic diseases, eye and musculoskeletal diseases*, and *vascular diseases* clusters, respectively.

Different multimorbidity clusters confer different mortality risks. The association between the multimorbidity clusters and mortality was tested in logistic regression models adjusted by age, sex, and education, taking the *unspecific* cluster as the reference group. As shown in Table 1, at baseline the *heart diseases* (OR 3.07; 95% CI 2.26–4.19), the *cognitive and sensory impairment* (OR 6.00; 95% CI 4.21–8.54), and the *psychiatric and respiratory diseases* (OR 1.60; 95% CI 1.02–2.51) clusters were significantly associated with a higher six-year mortality, compared with the people in the *unspecific* cluster. These clusters accounted for 51% of deaths. At first follow-up, the *heart and vascular diseases* (OR 3.78; 95% CI 2.13–6.70), the *heart diseases and cognitive impairment* (OR 3.73; 95% CI 2.41–5.79), and *neuropsychiatric and respiratory diseases* (OR 4.30; 95% CI 2.95–6.27) clusters had the highest OR for 6-year mortality, compared with the group of people in the *unspecific* cluster. These clusters accounted for 57% of deaths in the following 6 years.

Discussion

Tracing the evolution of multimorbidity clusters and the clinical trajectories of older adults with multimorbidity overtime led to two major findings. The first was a high heterogeneity in the multimorbidity clustering at baseline. Only half of the participants could be grouped into a well-characterized cluster: *psychiatric and respiratory diseases, heart diseases, respiratory and musculoskeletal diseases, cognitive and sensory impairment*, and *eye diseases and cancer*. The other half of the participants were sorted into an unspecific cluster and were characterized by having a younger age, lower numbers of co-occurring diseases and drugs, good functional and cognitive abilities, and a high percentage of cardiovascular risk factors. The second major finding was a highly dynamic evolution of multimorbidity clusters at both 6 and 12 years. Over 12 years, changes in cluster composition, participants' transitions from one cluster to another, and participant mortality generated a dynamic but well-defined clinical picture. The first remarkable trajectory involved the group of people part of the unspecific cluster at baseline. The number of participants grouped in this cluster halved at the 6- and 12-year follow-ups as the majority transitioned toward the specific multimorbidity clusters identified at follow-ups. Given the young age and less complex clinical picture of these individuals, they may be considered an

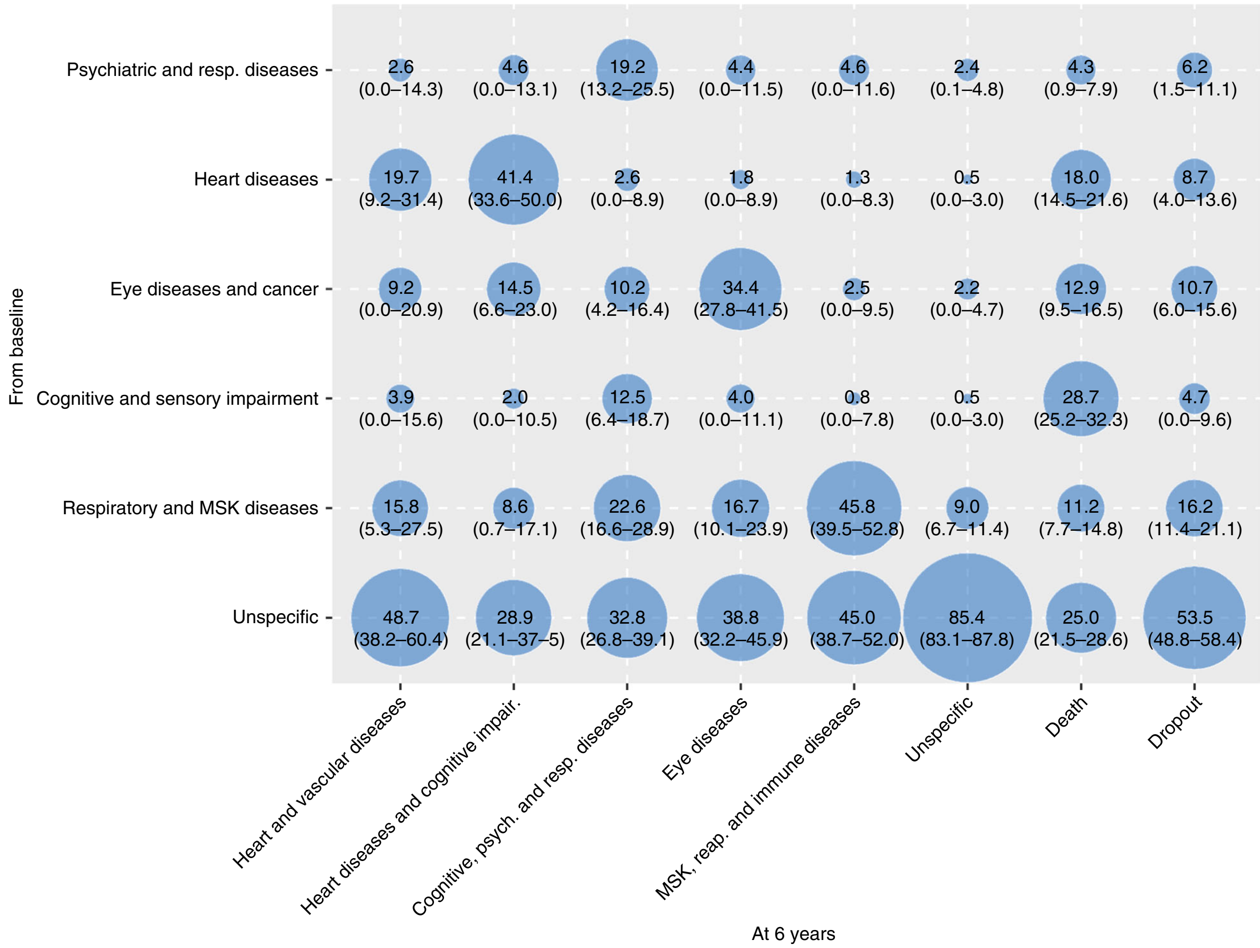


Fig. 2 Contribution of the baseline multimorbidity clusters to the 6-year follow-up clusters. Numbers indicate the percentage (%) of people belonging to the 6-year follow-up clusters that moved from the baseline clusters. To note, for this analysis participants were assigned to the cluster they were more likely to belong.

at-risk population for developing more complex multimorbidity and as such potentially susceptible to preventive intervention. The second relevant trajectory was the high mortality of individuals in clusters characterized by cardiovascular and neuropsychiatric diseases, which, despite representing only 25%, 28%, and 29% of the participants at baseline, 6 years, and 12 years, respectively, accounted for 51% and 57% of deaths during the first and second follow-up periods, respectively.

Increasingly, studies are analyzing clusters of multimorbidity across different populations, settings, and countries, but most studies have had a cross-sectional design or focused on the progression of co-morbidities of index diseases[12,16,17]. There is scanty evidence of how clusters of multimorbidity change overtime. The comparison is also limited by the fact that previous studies have used primary care, hospital-based registries or self-reported diagnoses, included only middle-aged people, or examined both acute and chronic conditions. A study from Spain that used a similar analytical strategy on large data from electronic primary health care records identified six clusters of multimorbidity: musculoskeletal, endocrine-metabolic, digestive/respiratory, neuropsychiatric, cardiovascular, and an unspecific group. These clusters exhibited less variation during the 6 years of follow-up than the patterns identified in our study, which could be explained by our longer follow-up period[18]. The use of electronic health records may also have led to an under detection of less severe diseases and multimorbidity[19]. A study from the Netherlands focused on six cardiovascular conditions. Clinical data from a large sample of general practice showed that the more diseases present at baseline, the higher the cumulative incidence rates of one or more new diseases (up to 47% at the 3-year follow-up and 76% at the 5-year follow-up)[20]. Another study of a population-wide registry of more than six million patients in Denmark showed more than a thousand significant longitudinal disease trajectories and some major multimorbidity clusters characterized by diseases of the prostate, chronic obstructive pulmonary disease, cerebrovascular disease, cardiovascular disease, and diabetes mellitus. The study had the limitation of data drawn retrospectively from a hospital registry of primary and secondary diagnostic codes. Further, both chronic and acute diseases were included[21], making the findings difficult to compare with ours. Finally, in an Australian study more than 13,000 middle-aged women with no history of diabetes, heart disease, or stroke at baseline were followed for 20 years in order to evaluate the longitudinal progression of the three conditions. Over 20 years, 18% of the women progressed to at least one condition, and 16.8% had two or three of these conditions; moreover, the onset of stroke was more strongly associated with an increased risk of progressing to the other two diseases. This is in contrast with what we observed in our study, which showed an opposite transition, from cardiovascular risk factors (e.g., diabetes) to overt cardiovascular and neuropsychiatric diseases. In the same Australian study, social inequality, obesity, hypertension, physical inactivity, smoking, and other chronic conditions were significantly associated with the three diseases independently but

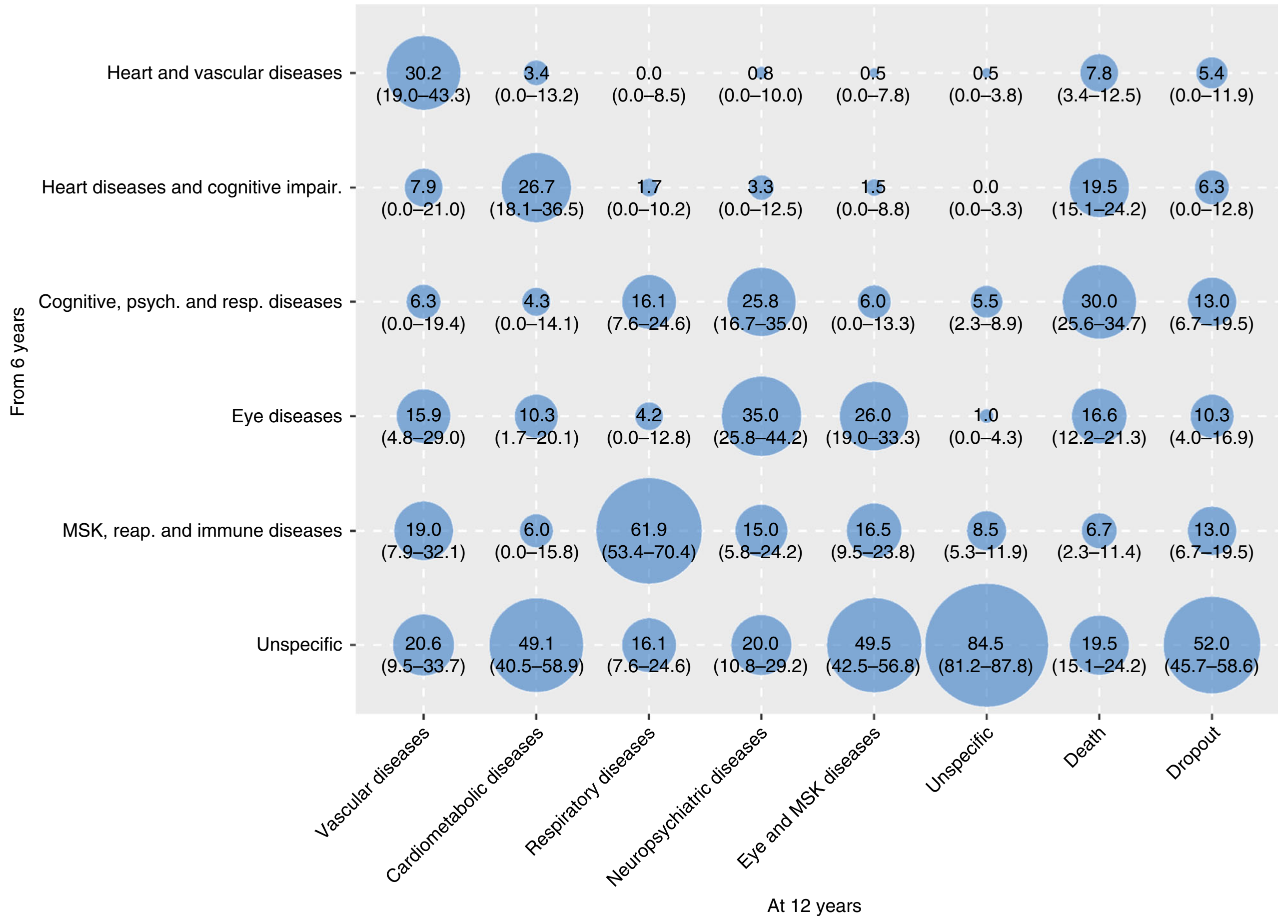


Fig. 3 Contribution of the 6-year follow-up multimorbidity clusters to the 12-year follow-up clusters. Numbers indicate the percentage (%) of people belonging to the 12-year follow-up clusters that moved from the 6-year follow-up clusters. To note, for this analysis participants were assigned to the cluster they were more likely to belong.

also with their co-occurrence. The study used self-reported diagnoses[22].

Some diseases may not be as independent of each other as we have previously thought. Biological, health-care related (e.g., pharmacological treatment), and psychosocial factors may increase susceptibility to a specific disease or to diseases in general in an individual[1,23]. Such factors can systematically drive diseases clustering beyond chance as well as their evolution to other clusters over time. First, direct consequences may explain why a large number of people in the *heart diseases* cluster at baseline became part of the *heart diseases and cognitive impairment* cluster at 6 years. Extensive scientific evidence supports the association between heart disease and cognitive decline through different mechanisms such as emboli, ischemic events, small vessel disease, cerebral hypoperfusion, and hypoxia. Indeed, mixed dementia, resulting from both cerebrovascular lesions and neurodegeneration, accounts for the majority of dementia cases among very old individuals[24]. Second, treatment consequences are another possible pathway when a disease occurs as the result of the pharmacological or surgical treatment of another condition. For example, part of the *neuropsychiatric and respiratory diseases* cluster, an association that remained over the entire course of our study, may be linked to the steroid treatment of respiratory diseases. Steroid treatment can often cause neurotic disorders and depression[25]. Third, overlapping symptomatology may result in diseases being misdiagnosed in an initial phase. This may have occurred with some baseline psychiatric conditions in the *psychiatric and respiratory diseases* cluster, which by 6 or 12 years may have evolved into, or been correctly classified as, cognitive impairment and dementia, putting them in the *cognitive impairment, psychiatric and respiratory diseases* cluster.

Finally, the *unspecific* cluster deserves special attention. These participants were characterized by diseases that were not overrepresented. However, despite their younger age and better physical and mental fitness, they had a high prevalence of cardiovascular and metabolic risk factors (diabetes, obesity, dyslipidemia, and hypertension). At baseline, almost half of the sample was part of this group. These people contributed to 29–49% of the multimorbidity clusters at the 6-year follow-up and to 16–50% of the multimorbidity clusters at 12 years, especially to those characterized by cardiovascular, eye, respiratory, and musculoskeletal diseases. Despite it is now well established that cardiometabolic conditions such as diabetes, obesity, dyslipidemia, and hypertension are important risk factors for the development of several cardiovascular diseases, less is known about the same risk factors, and the risk of other chronic conditions[26,27]. A few individuals moved from a specific cluster to the unspecific cluster over time. This may be explained by the fact that the progressive accrual of new diseases and the mortality (or dropout) of participants included in any of the specific clusters changed the reciprocal relation among diseases in survivors—in terms of prevalence, O/E ratio and exclusivity—making some of the subjects no longer classifiable into a specific cluster.

At least four out of ten participants died over the course of the study. Both at baseline and at 6-year follow-up, individuals with multimorbidity patterns characterized by cardiovascular and

Table 1 Association between clusters and mortality during the first (0–6 years) and second (6–12 years) follow-up.

Multimorbidity clusters at baseline	Events/at risk (%)	OR (95% CI)* 0–6 years mortality	Multimorbidity clusters at 6 years	Events/at risk	OR (95% CI)* 6–12 years mortality
Psychiatric and respiratory diseases	35/159 (22)	1.60 (1.02–2.51)	Heart and vascular diseases	37/76 (49)	3.78 (2.13–6.70)
Heart diseases	146/277 (53)	3.07 (2.26–4.19)	Heart diseases and cognitive impairment	93/152 (61)	3.73 (2.41–5.79)
Eye diseases and cancer	105/305 (34)	1.23 (0.90–1.68)	Neuropsychiatric and respiratory dis.	143/265 (54)	4.30 (2.95–6.27)
Cognitive and sensory impair.	233/306 (76)	6.00 (4.21–8.54)	Eye diseases	79/227 (35)	1.33 (0.89–2.00)
Respiratory and MSK diseases	91/456 (20)	1.29 (0.96–1.74)	MSK, respiratory, and immune diseases	32/238 (13)	1.06 (0.67–1.70)
Unspecific group	203/1428 (14)	Ref.	Unspecific group	93/758 (12)	Ref.

To note, for this analysis participants were assigned to the cluster they were more likely to belong.
Asterisk adjusted for age, sex, and education.
OR odds ratio; *CI* confidence interval; *MSK* musculoskeletal.

neuropsychiatric diseases had the highest mortality; with adjusted odds ratios ranging between 1.60 and 6.00 (taking people in the *unspecified* cluster as the reference). Those clusters accounted for 51% of deaths during the first follow-up and for 57% of deaths during the second follow-up. Notably, at 6 years there were two clusters characterized by cardiovascular diseases. Cardiovascular and neuropsychiatric diseases—the former including diseases such as heart failure and coronary diseases and the latter including diseases such as dementia and depression—are frequent and burdensome chronic conditions in older adults and are among the most important determinants of years of life spent with disability[28]. This is in line with a previous study from our group, showing that neuropsychiatric disease clusters, especially when combined with one or multiple cardiovascular diseases, have the highest impact on function decline in older persons[5]. Such findings were confirmed in other studies as well[29–31]. Indeed, the high mortality of people belonging to neuropsychiatric and heart disease clusters was not surprising as those clusters had the highest functional disability and lowest walking speed both at baseline and at first follow-up. Similar findings were reported also in studies from Spain[13] and from the United Kingdom[4]. The authors of the first report found that, compared with those subjects part of the musculoskeletal cluster, women in the cardiovascular clusters had the highest risk of dying. In the latter, co-occurring cardiometabolic disorders, unlike single disorders, decreased survival in a multiplicative way. It can be argued that not all diseases included in the cardiovascular or neuropsychiatric clusters transmit the same mortality risk. In fact, the nature of diseases, their impact at the organism level, and their severity may play major prognostic roles[13]. However, previous studies conducted in the field of associative multimorbidity have shown that the group-specific effect of clusters of diseases remains regardless of the role played by single diseases[5].

The main strength of this study was the thorough clinical evaluation that underlay disease assessment. Each participant in SNAC-K undergoes a 5 h comprehensive assessment that follows a standard protocol and is carried out by a physician, a nurse, and a psychologist. We then categorized diseases using a strict clinically driven method developed and tested by our group[32]. Furthermore, the lack of missing information on disease status increases the internal validity of our study. Another major strength of this study was the statistical method, which allowed us to cluster people by their co-occurring diseases. We took advantage of the method to follow individuals overtime and track their trajectories. The fuzzy c-means cluster algorithm is the choice method for pattern recognition when clusters tend to overlap, which is often the case as older adults present high prevalence of co-occurring conditions. In contrast to previous studies, each participant was assigned a probability of belonging to a cluster without being forced to be part of it exclusively. Other strengths included the long follow-up time, the high number of very old people, and the large age span of the participants (60–104 years). Moreover, including both mental and physical conditions in the analyses gave us the opportunity to investigate the interplay, potentially bidirectional, between mental health problems and chronic physical conditions. Several limitations of the present study should be mentioned. First, diseases were considered regardless their severity. Disease severity may indeed partially explain the clinical trajectories described in the present study. However, the interaction among different comorbidities still seems to play a major role—as it has been shown by us and others in previous studies—even when measures of disease severity are taken into account[4,5,31,33]. Moreover, in our opinion, independently from disease severity, the insights on the natural evolution of multimorbidity provided in this study are highly valuable and cover an important knowledge gap left by previous cross-sectional studies. Further, there is evidence that the burden of specific conditions changes depending on the overall multimorbidity status of one individual, making it difficult—especially in older individuals—to ascertain the relevance of single disease severity. Second, the dropout rate of participants (14% at 6 years and 8% at 12 years) may have affected cluster definition. However, to the best of our knowledge, this is an exceptionally low figure compared with studies of this type. Third, the discontinuous follow-up carried out in SNAC-K—every 3 or 6 years —may have affected disease detection and consequently the cluster analysis, especially among people who died or dropped out during the observation period. Finally, the average high socioeconomic status of participants in SNAC-K may potentially limit the generalizability of the findings.

Over their life course, individuals develop multiple diseases. This challenges the current organization of medical care services and the traditional research approach based on single diseases. Programs that bridge multiple clinical specialties and health care units should be developed to focus on single individuals, their specific clinical profiles, and their specific clinical trajectories[34]. Knowing how diseases cluster together, and importantly, how the clinical status of people with multimorbidity can change over subsequent years helps not only in understanding the complexity and dynamic evolution of multimorbidity clusters but also in supporting clinicians who manage co-occurring chronic diseases and health policy makers who plan care resources use. The findings from our study contribute in many ways. Firstly, they help identify people at high risk of progressing to severe disease clusters with worse prognosis. The people who could not be grouped in any specific cluster are at risk of cumulating further chronic disorders and increasing the severity of their multimorbidity profile. However, 28% of the people in this group remained relatively healthy during follow-ups. They had the lowest numbers of co-occurring chronic diseases and drugs and a better functional status than people in specific multimorbidity clusters, providing a large time window for preventive intervention. Future studies should focus on promotion of healthy aging in this group of individuals. Our findings contribute secondly to the development of personalized medicine in multimorbidity as our analysis is based on individuals and not diseases. There is solid evidence that persons who are affected by multimorbidity, face complex treatments, and require continuous monitoring far better from primary care with a patient-centered approach[35]. The strong transition we found from heart to brain diseases gives impetus to efforts in primary care to treat and monitor patients affected by heart disease. Treatment adherence is very low among older people with multimorbidity and heart diseases in particular[36]. Thirdly, our findings support prognostic counseling for patients and caregivers, given the high mortality of people with co-occurring mental and cardiovascular disorders. Fourthly, our findings encourage the planning of future randomized clinical trials toward the better management of multimorbidity. The 3D approach recently proposed by Salisbury et al. is an example of an intervention that could have focused on those multimorbidity clusters that may most likely lead to negative health outcomes (neuropsychiatric and cardiovascular clusters)[37]. In this pragmatic trial, the target population was selected based exclusively on the number of diseases and did not take into account specific groups of diseases. This may explain why the intervention was not able to improve participants' quality of life[38].

In conclusion, clinical trajectories of older adults with multimorbidity are characterized by great dynamism and complexity but can still be tracked over time. By analyzing data from a large population-based study of people aged 60+ years, we were able to identify multimorbidity clusters, trace their evolution overtime, and follow individuals' trajectories over 12 years. Shared risk factors and pathophysiology, development of diseases as a consequence of other conditions or treatments, and symptomatic overlap among diseases underlie most of the trajectories identified. Although the ability to discriminate among the potential mechanisms underlying the co-occurrence of multiple chronic diseases needs further improvement, taking into account multimorbidity clusters, and their evolution overtime may enable better decisions for patients with multimorbidity at every health care level and better tailoring of the target population in future interventions.

Methods

Study population. We used longitudinal data from the population-based SNAC-K[39]. The study population consists of adults ≥60 years living in the community or in institutions in the Kungsholmen district of Stockholm, Sweden. A random sample of 11 age cohorts born between 1892 and 1939 (the youngest and oldest age cohorts were oversampled) was invited to participate in the study. People who agreed to participate were evaluated for the first time between 2001 and 2004. Participants who were <78 years of age were then followed up every 6 years and participants ≥78 years every three years. The present study is based on data collected at baseline, 6 years, and 12 years. At baseline, 3363 people were examined (participation rate 73%). Overall, 432 participants were excluded because they did not have multimorbidity (≥2 chronic diseases) at baseline. The study was approved by the Regional Ethics Review Board in Stockholm. Participants in the study provided written informed consent. For participants with prevalent or incident cognitive impairment, written informed consent was obtained from the next of kin. The present study was reported in keeping with the STrengthening the Reporting of OBservational studies in Epidemiology recommendations.

Chronic diseases. At each study wave, SNAC-K participants undergo an ~5 h-long comprehensive clinical and functional assessment carried out by trained physicians, nurses, and neuropsychologists. Physicians collect information on diagnoses via physical examination, medical history, examination of medical charts, self-reported information, and/or proxy interviews. Clinical parameters, lab tests, drug information, and inpatient and outpatient care data are also used to identify specific conditions. All diagnoses are coded in accordance with the International Classification of Diseases, 10th revision (ICD-10). In the current study we sorted the ICD-10 codes into 60 chronic disease categories in accordance with a clinically driven methodology (Tables S2 and S3)[32]. To avoid statistical noise and the resulting spurious findings in the models, we excluded diseases with a prevalence of <2%. In SNAC-K at each study wave, drugs are coded in accordance with the Anatomical Therapeutic Chemical classification.

Vital status and loss to follow-up. Information about vital status was derived from death certificates provided by Statistics Sweden, the Swedish governmental statistics agency. Survival status was assessed throughout the follow-up period. Participants were considered lost to follow up if they or a proxy declined to participate, could not be contacted, had moved out of the study area, or canceled an assessment.

Other variables. Information on demographics (age, sex, and education) was collected during nurse interviews. We divided education into elementary, secondary, university, or higher. Level of disability was measured as the sum of the basic and instrumental activities of daily living (ADL and IADL) a person was unable to perform independently[40]. The six ADLs were bathing, dressing, toileting, continence, transferring, and eating. The eight IADLs were grocery shopping, meal preparation, housekeeping, doing laundry, managing money, using the telephone, taking medications, and using public transportation. Walking speed (m/s) was assessed by asking participants to walk 6 m at their usual speed or 2.44 m if the participant reported walking quite slowly. Speeds of <0.8 m/s were categorized as impaired[41]. Cognitive status was assessed by physicians using the MMSE, with a score range of 30 at best to 0 at worst[42].

Statistical analysis. Sample characteristics at baseline, 6-year follow-up, and 12-year follow-up were described for each multimorbidity cluster using weighted means and proportions obtained by the membership matrix (see below). At each study wave, clusters of older adults who shared patterns of multimorbidity were independently identified using the fuzzy c-means cluster analysis algorithm, which belongs to the family of *soft* clustering algorithms. The algorithm estimates c cluster centers (similar to k-means) but with fuzziness so that individuals may belong to more than one cluster. The use of a fuzzy cluster analysis over a hard cluster analysis helps to better handle the stochastic nature of some disease association, the potential noise stemming from the measurement (e.g., disease assessment), and the variance due to between-individual differences. Through this technique, we obtained clusters of individuals and a membership matrix that indicated the degree of participation of each subject in each cluster. In a second step, to evaluate the most likely clinical trajectories of the participants as they moved among clusters over time, each individual was assigned to the cluster with the highest membership score at each time point. We used dimensionality reduction techniques (multiple correspondence analysis) to obtain the input data for clustering the participants. The Karlis–Saporta–Spinaki rule was used to decide how many dimensions to retain[43]. The main parameters used during our cluster analysis were the number of clusters and a fuzziness parameter, denoted as "m", which ranges from just above 1 to infinity. High m values produce a fuzzy set of c clusters, so that individuals are equally distributed across clusters, whereas lower ones generate non-overlapped clusters. In our study we checked $m = 1.1$, 1.2, 1.4, 1.5, 2, 4 over 1 to 20 cluster combinations; the value $m = 1.1$ over performed the rest of values. Since clustering algorithms are unsupervised techniques, the model fitting the dataset is traditionally computed through cost functions that depend on both the dataset and the clustering parameters and are denoted as validation indices. We computed different validation indices to obtain the optimal number of clusters c and the optimal value of the fuzziness parameter m. Included parameters were: the Fukuyama index (optimal when presenting low values), the Xie–Beni index (optimal when presenting low values), the Partition coefficient index (optimal when presenting high values), the Partition entropy index (optimal when presenting low values), and the Calinski–Harabasz index (optimal when presenting high values; Supplementary Figs. 1–15)[44]. Given the stochastic nature of the clusters, we ran 100 independent clustering repetitions to obtain the average final solution. We based our evaluation of the consistency and significance of the final solution on clinical criteria. To cross-validate the model, we randomly split the individuals into two independent data sets and compared their validation indices. Indices were computed and averaged over 100 repetitions. To characterize the clusters of multimorbidity that corresponded to each cluster of individuals, we calculated the frequency of chronic diseases in each cluster. Observed/expected ratios (O/E-ratios) were calculated by dividing the prevalence of a given disease within a cluster by its prevalence in the overall population. The exclusivity of different diseases, defined as the fraction of participants with the disease in the cluster over the total number of participants with the disease, was also calculated. We considered a disease to be associated with a given cluster of individuals when the O/E ratio was ≥2 or the exclusivity was ≥25%[18]. Such criteria were used to name multimorbidity clusters after the diseases that mostly characterized them. To note, due to the dynamism of the phenomenon, the names of the clusters change overtime, reflecting the evolving combinations of diseases that characterize them at each time point. Shifts between clusters were computed by cross-tabulating individuals between each wave (baseline to 6-year follow-up and 6-year to 12-year follow-up) after assigning them individuals to the cluster where they were more likely to belong. In this way, we analyzed the most likely individual trajectories. Frequencies (percentages) of participants who changed from one cluster to another were computed to assess the overlap between waves. Both column percentages and row percentages are provided in Supplementary Tables. Mortality and dropout status were considered as fixed clusters in both 6-year and 12-year follow-ups. Logistic regression models adjusted by age, sex and education were fitted to estimate the association between clusters and mortality, using the *unspecific* cluster as the reference group. Also in this case, participants were assigned to the cluster where they were more likely to belong. Odd ratios (OR) and 95% confidence intervals (CI) were adjusted for age, sex, and education. All comparisons were adjusted for multiplicity. Pairwise comparison of p values, corrected for multiple comparisons, was calculated. Tukey method were used when the explanatory variable was normal-distributed or Benjamini and Hochberg method otherwise[45]. The significance level was set at $p = 0.05$. Although the overall number of significant tests between clusters at each follow-up remained stable at each follow-up, the number of highly significant pairwise statistical test (i.e., $p < 0.001$) decreased from 60.0 to 36.7%. Statistical analyses were performed using R 3.5.1 and Stata 15. Codes are available on demand.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The source data underlying all the figures and tables (including supplementary ones) is represented by the SNAC-K project, a population-based study on aging and dementia (http://www.snac-k.se/). Access to these original data is available to the research community upon approval by the SNAC-K data management and maintenance committee. Applications for accessing these data can be submitted to Maria Wahlberg (Maria.Wahlberg@ki.se) at the Aging Research Center, Karolinska Institutet.

Received: 14 February 2019; Accepted: 22 May 2020;
Published online: 26 June 2020

References

1. Calderon-Larranaga, A. et al. Multimorbidity and functional impairment-bidirectional interplay, synergistic effects and common pathways. *J. Intern. Med.* **285**, 255–271 (2019).
2. Marengoni, A. et al. Aging with multimorbidity: a systematic review of the literature. *Ageing Res. Rev.* **10**, 430–439 (2011).
3. Salisbury, C. et al. Epidemiology and impact of multimorbidity in primary care: a retrospective cohort study. *Br. J. Gen. Pract.* **61**, e12–e21 (2011).
4. Di Angelantonio, E. et al. Association of cardiometabolic multimorbidity with mortality. *Jama* **314**, 52–60 (2015).
5. Vetrano, D. L. et al. Trajectories of functional decline in older adults with neuropsychiatric and cardiovascular multimorbidity: a Swedish cohort study. *PLoS Med.* **15**, e1002503 (2018).
6. Dumbreck, S. et al. Drug-disease and drug-drug interactions: systematic examination of recommendations in 12 UK national clinical guidelines. *BMJ* **350**, h949 (2015).
7. Vetrano, D. L. et al. An international perspective on chronic multimorbidity: approaching the elephant in the room. *J. Gerontol. A Biol. Sci. Med. Sci.* **73**, 1350–1356 (2018).
8. Farmer, C. et al. Clinical assessment and management of multimorbidity: summary of NICE guidance. *BMJ* **354**, i4843 (2016).
9. Muth, C. et al. Evidence supporting the best clinical management of patients with multimorbidity and polypharmacy: a systematic guideline review and expert consensus. *J. Intern. Med.* **285**, 272–288 (2019).
10. Marengoni, A. & Onder, G. Guidelines, polypharmacy, and drug-drug interactions in patients with multimorbidity. *BMJ* **350**, h1059 (2015).
11. Vetrano D. L., et al. Walking Speed Drives the Prognosis of Older Adults with Cardiovascular and Neuropsychiatric Multimorbidity. *Am. J. Med.* https://doi.org/10.1016/j.amjmed.2019.05.005. (2019).
12. Prados-Torres, A. et al. Multimorbidity patterns: a systematic review. *J. Clin. Epidemiol.* **67**, 254–266 (2014).
13. Ibarra-Castillo, C. et al. Survival in relation to multimorbidity patterns in older adults in primary care in Barcelona, Spain (2010–2014): a longitudinal study based on electronic health records. *J Epidemiol Community Health* **72**, 185–192 (2018).
14. Westergaard, D. et al. Population-wide analysis of differences in disease progression patterns in men and women. *Nat. Commun.* **10**, 666 (2019).
15. Bierman, A. S. & Tinetti, M. E. Precision medicine to precision care: managing multimorbidity. *Lancet* **388**, 2721–2723 (2016).
16. Goodman, R. A. et al. Multimorbidity patterns in the United States: implications for research and clinical practice. *J. Gerontol. A Biol. Sci. Med. Sci.* **71**, 215–220 (2016).
17. Petri, H., Maldonato, D. & Robinson, N. J. Data-driven identification of co-morbidities associated with rheumatoid arthritis in a large US health plan claims database. *BMC Musculoskelet. Disord.* **11**, 247 (2010).
18. Guisado-Clavero, M. et al. Multimorbidity patterns in the elderly: a prospective cohort study with cluster analysis. *BMC Geriatr.* **18**, 16 (2018).
19. Rizzuto, D. et al. Detection of Dementia cases in two Swedish health registers: a validation study. *J. Alzheimers Dis.* **61**, 1301–1310 (2018).
20. Lappenschaar, M. et al. Multilevel temporal Bayesian networks can model longitudinal change in multimorbidity. *J. Clin. Epidemiol.* **66**, 1405–1416 (2013).
21. Jensen, A. B. et al. Temporal disease trajectories condensed from population-wide registry data covering 6.2 million patients. *Nat. Commun.* **5**, 4022 (2014).
22. Xu X, Mishra GD, Dobson AJ, & Jones M. Progression of diabetes, heart disease, and stroke multimorbidity in middle-aged women: A 20-year cohort study. *PLoS Med.* **15**, e1002516 (2018).
23. van den Akker, M. et al. Multimorbidity in general practice: prevalence, incidence, and determinants of co-occurring chronic and recurrent diseases. *J. Clin. Epidemiol.* **51**, 367–375 (1998).
24. Qiu, C. & Fratiglioni, L. A major role for cardiovascular burden in age-related cognitive decline. *Nat. Rev. Cardiol.* **12**, 267–277 (2015).
25. Herbert, J. Cortisol and depression: three questions for psychiatry. *Psychol. Med.* **43**, 449–469 (2013).
26. Marseglia, A. et al. Prediabetes and diabetes accelerate cognitive decline and predict microvascular lesions: a population-based cohort study. *Alzheimers Dement.* https://doi.org/10.1016/j.jalz.2018.06.3060 (2018).
27. Liu, Y. et al. Hypertension-induced cerebral small vessel disease leading to cognitive impairment. *Chin. Med. J.* **131**, 615–619 (2018).
28. Rizzuto, D. et al. Effect of chronic diseases and multimorbidity on survival and functioning in elderly adults. *J. Am. Geriatr. Soc.* **65**, 1056–1060 (2017).
29. Quinones, A. R., Markwardt, S. & Botoseneanu, A. Multimorbidity combinations and disability in older adults. *J. Gerontol. A Biol. Sci. Med. Sci.* **71**, 823–830 (2016).
30. Quinones, A. R. et al. Prospective disability in different combinations of somatic and mental multimorbidity. *J. Gerontol. A Biol. Sci. Med. Sci.* **73**, 204–210 (2018).
31. Jindai, K. et al. Multimorbidity and functional limitations among adults 65 or older, NHANES 2005-2012. *Prev Chronic Dis.* **13**, E151 (2016).
32. Calderon-Larranaga, A. et al. Assessing and measuring chronic multimorbidity in the older population: a proposal for its operationalization. *J. Gerontol. A Biol. Sci. Med. Sci.* **72**, 1417–1423 (2017).

33. Jackson, C. A. et al. Multimorbidity patterns are differentially associated with functional ability and decline in a longitudinal cohort of older women. *Age Ageing* **44**, 810–816 (2015).
34. Hemingway, H. et al. Prognosis research strategy (PROGRESS) 1: a framework for researching clinical outcomes. *BMJ* **346**, e5595 (2013).
35. National Institute for Health and Care Excellence (NICE). Multimorbidity: clinical assessment and management. (2016). https://www.nice.org.uk/guidance/ng56/resources/multimorbidity-clinical-assessment-and-management-pdf-1837516654789 Accessed on 1st January 2020.
36. Ruppar, T. M. et al. Medication adherence interventions improve heart failure mortality and readmission rates: systematic review and meta-analysis of controlled trials. *J Am. Heart Assoc.* **5**. https://doi.org/10.1161/jaha.115.002606. (2016).
37. Salisbury, C. et al. Management of multimorbidity using a patient-centred care model: a pragmatic cluster-randomised trial of the 3D approach. *Lancet* **392**, 41–50 (2018).
38. Marengoni, A., Vetrano, D. L., & Onder, G. Target population for clinical trials on multimorbidity: is disease count enough? *J. Am. Med. Dir. Assoc.* https://doi.org/10.1016/j.jamda.2018.10.012. (2018)
39. Lagergren, M. et al. A longitudinal study integrating population, care and social services data. The Swedish National study on Aging and Care (SNAC). *Aging Clin. Exp. Res.* **16**, 158–168 (2004).
40. Spector, W. D. & Fleishman, J. A. Combining activities of daily living with instrumental activities of daily living to measure functional disability. *J. Gerontol. B Psychol. Sci. Soc. Sci.* **53**, S46–S57 (1998).
41. Studenski, S. et al. Gait speed and survival in older adults. *Jama* **305**, 50–58 (2011).
42. Folstein, M. F., Folstein, S. E. & McHugh, P. R. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *J. Psychiatr. Res.* **12**, 189–198 (1975).
43. Sourial, N. et al. Correspondence analysis is a useful tool to uncover the relationships among categorical variables. *J. Clin. Epidemiol.* **63**, 638–646 (2010).
44. Zhao Q. Cluster validity in clustering methods. Publications of the University of Eastern Finland. (2012) http://cs.joensuu.fi/sipu/pub/qinpei-thesis.pdf. Accessed on 1st January 2020.
45. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B.* **57**, 289–300 (1995).

Acknowledgements

We thank the SNAC-K participants and the SNAC-K Group for their collaboration in data collection and management, and scientific editors Kimberly Kane and Karen Hagersten for useful comments on the text. This work was supported by the funders of the Swedish National study on Aging and Care (SNAC): the Ministry of Health and Social Affairs, Sweden; the participating County Councils and Municipalities; and the Swedish Research Council. Specific grants were received from The Swedish Research Council for Medicine (VR; 521-2013-8676; 2017-06088; 2016-00981); the Swedish Research Council for Health, Working life and Welfare (Forte; 2016-07175; 2017-01764); Gamla Tjanarinnor (2019-00897), and the Ermenegildo Zegna Foundation. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the paper. Open access funding provided by Karolinska Institute.

Author contributions

Conception or design of the work: D.L.V., A.R.L., A.C.L., S.F., C.V., A.M. Data analysis: A.R.L., S.F., D.L.V., A.C.L. Interpretation of the results: D.L.V., A.R.L., A.C.L., S.F., C.V., A.M., M.G.C., G.O., L.F. Drafting the article: D.L.V., A.R.L., A.C.L., A.M. Critical revision of the paper: D.L.V., A.R.L., A.C.L., S.F., C.V., A.M., M.G.C., G.O., L.F. Final approval of the paper: D.L.V., A.R.L., A.C.L., S.F., C.V., A.M., M.G.C., G.O., L.F. All the authors fulfill the ICMJE criteria for authorship.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41467-020-16780-x.

Correspondence and requests for materials should be addressed to D.L.V.

Peer review information *Nature Communications* thanks Anders Jensen, Marjan Van Den Akker, and the other, anonymous, reviewers for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permission information is available at http://www.nature.com/reprints

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

6.2.3. Study 3

Roso-Llorach A, Vetrano DL, Trevisan C, Fernández S, Guisado-Clavero M, Carrasco-Ribelles LA, Fratiglioni L, Violán C, Calderón-Larrañaga A. 12-year evolution of multimorbidity patterns among older adults based on Hidden Markov Models. Aging (Albany NY). 2022 Nov 23;14. doi: 10.18632/aging.204395. Epub ahead of print. PMID: 36435509.

Research Paper

12-year evolution of multimorbidity patterns among older adults based on Hidden Markov Models

Albert Roso-Llorach[1,2,3,*], Davide L. Vetrano[4,5,*], Caterina Trevisan[4,6], Sergio Fernández[1,2], Marina Guisado-Clavero[7], Lucía A. Carrasco-Ribelles[1,8], Laura Fratiglioni[4,5], Concepción Violán[2,9,#], Amaia Calderón-Larrañaga[4,5,#]

[1]Fundació Institut Universitari per a la recerca a l'Atenció Primària de Salut Jordi Gol i Gurina (IDIAPJGol), Barcelona, Spain
[2]Universitat Autònoma de Barcelona, Bellaterra (Cerdanyola de Vallès), Spain
[3]Programa de Doctorat en Metodologia de la Recerca Biomèdica i Salut Pública, Universitat Autònoma de Barcelona, Bellaterra (Cerdanyola del Vallès), Spain
[4]Aging Research Center, Department of Neurobiology, Care Sciences and Society, Karolinska Institutet and Stockholm University, Stockholm, Sweden
[5]Stockholm Gerontology Research Center, Stockholm, Sweden
[6]Department of Medical Sciences, University of Ferrara, Ferrara, Italy
[7]Unidad Docente Multiprofesional de Atención Familiar y Comunitaria Norte, Gerencia Asistencial Atención Primaria, Madrid Health Service, Madrid, Spain
[8]Signal Theory and Communications Department, Universitat Politecnica de Catalunya, Barcelona, Spain
[9]Unitat de Suport a la Recerca Metropolitana Nord, Fundació Institut Universitaria per a la recerca a l'Atenció Primària de Salut Jordi Gol i Gurina (IDIAPJGol), Mataró, Barcelona, Spain
*Shared first authorship
#Shared last authorship

Correspondence to: Albert Roso-Llorach, Amaia Calderón-Larrañaga; **email:** aroso@idiapjgol.org, amaia.calderon.larranaga@ki.se
Keywords: multimorbidity, older adults, longitudinal population-based study, aging, Hidden Markov Models
Received: July 13, 2022 **Accepted:** November 14, 2022 **Published:** November 23, 2022

ABSTRACT

Background: The evolution of multimorbidity patterns during aging is still an under-researched area. We lack evidence concerning the time spent by older adults within one same multimorbidity pattern, and their transitional probability across different patterns when further chronic diseases arise. The aim of this study is to fill this gap by exploring multimorbidity patterns across decades of age in older adults, and longitudinal dynamics among these patterns.
Methods: Longitudinal study based on the Swedish National study on Aging and Care in Kungsholmen (SNAC-K) on adults ≥60 years (N=3,363). Hidden Markov Models were applied to model the temporal evolution of both multimorbidity patterns and individuals' transitions over a 12-year follow-up.
Findings: Within the study population (mean age 76.1 years, 66.6% female), 87.2% had ≥2 chronic conditions at baseline. Four longitudinal multimorbidity patterns were identified for each decade. Individuals in all decades showed the shortest permanence time in an *Unspecific* pattern lacking any overrepresented diseases (range: 4.6-10.9 years), but the pattern with the longest permanence time varied by age. Sexagenarians remained longest in the *Psychiatric-endocrine and sensorial* pattern (15.4 years); septuagenarians in the *Neuro-vascular*

***and skin-sensorial* pattern (11.0 years); and octogenarians and beyond in the *Neuro-sensorial* pattern (8.9 years). Transition probabilities varied across decades, sexagenarians showing the highest levels of stability.**
Interpretation: Our findings highlight the dynamism and heterogeneity underlying multimorbidity by quantifying the varying permanence times and transition probabilities across patterns in different decades. With increasing age, older adults experience decreasing stability and progressively shorter permanence time within one same multimorbidity pattern.

INTRODUCTION

Extended human longevity is a goal achieved in the last century, and a reality in middle- and high-income countries [1]. Improvements in health resources and medical sciences, and decreases in preventable mortality have been key to living longer [2]. However, increasing life expectancy comes along with a higher burden of chronic diseases [3]. The coexistence of multiple chronic diseases in a single person is known as multimorbidity. Multimorbidity is associated with a higher risk of polypharmacy and decreased quality of life, and challenges the decision-making of clinicians that lack effective guidelines for the management and treatment of patients with cohexisting complex diseases [4].

In an attempt to understand how chronic diseases are inter-related, several studies have explored so-called multimorbidity patterns [5–7]. In a previous systematic review, three patterns of multimorbidity involving cardiometabolic diseases, mental health problems, and musculoskeletal disorders have been consistently suggested to be the most prevalent in the older population [5]. Diseases cluster in specific patterns due to common pathophysiological pathways and risk factors, or because they may be the cause or consequence of other coexisting diseases. Along with the above mentioned patterns, a high number of less reproducible and sparse disease combinations have been described, often inconsistently across studies. Several factors may explain such disparate observations: first, the use of cross-sectional designs, which do not account for the dynamic nature of multimorbidity in old age; second, the use of different disease lists, spanning from less than ten to more than two hundred conditions; and third, the employement of statistical methods that cannot properly manage the complexity of the phenomenon. Recently, several advanced machine-learning techniques such as non-hierarchical and hierarchical clustering tehcniques have been used to explore multimorbidity patterns.

Exploring how multimorbidity patterns evolve throughout people's lives and the time subjects remain within specific patterns is still an under-researched area [7, 8]. The understanding of how diseases cluster longitudinally in specific age groups would pave the way to the design of new prognostic tools, as well as new preventive and, eventually, therapeutic approaches. Hidden Markov Models (HMM) over-come several of the limitations of previously employed methods, which were unable to account for the variability in chronic disease interactions throughout time [9]. HMM consider diseases in each person to be random variables conditioned by a hidden state or cluster. Despite the technique's potential, only one previous register-based study has used HMM for the longitudinal study of multimorbidity [9], but the folllow-up time was insufficient to draw any relevant conclusions. Cohort studies with homogeneously collected data over long periods of time represent a unique resource for the longitudinal analysis of multimorbidity patterns, and their use for such a purpose is warranted.

The aims of this study were: 1) to explore longitudinal multimorbidity patterns across decades of age after 60 using HMM, and 2) to detect the dynamics underlying such patterns in terms of the time subjects remained within the same pattern, and the probability of transitioning across different patterns.

RESULTS

Multimorbidity patterns

The study population included 3,363 individuals aged 60+ of whom 87.2% had multimorbidity at baseline. Participants' mean age at baseline was 76.1 years, and 66.6% were female. Over the 12-year follow-up, 1346 (40%) deaths occurred (25% within the first 6 years and 15% within the next 6 years). Moreover, 719 (21.4%) individuals dropped out (13.7% within the first 6 years and 7.7% within the next 6 years). Descriptive statistics of each age cohort at each follow-up wave can be found in Table 1.

In the three age groups, a total of 44, 49 and 47 chronic disease categories, respectively, showed a median prevalence ≥2% during the study period, and were thus included in the HMM estimations (Supplementary Table 1). Overall, four multimorbidity patterns were identified for each age group, and two additional patterns were artificially added to account for death and dropout during the follow-up period (Supplementary Table 2).

Table 1. Sociodemographic, clinical, and functional characteristics of the study population by baseline age group (N=3,363).

	Sexagenarians			*Septuagenarians*			*Octogenarians and beyond*		
	Baseline N=1304	**6 years follow-up N=1045**	**12 years follow-up N=846**	**Baseline N=939**	**6 years follow-up N=639**	**12 years follow-up N=358**	**Baseline N=1120**	**6 years follow-up N=374**	**12 years follow-up N=94**
Age, mean (SD)	63.0 (2.91)	68.9 (2.89)	74.9 (2.88)	75.3 (3.00)	81.1 (2.98)	86.6 (2.89)	87.9 (5.10)	91.5 (4.11)	95.5 (2.84)
Female, n (%)	735 (56.4%)	603 (57.7%)	503 (59.5%)	598 (63.7%)	419 (65.6%)	245 (68.4%)	849 (75.8%)	276 (73.8%)	71 (75.5%)
Education, n (%)									
Elementary	93 (7.14%)	61 (5.84%)	45 (5.32%)	150 (16.1%)	95 (14.9%)	48 (13.4%)	347 (31.7%)	95 (25.7%)	22 (23.4%)
High school	561 (43.1%)	445 (42.6%)	346 (40.9%)	514 (55.1%)	343 (53.7%)	189 (52.8%)	576 (52.6%)	197 (53.4%)	53 (56.4%)
University	648 (49.8%)	539 (51.6%)	455 (53.8%)	269 (28.8%)	201 (31.5%)	121 (33.8%)	173 (15.8%)	77 (20.9%)	19 (20.2%)
# chronic diseases, mean (SD)	2.72 (1.78)	4.87 (2.78)	7.70 (3.57)	4.24 (2.28)	7.71 (3.46)	12.0 (4.56)	5.47 (2.51)	9.70 (3.58)	14.2 (4.41)
# drugs, mean (SD)	2.66 (2.77)	4.13 (3.37)	5.18 (3.92)	4.39 (3.42)	6.10 (3.92)	7.44 (4.50)	5.37 (3.48)	7.25 (3.97)	8.47 (4.46)
Walking speed, mean (SD)	1.26 (0.31)	1.20 (0.35)	1.08 (0.35)	1.00 (0.38)	0.79 (0.41)	0.66 (0.41)	0.54 (0.41)	0.43 (0.36)	0.37 (0.35)
MMSE, mean (SD)	29.3 (1.45)	28.7 (1.59)	28.5 (2.27)	28.4 (3.32)	26.9 (4.25)	25.5 (5.73)	24.8 (7.39)	24.1 (6.99)	21.9 (8.64)

Abbreviations: MMSE, Mini Mental State Examination; SD, standard deviation.

Among sexagenarians, subjects in the *Unspecific* pattern were the youngest across all follow-ups, while those in the *Cardiovascular and anemia* pattern were the oldest (Supplementary Table 3). Subjects in the *Cardio-metabolic* pattern were more frequently male while those in the *Psychiatric-endocrine and sensorial* pattern were more likely to be female. Subjects in the latter pattern showed the highest level of education.

Among septuagenarians, subjects in the *Unspecific* pattern were the youngest, while those in the *Neuro-vascular and skin-sensorial* pattern were the oldest. Subjects in the *Cardiovascular and diabetes* pattern were more frequently male while those in the *Neuro-vascular and skin-sensorial* and *Neuro-psychiatric and sensorial* patterns were more likely to be female. Subjects in the *Cardiovascular and diabetes* pattern had the lowest proportion of university education.

In the group of octogenarians and beyond, those in the *Respiratory-circulatory and skin* pattern were the youngest, while those in the *Cardio-respiratory and neurological* were the oldest. All patterns had a higher proportion of females. Subjects in the *Neuro-sensorial* pattern showed the highest level of education.

Evolution and transitions across multimorbidity patterns

The evolution and transitions of and among multimorbidity patterns are graphically represented through river plots in Figure 1. For all age groups, pattern prevalence varied over time, showing that people commonly transition from one pattern to another. A general trend was that the most represented patterns at baseline (i.e., containing the healthiest subjects) evolved towards smaller ones over time, and the smallest ones (i.e., presumably containing the sickest subjects) tended to become larger over time. For example, among sexagenarians, subjects in the *Unspecific* pattern represented 80% of the study population at baseline, but the figure went down to 52.4% after 6 years and to 22.6% after 12 years. The prevalence of the death and dropout patterns increased in older age groups; an important part of the transitions among octogenarians and beyond were in fact towards death.

The estimated mean permanence times were computed for each age group. As an example, for sexagenarians belonging to the *Cardiovascular and anemia* pattern at baseline, it was estimated that they would remain in the same pattern for a mean time of 14.9 years before transitioning to other patterns. In all age groups, the *Unspecific* patterns showed the shortest sojourn times, and the *Psychiatric-endocrine and sensorial*, *Neuro-vascular and skin-sensorial* and *Neuro-sensorial* were the patterns with the longest sojourn time for sexagenarians, septuagenarians and octogenarians and beyond, respectively.

The transition probability matrices by age group are shown in Figure 2. Regarding the interpretation of these probabilities, the models show that, for example, sexagenarians belonging to the *Unspecific* pattern at baseline had a probability of 0.9% of transitioning to the *Cardiovascular and anemia* pattern and of 20.0% of staying in the same pattern in the next 12 years. In general, sexagenarians showed the highest levels of stability, as the probabilities of staying in the same pattern were higher than in the other age groups. More specifically, among sexagenarians, the most likely transition between patterns was from the *Unspecific* to

Pattern	Baseline (%)	12 years (%)	Permanence time (years)
Unsp	84.1	22.6	10.9
CV & Anemia	0.4	16.9	14.9
Cardio-Meta	6.8	12.5	12.2
Psy-Endoc & Sens	8.7	48.0	15.4

Pattern	Baseline (%)	12 years (%)	Permanence time (years)
Unsp	69.7	5.9	7.7
CV & Diab	8.1	7.5	10.9
NeuroVasc & Skin	0.4	36.6	11.0
NeuroPsy & Sens	21.8	50.0	8.6

Pattern	Baseline (%)	12 years (%)	Permanence time (years)
Unsp	73.1	6.4	4.7
Resp-Circula & Skin	1.3	16.0	7.2
CardioResp & Neuro	22.8	52.1	5.2
Neuro-Sens	2.9	25.5	8.9

Figure 1. Evolution and transitions of multimorbidity patterns over time by age group (N=3,363). Sexagenarians: Unspecific (Unsp); Cardiovascular and anemia (CV and Anemia); Cardio-metabolic (Cardio-Meta) and Psychiatric-endocrine and sensorial (Psy-Endoc and Sens). Septuagenarians: Unspecific (Unsp); Cardiovascular and diabetes (CV and Diab); Neuro-vascular and skin-sensorial (NeuroVasc and Skin); and Neuro-psychiatric and sensorial (NeuroPsy and Sens). Octogenarians and beyond: Unspecific (Unsp); Respiratory-circulatory and skin (Resp-Circula and Skin); Cardio-respiratory and neurological (CardioResp and Neuro); and Neuro-sensorial (Neuro-Sens).

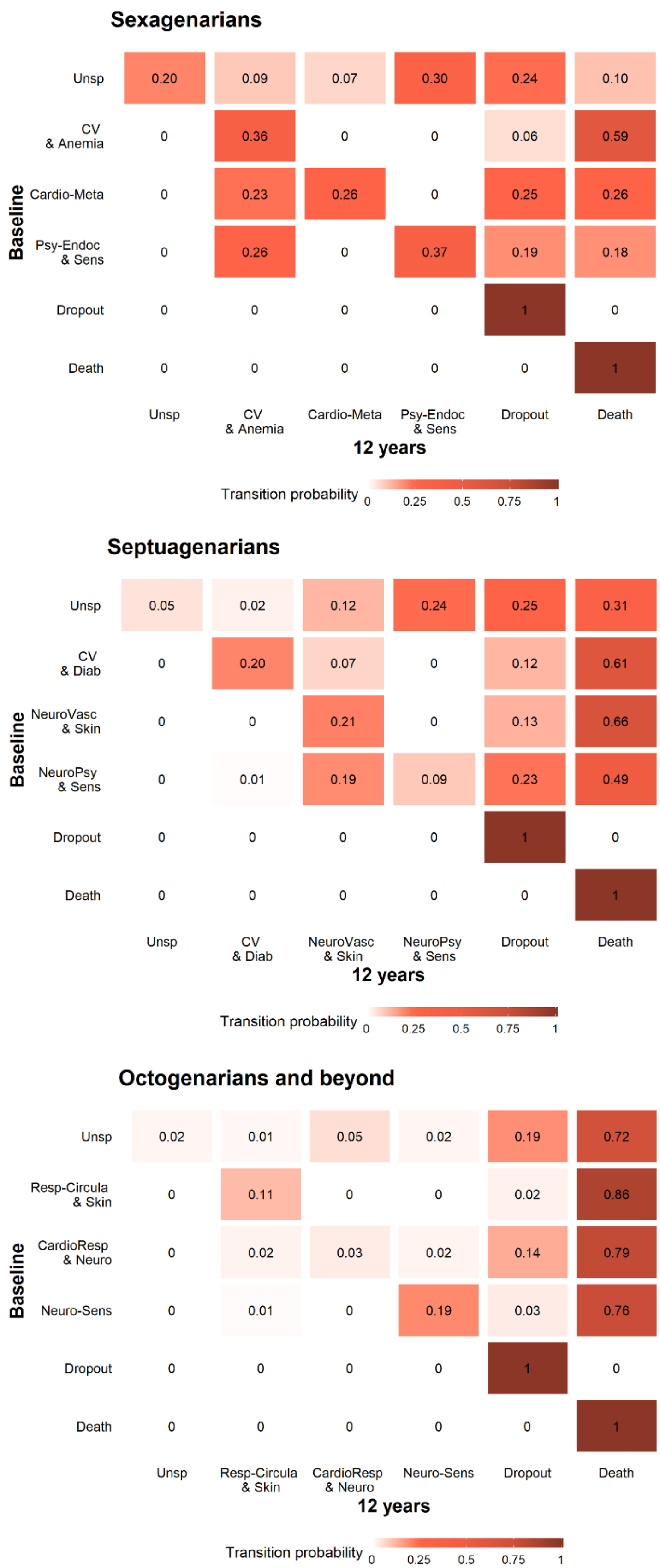


Figure 2. Transition probability matrices by age group from baseline to the 12-year follow-up (N=3,363). Sexagenarians: Unspecific (Unsp); Cardiovascular and anemia (CV and Anemia); Cardio-metabolic (Cardio-Meta) and Psychiatric-endocrine and

sensorial (Psy-Endoc and Sens). Septuagenarians: Unspecific (Unsp); Cardiovascular and diabetes (CV and Diab); Neuro-vascular and skin-sensorial (NeuroVasc and Skin); and Neuro-psychiatric and sensorial (NeuroPsy and Sens). Octogenarians and beyond: Unspecific (Unsp); Respiratory-circulatory and skin (Resp-Circula and Skin); Cardio-respiratory and neurological (CardioResp and Neuro); and Neuro-sensorial (Neuro-Sens).

the *Psychiatric-endocrine and sensorial* pattern (30.0%) after 12 years. Among septuagenarians, the most likely transition was from the *Unspecific* to the *Neuro-psychiatric and sensorial* pattern (24.0%) after 12 years. Finally, in octogenarians and beyond, the transition from the *Unspecific* to the *Cardio-respiratory and neurological* pattern (5.0%) after 12 years was the likeliest. The *Cardiovascular and anemia*, *Neuro-vascular and skin-sensorial*, and *Respiratory-circulatory and skin* patterns showed the highest probabilities of transitioning to death after 12 years in the three age groups, respectively.

Characterization of multimorbidity patterns

Estimations of the longitudinal trends (predicted values from linear mixed models) for different clinical and functional variables by patterns and for each age group are shown in Figure 3. An increasing trend was observed for the number of chronic conditions and drugs across age groups, with subjects in the *Unspecific* patterns consistently showing the lowest values. Conversely, a decreasing trend was observed for walking speed and MMSE in all age groups. While subjects in the *Unspecific* patterns showed the slowest changes over time, except for octogenarians, those in the patterns characterized by cardiovascular and/or neurological diseases showed the worse baseline values and fastest declines for all studied variables.

DISCUSSION

In this study we identified and characterized longitudinal multimorbidity patterns among older adults from a Swedish urban population, and estimated the time they spent in each pattern as well as the probability of transitioning across different patterns throughout a 12-year follow-up period.

Our findings highlight the dynamism and heterogeneity underlying multimorbidity. The dynamism among multimorbidity patterns was reflected by the varying sojourn times across patterns, which differed by age group, and the specific patterns people presented with. In sexagenarians, the average time was 13.3 years, while in octogenarians and beyond, it

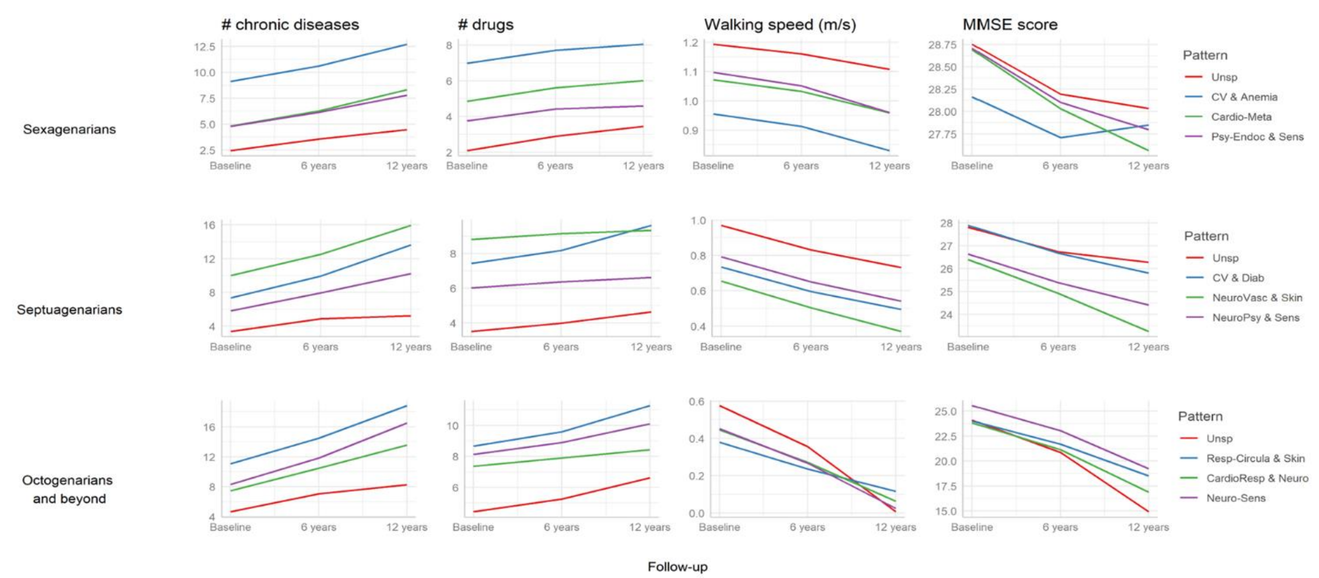


Figure 3. Longitudinal trends (predicted values from linear mixed models) in clinical and functional characteristics associated with the multimorbidity patterns by age group (N=3,363). Sexagenarians: Unspecific (Unsp); Cardiovascular and anemia (CV and Anemia); Cardio-metabolic (Cardio-Meta) and Psychiatric-endocrine and sensorial (Psy-Endoc and Sens). Septuagenarians: Unspecific (Unsp); Cardiovascular and diabetes (CV and Diab); Neuro-vascular and skin-sensorial (NeuroVasc and Skin); and Neuro-psychiatric and sensorial (NeuroPsy and Sens). Octogenarians and beyond: Unspecific (Unsp); Respiratory-circulatory and skin (Resp-Circula and Skin); Cardio-respiratory and neurological (CardioResp and Neuro); and Neuro-sensorial (Neuro-Sens). MMSE: Mini Mental State Examination.

was 6.5 years. This observation implies that, as expected, the time of permanence in each pattern is greater in the younger age groups, especially when less burdensome patterns are at play. For example, the *Unspecific* pattern was characterized in all age groups by a lack of overrepresentation of any of the low-severity chronic conditions the pattern was composed of (e.g., cardio-vascular risk factors, osteoarthritis, hearing impairment, etc.). Consequently, people belonging to this pattern could be regarded as being the healthiest, and thus the target for primary and secondary preventive strategies. Indeed, almost one third of sexagenarians in the *Unspecific* pattern at baseline transitioned to the *Psychiatric and sensorial* pattern, and almost one in ten to the *Cardio-metabolic* pattern during the follow-up. The heterogeneity of multimorbidity was evidenced by the different patterns obtained within, but especially, across age groups. Despite being similar, patterns at different ages represent different states of the disease severity continuum. These different stages may be associated with differential probabilities of developing complications and functional decline, and may trigger different pharmacological and non-pharmacological treatments. In relation to mortality, trajectories characterized by cardiovascular and circulatory diseases were found to concentrate the highest death probabilities. All these aspects may contribute to increase the heterogeneity of the multimorbidity landscape.

Moreover, our study serves as an example of how longitudinal data may be used to explore the trajectories of multimorbidity – that is, the evolution of and transitions among patterns of diseases. To date, studies on patterns of multimorbidity have predominantly focused on analyzing the association between diseases, paying less attention to individuals' "journeys" in and out of these patterns [5, 10]. This is mainly because most studies, even those using longitudinal data [11], were based on cross-sectional designs. Indeed, studies incorporating the entire longitudinal structure of the data are scarce [12–14]. Studying patterns of multimorbidity longitudinally is a challenging endeavor given that the heterogeneity in disease clustering originates both from the cross-sectional and longitudinal axes. Therefore, to understand the interdependence among diseases when looking at longitudinal multimorbidity patterns, dynamic machine learning methodologies such as the HMM are required. These models integrate a dynamic Bayesian network that accounts for the temporal sequence of the person-level data observed. This allows considering the longitudinal structure of the data (i.e., time series) and the correlations among observations.

Comparing our results with those from previous studies is difficult for the reasons mentioned above. Nevertheless, two previous studies analyzed disease progression and multimorbidity pattern trajectories using primary care electronic health records in the United Kingdom [15] and the Netherlands [16]. The studies by Strauss et al. [15] and Lappenschaar et al. [16] were carried out on adult populations, older than 35 and 50, and with a follow-up period of 3 and 5 years, respectively; and both included a lower number of chronic diseases than that used in this study. In terms of the analytical approach, the latent class growth models employed by Strauss et al. are designed to identify longitudinal trajectories, but one cannot infer transitions among classes. Also, Lappenschaar et al. used multilevel temporal Bayesian networks, which are aimed at analyzing the relationships between diseases (i.e., networks) but not the transitions across clusters. Other studies [17, 18] have focused on the incidence of new chronic diseases across time, but failed to examine patterns of multimorbidity. In brief, none of the previously applied statistical methods makes it possible to study the evolution and transitions between patterns of multimorbidity. In contrast, when applying HMM, one can explore the variability of chronic disease evolution over time by considering each subject's diseases as random variables conditioned by a hidden or conglomerate state, which further enables depicting people's transitions among different patterns of multimorbidity. Other studies looking at multimorbidity patterns within large databases have considered disease trajectories rather than individual trajectories as the main axis of interest [19]. This approach, which is somewhat disease- rather than person-oriented, is limited by the inability to identify homogenous groups of patients. Another example is the work by Giannoula et al., which focused on the identification of complex time-dependent disease associations using dynamic time warping, a machine learning technique [20]. Similar problems are present in the study by Xu et al., which moreover only considered three pathologies [21].

This study has several strengths. First, thanks to the exhaustive clinical evaluation that SNAC-K participants undergo in each follow-up wave, the reliability of the diagnostic data, which moreover integrates data from electronic health records, lab tests and drug use, is optimal. Second, the statistical methods applied allowed us to cluster people by their co-occurring diseases taking both the cross-sectional and longitudinal axes into account: HMM and the fuzzy c-means cluster algorithm. The latter is the choice method for pattern recognition when clusters tend to overlap, which is often the case as older adults show a high prevalence of co-occurring conditions. Furthermore, in this study we were able to explore longitudinal multimorbidity patterns by age group and the time that people remained in each pattern. As far as we know, these aspects have not been previously studied and are key to personalized clinical decision-making. Moreover, by stratifying our study sample by decade age groups, we were able to account for the selection bias inherent to aging cohorts, whereby the oldest age groups tend to represent healthier individuals characterized by better biological and environmental living conditions.

Some limitations must also be considered. First, the relatively small size of the SNAC-K cohort and the further stratification of the study sample into three different age groups led to some of the patterns including few people (i.e., <14 people). However, the methods applied have been shown to be responsive enough for the identification of subgroups of people even in small samples. Additionally, the iterative estimation process and the number of realizations allowed us to maximize the likelihood of the models applied given the data. Second, participant dropout (14% within the first 6 years and 8% within the next 6 years) may have affected the cluster definition process. Still, to the best of our knowledge, this is an exceptionally low figure compared with studies of this type. Third, the discontinuous follow-up carried out in SNAC-K (i.e., every 3 or 6 years depending on the age of participants) may have affected the rate of disease detection and, consequently, the longitudinal cluster analysis, especially among people who died or dropped out during the observation period. To adapt to the assumptions of our study design, participant data were analyzed in accordance with the available follow-up waves, avoiding any data extrapolation. Last, differences in the baseline composition and evolution of patterns across age groups could be due to variations in exposure history, and not only to age, given that there is up to 40 years of a gap between the youngest and oldest subjects at study baseline.

The analysis of longitudinal multimorbidity patterns is fundamental for the provision of personalized medical care that is not based merely on the application of guidelines targeting each chronic condition individually. While some of our findings can be explained through known pathophysiological mechanisms, others may serve to generate new hypotheses worth exploring in future studies. Our statistical approach enabled us to model the evolution and transitions of multimorbidity over time, and the results of this could be applied in the interests of healthier aging. Moreover, the age-stratified analyses allowed us to identify which disease combinations and transitions were more prevalent in each decade. This information is key to defining specific care plans to prevent or delay the negative consequences of the most frequent diseases identified. The characterization of multimorbidity patterns using HMM could moreover be expanded, for instance, by aggregating information on complementary health indicators such as frailty and biological and physiological variables, which could further optimize patient stratification and management efforts.

Our study provides evidence that multimorbidity is dynamic and heterogeneous in old age. With increasing age, older adults experience decreasing clinical stability and progressively shorter permanence time within one same multimorbidity pattern. Moreover, a significant proportion ranging between 5.9%-22.6% belongs to an *Unspecific* pattern with a low burden of diseases and a promising preventive potential. Adding new variables related to drug use, environmental and genetic factors, and/or frailty to the longitudinal analysis of multimorbidity patterns may allow optimizing the epidemiological understanding and applicability of these models for patient-tailored prevention and management strategies.

MATERIALS AND METHODS

Study population

Longitudinal data from the population-based Swedish National study on Aging and Care in Kungsholmen (SNAC-K) was used [22]. The study population consisted of adults ≥60 years of age living in the community or in institutions in the Kungsholmen district of Stockholm, Sweden. A random sample of 11 age cohorts (ages 60, 66, 72, 78, 81, 84, 87, 90, 93, 96 and ≥99) born between 1898 and 1943 (the youngest and oldest age cohorts were oversampled) was invited to participate in the study. People who agreed to participate were evaluated for the first time between 2001 and 2004. Participants who were <78 years of age were then followed up every six years and participants ≥78 every three years. The present study is based on data collected at baseline, the six-year follow-up, and the 12-year follow-up. At baseline, 3363 people were examined (participation rate: 73%). For our study, the sample was stratified into three age groups: sexagenarians (age cohorts of 60 and 66 years), septuagenarians (age cohorts of 72 and 78 years) and octogenarians and beyond (age cohorts of 81 years and over).

Chronic diseases

At each follow-up wave, SNAC-K participants undergo an approximately five-hour-long comprehensive clinical and functional assessment carried out by trained physicians, nurses, and neuropsychologists. Physicians collect information on diagnoses via physical examination, medical history, examination of medical charts, self-reported information, and/or proxy interviews. Clinical parameters, lab tests, drug information, and inpatient and outpatient care data are also used to identify specific conditions. All diagnoses are coded in accordance with the International Classification of Diseases, 10th revision (ICD-10). In the current study we classified all the ICD-10 codes into 60 chronic disease categories in accordance with a clinically driven methodology [23]. In SNAC-K, drugs are coded in accordance with the Anatomical Therapeutic Chemical (ATC) classification.

Covariates

Information on demographics (age, sex, education) was collected during nurse interviews. We divided education into elementary, secondary, university or higher. Information about vital status was derived from death certificates provided by Statistics Sweden, the Swedish governmental statistics agency. Survival status was assessed throughout the follow-up period. Participants were considered lost to follow-up if they or a proxy declined to participate, could not be contacted, had moved out of the study area, or cancelled an assessment. Walking speed (m/s) was assessed by asking participants to walk 6 m at their usual speed or 2.44 m if the participant reported walking quite slowly [24]. Cognitive status was assessed by physicians using the Mini-Mental State Examination (MMSE), with a score range of 30 at best to 0 at worst [25].

Statistical analysis

The sample characteristics at baseline, the 6-year follow-up and the 12-year follow-up for all age groups were described as appropriate. Additionally, 3-year and 9-year follow-up data was considered for the group of octogenarians and beyond.

To model the temporal evolution of multimorbidity patterns and individuals' transitions across these patterns, a dynamic random process represented by a HMM was assumed [9]. Disease information from all individuals and across all follow-up waves is used by the HMM to identify so-called hidden states (i.e., longitudinal multimorbidity pattern). HMM estimates the transition probabilities between patterns, i.e., the probability that any individual moves from one pattern to another in a given time-frame. Furthermore, by using HMM, one can examine individuals' probability of following different longitudinal multimorbidity patterns, and subsequently identify the one that is most likely to happen.

The dataset was pre-processed by applying a Multiple Correspondence Analysis (MCA) to the categorical features (i.e., diseases), in order to reduce the dimensionality of the longitudinal dataset. To prevent statistical noise and spurious findings from the models, only diseases that achieved a median prevalence of 2% across all follow-up waves were included (Supplementary Table 1). Afterwards, a fuzzy segmentation procedure (Fuzzy C-means algorithm, FCM) [11] was applied on the new dataset to identify an initial set of clusters, which was used to initialize some of the HMM parameters in the next stage. Finally, two more clusters were added in order to account for dropout and/or death.

The set of HMM parameters, composed of the initial cluster probabilities, the inter-cluster transition probabilities and the emission distributions provided by the FCM, were fitted into the observation dataset by applying the Baum-Welch (BW) algorithm. This made it possible to infer the longitudinal trajectories followed by each individual. The best cluster trajectory was identified by maximizing the probability of the observed sequence conditioned to the computed model parameters (Viterbi Algorithm). To validate the model, a comparison between BW and Viterbi transition probability matrices was conducted, showing a good agreement between theoretical and observed values [26].

The time unit considered for each transition across clusters/states was the time between follow-up waves, 6 years for sexagenarians and septuagenarians and 3 years for octogenarians and beyond. The time spent in a specific cluster/state before moving to other clusters/ states was assumed to follow a geometric distribution. Subsequently, the expected average time spent or mean sojourn (permanence) time was computed.

To optimize the performance of the selected mathematical model, the iterative process involved in the application of the BW algorithm was initialized using a range of 100 different values of the parameters to be learned. The best model was selected using a procedure that is equivalent to applying the Bayes Information Criterion to choose the best set of HMM parameters [9].

Multimorbidity patterns

For each age group, a final number of longitudinal patterns was selected. To evaluate the consistency and utility of the final clusters, we contrasted the clinical relevance of our findings in the context of previous literature, and we dicussed the findings within the research team (2 GPs, 2 geriatricians, 3 epidemiologists and 2 statisticians).

To characterize the multimorbidity patterns, we calculated the frequency of chronic diseases in each cluster. Observed/expected ratios (O/E-ratios) were calculated by dividing the prevalence of a given disease within a cluster by its prevalence in the overall population. The exclusivity of different diseases, defined as the fraction of participants with the disease in the cluster over the total number of participants with the disease, was also calculated. We considered a disease to be associated with a given cluster of individuals when the O/E ratio was ≥2 or the exclusivity was ≥ 20% [12]. Such criteria were used to name multimorbidity patterns after the diseases that predominantly characterized them.

The longitudinal trends of clinical and functional characteristics (no. of chronic diseases, no. of drugs, walking speed and MMSE) associated with the multimorbidity patterns were estimated through linear mixed models, assuming a random intercept and including an interaction between the patterns and follow-up time, both as linear and quadratic. The models were additionally adjusted by age, sex and education.

The analyses were carried out using Stata version 17 and R version 4.1.2. The significance level was set at $\alpha=0.05$.

AUTHOR CONTRIBUTIONS

ARL, DLV, CV and ACL developed the study concept and design. ARL performed the data analysis, and DLV, CV and ACL contributed to the interpretation of the results. ARL drafted the manuscript. All authors provided critical revisions and approved the final version of the manuscript for submission.

ACKNOWLEDGMENTS

We thank the SNAC-K participants and the SNAC-K Group for their collaboration in data collection and management.

CONFLICTS OF INTEREST

The authors have no conflicts.

ETHICAL STATEMENT AND CONSENT

The study was approved by the Regional Ethics Review Board in Stockholm. Participants in the study completed a written informed consent form as stipulated by the ethics board. For participants with prevalent or incident cognitive impairment, consent was obtained from next of kin.

FUNDING

This work was supported by the funders of the Swedish National study on Aging and Care (SNAC): the Ministry of Health and Social Affairs, Sweden; the participating county councils and municipalities; and the Swedish Research Council. Specific grants were obtained from the Swedish Research Council for Medicine (VR; 521-2013-8676; 2016-00981; 2017-06088; 2021-03324) and the Swedish Research Council for Health, Working Life and Welfare (Forte; 2016-07175; 2017-01764). This work was supported by Spain’s Ministry of Science and Innovation through the Carlos III Health Institute and by European Union ERDF (European Regional Development Fund) funds through the Research Network in Preventive Activities and Health Promotion in Primary Care (redIAPP, RD16/0007/0001). This study has also been funded by the Carlos III Health Institute through the project “PI19/00535” (co-funded by the ERDF; “A way to make Europe”). The IDIAP Jordi Gol also supported this work by offering the first author office space and access to the institution’s resources throughout the research process.

REFERENCES

1. Kontis V, Bennett JE, Mathers CD, Li G, Foreman K, Ezzati M. Future life expectancy in 35 industrialised countries: projections with a Bayesian model ensemble. Lancet. 2017; 389:1323–35. https://doi.org/10.1016/S0140-6736(16)32381-9 PMID:28236464
2. Oeppen J, Vaupel JW. Demography. Broken limits to life expectancy. Science. 2002; 296:1029–31. https://doi.org/10.1126/science.1069675 PMID:12004104
3. Chudasama YV, Khunti K, Gillies CL, Dhalwani NN, Davies MJ, Yates T, Zaccardi F. Healthy lifestyle and life expectancy in people with multimorbidity in the UK Biobank: A longitudinal cohort study. PLoS Med. 2020; 17:e1003332. https://doi.org/10.1371/journal.pmed.1003332 PMID:32960883
4. González-González AI, Meid AD, Dinh TS, Blom JW, van den Akker M, Elders PJ, Thiem U, De Gaudry DK, Swart KM, Rudolf H, Bosch-Lenders D, Trampisch HJ, Meerpohl JJ, et al. A prognostic model predicted deterioration in health-related quality of life in older patients with multimorbidity and polypharmacy. J Clin Epidemiol. 2021; 130:1–12. https://doi.org/10.1016/j.jclinepi.2020.10.006 PMID:33065164

5. Prados-Torres A, Calderón-Larrañaga A, Hancco-Saavedra J, Poblador-Plou B, van den Akker M. Multimorbidity patterns: a systematic review. J Clin Epidemiol. 2014; 67:254–66.
https://doi.org/10.1016/j.jclinepi.2013.09.021
PMID:24472295

6. Violan C, Foguet-Boreu Q, Flores-Mateo G, Salisbury C, Blom J, Freitag M, Glynn L, Muth C, Valderas JM. Prevalence, determinants and patterns of multimorbidity in primary care: a systematic review of observational studies. PLoS One. 2014; 9:e102149.
https://doi.org/10.1371/journal.pone.0102149
PMID:25048354

7. Ho IS, Azcoaga-Lorenzo A, Akbari A, Black C, Davies J, Hodgins P, Khunti K, Kadam U, Lyons RA, McCowan C, Mercer S, Nirantharakumar K, Guthrie B. Examining variation in the measurement of multimorbidity in research: a systematic review of 566 studies. Lancet Public Health. 2021; 6:e587–97.
https://doi.org/10.1016/S2468-2667(21)00107-9
PMID:34166630

8. Cezard G, McHale CT, Sullivan F, Bowles JK, Keenan K. Studying trajectories of multimorbidity: a systematic scoping review of longitudinal approaches and evidence. BMJ Open. 2021; 11:e048485.
https://doi.org/10.1136/bmjopen-2020-048485
PMID:34810182

9. Violán C, Fernández-Bertolín S, Guisado-Clavero M, Foguet-Boreu Q, Valderas JM, Vidal Manzano J, Roso-Llorach A, Cabrera-Bean M. Five-year trajectories of multimorbidity patterns in an elderly Mediterranean population using Hidden Markov Models. Sci Rep. 2020; 10:16879.
https://doi.org/10.1038/s41598-020-73231-9
PMID:33037233

10. Ng SK, Tawiah R, Sawyer M, Scuffham P. Patterns of multimorbid health conditions: a systematic review of analytical methods and comparison analysis. Int J Epidemiol. 2018; 47:1687–704.
https://doi.org/10.1093/ije/dyy134
PMID:30016472

11. Vetrano DL, Roso-Llorach A, Fernández S, Guisado-Clavero M, Violán C, Onder G, Fratiglioni L, Calderón-Larrañaga A, Marengoni A. Twelve-year clinical trajectories of multimorbidity in a population of older adults. Nat Commun. 2020; 11:3223.
https://doi.org/10.1038/s41467-020-16780-x
PMID:32591506

12. Guisado-Clavero M, Roso-Llorach A, López-Jimenez T, Pons-Vigués M, Foguet-Boreu Q, Muñoz MA, Violán C. Multimorbidity patterns in the elderly: a prospective cohort study with cluster analysis. BMC Geriatr. 2018; 18:16.
https://doi.org/10.1186/s12877-018-0705-7
PMID:29338690

13. Violán C, Foguet-Boreu Q, Fernández-Bertolín S, Guisado-Clavero M, Cabrera-Bean M, Formiga F, Valderas JM, Roso-Llorach A. Soft clustering using real-world data for the identification of multimorbidity patterns in an elderly population: cross-sectional study in a Mediterranean population. BMJ Open. 2019; 9:e029594.
https://doi.org/10.1136/bmjopen-2019-029594
PMID:31471439

14. France EF, Wyke S, Gunn JM, Mair FS, McLean G, Mercer SW. Multimorbidity in primary care: a systematic review of prospective cohort studies. Br J Gen Pract. 2012; 62:e297–307.
https://doi.org/10.3399/bjgp12X636146
PMID:22520918

15. Strauss VY, Jones PW, Kadam UT, Jordan KP. Distinct trajectories of multimorbidity in primary care were identified using latent class growth analysis. J Clin Epidemiol. 2014; 67:1163–71.
https://doi.org/10.1016/j.jclinepi.2014.06.003
PMID:25063556

16. Lappenschaar M, Hommersom A, Lucas PJ, Lagro J, Visscher S, Korevaar JC, Schellevis FG. Multilevel temporal Bayesian networks can model longitudinal change in multimorbidity. J Clin Epidemiol. 2013; 66:1405–16.
https://doi.org/10.1016/j.jclinepi.2013.06.018
PMID:24035172

17. Vos R, van den Akker M, Boesten J, Robertson C, Metsemakers J. Trajectories of multimorbidity: exploring patterns of multimorbidity in patients with more than ten chronic health problems in life course. BMC Fam Pract. 2015; 16:2.
https://doi.org/10.1186/s12875-014-0213-6
PMID:25608728

18. Ruel G, Lévesque JF, Stocks N, Sirois C, Kroger E, Adams RJ, Doucet M, Taylor AW. Understanding the evolution of multimorbidity: evidences from the North West Adelaide Health Longitudinal Study (NWAHS). PLoS One. 2014; 9:e96291.
https://doi.org/10.1371/journal.pone.0096291
PMID:24798485

19. Jensen AB, Moseley PL, Oprea TI, Ellesøe SG, Eriksson R, Schmock H, Jensen PB, Jensen LJ, Brunak S. Temporal disease trajectories condensed from population-wide registry data covering 6.2 million patients. Nat Commun. 2014; 5:4022.
https://doi.org/10.1038/ncomms5022 PMID:24959948

20. Giannoula A, Gutierrez-Sacristán A, Bravo Á, Sanz F, Furlong LI. Identifying temporal patterns in

patient disease trajectories using dynamic time warping: A population-based study. Sci Rep. 2018; 8:4216.
https://doi.org/10.1038/s41598-018-22578-1
PMID:29523868

21. Xu X, Mishra GD, Dobson AJ, Jones M. Progression of diabetes, heart disease, and stroke multimorbidity in middle-aged women: A 20-year cohort study. PLoS Med. 2018; 15:e1002516.
https://doi.org/10.1371/journal.pmed.1002516
PMID:29534066

22. Lagergren M, Fratiglioni L, Hallberg IR, Berglund J, Elmståhl S, Hagberg B, Holst G, Rennemark M, Sjölund BM, Thorslund M, Wiberg I, Winblad B, Wimo A. A longitudinal study integrating population, care and social services data. The Swedish National study on Aging and Care (SNAC). Aging Clin Exp Res. 2004; 16:158–68.
https://doi.org/10.1007/BF03324546
PMID:15195992

23. Calderón-Larrañaga A, Vetrano DL, Onder G, Gimeno-Feliu LA, Coscollar-Santaliestra C, Carfí A, Pisciotta MS, Angleman S, Melis RJ, Santoni G, Mangialasche F, Rizzuto D, Welmer AK, et al. Assessing and Measuring Chronic Multimorbidity in the Older Population: A Proposal for Its Operationalization. J Gerontol A Biol Sci Med Sci. 2017; 72:1417–23.
https://doi.org/10.1093/gerona/glw233
PMID:28003375

24. Studenski S, Perera S, Patel K, Rosano C, Faulkner K, Inzitari M, Brach J, Chandler J, Cawthon P, Connor EB, Nevitt M, Visser M, Kritchevsky S, et al. Gait speed and survival in older adults. JAMA. 2011; 305:50–8.
https://doi.org/10.1001/jama.2010.1923
PMID:21205966

25. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res. 1975; 12:189–98.
https://doi.org/10.1016/0022-3956(75)90026-6
PMID:1202204

26. Rabiner LR. A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition. Proc IEEE. 1989; 77:257–86.
https://doi.org/10.1109/5.18626

SUPPLEMENTARY MATERIALS

Supplementary Tables

Please browse Full Text version to see the data of Supplementary Tables 1–3.

Supplementary Table 1. Disease prevalence by age group and follow-up wave.

Supplementary Table 2. Description of multimorbidity patterns in terms of the top 10 diseases characterizing them by age group and follow-up wave.

Supplementary Table 3. Description of multimorbidity patterns in terms of sociodemographic, clinical and functional characteristics by age group and follow-up wave.

7. Discussion

7.1. Main findings

The aim of this doctoral thesis was to propose a statistical and machine learning methodology to estimate the longitudinal nature of multimorbidity patterns in a Swedish population-based cohort of older adults followed up for 12 years.

Our results were reported in the articles described in the previous section. The main findings can be summarized as follows:

1) When we applied a soft clustering technique (fuzzy c-means), around half of participants could be classified into five clinically meaningful clusters: respiratory and musculoskeletal diseases (RESP-MSK; 15.7%), eye diseases and cancer (EYE-CANCER; 10.7%), cognitive and sensory impairment (CNS-IMP; 10.6%), heart diseases (HEART; 9.3%), and psychiatric and respiratory diseases (PSY-RESP; 5.4%). Individuals in the CNS-IMP cluster were the oldest, had the greatest functional disability and were more likely to live in a nursing home. Participants in the HEART cluster had the highest number of co-occurring diseases and drugs, high values of inadequate physical activity and the highest mean values of serum creatinine and CRP. The PSY-RESP cluster was associated with higher levels of smoking, alcoholism and neuroticism. The other half of the cohort was grouped in an UNSPECIFIC cluster, which had the youngest individuals, the lowest number of co-occurring diseases and the best functional and cognitive status.
2) Clinical trajectories of older adults with multimorbidity are characterized by great dynamism and complexity but can still be tracked over time. At baseline, 52% of participants could be classified into five clinically meaningful clusters: psychiatric and respiratory diseases (5%), heart diseases (9%), respiratory and musculoskeletal diseases (16%), cognitive and sensory impairment (10%) and eye diseases and cancer (11%). The remaining 48% of participants (unspecified group) could not be grouped in any cluster at baseline but greatly contributed to the other clusters at follow-up assessments. Participants in this unspecific group were the youngest and healthiest and presented a high prevalence of cardiovascular risk factors; they were also the most likely to shift between clusters during follow-up periods, moving primarily to clusters characterized by cardiovascular, eye, respiratory and musculoskeletal diseases. Multimorbidity clusters that included cardiovascular and neuropsychiatric diseases (three at baseline and three at six years) presented a higher mortality risk (ORs ranging from 1.60 to 6.00;

p < 0.05 for all) than the group of participants who were not part of any cluster. Clusters characterized by cardiovascular and neuropsychiatric diseases included 25% of the study population at baseline and 28% of participants at six years, and they accounted for 51% of deaths at six years and 57% of deaths at twelve years.

3) When applying HMMs to model the longitudinal nature of multimorbidity, we identified four longitudinal multimorbidity patterns for each decade. An unspecific pattern lacking any overrepresented diseases had the shortest permanence time for all age groups (range: 4.6 years to 10.9 years), but the pattern with the longest permanence time varied by age. Sexagenarians remained longest in the Psychiatric-endocrine and sensorial pattern (15.4 years), septuagenarians in the Neuro-vascular and skin-sensorial pattern (11.0 years) and octogenarians and beyond in the Neuro-sensorial pattern (8.9 years). Transition probabilities varied across age groups, with sexagenarians showing the highest levels of stability. In relation to mortality, trajectories characterized by cardiovascular and circulatory diseases were associated with the highest death probabilities.

7.2. Discussion by aims

This section will explore to what extent the three studies met their specific aims (see section 4), and how the results compare with those of previous studies in the field.

7.2.1. Aim 1

Aim 1 was to identify clusters of older people based on their multimorbidity patterns, and to analyze differences among clusters according to sociodemographic, lifestyle, clinical and functional characteristics. The results of Study 1 validate the hypothesis 'Multimorbidity patterns differ according to sociodemographic, lifestyle, clinical, and functional characteristics'.

By applying fuzzy c-means, we identified six clusters of multimorbidity. Each cluster differed in terms of overrepresented diseases, as expected. Furthermore, each cluster showed significant differences in non-clustered variables.

For example, people included in the UNSPECIFIC cluster were younger and had fewer co-occurring diseases, lower drug usage and good functional and cognitive abilities. Other studies have shown similar results (33).

The PSY-RESP cluster included people with asthma and psychiatric conditions. The co-occurrence of these diseases could be a result of chronic drug treatment with steroids, which can increase neuroticism, depression and sleep disorders (142). In addition, asthma symptoms have been associated with depression, also in older adults (143). This cluster included relatively young people with alcohol abuse problems and low life satisfaction. The association between alcohol use and psychiatric disorders is well known (144), and this first study confirms this association in older people. Previous studies have also reported poor quality of life in people affected by psychiatric and respiratory disorders (145,146).

The HEART cluster illustrates the established link between cardio- and cerebrovascular diseases. Atrial fibrillation and heart failure are both risk factors for stroke (147), and diabetes is a risk factor for stroke and coronary heart disease (148). The high prevalence of migraine may be related to cerebrovascular pathology or the drugs prescribed for cardiac diseases, such as nitrates (149,150). Individuals in the HEART cluster were characterized by having a high number of co-occurring chronic conditions and using several drugs. This cluster had the second highest percentage of individuals, after the CNS-IMP cluster, with limitations in activities of daily living and instrumental activities of daily living; and the highest serum creatinine and CRP levels. Expression of pro-inflammatory cytokines increases throughout the human lifespan, and this increase is correlated with cardiovascular health (151). Chronic low-grade inflammation, in turn, promotes autonomic imbalance, stimulates remodeling, depresses cardiac function, prompts endothelial dysfunction and leads to a progression of atherosclerosis and impaired renal function (152).

The RESP-MSK cluster included osteoporosis, which may be related to chronic steroid treatment for asthma and chronic obstructive pulmonary disease (COPD) (153). Vitamin D deficiency is a common factor in respiratory and skeletal disorders; vitamin D supplements are beneficial both in preventing exacerbations of COPD and improving bone density measures (154,155). The presence of upper gastrointestinal system disorders may also be related to the treatment of respiratory diseases and the use of diphosphonates in osteoporosis (156), while thyroid and other autoimmune disease are often correlated (157).

The EYE-CANCER cluster included several eye impairments and solid cancers. A high percentage of participants in this cluster were widowed, which is explained by their higher age. Old age may also explain why they had the lowest grip strength (158). The CNS-IMP cluster illustrated the association recently found between sensorial impairment and dementia. Hearing deficits have attracted much interest, owing to the strong evidence that impaired hearing is a risk factor for cognitive decline and dementia (159,160). Cross-sectional and longitudinal studies have evaluated the relationship between vision loss and dementia; the 3C cohort study suggested that poor vision, in particular near vision loss, may be an indicator of dementia risk in the short- and mid-term (161). Retinal microvasculature pathology has been associated with vascular dementia, especially in people with diabetes (162). Multiple sensorial impairments have also been found to increase dementia risk (163). Individuals in the CNS-IMP cluster were very old and had the worst levels of physical and cognitive function; these factors explain why 43% of them were living in a nursing home (164). Any disease in the cluster could explain the functional impairment, particularly dementia and cerebrovascular diseases (165). This cluster included the highest percentage of manual workers with low education. Low educational attainment and a manual occupation during early life have been consistently associated with an increased risk of dementia (166) and poor income in later life. Finally, this cluster was also characterized by the highest percentage of people with a poor social network and inadequate physical activity levels.

There are few similarities between the findings of this first study and the literature in the field, because of differing methodological approaches. Previous studies have focused on the clustering of diseases rather than individuals, finding that the most consistent patterns were of cardiovascular, neuropsychiatric and musculoskeletal diseases (167). These patterns were mainly examined from the perspective of etiopathological pathways underlying disease coexistence. They have never before been analyzed in the context of sociodemographic, lifestyle, clinical and functional characteristics.

Two recent studies applied the fuzzy c-means technique in different settings. Violan et al. (79) identified eight multimorbidity patterns in a large primary care database from Catalonia with almost one million people aged 65 and older. The model identified clusters like HEART, CNS-IMP and UNSPECIFIC, as in our first study, although the authors of the Catalonian study included age and sex as clustering variables. The differences in the remaining clusters between the Catalan general population and the SNAC-K cohort may be due to sociodemographic characteristics as well the methodological approach. The source of information may also play a role, as the variables and diagnoses in our study may be more curated than those in the Catalan study, which was based on real-world data.

Bare et al. analyzed a small sample of patients aged 65 and older who had been hospitalized following an exacerbation of their chronic conditions (168). Clustering variables included active chronic conditions and geriatric syndromes, and the analysis produced four statistically and clinically significant multimorbidity patterns: osteoarticular, psychogeriatric, cardiorespiratory and unspecific. Despite differences in study design and setting between Bare el al. and our first study, some of the resulting clusters were comparable, especially the cardio and unspecific clusters.

7.2.2. Aim 2

Aim 2 was to identify multimorbidity patterns, trace their evolution and detect clinical trajectories and mortality over time. The results of Study 2 confirm the hypothesis ‘Multimorbidity patterns change over time. Clinical trajectories and mortality depend on the longitudinal multimorbidity pattern’.

After applying fuzzy c-means, we identified six clusters of multimorbidity at each follow-up period. Cluster composition varied at each timepoint, and mortality was dependent on the cluster trajectory of each individual.

Increasingly, studies are analyzing clusters of multimorbidity across different populations, settings and countries, but most studies have a cross-sectional design or focus on the progression of comorbidities of index diseases (104,105). There is scarce evidence on how clusters of multimorbidity change over time. Previous studies have used primary care records, hospital-based registries or self-reported diagnoses; included only middle-aged people; or examined both acute and chronic conditions. All these factors limit the possibility of comparing our findings with those published in the literature.

One study from Catalonia used a similar analytical strategy to ours on a large data set extracted from electronic primary health care records (33). It identified six multimorbidity clusters: musculoskeletal, endocrine-metabolic, digestive/respiratory, neuropsychiatric, cardiovascular and an unspecific group. These clusters exhibited less variation over the six years of follow-up than the patterns identified in our second study, possibly because our follow-up period was longer. The use of electronic health records in the Catalan study may have led to underdetection of less severe diseases and multimorbidity (169).

One study from the Netherlands used a large data set from primary care records and focused on six cardiovascular conditions. The authors concluded that the more diseases present at baseline, the higher the cumulative incidence rates of one or more new diseases (up to 47% at three-year follow-up and up to 76% at five-year follow-up) (110).

Another study of a population-wide registry in Denmark including more than six million patients showed more than 1000 significant longitudinal disease trajectories and some major multimorbidity clusters characterized by prostate disease, chronic obstructive pulmonary disease, cerebrovascular disease, cardiovascular disease and diabetes mellitus. The study was limited by the retrospective collection of data from a registry of hospital primary and secondary diagnostic codes. Because the authors included both chronic and acute diseases, their findings are not readily comparable with ours (111).

Finally, one Australian study followed up more than 13,000 middle-aged women with no history of diabetes, heart disease or stroke at baseline for 20 years to evaluate the longitudinal progression of the three conditions. Over 20 years, 18% of the women developed at least one condition, and 16.8% had two or three. Moreover, the onset of stroke was strongly associated with an increased risk of progressing to the other two diseases (115). In contrast, our study showed the opposite transition, from cardiovascular risk factors such as diabetes to overt cardiovascular and neuropsychiatric diseases. In the Australian study, social inequality, obesity, hypertension, physical inactivity, smoking, and other chronic conditions were significantly associated with each of the three diseases, but also with their co-occurrence. The study used self-reported diagnoses (115).

Regarding the analytical approach, studies that have used multistate models to define transitions between chronic disease population have included a fixed and small number of different chronic conditions (117, 118). Moreover, multistate model studies have considered only single diseases or small combinations of diseases rather than multimorbidity patterns that include an exhaustive list of chronic conditions. For example, the study of Freisling et al. assumed a multistate modelling for transitions to cancer, CVD, type 2 diabetes and subsequently to multimorbidity state (118).

In summary, sample selection, the lack of clinical assessment of disease, the use of electronic health records and different analytical approaches in previous studies mean their results cannot be easily compared with ours.

7.2.3. Aim 3

Aim 3 was to estimate the longitudinal evolution of older individuals as they move among patterns, using statistical and machine learning methods to detect the dynamics underlying such patterns**.** The results of Study 3 confirmed the hypothesis ‘People’s longitudinal shifts from one pattern to another over time depend on individual characteristics and multimorbidity evolution’.

After applying HMMs, we identified four clusters of multimorbidity for each group during the follow-up period. We calculated transitions to other clusters that depended on the individual multimorbidity patterns, and we estimated the expected time in each pattern. This sojourn time varied between patterns. We also assessed frailty evolution, finding differences in rate of decline between longitudinal patterns.

Comparing our results with those of previous studies is difficult for the reasons mentioned in A1 and A2. Nevertheless, two previous studies analyzed disease progression and multimorbidity pattern trajectories using primary care electronic health records in the UK (108) and the Netherlands (110,170).

Strauss et al. (108) followed up adults aged 35 years or older for three years, while Lappenschaar et al. (110) followed up adults aged 50 years and older for five years. Both studies included fewer chronic diseases than the SNAC-K study. In terms of the analytical approach, the latent class growth models employed by Strauss et al. were designed to identify longitudinal trajectories, but cannot identify transitions among classes. Lappenschaar et al. used multilevel temporal Bayesian networks, which are aimed at analyzing relationships between diseases (i.e., networks) but not transitions across clusters. Other studies (91,170) have focused on the incidence of new chronic diseases over time without examining patterns of multimorbidity.

Other studies that have investigated multimorbidity patterns within large databases have considered disease trajectories rather than patient trajectories as the main axis of interest (111). The limitation of this approach is that it cannot identify homogenous groups of patients. Another example is the work by Giannoula et al., which focused on the identification of complex time-dependent disease associations using dynamic time warping, a machine learning technique (113). Similar problems are present in the study by Xu et al., which had the additional limitation of including only three pathologies (115), and the study by Alaeddini et al., which modelled disease transitions with Markov chain models placed in a latent regression Markov mixture model to incorporate subject–specific covariates (121). The authors used the Markov clustering algorithm to identify patterns of disease progression rather than obtaining longitudinal multimorbidity patterns. In brief, no previously applied statistical methods are suitable for studying the evolution of and transitions between patterns of multimorbidity. In contrast, when applying HMM, researchers can explore the variability of chronic disease evolution over time by considering each individual's diseases as random variables conditioned by a hidden or conglomerate state, which further enables depiction of transitions between different patterns of multimorbidity.

Of the few publications with similar methods to our third article, the most direct comparisons can be drawn with those of Violan and Villén (122,127), which were based on a large Catalan primary care sample of people aged 65 years and older, followed up for five years. The authors applied an HMM to identify ten multimorbidity patterns, considering two additional clusters for death and dropout. Although our Study 3 sample was stratified into 10-year age groups, we identified unspecific clusters of younger people with low burden of disease, similar to the Catalan study. In addition, the cardiovascular and neurologic patterns were present across all age groups. Nevertheless, there were some methodological differences, as the Catalan study had a shorter follow-up but with more observation time points, and a larger sample size. The smaller sample size in SNAC-K may have conditioned the algorithm performance to obtain a larger optimal number of clusters. In contrast, the use of electronic health records in the Catalan study may have led to underdetection of less severe diseases and multimorbidity variables (166).

7.3. Discussion of general aspects

7.3.1. Population

Cohort-based studies usually focus on a specific topic of interest, such as health examination; biological indicators; socioeconomic information; lifestyle information, including income, education, exercise and diet; or other qualitative data from questionnaires or interviews. However, they usually have limited years of follow-up with a suboptimal follow-up rate (171).

This thesis focuses on the SNAC-K cohort, which included individuals aged over 60 years from the Kungsholmen area of Stockholm who were followed over 12 years. We believe our studies, based on this high-quality population-based cohort data, represents an important scientific step within the field. Compared to data sets produced through routine collection from electronic health records, our cohort was relatively small. On the other hand, it included a comprehensive list of conditions, and the quality of data registration was high, which is not always possible in electronic health record databases. Cohort studies can obtain more detailed and customized variables while electronic health records can provide more data that are less subject to attrition or response bias (172).

7.3.2. Age

Most studies included in the review by Ho et al. examined multimorbidity in adults of any age (42·4%), in older adults (38·2%) and middle-aged and older adults (14·1%) (16). The most common age range is 65 years and older. Some authors consider it important to start studying multimorbidity from its onset around the age of 40 years.

The age range of our population was slightly different from that of other multimorbidity pattern studies. In addition, the SNAC-K investigators oversampled individuals from the oldest and the youngest birth cohorts. Stratifying the cohort into ten-year age groups represented a new approach to the epidemiological study of longitudinal multimorbidity patterns. With this approach, we were able to extract more detailed information from our analyses.

7.3.3. List of diseases

There is a clear lack of consensus on the operationalization of chronic diseases and multimorbidity, as highlighted by Ho et al. (16) and other groups, including ours. As a result, studies can use very different underlying measures, which makes it difficult to draw comparisons between them

To maximize the reproducibility of our study, we used a validated operational definition of chronic disease and multimorbidity. This methodology is based on a consensus definition of chronic disease, whereby an international multidisciplinary team classified all four-digit level ICD-10 codes as chronic or non-chronic, before grouping the chronic codes into broader categories according to clinical criteria.

This operational list can be used in most countries and settings. The full list can be found in the paper by Calderón-Larrañaga (27). More than 250 papers have adopted this approach, in Sweden (where it was originally developed) and beyond (Spain, Germany, etc. (28–31)).

7.3.4. Analytical approaches

By using the soft cluster algorithm (fuzzy c-means), we were able to identify the optimal number of clusters in our population following a robust methodology (78). Most previous studies have focused on diseases rather than individuals as the unit of analysis when assessing multimorbidity patterns. Compared with hierarchical clustering, fuzzy c-means cluster analysis is less susceptible to outliers in the data, to the choice of distance measure and to the inclusion of inappropriate or irrelevant variables. Moreover, hard clustering forces each person into a single cluster, whereas soft clustering assigns each individual a probability of membership to all identified clusters, which makes more sense from a biological perspective. In particular, soft clustering analysis allows simultaneous linking of individuals and diseases to multiple clusters and is more consistent with clinical experience than other approaches frequently found in the literature (79).

Unlike other statistical and machine learning methods formerly employed in the study of multimorbidity, HMMs account for the variability in chronic disease interactions over time (123,124). The longitudinal multimorbidity patterns obtained with HMM methods provide a comprehensive approach to the evolution of multimorbidity over a patient's lifetime. All longitudinal information is used in the model's estimation. The model assumes that the sequential individual observations follow a dynamic random process represented by an HMM, so that each cluster is associated with a hidden state or multimorbidity pattern. This assumption is crucial, because it allows a complete characterization of the evolution of the individual, all their transitions between clusters and their permanence time. Transition to other clusters depends on the evolution of the chronic diseases burden that an individual is accumulating longitudinally. The model can predict the pattern in which a person will be in the next few years, for example at six or 12 years, taking into account these diseases variables. By refocusing the analysis on individuals and considering all their longitudinal information, we can obtain a better characterization of the population groups with multimorbidity. Importantly, many diseases identified in the multimorbidity patterns have shared risk factors; consequently, preventive interventions in these chronic diseases could alter many trajectories and even shift causes of mortality (122).

7.3.5. Generalizability

The average higher socioeconomic background of participants in the SNAC-K study may limit the generalizability of our findings. The aim of scientific research is to supply generalizable results (i.e. results that can be applied to different populations). However, the SNAC-K population was found to be healthier and wealthier than the general population living in the Kungsholmen district of Stockholm, and there is likely to be an even greater disparity with the remainder of the Swedish population or the European population (128).

For these reasons, we advise caution when generalizing the results of our studies to other settings. Nevertheless, while generalizability may be an issue when inferring population epidemiological data (prevalence, incidence, etc.), it is less likely to affect associations between variables. We have demonstrated the biological plausibility of our findings and identified some well-known underlying biological mechanisms. For example, the Catalan population in Violan et al. and the SNAC-K population showed some similar patterns and transitions (122).

To summarize, if we understand the biological basis of a determined phenomenon, we can design better studies and account for specific confounders. Therefore, generalizability depends not only on sample characteristics, but also on good biological plausibility and knowledge of the underlying mechanisms.

7.4. Strengths

The main strength of this thesis was the high number of older people in the SNAC-K cohort and the comprehensive list of both mental and physical chronic conditions included in the analyses. Each participant in SNAC-K underwent a six-hour comprehensive assessment that followed a standard protocol and was carried out by a physician, a nurse, and a psychologist. Diseases were categorized using a strict clinically driven method developed and tested by our group (27). Moreover, by including both mental and physical conditions in the analyses, we were able to investigate the interplay – potentially bidirectional – between mental health problems and chronic physical conditions. Furthermore, the lack of missing information on disease status increases the internal validity of our study.

Other strengths included the long follow-up time and the large age range of the participants (60 years to 104 years). Regarding Study 3 design, by stratifying the study sample into ten-year age groups, we were able to account for the selection bias inherent to aging cohorts, whereby the oldest age groups tend to include healthier individuals characterized by better biological and environmental living conditions.

The statistical and machine learning methods applied in our studies constitute their main methodological strength. The fuzzy c-means cluster algorithm and HMM cluster people by their co-occurring diseases, taking both the cross-sectional and longitudinal axes into account. These methods make us of each individual's information over time and track their trajectories. The fuzzy c-means cluster algorithm is the method of choice for pattern recognition when clusters tend to overlap, which they often do in multimorbidity analysis, as older adults present high prevalence of co-occurring conditions. Furthermore, we were able to explore longitudinal multimorbidity patterns by age group, and to measure the time that people remained in each pattern. To the best of our knowledge, no previous study has examined these aspects, although they are key to personalized clinical decision-making. Another major strength is that the final clustering solution presented in each study was obtained through a systematic and rigorous process, which included comparing the results from a randomly split data set, testing different clustering algorithms, using different objective numeric criteria to decide the number of clusters, and applying subjective clinical criteria to assess whether the groupings were clinically interpretable.

7.5. Limitations

The main limitation of this thesis is inherent to the population-based cohort of individuals participating in the SNAC-K study. The investigators applied few exclusion criteria at baseline: age under 60 years, nonproficiency in the Swedish language and residency outside the Stockholm district of Kungsholmen. The final group of participants had better health and higher income compared to excluded people and compared to the Swedish population as a whole. The high socioeconomic status of SNAC-K participants may limit the generalizability of the findings of each study.

In Study 1, the cross-sectional design limited the analysis of multimorbidity pattern evolution, as some of the included sociodemographic, lifestyle, clinical and functional profiles were only measured at baseline.

The first important limitation of Study 2 and Study 3 is that diseases entered the model regardless of their severity. Disease severity may partially explain the clinical trajectories identified in the studies. Second, the dropout rates (14% at six years and 8% at 12 years) may have affected cluster definition. Third, the discontinuous follow-up in SNAC-K (every three or six years) may have affected disease detection and consequently the cluster analysis, especially among people who died or dropped out during the observation period.

We identified two methodological limitations related specifically to Study 2. First, although we defined longitudinal patterns, we performed cross-sectional cluster analysis at each timepoint. And second, an important disadvantage of fuzzy c-means is that different solutions can occur for each set of seed points, and there is no guarantee of optimal clustering.

Regarding Study 3, the relatively small size of the SNAC-K cohort and the further stratification of the study sample into three different age groups led to small numbers of participants (i.e. fewer than 14) in some patterns. In general, it was impossible to stratify by sex owing to the great imbalance in the oldest age groups, which comprised 64% to76% women. In addition, different initializations can be considered in the HMM and there is no guarantee of reaching a global optimum solution, since HMM obtains a local optimum instead.

7.6. Future research

The present thesis contributes to a better description of the nature of longitudinal multimorbidity patterns and their characteristics in older individuals. However, further research is needed to better understand the complexity of multimorbidity and its evolution.

First, future studies on the trajectory of multimorbidity patterns should follow up a younger population for a longer period of time, as the onset of certain chronic conditions occurs between 40 and 60 years of age (173). It is important to examine the time sequence of disease onset to help determine clinical signs that could lead to early diagnosis.

Second, some of the findings of this thesis can be confirmed in a large high-quality cohort using a more representative sample of the general population. In addition, large databases and longer follow-up periods with more observation time points could help to optimize clustering algorithm performance (174).

Third, researchers should make use of genetic databases to further investigate the evolution of multimorbidity. Multimorbidity patterns and trajectories are conditioned by genetic and non-genetic factors of the individual exposome, defined as the measure of all the exposures of an individual in a lifetime and how those exposures relate to health (175). Genetic profiling of disease and individuals could disentangle which disease or multimorbidity patterns account for the causal relationship of a risk disease trajectory. With further research, practices can begin to shift to a new paradigm of personalized medicine segmented by groups, where multimorbidity is a target of preventive activities and therapeutic guidelines.

Fourth, future studies must examine the interplay and dynamics of frailty and multimorbidity. Multimorbidity and frailty are characteristics of ageing that need to be assessed at the individual level (63,70). There is a need for research on multimorbidity patterns that considers frailty indexes and variables to help identify people with specific needs related to their chronic diseases and frailty deficits.

Finally, statistical and machine learning methodologies are in constant evolution. The modern techniques in multimorbidity applied in this thesis have their own strengths and limitations.

Cutting-edge new methodologies may help to improve the characterization of longitudinal multimorbidity patterns and overcome some of the limitations of fuzzy c-means and Hidden Markov Models. In particular, deep learning-based solutions can involve architectures based on recurrent neural networks, contextual embeddings (i.e., transformers) or other architectures (convolutional neural networks, fully connected networks, etc.) that can model the longitudinal evolution of multimorbidity patterns (176).

8. Conclusions

1) Multimorbidity patterns showed significantly different sociodemographic, lifestyle, clinical and functional profiles.

2) The younger and healthier half of the cohort was grouped into one unspecific cluster, while the other half was classified into clinically meaningful clusters.

3) Clinical trajectories of older adults with multimorbidity are characterized by great dynamism and complexity but can be tracked over time.

4) Different clusters contributed differentially to the longitudinal development of other clusters and were differentially associated with mortality.

5) With increasing age, multimorbidity patterns showed decreasing clinical stability.

6) Participants in the older age groups spent less time within a single multimorbidity pattern.

7) Walking speed and mental function evolved differently between longitudinal patterns, showing stable or fast declines.

8) Fuzzy c-means clustering, a soft clustering technique, was sufficiently flexible to assign people to more than one pattern.

9) Through Hidden Markov Models, a machine learning technique to model dynamic processes, we were able to track people's longitudinal shifts from one pattern to another over long periods of time.

10) Our results may help to clarify the complex interactions among co-occurring diseases over time and, more importantly, may help to improve preventive interventions and optimally address individuals' care needs and risk of adverse outcomes.

9. References

1. Kontis V, Bennett JE, Mathers CD, Li G, Foreman K, Ezzati M. Future life expectancy in 35 industrialised countries: projections with a Bayesian model ensemble. Lancet [Internet]. 2017;389(10076):1323–35. Available from: http://dx.doi.org/10.1016/S0140-6736(16)32381-9
2. Oeppen J, Vaupel JW. 10.1126_Science.1069675. Science (80-). 2002;296(5570):1029–31.
3. Partridge L, Deelen J, Slagboom PE. Facing up to the global challenges of ageing. Nature [Internet]. 2018;561(7721):45–56. Available from: http://dx.doi.org/10.1038/s41586-018-0457-8
4. Murray CJL, Barber RM, Foreman KJ, Ozgoren AA, Abd-Allah F, Abera SF, et al. Global, regional, and national disability-adjusted life years (DALYs) for 306 diseases and injuries and healthy life expectancy (HALE) for 188 countries, 1990-2013: Quantifying the epidemiological transition. Lancet. 2015;386(10009):2145–91.
5. Chudasama Y V., Khunti K, Gillies CL, Dhalwani NN, Davies MJ, Yates T, et al. Healthy lifestyle and life expectancy in people with multimorbidity in the UK Biobank: A longitudinal cohort study. PLoS Med [Internet]. 2020;17(9):1–18. Available from: http://dx.doi.org/10.1371/journal.pmed.1003332
6. MacMahon S, The Academy of Medical Sciences. Multimorbidity: a priority for global health research. Acad Med Sci [Internet]. 2018;(April). Available from: https://acmedsci.ac.uk/file-download/82222577
7. Fortin M, Stewart M, Poitras ME, Almirall J, Maddocks H. A systematic review of prevalence studies on multimorbidity: Toward a more uniform methodology. Ann Fam Med. 2012;10(2):142–51.
8. Uijen A, van de Lisdonk E. Multimorbidity in primary care: Prevalence and trend over the last 20 years. Eur J Gen Pract. 2008;14(SUPPL. 1):28–32.
9. Koné Pefoyo AJ, Bronskill SE, Gruneir A, Calzavara A, Thavorn K, Petrosyan Y, et al. The increasing burden and complexity of multimorbidity disease epidemiology - Chronic. BMC Public Health. 2015;15(1):1–11.
10. King DE, Xiang J, Pilkerton CS. Multimorbidity Trends in United States Adults, 1988–2014. J Am Board Fam Med [Internet]. 2018 Jul 9;31(4):503–13. Available from: http://www.jabfm.org/lookup/doi/10.3122/jabfm.2018.04.180008
11. Garin N, Koyanagi A, Chatterji S, Tyrovolas S, Olaya B, Leonardi M, et al. Global Multimorbidity Patterns: A Cross-Sectional, Population-Based, Multi-Country Study. Journals Gerontol - Ser A Biol Sci Med Sci. 2016;71(2):205–14.
12. Rivera-Almaraz A, Manrique-Espinoza B, Ávila-Funes JA, Chatterji S, Naidoo N, Kowal P, et al. Disability, quality of life and all-cause mortality in older Mexican adults: Association with multimorbidity and frailty. BMC Geriatr. 2018;18(1):1–9.
13. Arokiasamy P, Uttamacharya U, Jain K, Biritwum RB, Yawson AE, Wu F, et al. The impact of multimorbidity on adult physical and mental health in low- and middle-income countries: What does the study on global ageing and adult health (SAGE) reveal? BMC Med [Internet]. 2015;13(1):1–16. Available from: http://dx.doi.org/10.1186/s12916-

015-0402-8

14. Pearson-Stuttard J, Ezzati M, Gregg EW. Multimorbidity—a defining challenge for health systems. Lancet Public Heal. 2019;4(12):e599–600.

15. Moffat K, Mercer SW. Challenges of managing people with multimorbidity in today's healthcare systems. BMC Fam Pract [Internet]. 2015;16(1):15–7. Available from: http://dx.doi.org/10.1186/s12875-015-0344-4

16. Ho ISS, Azcoaga-Lorenzo A, Akbari A, Black C, Davies J, Hodgins P, et al. Examining variation in the measurement of multimorbidity in research: a systematic review of 566 studies. Lancet Public Heal [Internet]. 2021;6(8):e587–97. Available from: http://dx.doi.org/10.1016/S2468-2667(21)00107-9

17. Valderas JM, Starfi B, Sibbald B. Understanding Health and Health Services. Ann Fam Med. 2009;357–63.

18. Boyd CM, Fortin M. Future of multimorbidity research: How should understanding of multimorbidity inform health system design? Public Health Rev. 2011;33(2):451–74.

19. Xu X, Mishra GD, Jones M. Evidence on multimorbidity from definition to intervention: An overview of systematic reviews. Ageing Res Rev [Internet]. 2017;37:53–68. Available from: http://dx.doi.org/10.1016/j.arr.2017.05.003

20. Lefèvre T, d'Ivernois JF, De Andrade V, Crozet C, Lombrail P, Gagnayre R. What do we mean by multimorbidity? An analysis of the literature on multimorbidity measures, associated factors, and impact on health services organization. Rev Epidemiol Sante Publique. 2014;62(5):305–14.

21. Harrison C, Britt H, Miller G, Henderson J. Examining different measures of multimorbidity, using a large prospective cross-sectional study in Australian general practice. BMJ Open. 2014;4(7):1–9.

22. Violan C, Foguet-Boreu Q, Flores-Mateo G, Salisbury C, Blom J, Freitag M, et al. Prevalence, determinants and patterns of multimorbidity in primary care: A systematic review of observational studies. PLoS One. 2014;

23. Diederichs C, Berger K, Bartels DB. The measurement of multiple chronic diseases - A systematic review on existing multimorbidity indices. Journals Gerontol - Ser A Biol Sci Med Sci. 2011;66 A(3):301–11.

24. Van Den Brink AMA, Gerritsen DL, Voshaar RCO, Koopmans RTCM. Residents with mental-physical multimorbidity living in long-term care facilities: Prevalence and characteristics. A systematic review. Int Psychogeriatrics. 2013;25(4):531–48.

25. Nowels MA, Vanderwielen LM. Comorbidity indices: A call for the integration of physical and mental health. Prim Heal Care Res Dev. 2018;19(1):96–8.

26. Stirland LE, González-Saavedra L, Mullin DS, Ritchie CW, Muniz-Terrera G, Russ TC. Measuring multimorbidity beyond counting diseases: systematic review of community and population studies and guide to index choice. BMJ [Internet]. 2020 Feb 18;368:m160. Available from: https://www.bmj.com/lookup/doi/10.1136/bmj.m160

27. Calderón-Larrañaga A, Vetrano DL, Onder G, Gimeno-Feliu LA, Coscollar-Santaliestra C, Carfí A, et al. Assessing and Measuring Chronic Multimorbidity in the Older Population: A Proposal for Its Operationalization. Journals Gerontol - Ser A Biol Sci Med Sci. 2017;

28. Troncoso-Mariño A, Roso-Llorach A, López-Jiménez T, Villen N, Amado-Guirado E,

Fernández-Bertolin S, et al. Medication-related problems in older people with multimorbidity in Catalonia: A real-world data study with 5 years' follow-up. J Clin Med. 2021;10(4):1–14.

29. Grande G, Marengoni A, Vetrano DL, Roso-Llorach A, Rizzuto D, Zucchelli A, et al. Multimorbidity burden and dementia risk in older adults: The role of inflammation and genetics. Alzheimer's Dement. 2021;17(5):768–76.

30. Kim-Dorner S-J, Schmidt T, Kuhlmann A, Graf von der Schulenburg J-M, Welte T, Lingner H. Age- and gender-based comorbidity categories in general practitioner and pulmonology patients with COPD. NPJ Prim care Respir Med [Internet]. 2022;32(1):17. Available from: http://www.ncbi.nlm.nih.gov/pubmed/35501357%0Ahttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC9061861

31. Scholten M, Midlöv P, Halling A. Disparities in prevalence of heart failure according to age, multimorbidity level and socioeconomic status in southern Sweden: a cross-sectional study. BMJ Open. 2022;12(3):e051997.

32. Marengoni A, Angleman S, Melis R, Mangialasche F, Karp A, Garmen A, et al. Aging with multimorbidity: A systematic review of the literature. Ageing Research Reviews. 2011.

33. Guisado-Clavero M, Roso-Llorach A, López-Jimenez T, Pons-Vigués M, Foguet-Boreu Q, Muñoz MA, et al. Multimorbidity patterns in the elderly: A prospective cohort study with cluster analysis. BMC Geriatr. 2018;

34. Prados-Torres A, Calderón-Larrañaga A, Hancco-Saavedra J, Poblador-Plou B, Van Den Akker M. Multimorbidity patterns: A systematic review. J Clin Epidemiol. 2014;67(3):254–66.

35. Strom BL, Kimmel SE, Hennessy S. Pharmacoepidemiology [Internet]. Strom BL, Kimmel SE, Hennessy S, editors. Wiley; 2019. 1–7 p. Available from: https://onlinelibrary.wiley.com/doi/book/10.1002/9781119413431

36. The World Bank. Life expectancy at birth, total (years) [Internet]. 2019. Available from: https://data.worldbank.org/indicator/SP.DYN.LE00.IN

37. Thompson AE, Anisimowicz Y, Miedema B, Hogg W, Wodchis WP, Aubrey-Bassler K. The influence of gender and other patient characteristics on health care-seeking behaviour: A QUALICOPC study. BMC Fam Pract [Internet]. 2016;17(1):1–7. Available from: http://dx.doi.org/10.1186/s12875-016-0440-0

38. Vos R, Boesten J, van den Akker M. Fifteen-year trajectories of multimorbidity and polypharmacy in Dutch primary care—A longitudinal analysis of age and sex patterns. PLoS One [Internet]. 2022;17(2 February):1–23. Available from: http://dx.doi.org/10.1371/journal.pone.0264343

39. Barnett K, Mercer SW, Norbury M, Watt G, Wyke S, Guthrie B. Epidemiology of multimorbidity and implications for health care, research, and medical education: A cross-sectional study. Lancet. 2012;

40. Marengoni A, Vetrano DL. Multimorbidity: disease of society? Lancet Heal Longev [Internet]. 2021;2(8):e451–2. Available from: http://dx.doi.org/10.1016/S2666-7568(21)00167-7

41. Schramm S, Møller SP, Tolstrup JS, Laursen B. Effects of individual and parental educational levels on multimorbidity classes: A register-based longitudinal study in a

Danish population. BMJ Open. 2022;12(2):1–8.

42. Ryan A, Murphy C, Boland F, Galvin R, Smith SM. What is the impact of physical activity and physical function on the development of multimorbidity in older adults over time? A population-based cohort study. Journals Gerontol - Ser A Biol Sci Med Sci. 2018;73(11):1538–44.

43. Cimarras-Otal C, Calderón-Larrañaga A, Poblador-Plou B, González-Rubio F, Gimeno-Feliu LA, Arjol-Serrano JL, et al. Association between physical activity, multimorbidity, self-rated health and functional limitation in the Spanish population. BMC Public Health. 2014;14(1):1–10.

44. Dankel SJ, Loenneke JP, Loprinzi PD. Participation in muscle-strengthening activities as an alternative method for the prevention of multimorbidity. Prev Med (Baltim) [Internet]. 2015;81:54–7. Available from: http://dx.doi.org/10.1016/j.ypmed.2015.08.002

45. Dankel SJ, Loenneke JP, Loprinzi PD. Combined Associations of Muscle-Strengthening Activities and Accelerometer-Assessed Physical Activity on Multimorbidity: Findings from NHANES. Am J Heal Promot. 2017;31(4):274–7.

46. Dhalwani NN, O'Donovan G, Zaccardi F, Hamer M, Yates T, Davies M, et al. Long terms trends of multimorbidity and association with physical activity in older English population. Int J Behav Nutr Phys Act [Internet]. 2016;13(1):1–9. Available from: http://dx.doi.org/10.1186/s12966-016-0330-9

47. Calderón-Larrañaga A, Vetrano DL, Ferrucci L, Mercer SW, Marengoni A, Onder G, et al. Multimorbidity and functional impairment–bidirectional interplay, synergistic effects and common pathways. J Intern Med. 2019;285(3):255–71.

48. Brownlee S, Chalkidou K, Doust J, Elshaug AG, Glasziou P, Heath I, et al. Evidence for overuse of medical services around the world. Lancet [Internet]. 2017 Jul;390(10090):156–68. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0140673616325855

49. Foguet-Boreu Q, Violan C, Roso-Llorach A, Rodriguez-Blanco T, Pons-Vigués M, Muñoz-Pérez MA, et al. Impact of multimorbidity: Acute morbidity, area of residency and use of health services across the life span in a region of south Europe. BMC Fam Pract. 2014;15(1).

50. Khezrian M, McNeil CJ, Murray AD, Myint PK. An overview of prevalence, determinants and health outcomes of polypharmacy. Ther Adv Drug Saf. 2020;11:1–10.

51. González-González AI, Meid AD, Dinh TS, Blom JW, van den Akker M, Elders PJM, et al. A prognostic model predicted deterioration in health-related quality of life in older patients with multimorbidity and polypharmacy. J Clin Epidemiol. 2021;130:1–12.

52. Pazan F, Wehling M. Polypharmacy in older adults: a narrative review of definitions, epidemiology and consequences. Eur Geriatr Med [Internet]. 2021;12(3):443–52. Available from: https://doi.org/10.1007/s41999-021-00479-3

53. Varghese D, Ishida C, Haseer Koya H. Polypharmacy [Internet]. StatPearls. 2022. Available from: http://www.ncbi.nlm.nih.gov/pubmed/30422548

54. Report T. World Health Organization Technical Report; Medication Safety in Polypharmacy. 2019;

55. Stafford G, Villén N, Roso-Llorach A, Troncoso-Mariño A, Monteagudo M, Violán C. Combined multimorbidity and polypharmacy patterns in the elderly: A cross-sectional study in primary health care. Int J Environ Res Public Health. 2021;18(17):1–22.

56. Larkin J, Foley L, Smith SM, Harrington P, Clyne B. The experience of financial burden for people with multimorbidity: A systematic review of qualitative research. Heal Expect. 2020;(November 2020):282–95.

57. Clegg A, Young J, Iliffe S, Rikkert MO, Rockwood K. Frailty in elderly people. Lancet [Internet]. 2013 Mar;381(9868):752–62. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0140673612621679

58. Morley JE, Vellas B, Abellan van Kan G, Anker SD, Bauer JM, Bernabei R, et al. Frailty Consensus: A Call to Action. J Am Med Dir Assoc [Internet]. 2013 Jun;14(6):392–7. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1525861013001825

59. Hanlon P, Nicholl BI, Jani BD, Lee D, McQueenie R, Mair FS. Frailty and pre-frailty in middle-aged and older adults and its association with multimorbidity and mortality: a prospective analysis of 493 737 UK Biobank participants. Lancet Public Heal [Internet]. 2018;3(7):e323–32. Available from: http://dx.doi.org/10.1016/S2468-2667(18)30091-4

60. Fried LP, Tangen CM, Walston J, Newman AB, Hirsch C, Gottdiener J, et al. Frailty in older adults: Evidence for a phenotype. Journals Gerontol - Ser A Biol Sci Med Sci. 2001;56(3):146–57.

61. Walston J, Hadley EC, Ferrucci L, Guralnik JM, Newman AB, Studenski SA, et al. Research agenda for frailty in older adults: Toward a better understanding of physiology and etiology: Summary from the American Geriatrics Society/National Institute on Aging research conference on frailty in older adults. J Am Geriatr Soc. 2006;54(6):991–1001.

62. Rockwood K, Song X, MacKnight C, Bergman H, Hogan DB, McDowell I, et al. A global clinical measure of fitness and frailty in elderly people. Cmaj. 2005;173(5):489–95.

63. Orfila F, Carrasco-Ribelles LA, Abellana R, Roso-Llorach A, Cegri F, Reyes C, et al. Validation of an electronic frailty index with electronic health records: eFRAGICAP index. BMC Geriatr [Internet]. 2022 Dec 7;22(1):404. Available from: https://doi.org/10.21203/rs.3.rs-1368757/v1

64. Vetrano DL, Palmer K, Marengoni A, Marzetti E, Lattanzio F, Roller-Wirnsberger R, et al. Frailty and multimorbidity: A systematic review and meta-analysis. Journals Gerontol - Ser A Biol Sci Med Sci. 2019;74(5):659–66.

65. Vetrano DL, Calderón-Larrañaga A, Marengoni A, Onder G, Bauer JM, Cesari M, et al. An International Perspective on Chronic Multimorbidity: Approaching the Elephant in the Room. Journals Gerontol - Ser A Biol Sci Med Sci. 2018;73(10):1350–6.

66. Chen C, Park JY, Wu C, Xue QL, Agogo G, Han L, et al. Cognitive frailty in relation to adverse health outcomes independent of multimorbidity: results from the China health and retirement longitudinal study. Aging (Albany NY). 2020;12(22):23129–45.

67. Nicholson K, Griffith LE, Sohel N, Raina P. Examining early and late onset of multimorbidity in the Canadian Longitudinal Study on Aging. J Am Geriatr Soc. 2021;69(6):1579–91.

68. Villacampa-Fernández P, Navarro-Pardo E, Tarín JJ, Cano A. Frailty and multimorbidity: Two related yet different concepts. Maturitas [Internet]. 2017;95:31–5. Available from: http://dx.doi.org/10.1016/j.maturitas.2016.10.008

69. Guaraldi G, Brothers TD, Zona S, Stentarelli C, Carli F, Malagoli A, et al. A frailty index predicts survival and incident multimorbidity independent of markers of HIV disease severity. Aids. 2015;29(13):1633–41.

70. Carrasco-Ribelles LA, Roso-Llorach A, Cabrera-Bean M, Costa-Garrido A, Zabaleta-del-Olmo E, Toran-Monserrat P, et al. Dynamics of multimorbidity and frailty, and their contribution to mortality, nursing home and home care need: A primary care cohort of 1 456 052 ageing people. eClinicalMedicine. 2022;52.

71. Ng SK, Tawiah R, Sawyer M, Scuffham P. Patterns of multimorbid health conditions: A systematic review of analytical methods and comparison analysis. International Journal of Epidemiology. 2018.

72. Prados-Torres A, Poblador-Plou B, Calderón-Larrañaga A, Gimeno-Feliu LA, González-Rubio F, Poncel-Falcó A, et al. Multimorbidity patterns in primary care: Interactions among chronic diseases using factor analysis. PLoS One. 2012;

73. Ng SK. A two-way clustering framework to identify disparities in multimorbidity patterns of mental and physical health conditions among Australians. Stat Med. 2015;34(26):3444–60.

74. Ng SK, Holden L, Sun J. Identifying comorbidity patterns of health conditions via cluster analysis of pairwise concordance statistics. Stat Med. 2012;31(27):3393–405.

75. Sourial N, Wolfson C, Zhu B, Quail J, Fletcher J, Karunananthan S, et al. Correspondence analysis is a useful tool to uncover the relationships among categorical variables. J Clin Epidemiol. 2010;

76. Held FP, Blyth F, Gnjidic D, Hirani V, Naganathan V, Waite LM, et al. Association Rules Analysis of Comorbidity and Multimorbidity: The Concord Health and Aging in Men Project. Journals Gerontol - Ser A Biol Sci Med Sci. 2016;71(5):625–31.

77. Roso-Llorach A, Violán C, Foguet-Boreu Q, Rodriguez-Blanco T, Pons-Vigués M, Pujol-Ribera E, et al. Comparative analysis of methods for identifying multimorbidity patterns: A study of 'real-world' data. BMJ Open. 2018;8(3):1–12.

78. Everitt BS, Landau S, Leese M, Stahl D. Cluster analysis: Fifth edition. Cluster Analysis: Fifth Edition. 2011.

79. Violán C, Foguet-Boreu Q, Fernández-Bertolín S, Guisado-Clavero M, Cabrera-Bean M, Formiga F, et al. Soft clustering using real-world data for the identification of multimorbidity patterns in an elderly population: Cross-sectional study in a Mediterranean population. BMJ Open. 2019;9(8).

80. Zhu Y, Edwards D, Mant J, Payne RA, Kiddle S. Characteristics, service use and mortality of clusters of multimorbid patients in England: A population-based study. BMC Med. 2020;18(1):1–11.

81. Larsen FB, Pedersen MH, Friis K, Gluèmer C, Lasgaard M. A Latent class analysis of multimorbidity and the relationship to socio-demographic factors and health-related quality of life. A national population-based study of 162,283 Danish Adults. PLoS One. 2017;12(1):1–17.

82. Hall M, Dondo TB, Yan AT, Mamas MA, Timmis AD, Deanfield JE, et al. Multimorbidity and survival for patients with acute myocardial infarction in England and Wales: Latent class analysis of a nationwide population-based cohort. PLoS Med. 2018;15(3):1–18.

83. Sinha P, Calfee CS, Delucchi KL. Practitioner's Guide to Latent Class Analysis: Methodological Considerations and Common Pitfalls [Internet]. Vol. 49, Critical Care Medicine. 2020. p. E63–79. Available from: https://journals.lww.com/10.1097/CCM.0000000000004710

84. Collerton J, Jagger C, Yadegarfar ME, Davies K, Parker SG, Robinson L, et al. Deconstructing Complex Multimorbidity in the Very Old: Findings from the Newcastle 85+ Study. Biomed Res Int. 2016;

85. Lacedonia D, Carpagnano GE, Sabato R, Storto MM lo, Palmiotti GA, Capozzi V, et al. Characterization of obstructive sleep apnea–hypopnea syndrome (OSA) population by means of cluster analysis. J Sleep Res. 2016;25(6):724–30.

86. Vanfleteren LEGW, Spruit MA, Groenen M, Gaffron S, Van Empel VPM, Bruijnzeel PLB, et al. Clusters of comorbidities based on validated objective measurements and systemic inflammation in patients with chronic obstructive pulmonary disease. Am J Respir Crit Care Med. 2013;187(7):728–35.

87. St Sauver JL, Boyd CM, Grossardt BR, Bobo W V., Rutten LJF, Roger VL, et al. Risk of developing multimorbidity across all ages in an historical cohort study: Differences by sex and ethnicity. BMJ Open. 2015;5(2).

88. Ruel G, Shi Z, Zhen S, Zuo H, Kröger E, Sirois C, et al. Association between nutrition and the evolution of multimorbidity: The importance of fruits and vegetables and whole grain products. Clin Nutr [Internet]. 2014;33(3):513–20. Available from: http://dx.doi.org/10.1016/j.clnu.2013.07.009

89. Weimann A, Dai D, Oni T. A cross-sectional and spatial analysis of the prevalence of multimorbidity and its association with socioeconomic disadvantage in South Africa: A comparison between 2008 and 2012. Soc Sci Med [Internet]. 2016;163:144–56. Available from: http://dx.doi.org/10.1016/j.socscimed.2016.06.055

90. Melis R, Marengoni A, Angleman S, Fratiglioni L. Incidence and predictors of multimorbidity in the elderly: A population-based longitudinal study. PLoS One. 2014;9(7).

91. Ruel G, Lévesque JF, Stocks N, Sirois C, Kroger E, Adams RJ, et al. Understanding the evolution of multimorbidity: Evidences from the North West Adelaide Health Longitudinal Study (NWAHS). PLoS One. 2014;9(5).

92. Singer L, Green M, Rowe F, Ben-Shlomo Y, Kulu H, Morrissey K. Trends in multimorbidity, complex multimorbidity and multiple functional limitations in the ageing population of England, 2002–2015. J Comorbidity. 2019;9:2235042X1987203.

93. Ashworth M, Durbaba S, Whitney D, Crompton J, Wright M, Dodhia H. Journey to multimorbidity: Longitudinal analysis exploring cardiovascular risk factors and sociodemographic determinants in an urban setting. BMJ Open. 2019;9(12):1–8.

94. Canizares M, Hogg-Johnson S, Gignac MAM, Glazier RH, Badley EM. Increasing trajectories of multimorbidity over time: Birth cohort differences and the role of changes in obesity and income. Journals Gerontol - Ser B Psychol Sci Soc Sci. 2018;73(7):1303–14.

95. Fraccaro P, Kontopantelis E, Sperrin M, Peek N, Mallen C, Urban P, et al. Predicting mortality from change-over-time in the Charlson Comorbidity Index. Medicine (Baltimore). 2016;95(43):e4973.

96. Calderón-Larrañaga A, Santoni G, Wang HX, Welmer AK, Rizzuto D, Vetrano DL, et al. Rapidly developing multimorbidity and disability in older adults: does social background matter? J Intern Med. 2018;

97. Pérez LM, Hooshmand B, Mangialasche F, Mecocci P, Smith AD, Refsum H, et al. Glutathione serum levels and rate of multimorbidity development in older adults. Journals Gerontol - Ser A Biol Sci Med Sci. 2020;75(6):1089–94.

98. Quiñones AR, Liang J, Bennett JM, Xu X, Ye W. How does the trajectory of multimorbidity vary across black, white, and mexican americans in middle and old age? Journals Gerontol - Ser B Psychol Sci Soc Sci. 2011;66 B(6):739–49.

99. Quiñones AR, Botoseneanu A, Markwardt S, Nagel CL, Newsom JT, Dorr DA, et al. Racial/ethnic differences in multimorbidity development and chronic disease accumulation for middle-aged adults. PLoS One. 2019;14(6):1–13.

100. Mondor L, Cohen D, Khan AI, Wodchis WP. Income inequalities in multimorbidity prevalence in Ontario, Canada: A decomposition analysis of linked survey and health administrative data. Int J Equity Health. 2018;17(1):1–13.

101. Yang S, Hoshi T, Wang S, Nakayama N, Kong F. Socioeconomic status, comorbidity, activity limitation, and healthy life expectancy in older men and women: A 6-year follow-up study in Japan. J Appl Gerontol. 2014;33(7):831–47.

102. Kingston A, Robinson L, Booth H, Knapp M, Jagger C, Adelaja B, et al. Projections of multi-morbidity in the older population in England to 2035: Estimates from the Population Ageing and Care Simulation (PACSim) model. Age Ageing. 2018;47(3):374–80.

103. Jackson CA, Dobson A, Tooth L, Mishra GD. Body mass index and socioeconomic position are associated with 9-year trajectories of multimorbidity: A population-based study. Prev Med (Baltim) [Internet]. 2015;81:92–8. Available from: http://dx.doi.org/10.1016/j.ypmed.2015.08.013

104. Goodman RA, Ling SM, Briss PA, Parrish RG, Salive ME, Finke BS. Multimorbidity Patterns in the United States: Implications for Research and Clinical Practice. Journals Gerontol - Ser A Biol Sci Med Sci. 2016;71(2):215–20.

105. Petri H, Maldonato D, Robinson NJ. Data-driven identification of co-morbidities associated with rheumatoid arthritis in a large US health plan claims database. BMC Musculoskelet Disord. 2010;11.

106. Ibarra-Castillo C, Guisado-Clavero M, Violan-Fors C, Pons-Vigués M, López-Jiménez T, Roso-Llorach A. Survival in relation to multimorbidity patterns in older adults in primary care in Barcelona, Spain (2010-2014): A longitudinal study based on electronic health records. J Epidemiol Community Health. 2018;72(3).

107. Cezard G, McHale CT, Sullivan F, Bowles JKF, Keenan K. Studying trajectories of multimorbidity: A systematic scoping review of longitudinal approaches and evidence. BMJ Open. 2021;11(11):1–19.

108. Strauss VY, Jones PW, Kadam UT, Jordan KP. Distinct trajectories of multimorbidity in primary care were identified using latent class growth analysis. J Clin Epidemiol [Internet]. 2014;67(10):1163–71. Available from: http://dx.doi.org/10.1016/j.jclinepi.2014.06.003

109. Pugh MJ, Finley EP, Wang CP, Copeland LA, Jaramillo CA, Swan AA, et al. A retrospective

cohort study of comorbidity trajectories associated with traumatic brain injury in veterans of the Iraq and Afghanistan wars. Brain Inj [Internet]. 2016;30(12):1481–90. Available from: http://dx.doi.org/10.1080/02699052.2016.1219055

110. Lappenschaar M, Hommersom A, Lucas PJF, Lagro J, Visscher S, Korevaar JC, et al. Multilevel temporal Bayesian networks can model longitudinal change in multimorbidity. J Clin Epidemiol. 2013;66(12):1405–16.

111. Jensen AB, Moseley PL, Oprea TI, Ellesøe SG, Eriksson R, Schmock H, et al. Temporal disease trajectories condensed from population-wide registry data covering 6.2 million patients. Nat Commun. 2014;5(May):1–10.

112. Faruqui SHA, Alaeddini A, Jaramillo CA, Potter JS, Pugh MJ. Mining patterns of comorbidity evolution in patients with multiple chronic conditions using unsupervised multi-level temporal Bayesian network. PLoS One. 2018;13(7):1–22.

113. Giannoula A, Gutierrez-Sacristán A, Bravo Á, Sanz F, Furlong LI. Identifying temporal patterns in patient disease trajectories using dynamic time warping: A population-based study. Sci Rep. 2018;8(1):1–14.

114. Giannoula A, Centeno E, Mayer MA, Sanz F, Furlong LI. A system-level analysis of patient disease trajectories based on clinical, phenotypic and molecular similarities. Bioinformatics. 2021;37(10):1435–43.

115. Xu X, Mishra GD, Dobson AJ, Jones M. Progression of diabetes, heart disease, and stroke multimorbidity in middle-aged women: A 20-year cohort study. PLoS Med. 2018;15(3):1–18.

116. Meira-Machado LF, de Uña-Álvarez J, Cadarso-Suárez C, Andersen PK. Multi-state models for the analysis of time-to-event data. Vol. 18, Statistical Methods in Medical Research. 2009. 195–222 p.

117. Jackson CH, Sharples LD, Thompson SG, Duffy SW, Couto E. Multistate Markov models for disease progression with classification error. J R Stat Soc Ser D (The Stat [Internet]. 2003 Jul;52(2):193–209. Available from: https://onlinelibrary.wiley.com/doi/10.1111/1467-9884.00351

118. Hougaard P. Multi-state Models : A Review. 1999;264:239–64.

119. Siriwardhana C, Lim E, Davis J, Chen JJ. Progression of diabetes, ischemic heart disease, and chronic kidney disease in a three chronic conditions multistate model. BMC Public Health. 2018;18(1):1–14.

120. Freisling H, Viallon V, Lennon H, Bagnardi V, Ricci C, Butterworth AS, et al. Lifestyle factors and risk of multimorbidity of cancer and cardiometabolic diseases: A multinational cohort study. BMC Med. 2020;18(1):1–11.

121. Alaeddini A, Jaramillo CA, Faruqui SHA, Pugh MJ. Mining major transitions of chronic conditions in patients with multiple chronic conditions. Methods Inf Med. 2017;56(5):391–400.

122. Violán C, Fernández-Bertolín S, Guisado-Clavero M, Foguet-Boreu Q, Valderas JM, Vidal Manzano J, et al. Five-year trajectories of multimorbidity patterns in an elderly Mediterranean population using Hidden Markov Models. Sci Rep. 2020;10(1).

123. Rabiner LR. A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition. Proc IEEE. 1989;

124. Barber D. Bayesian Reasoning and Machine Learning. Bayesian Reasoning and Machine Learning. 2011.

125. Marx P, Antal P. Decomposition of Shared Latent Factors Using Bayesian Multi-morbidity Dependency Maps. In: IFMBE Proceedings [Internet]. 2015. p. 40–3. Available from: http://link.springer.com/10.1007/978-981-287-573-0_10

126. Mesgarpour M, Chaussalet T, Chahed S. A review of Dynamic Bayesian Network techniques with applications in healthcare risk modelling. OpenAccess Ser Informatics. 2014;37:89–100.

127. Villén N, Guisado-Clavero M, Guisado-Clavero M, Fernández-Bertolín S, Fernández-Bertolín S, Troncoso-Mariño A, et al. Multimorbidity patterns, polypharmacy and their association with liver and kidney abnormalities in people over 65 years of age: A longitudinal study. BMC Geriatr. 2020;20(1):1–13.

128. Lagergren M, Fratiglioni L, Hallberg IR, Berglund J, Elmståhl S, Hagberg B, et al. A longitudinal study integrating population, care and social services data. The Swedish National study on Aging and Care (SNAC). Aging Clin Exp Res. 2004;

129. Fratiglioni L, Ding M, Santoni G, Berglund J, Elmståhl S, Fagerström C, et al. Demensförekomst i Sverige: geografiska och tidsmässiga trender 2001-2013. 2017; Available from: www.snac.org

130. Rydwik E, Welmer AK, Kåreholt I, Angleman S, Fratiglioni L, Wang HX. Adherence to physical exercise recommendations in people over 65 - The SNAC-Kungsholmen study. Eur J Public Health. 2013;23(5):799–804.

131. Nelson ME, Rejeski WJ, Blair SN, Duncan PW, Judge JO, King AC, et al. Physical activity and public health in older adults: Recommendation from the American College of Sports Medicine and the American Heart Association. Medicine and Science in Sports and Exercise. 2007.

132. NEUGARTEN BL, HAVIGHURST RJ, TOBIN SS. The measurement of life satisfaction. J Gerontol. 1961;16:134–43.

133. Okun MA, Stock WA. The construct validity of subjective well-being measures: An assessment via quantitative research syntheses. J Community Psychol. 1987;

134. Cornwell EY, Waite LJ. Measuring social isolation among older adults using multiple indicators from the nshap study. Journals Gerontol - Ser B Psychol Sci Soc Sci. 2009;

135. Karlis D, Saporta G, Spinakis A. A simple rule for the selection of principal components. Commun Stat - Theory Methods. 2003;32(3):643–66.

136. Chavent M, Kuentz-Simonet V, Labenne A, Saracco J. Multivariate Analysis of Mixed Data: The R Package PCAmixdata. 2014;4. Available from: http://arxiv.org/abs/1411.4911

137. Bezdek JC. Pattern Recognition with Fuzzy Objective Function Algorithms. Pattern Recognit with Fuzzy Object Funct Algorithms. 1981;

138. Zhao Q. Cluster Validity in Clustering Methods. 2012. 189 p.

139. Violán C, Roso-Llorach A, Foguet-Boreu Q, Guisado-Clavero M, Pons-Vigués M, Pujol-Ribera E, et al. Multimorbidity patterns with K-means nonhierarchical cluster analysis. BMC Fam Pract. 2018;

140. Theodoridis S. Machine learning A Bayesian. Vol. 53, Journal of Chemical Information and Modeling. 2019. 1689–1699 p.

141. Bishop CM. Pattern Recognition and Machine Learning. In: Jordan M, Kleinberg J, Scho¨lkopf B, editors. Information Science and Statistics [Internet]. Springer Science+Business Media, LLC; 2006. Available from: http://users.isr.ist.utl.pt/~wurmd/Livros/school/Bishop - Pattern Recognition And Machine Learning - Springer 2006.pdf

142. Lebedeva KA, Allen J, Kulhawy EY, Caruncho HJ, Kalynchuk LE. Cyclical administration of corticosterone results in aggravation of depression-like behaviors and accompanying downregulations in reelin in an animal model of chronic stress relevant to human recurrent depression. Physiol Behav [Internet]. 2020;224(May):113070. Available from: https://doi.org/10.1016/j.physbeh.2020.113070

143. de Roos EW, Lahousse L, Verhamme KMC, Braunstahl GJ, Ikram MA, in 't Veen JCCM, et al. Asthma and its comorbidities in middle-aged and older adults; the Rotterdam Study. Respir Med. 2018;

144. Carton L, Pignon B, Baguet A, Benradia I, Roelandt JL, Vaiva G, et al. Influence of comorbid alcohol use disorders on the clinical patterns of major depressive disorder: A general population-based study. Drug Alcohol Depend. 2018;

145. Boyer L, Baumstarck K, Boucekine M, Blanc J, Lançon C, Auquier P. Measuring quality of life in patients with schizophrenia: An overview. Expert Review of Pharmacoeconomics and Outcomes Research. 2013.

146. Oga T, Windisch W, Handa T, Hirai T, Chin K. Health-related quality of life measurement in patients with chronic respiratory failure. Respir Investig [Internet]. 2018;56(3):214–21. Available from: https://linkinghub.elsevier.com/retrieve/pii/S221253451830011X

147. Haeusler KG, Gröschel K, Köhrmann M, Anker SD, Brachmann J, Böhm M, et al. Expert opinion paper on atrial fibrillation detection after ischemic stroke. Clinical Research in Cardiology. 2018.

148. Mitsios JP, Ekinci EI, Mitsios GP, Churilov L, Thijs V. Relationship between glycated hemoglobin and stroke risk: A systematic review and meta-analysis. J Am Heart Assoc. 2018;

149. Bektas F, Soyuncu S. Nitroglycerin-induced migraine type headache: Bilaterally visible temporal arteries. Int J Emerg Med. 2010;

150. Mason BN, Russo AF. Vascular Contributions to Migraine: Time to Revisit? Front Cell Neurosci [Internet]. 2018 Aug 3;12. Available from: https://www.frontiersin.org/article/10.3389/fncel.2018.00233/full

151. Ferrucci L, Fabbri E. Inflammageing: chronic inflammation in ageing, cardiovascular disease, and frailty. Nature Reviews Cardiology. 2018.

152. Smykiewicz P, Segiet A, Keag M, Żera T. Proinflammatory cytokines and ageing of the cardiovascular-renal system. Mech Ageing Dev [Internet]. 2018 Oct;175:35–45. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0047637418300605

153. Kanis JA, Stevenson M, McCloskey E V., Davis S, Lloyd-Jones M. Glucocorticoid-induced osteoporosis: A systematic review and cost-utility analysis. Health Technology Assessment. 2007.

154. Zografos GN, Georgiadou D, Thomas D, Kaltsas G, Digalakis M. Drug-induced esophagitis. Diseases of the Esophagus. 2009.

155. Jolliffe DA, Greenberg L, Hooper RL, Mathyssen C, Rafiq R, De Jongh RT, et al. Vitamin D to prevent exacerbations of COPD: Systematic review and meta-analysis of individual participant data from randomised controlled trials. Thorax. 2019;74(4):337–45.

156. Shahnazari B, Moghimi J, Foroutan M, Mirmohammadkhani M, Ghorbani A. Comparison of the effect of Vitamin D on osteoporosis and osteoporotic patients with healthy individuals referred to the Bone Density Measurement Center. Biomol Concepts. 2019;10(1):44–50.

157. Luo W, Mao P, Zhang L, Yang Z. Association between systemic lupus erythematosus and thyroid dysfunction: a meta-analysis. Lupus. 2018.

158. Dodds RM, Syddall HE, Cooper R, Kuh D, Cooper C, Avan Aihie Sayer. Global variation in grip strength: A systematic review and meta-analysis of normative data. Age Ageing. 2016;

159. Shen Y, Ye B, Chen P, Wang Q, Fan C, Shu Y, et al. Cognitive Decline, Dementia, Alzheimer’s Disease and Presbycusis: Examination of the Possible Molecular Mechanism. Front Neurosci [Internet]. 2018;12:394. Available from: http://www.ncbi.nlm.nih.gov/pubmed/29937713

160. Livingston G, Huntley J, Sommerlad A, Ames D, Ballard C, Banerjee S, et al. Dementia prevention, intervention, and care: 2020 report of the Lancet Commission. Lancet. 2020;396(10248):413–46.

161. Naël V, Pérès K, Dartigues JF, Letenneur L, Amieva H, Arleo A, et al. Vision loss and 12-year risk of dementia in older adults: the 3C cohort study. Eur J Epidemiol. 2019;34(2):141–52.

162. Brenowitz WD, Kaup AR, Lin FR, Yaffe K. Multiple sensory impairment is associated with increased risk of dementia among black and white older adults. Journals Gerontol - Ser A Biol Sci Med Sci. 2019;74(6):890–6.

163. Deal JA, Sharrett AR, Albert M, Bandeen-Roche K, Burgard S, Thomas SD, et al. Retinal signs and risk of incident dementia in the Atherosclerosis Risk in Communities study. Alzheimer’s Dement [Internet]. 2019;15(3):477–86. Available from: https://doi.org/10.1016/j.jalz.2018.10.002

164. Toot S, Swinson T, Devine M, Challis D, Orrell M. Causes of nursing home placement for older people with dementia: A systematic review and meta-analysis. International Psychogeriatrics. 2017.

165. Mograbi DC, Morris RG, Fichman HC, Faria CA, Sanchez MA, Ribeiro PCC, et al. The impact of dementia, depression and awareness on activities of daily living in a sample from a middle-income country. In: International Journal of Geriatric Psychiatry. 2018.

166. Marengoni A, Fratiglioni L, Bandinelli S, Ferrucci L. Socioeconomic status during lifetime and cognitive impairment no-dementia in late life: The population-based aging in the Chianti area (InCHIANTI) study. J Alzheimer’s Dis. 2011;

167. Nobili A, Marengoni A, Tettamanti M, Salerno F, Pasina L, Franchi C, et al. Association between clusters of diseases and polypharmacy in hospitalized elderly patients: Results from the REPOSI study. Eur J Intern Med. 2011;

168. Baré M, Herranz S, Roso-Llorach A, Jordana R, Violán C, Lleal M, et al. Multimorbidity patterns of chronic conditions and geriatric syndromes in older patients from the MoPIM multicentre cohort study. BMJ Open. 2021;11(11):1–9.

169. Rizzuto D, Feldman AL, Karlsson IK, Dahl Aslan AK, Gatz M, Pedersen NL. Detection of dementia cases in two swedish health registers: A validation study. J Alzheimer's Dis. 2018;61(4):1301–10.

170. Vos R, van den Akker M, Boesten J, Robertson C, Metsemakers J. Trajectories of multimorbidity: exploring patterns of multimorbidity in patients with more than ten chronic health problems in life course. BMC Fam Pract. 2015 Dec;16(1):2.

171. Gao L, Leung MTY, Li X, Chui CSL, Wong RSM, Au Yeung SL, et al. Linking cohort-based data with electronic health records: A proof-of-concept methodological study in Hong Kong. BMJ Open. 2021;11(6):1–7.

172. Harron K, Doidge JC, Goldstein H. Assessing data linkage quality in cohort studies. Ann Hum Biol [Internet]. 2020;47(2):218–26. Available from: https://doi.org/10.1080/03014460.2020.1742379

173. Amell A, Roso-Llorach A, Palomero L, Cuadras D, Galván-Femenía I, Serra-Musach J, et al. Disease networks identify specific conditions and pleiotropy influencing multimorbidity in the general population. Sci Rep. 2018;8(1):1–16.

174. Van Der Ploeg T, Austin PC, Steyerberg EW. Modern modelling techniques are data hungry: A simulation study for predicting dichotomous endpoints. BMC Med Res Methodol. 2014;14(1):1–13.

175. Obón-Santacana M, Vilardell M, Carreras A, Duran X, Velasco J, Galván-Femenía I, et al. GCAT|Genomes for life: A prospective cohort study of the genomes of Catalonia. BMJ Open. 2018;8(3).

176. Silva JF, Matos S. Patient trajectory modelling in longitudinal data: A review on existing solutions. Proc - IEEE Symp Comput Med Syst. 2021;2021-June:480–5.

10. Appendix

10.1 Supplementary files Study 1

Supplementary material

Supplementary Table 1. Chronic diseases characterizing clusters of older people identified at baseline in the Swedish National Study on Aging and Care in Kungsholmen (SNAC-K) (N = 2931).

Psychiatric and respiratory diseases	**Prev (%)**	**O/E ratio**	**Exc (%)**	**Eye diseases and cancer**	**Prev (%)**	**O/E ratio**	**Exc (%)**	**Respiratory and musculoskeletal diseases**	**Prev (%)**	**O/E ratio**	**Exc (%)**
Neurotic, stress-related and somatoform diseases	**62.16**	**17.52**	**95.03**	Other eye diseases	39.67	**6.96**	**74.28**	Asthma	27.62	**4.03**	**63.22**
Depression and mood diseases	**80.32**	**7.69**	**41.73**	Cataract and other lens diseases	43.41	**6.95**	**74.18**	Osteoporosis	26.93	**3.51**	**55.06**
Sleep disorders	7.39	**3.09**	16.77	Glaucoma	30.35	**4.76**	**50.74**	COPD, emphysema, chronic bronchitis	18.24	**3.22**	**50.56**
Other neurological diseases	6.52	**2.99**	16.21	Blindness, visual impairment	18.67	**3.80**	**40.53**	Esophagus, stomach and duodenum diseases	14.46	**2.92**	**45.88**
Asthma	15.03	**2.19**	11.89	Solid neoplasms	23.99	**2.38**	**25.34**	Inflammatory arthropathies	11.74	**2.57**	**40.31**
Other psychiatric and behavioral diseases	5.46	**2.16**	11.73	Other genitourinary diseases	6.09	**2.13**	22.68	Sleep disorders	5.73	**2.40**	**37.66**
Colitis and related diseases	28.39	1.97	10.67	Deafness, hearing impairment	25.78	1.97	21.05	Dorsopathies	17.49	**2.39**	**37.59**
Migraine and facial pain syndromes	5.00	1.93	10.46	Chronic kidney diseases	**68.69**	1.82	19.37	Autoimmune diseases	11.58	**2.26**	**35.52**
Other musculoskeletal and joint diseases	12.60	1.69	9.19	Anemia	22.23	1.62	17.33	Osteoarthritis and other degenerative joint diseases	29.61	**2.08**	**32.66**
COPD, emphysema, chronic bronchitis	9.46	1.67	9.06	Ischemic heart disease	22.41	1.28	13.66	Thyroid diseases	21.98	1.86	**29.14**
Heart diseases	**Prev (%)**	**O/E ratio**	**Exc (%)**	**Cognitive and sensory impairment**	**Prev (%)**	**O/E ratio**	**Exc (%)**	**Unspecific**	**Prev (%)**	**O/E ratio**	**Exc (%)**
Bradycardias and conduction diseases	19.93	**9.42**	**87.31**	Dementia	**74.90**	**6.90**	**72.84**	Dyslipidemia	**64.05**	1.28	**61.85**
Other cardiovascular diseases	29.72	**7.57**	**70.18**	Other psychiatric and	14.72	**5.83**	**61.50**	Obesity	15.44	1.18	**57.18**

				behavioral diseases							
Cardiac valve diseases	18.20	**6.43**	**59.56**	Blindness, visual impairment	16.29	**3.32**	**34.99**	Hypertension	**84.28**	1.15	**55.66**
Heart failure	**67.22**	**5.58**	**51.71**	Colitis and related diseases	45.56	**3.16**	**33.31**	Prostate diseases	5.14	1.14	**55.20**
Atrial fibrillation	**52.92**	**4.80**	**44.49**	Deafness, hearing impairment	41.10	**3.14**	**33.18**	Other genitourinary diseases	2.48	0.86	**41.82**
Ischemic heart disease	**60.44**	**3.45**	**32.00**	Cerebrovascular disease	26.75	**2.99**	**31.58**	Solid neoplasms	8.46	0.84	**40.53**
Diabetes	30.62	**3.07**	**28.48**	Anemia	30.81	**2.25**	23.76	Diabetes	8.23	0.83	**39.98**
Inflammatory arthropathies	13.64	**2.98**	**27.64**	Heart failure	27.09	**2.25**	23.73	Other musculoskeletal and joint diseases	5.63	0.76	**36.60**
Cerebrovascular disease	23.15	**2.59**	24.00	Atrial fibrillation	20.75	1.88	19.87	Other neurological diseases	1.65	0.76	**36.58**
Migraine and facial pain syndromes	5.65	**2.18**	20.20	Other neurological diseases	3.94	1.80	19.04	Osteoarthritis and other degenerative joint diseases	10.55	0.74	**35.89**

Note: Prev: Prevalence within cluster; O/E ratio: Observed/Expected ratio; Exc: Exclusivity. Diseases with O/E ratio ≥2 OR exclusivity ≥25% OR prevalence ≥50% are boldfaced.

10.2 Supplementary files Study 2

Supplementary tables and figures

Twelve-year clinical trajectories of multimorbidity in a population of older adults

Davide L Vetrano et al.

Nature Communications 2020

7 Supplementary tables

15 Supplementary figures

Supplementary table 1. Sample characteristics (weighted means and proportions) by clusters of multimorbidity over time.

Abbreviations: MMSE = Mini Mental State Examination. Weighted means and proportions have been obtained by the membership matrix.

	Baseline N = 2931						6 years N = 1716						12 years N = 1016					
	Psych. and respiratory diseases	Heart diseases	Eye diseases and cancer	Cognitive and sensory impairment	Respiratory and MSK diseases	Unspecific group	Heart and vascular diseases	Heart diseases and cognitive impairment	Neuropsychiatric and respiratory diseases	Eye diseases	MSK, respiratory, and immune diseases	Unspecific group	Vascular diseases	Cardiometabolic diseases	Respiratory diseases	Neuropsychiatric diseases	Eye and MSK diseases	Unspecific group
Number (n)	159	272	313	309	460	1418	101	170	201	235	270	740	63	112	120	122	200	400
Age (mean)	73.3 (73.0-73.7)	82.3 (81.9-82.6)	83.2 (82.9-83.5)	88.2 (87.9-88.5)	73.9 (73.5-74.3)	71.8 (71.5-72.1)	82.0 (81.6-82.4)	85.5 (85.1-85.9)	80.8 (80.3-81.3)	84.6 (84.2-85)	75.5 (75.1-75.9)	74.0 (73.3-74.4)	84.0 (83.5-84.5)	84.0 (83.6-84.5)	79.8 (79.4-80.3)	87.4 (86.9-87.8)	81.8 (81.4-82.3)	76.8 (76.4-77.1)
Sex (female, %)	74.2 (66.9-80.3)	59.2 (53.3-64.9)	72.5 (37.3-77.2)	76.9 (71.9-81.2)	74.4 (70.2-78.1)	61.0 (58.5-63.5)	51.0 (41.4-60.6)	57.3 (49.8-64.5)	72.9 (66.4-78.6)	76.6 (70.8-81.6)	72.4 (66.8-77.4)	59.3 (55.7-62.8)	73.3 (61.3-81.7)	44.2 (35.4-53.5)	77.5 (69.2-84.0)	70.6 (62.0-77.9)	77.6 (71.3-82.8)	58.7 (53.8-63.4)
Education																		
Elementary (%)	16.9 (11.9-23.5)	25.1 (20.3-30.6)	21.8 (17.5-26.7)	34.8 (29.6-40.5)	19.3 (16.0-23.2)	14.4 (12.7-16.3)	19.1 (12.6-27.8)	19.6 (14.3-26.3)	15.5 (11.2-21.2)	20.5 (15.8-26.1)	8.6 (5.8-12.5)	10.4 (8.4-12.8)	21.6 (13.2-33.2)	11.7 (6.9-18.9)	8.9 (5.0-15.3)	20.0 (13.9-27.9)	8.4 (5.3-13.1)	5.4 (3.6-8.0)
Secondary (%)	47.6 (33.9-55.3)	56.1 (50.1-61.9)	56.5 (50.9-61.9)	47.4 (41.8-53.2)	49.5 (45.0-54.1)	50.1 (47.5-52.7)	51.4 (41.8-61.0)	50.9 (43.4-58.3)	50.3 (43.5-57.2)	53.2 (46.8-59.5)	49.4 (43.5-55.3)	48.0 (44.4-51.6)	52.0 (40.0-63.9)	47.2 (38.2-56.4)	40.9 (32.5-49.8)	52.7 (43.9-61.3)	49.0 (42.2-55.9)	46.4 (41.5-51.3)
University (%)	35.5 (28.5-43.2)	18.8 (14.6-	21.8 (27.6-26.7)	17.7 (13.8-22.5)	31.1 (27.1-	35.5 (33.0-38.0)	29.5 (21.5-	29.5 (23.1-	34.1 (27.9-41.0)	26.3 (21.1-	42.0 (36.3-	41.6 (38.1-45.2)	26.4 (17.1-38.4)	41.1 (32.4-50.4)	50.2 (41.4-59.0)	27.3 (20.2-35.8)	42.6 (35.9-49.5)	48.3 (43.4-

		23.9)			35.5)		39.0)	36.8)		32.3)	48.0)							53.2)
Number of diseases (mean)	5.7 (5.6-5.8)	7.7 (7.6-7.8)	6.0 (6.0-6.1)	5.5 (5.5-5.6)	4.7 (4.6-4.7)	3.2 (3.2-3.3)	11.5 (11.4-11.7)	10.9 (10.7-11.0)	8.9 (8.8-9.1)	9.3 (9.1-9.4)	7.8 (7.7-7.9)	4.8 (4.7-4.9)	14.9 (14.6-15.1)	13.3 (13.1-13.5)	12.2 (12.0-12.5)	15.0 (14.7-15.2)	10.7 (10.6-10.9)	6.6 (5.5-6.7)
Number of drugs (mean)	6.2 (6.1-6.4)	7.7 (7.6-7.8)	5.1 (4.9-5.2)	6.1 (5.9-6.2)	5.3 (5.2-5.4)	2.9 (2.8-2.9)	8.3 (8.1-8.5)	8.8 (8.6-9.0)	7.8 (7.6-8.0)	6.6 (6.4-6.8)	6.3 (6.2-6.5)	3.8 (3.7-3.9)	9.6 (9.3-9.9)	8.7 (8.4-9.0)	8.8 (8.5-9.0)	9.4 (9.2-9.7)	6.4 (6.2-6.7)	4.5 (4.3-4.6)
Walking speed <0.8 m/s (%)	35.7 (28.5-43.5)	63.1 (57.1-68.7)	51.0 (45.3-56.5)	83.8 (12.1-21.3)	31.2 (27.1-35.6)	16.0 (14.2-18.0)	51.2 (51.5-60.8)	67.7 (60.2-74.3)	54.4 (47.5-61.2)	59.5 (53.1-65.7)	30.3 (25.2-36.1)	15.5 (13.1-18.3)	63.9 (51.4-74.7)	55.9 (46.6-64.7)	39.3 (31.1-48.3)	79.6 (71.6-85.8)	33.9 (27.7-40.8)	15.0 (11.9-18.9)
Disability score ≥1 (%)	32.5 (17.9-31.6)	49.9 (30.1-41.8)	38.4 (24.0-34.2)	86.1 (77.9-86.6)	18.4 (9.0-15.0)	8.8 (3.5-5.7)	23.0 (7.3-20.4)	24.3 (11.5-22.9)	22.7 (11.3-21.8)	17.3 (7.1-15.1)	7.3 (1.4-5.7)	2.8 (0.5-2.0)	9.7 (2.4-15.3)	5.7 (0.3-5.8)	4.7 (0.3-5.4)	11.1 (2.4-10.7)	2.2 (0.0-2.1)	1.1 (0.0-1.0)
MMSE (mean)	27.7 (27.5-27.9)	27.6 (27.5-27.7)	27.7 (27.6-27.8)	16.6 (16.3-17.0)	28.6 (28.5-28.7)	28.8 (28.7-28.8)	26.7 (26.5-26.9)	24.3 (24.0-24.7)	24.7 (24.4-25.1)	26.4 (26.1-26.6)	28.2 (28.1-28.3)	28.2 (28.1-28.3)	26.5 (26.2-26.8)	26.3 (26.0-26.7)	27.2 (27.0-27.5)	22.1 (21.6-22.5)	27.9 (27.7-28.1)	28.1 (27.9-28.3)
Hypertension (%)	65.4 (57.7-72.3)	73.4 (67.9-78.3)	80.6 (75.8-84.6)	42.4 (37.0-48.0)	57.8 (53.3-62.3)	84.3 (82.3-86.1)	93.7 (87.2-97.0)	91.1 (85.8-94.5)	75.2 (68.8-80.6)	90.3 (85.9-93.5)	80.2 (75.-84.5)	90.2 (87.8-92.1)	94.1 (85.4-97.7)	96.3 (81.0-98.6)	81.5 (73.6-87.5)	96.6 (91.8-98.7)	89.7 (84.7-93.2)	89.7 (86.3-92.3)
Diabetes (%)	6.4 (3.5-11.3)	30.6 (25.4-36.3)	11.5 (8.4-15.5)	4.9 (3.0-7.9)	6.7 (4.8-9.4)	8.2 (6.9-9.8)	31.1 (22.9-40.7)	24.6 (18.7-31.6)	9.4 (6.1-14.3)	12.3 (8.7-17.1)	9.6 (6.6-13.7)	13.4 (11.1-16.0)	24.6 (15.6-36.4)	46.7 (37.7-55.9)	12.7 (7.8-19.8)	20.2 (14.0-28.1)	10.1 (6.6-15.1)	12.1 (9.3-15.7)
Dyslipidemia (%)	48.1 (40.5-55.9)	40.9 (35.2-46.8)	44.8 (39.4-50.3)	12.9 (9.6-17.1)	42.0 (37.5-46.5)	64.1 (61.5-66.5)	59.7 (50.0-68.8)	59.4 (51.8-66.5)	58.7 (51.8-65.3)	56.8 (50.4-62.9)	57.9 (52.0-63.7)	77.8 (74.7-80.6)	69.9 (57.7-79.8)	76.8 (68.1-83.6)	66.8 (58.0-74.6)	60.9 (52.1-69.1)	78.0 (71.8-83.2)	84.8 (81.0-88.0)

Obesity (%)	8.3 (5.0-13.7)	16.0 (12.2-20.9)	7.3 (4.9-10.8)	0.7 (0.2-2.4)	17.8 (14.6-21.6)	15.4 (13.7-17.4)	16.6 (10.6-25.1)	16.4 (11.6-22.8)	13.4 (9.4-18.8)	13.3 (9.6-18.3)	24.2 (19.5-29.7)	24.7 (21.7-27.9)	23.6 (14.8-35.4)	29.7 (22.0-38.7)	26.1 (19.1-34.6)	26.4 (19.4-34.9)	18.9 (14.1-24.9)	27.2 (23.0-31.7)

Supplementary table 2. Additional clinical and drug-related* parameters used in SNAC-K for specific chronic conditions.

*The ATC codes corresponding to each drug are shown in brackets. Only those drugs that can be unequivocally linked to chronic conditions were considered. That is, drugs with more than one indication were excluded from the list. The selection of ATC codes was based on a literature review and the clinical judgement of physicians.

NOTE: The criteria presented in this table were used in addition to the diagnoses assigned in SNAC-K. For example, use of dopaminergic agents was considered to indicate presence of Parkinson syndrome, even in the absence of other diagnostic information.

Condition	Clinical and drug-related parameters
Anemia	Hemoglobin <13 g/dl in men and <12 g/dl in women Use of iron preparations (B03A) or other antianemic preparations (B03XA)
Asthma	Use of leukotriene receptor antagonists (R03DC) or antiallergic agents, excl. corticosteroids (R03BC)
Atrial fibrillation	Discrete P wave undetectable and irregular ventricular rate (12-lead electrocardiogram)
Autoimmune diseases	Use of antipsoriatics (D05)
Bradycardias and conduction diseases	Presence of a cardiac pacemaker (12-lead electrocardiogram)
Chronic infectious diseases	Use of drugs for treatment of tuberculosis excluding cycloserine, rifampicine, rifamicyne and hydrazides (J04A excl. J04AB01, J04AB02, J04AB03 and J04AC)
Chronic kidney diseases	Glomerular filtration rate <60 ml/min/1.73m2 (assessed using the CKD-EPI equation)
Chronic pancreas, biliary tract and gallbladder diseases	Use of multienzymes (lipase, protease etc.) (A09AA02)
Colitis and related diseases	Use of drugs for constipation (A06A)
COPD, Emphysema, Chronic Bronchitis	Use of anticholinergics (R03BB)
Dementia	Diagnostic and Statistical Manual of Mental Disorders, Third Edition, Revised (assessed by two different physicians, and a third one in case of disagreement)

	Use of anticholinesterases (N06DA) or memantine (N06DX01)
Diabetes	Glycated hemoglobin (A1C) ≥6.5% Use of antidiabetics (A10)
Esophagus, stomach and duodenum diseases	Use of other drugs for peptic ulcer and gastro-oesophageal reflux disease (A02BX)
Glaucoma	Use of beta blocking agents (S01ED)
Hearing impairment	Unable to hear the interviewer's voice at a normal volume (assessed by a nurse)
Hypercholesterolemia	Serum total cholesterol ≥6.22 mmol/L
Hypertension	Blood pressure ≥140/90 mmHg
Inflammatory arthropathies	Use of gold preparations (M01CB)
Inflammatory bowel diseases	Use of intestinal antiinflammatory agents (A07E)
Ischemic heart disease	Use of organic nitrates (C01DA) or ranolazine (C01EB18)
Migraine and facial pain syndromes	Use of antimigraine preparations (N02C)
Obesity	Body Mass Index ≥30 kg/m^2
Osteoporosis	Use of bisphosphonates (M05BA), bisphosphonate combinations (M05BB), strontium ranelate (M05BX03) or strontium ranelate and colecalciferol (M05BX53)
Other psychiatric and behavioral diseases	Use of drugs for alcohol dependence (N07BB)
Parkinson and parkinsonism	Use of dopa and dopa derivatives (N04BA), dopamine agonists (N04BC), or other dopaminergic agents (N04BX)
Peripheral vascular disease	Use of cilostazol (B01AC23)
Prostate diseases	Use of drugs for benign prostatic hypertrophy excluding testosterone-5-alpha reductase inhibitors (G04C excl. G04CB)
Thyroid diseases	Use of thyroid hormones (H03AA) or antithyroid preparations (H03B)
Visual impairment	Unable to see the physician at a close distance with or without aid (assessed by a nurse)

1. **Supplementary table 3. Descriptors of ICD-10 codes included and excluded in each chronic disease category.**

NOTE: When all sub-codes within a given ICD-10 code were classified as chronic, the highest possible level of aggregation of the hierarchy was included in the list (e.g. three-digit code for asthma (J45), one-digit code for malignant neoplasms (C), etc.

ALLERGY	
Included ICD-10 codes and labels	
J301	Allergic rhinitis due to pollen
J302	Other seasonal allergic rhinitis
J303	Other allergic rhinitis
J304	Allergic rhinitis, unspecified
J450	Predominantly allergic asthma
K522	Allergic and dietetic gastroenteritis and colitis
L20	Atopic dermatitis
L23	Allergic contact dermatitis
L500	Allergic urticaria
Z516	Desensitization to allergens
ANEMIA	
Included ICD-10 codes and labels	
D50	Iron deficiency anaemia
D51	Vitamin B12 deficiency anaemia
D52	Folate deficiency anaemia
D53	Other nutritional anaemias
D55	Anaemia due to enzyme disorders
D56	Thalassaemia
D57	Sickle-cell disorders
D58	Other hereditary haemolytic anaemias
D59	Acquired haemolytic anaemia
D60	Acquired pure red cell aplasia [erythroblastopenia]
D61	Other aplastic anaemias
D63	Anaemia in chronic diseases classified elsewhere
D64	Other anaemias
Excluded ICD-10 codes and labels	
D563	Thalassaemia trait
D590	Drug-induced autoimmune haemolytic anaemia
D592	Drug-induced nonautoimmune haemolytic anaemia
D593	Haemolytic-uraemic syndrome
D596	Haemoglobinuria due to haemolysis from other external causes
D601	Transient acquired pure red cell aplasia
D611	Drug-induced aplastic anaemia
D612	Aplastic anaemia due to other external agents
D642	Secondary sideroblastic anaemia due to drugs and toxins

ASTHMA	
Included ICD-10 codes and labels	
J45	Asthma
ATRIAL FIBRILLATION	
Included ICD-10 codes and labels	
I48	Atrial fibrillation and flutter
AUTOIMMUNE DISEASES	
Included ICD-10 codes and labels	
I731	Thromboangiitis obliterans [Buerger]
L10	Pemphigus
L12	Pemphigoid
L40	Psoriasis
L41	Parapsoriasis
L93	Lupus erythematosus
L94	Other localized connective tissue disorders
L95	Vasculitis limited to skin, not elsewhere classified
M30	Polyarteritis nodosa and related conditions
M31	Other necrotizing vasculopathies
M32	Systemic lupus erythematosus
M33	Dermatopolymyositis
M34	Systemic sclerosis
M35	Other systemic involvement of connective tissue
M36	Systemic disorders of connective tissue in diseases classified elsewhere
Excluded ICD-10 codes and labels	
L105	Drug-induced pemphigus
M320	Drug-induced systemic lupus erythematosus
M342	Systemic sclerosis induced by drugs and chemicals
M357	Hypermobility syndrome
M358	Other specified systemic involvement of connective tissue
M359	Systemic involvement of connective tissue, unspecified
M360	Dermato(poly)myositis in neoplastic disease
M361	Arthropathy in neoplastic disease
M362	Haemophilic arthropathy
M363	Arthropathy in other blood disorders
BLINDNESS, VISUAL IMPAIRMENT	
Included ICD-10 codes and labels	
H54	Visual impairment including blindness (binocular or monocular)
Z442	Fitting and adjustment of artificial eye
Z970	Presence of artificial eye
Excluded ICD-10 codes and labels	
H543	Mild or no visual impairment, binocular
BLOOD AND BLOOD FORMING ORGAN DISEASES	
Included ICD-10 codes and labels	

D66	Hereditary factor VIII deficiency
D67	Hereditary factor IX deficiency
D68	Other coagulation defects
D69	Purpura and other haemorrhagic conditions
D71	Functional disorders of polymorphonuclear neutrophils
D720	Genetic anomalies of leukocytes
D730	Hyposplenism
D731	Hypersplenism
D732	Chronic congestive splenomegaly
D74	Methaemoglobinaemia
D750	Familial erythrocytosis
D761	Haemophagocytic lymphohistiocytosis
D763	Other histiocytosis syndromes
D77	Other disorders of blood and blood-forming organs in diseases classified elsewhere
D80	Immunodeficiency with predominantly antibody defects
D81	Combined immunodeficiencies
D82	Immunodeficiency associated with other major defects
D83	Common variable immunodeficiency
D84	Other immunodeficiencies
D86	Sarcoidosis
D89	Other disorders involving the immune mechanism, not elsewhere classified
Excluded ICD-10 codes and labels	
D683	Haemorrhagic disorder due to circulating anticoagulants
D684	Acquired coagulation factor deficiency
D695	Secondary thrombocytopenia
D748	Other methaemoglobinaemias
D807	Transient hypogammaglobulinaemia of infancy
D891	Cryoglobulinaemia
D893	Immune reconstitution syndrome
BRADYCARDIAS AND CONDUCTION DISEASES	
Included ICD-10 codes and labels	
I441	Atrioventricular block, second degree
I442	Atrioventricular block, complete
I443	Other and unspecified atrioventricular block
I453	Trifascicular block
I455	Other specified heart block
Z950	Presence of cardiac pacemaker
CARDIAC VALVE DISEASES	
Included ICD-10 codes and labels	
I05	Rheumatic mitral valve diseases
I06	Rheumatic aortic valve diseases
I07	Rheumatic tricuspid valve diseases
I08	Multiple valve diseases

I091	Rheumatic diseases of endocardium, valve unspecified
I098	Other specified rheumatic heart diseases
I34	Nonrheumatic mitral valve disorders
I35	Nonrheumatic aortic valve disorders
I36	Nonrheumatic tricuspid valve disorders
I37	Pulmonary valve disorders
I38	Endocarditis, valve unspecified
I390	Mitral valve disorders in diseases classified elsewhere
I391	Aortic valve disorders in diseases classified elsewhere
I392	Tricuspid valve disorders in diseases classified elsewhere
I393	Pulmonary valve disorders in diseases classified elsewhere
I394	Multiple valve disorders in diseases classified elsewhere
Q22	Congenital malformations of pulmonary and tricuspid valves
Q23	Congenital malformations of aortic and mitral valves
Z952	Presence of prosthetic heart valve
Z953	Presence of xenogenic heart valve
Z954	Presence of other heart-valve replacement
CATARACT AND OTHER LENS DISEASES	
Included ICD-10 codes and labels	
H25	Senile cataract
H26	Other cataract
H27	Other disorders of lens
H28	Cataract and other disorders of lens in diseases classified elsewhere
Q12	Congenital lens malformations
Z961	Presence of intraocular lens
CEREBROVASCULAR DISEASE	
Included ICD-10 codes and labels	
G45	Transient cerebral ischaemic attacks and related syndromes
G46	Vascular syndromes of brain in cerebrovascular diseases
I60	Subarachnoid haemorrhage
I61	Intracerebral haemorrhage
I62	Other nontraumatic intracranial haemorrhage
I63	Cerebral infarction
I64	Stroke, not specified as haemorrhage or infarction
I67	Other cerebrovascular diseases
I69	Sequelae of cerebrovascular disease
CHROMOSOMAL ABNORMALITIES	
Included ICD-10 codes and labels	
Q90	Down syndrome
Q91	Edwards syndrome and Patau syndrome
Q92	Other trisomies and partial trisomies of the autosomes, not elsewhere classified
Q93	Monosomies and deletions from the autosomes, not elsewhere classified
Q95	Balanced rearrangements and structural markers, not elsewhere classified

Q96	Turner syndrome
Q97	Other sex chromosome abnormalities, female phenotype, not elsewhere classified
Q98	Other sex chromosome abnormalities, male phenotype, not elsewhere classified
Q99	Other chromosome abnormalities, not elsewhere classified
CHRONIC INFECTIOUS DISEASES	
Included ICD-10 codes and labels	
A15	Respiratory tuberculosis, bacteriologically and histologically confirmed
A16	Respiratory tuberculosis, not confirmed bacteriologically or histologically
A17	Tuberculosis of nervous system
A18	Tuberculosis of other organs
A19	Miliary tuberculosis
A30	Leprosy [Hansen disease]
A31	Infection due to other mycobacteria
A50	Congenital syphilis
A52	Late syphilis
A53	Other and unspecified syphilis
A65	Nonvenereal syphilis
A66	Yaws
A67	Pinta [carate]
A692	Lyme disease
A81	Atypical virus infections of central nervous system
B20	Human immunodeficiency virus [HIV] disease resulting in infectious and parasitic diseases
B21	Human immunodeficiency virus [HIV] disease resulting in malignant neoplasms
B22	Human immunodeficiency virus [HIV] disease resulting in other specified diseases
B23	Human immunodeficiency virus [HIV] disease resulting in other conditions
B24	Unspecified human immunodeficiency virus [HIV] disease
B381	Chronic pulmonary coccidioidomycosis
B391	Chronic pulmonary histoplasmosis capsulati
B401	Chronic pulmonary blastomycosis
B572	Chagas disease (chronic) with heart involvement
B573	Chagas disease (chronic) with digestive system involvement
B574	Chagas disease (chronic) with nervous system involvement
B575	Chagas disease (chronic) with other organ involvement
B65	Schistosomiasis [bilharziasis]
B92	Sequelae of leprosy
B94	Sequelae of other and unspecified infectious and parasitic diseases
J65	Pneumoconiosis associated with tuberculosis
M863	Chronic multifocal osteomyelitis
M864	Chronic osteomyelitis with draining sinus
M865	Other chronic haematogenous osteomyelitis
M866	Other chronic osteomyelitis
CHRONIC KIDNEY DISEASES	
Included ICD-10 codes and labels	

I120	Hypertensive renal disease with renal failure
I130	Hypertensive heart and renal disease with (congestive) heart failure
I131	Hypertensive heart and renal disease with renal failure
I132	Hypertensive heart and renal disease with both (congestive) heart failure and renal failure
I139	Hypertensive heart and renal disease, unspecified
N01	Rapidly progressive nephritic syndrome
N03	Chronic nephritic syndrome
N04	Nephrotic syndrome
N05	Unspecified nephritic syndrome
N07	Hereditary nephropathy, not elsewhere classified
N08	Glomerular disorders in diseases classified elsewhere
N11	Chronic tubulo-interstitial nephritis
N183	Chronic kidney disease, stage 3
N184	Chronic kidney disease, stage 4
N185	Chronic kidney disease, stage 5
N189	Chronic kidney disease, unspecified
Q60	Renal agenesis and other reduction defects of kidney
Q611	Polycystic kidney, autosomal recessive
Q612	Polycystic kidney, autosomal dominant
Q613	Polycystic kidney, unspecified
Q614	Renal dysplasia
Q615	Medullary cystic kidney
Q618	Other cystic kidney diseases
Q619	Cystic kidney disease, unspecified
Z905	Acquired absence of kidney
Z940	Kidney transplant status
CHRONIC LIVER DISEASES	
Included ICD-10 codes and labels	
B18	Chronic viral hepatitis
K70	Alcoholic liver disease
K713	Toxic liver disease with chronic persistent hepatitis
K714	Toxic liver disease with chronic lobular hepatitis
K715	Toxic liver disease with chronic active hepatitis
K717	Toxic liver disease with fibrosis and cirrhosis of liver
K721	Chronic hepatic failure
K73	Chronic hepatitis, not elsewhere classified
K74	Fibrosis and cirrhosis of liver
K753	Granulomatous hepatitis, not elsewhere classified
K754	Autoimmune hepatitis
K758	Other specified inflammatory liver diseases
K761	Chronic passive congestion of liver
K766	Portal hypertension
K767	Hepatorenal syndrome

K778	Liver disorders in other diseases classified elsewhere
Q446	Cystic disease of liver
Z944	Liver transplant status
Excluded ICD-10 codes and labels	
K700	Alcoholic fatty liver
K701	Alcoholic hepatitis
CHRONIC PANCREAS, BILIARY TRACT AND GALLBLADDER DISEASES	
Included ICD-10 codes and labels	
K800	Calculus of gallbladder with acute cholecystitis
K801	Calculus of gallbladder with other cholecystitis
K802	Calculus of gallbladder without cholecystitis
K808	Other cholelithiasis
K811	Chronic cholecystitis
K86	Other diseases of pancreas
Q440	Agenesis, aplasia and hypoplasia of gallbladder
Q441	Other congenital malformations of gallbladder
Q442	Atresia of bile ducts
Q443	Congenital stenosis and stricture of bile ducts
Q444	Choledochal cyst
Q445	Other congenital malformations of bile ducts
Q450	Agenesis, aplasia and hypoplasia of pancreas
Excluded ICD-10 codes and labels	
K862	Cyst of pancreas
K863	Pseudocyst of pancreas
K869	Disease of pancreas, unspecified
CHRONIC ULCER OF THE SKIN	
Included ICD-10 codes and labels	
I830	Varicose veins of lower extremities with ulcer
I832	Varicose veins of lower extremities with both ulcer and inflammation
L89	Decubitus ulcer and pressure area
L97	Ulcer of lower limb, not elsewhere classified
L984	Chronic ulcer of skin, not elsewhere classified
COLITIS AND RELATED DISEASES	
Included ICD-10 codes and labels	
K520	Gastroenteritis and colitis due to radiation
K528	Other specified noninfective gastroenteritis and colitis
K551	Chronic vascular disorders of intestine
K552	Angiodysplasia of colon
K572	Diverticular disease of large intestine with perforation and abscess
K573	Diverticular disease of large intestine without perforation or abscess
K574	Diverticular disease of both small and large intestine with perforation and abscess
K575	Diverticular disease of both small and large intestine without perforation or abscess
K578	Diverticular disease of intestine, part unspecified, with perforation and abscess

K579	Diverticular disease of intestine, part unspecified, without perforation or abscess
K58	Irritable bowel syndrome
K590	Constipation
K592	Neurogenic bowel, not elsewhere classified
K62	Other diseases of anus and rectum
K634	Enteroptosis
K64	Haemorrhoids and perianal venous thrombosis
Excluded ICD-10 codes and labels	
K620	Anal polyp
K621	Rectal polyp
K625	Haemorrhage of anus and rectum
K626	Ulcer of anus and rectum
K645	Perianal venous thrombosis
COPD, EMPHYSEMA, CHRONIC BRONCHITIS	
Included ICD-10 codes and labels	
J41	Simple and mucopurulent chronic bronchitis
J42	Unspecified chronic bronchitis
J43	Emphysema
J44	Other chronic obstructive pulmonary disease
J47	Bronchiectasis
DEAFNESS, HEARING IMPAIRMENT	
Included ICD-10 codes and labels	
H80	Otosclerosis
H90	Conductive and sensorineural hearing loss
H911	Presbycusis
H913	Deaf mutism, not elsewhere classified
H919	Hearing loss, unspecified
Q16	Congenital malformations of ear causing impairment of hearing
Z453	Adjustment and management of implanted hearing device
Z461	Fitting and adjustment of hearing aid
Z962	Presence of otological and audiological implants
Z974	Presence of external hearing-aid
DEMENTIA	
Included ICD-10 codes and labels	
F00	Dementia in Alzheimer disease
F01	Vascular dementia
F02	Dementia in other diseases classified elsewhere
F03	Unspecified dementia
F051	Delirium superimposed on dementia
G30	Alzheimer disease
G31	Other degenerative diseases of nervous system, not elsewhere classified
DEPRESSION AND MOOD DISEASES	
Included ICD-10 codes and labels	

F30	Manic episode
F31	Bipolar affective disorder
F32	Depressive episode
F33	Recurrent depressive disorder
F34	Persistent mood [affective] disorders
F38	Other mood [affective] disorders
F39	Unspecified mood [affective] disorder
F412	Mixed anxiety and depressive disorder
DIABETES	
Included ICD-10 codes and labels	
E10	Insulin-dependent diabetes mellitus
E11	Non-insulin-dependent diabetes mellitus
E13	Other specified diabetes mellitus
E14	Unspecified diabetes mellitus
E891	Postprocedural hypoinsulinaemia
DORSOPATHIES	
Included ICD-10 codes and labels	
M40	Kyphosis and lordosis
M41	Scoliosis
M42	Spinal osteochondrosis
M43	Other deforming dorsopathies
M47	Spondylosis
M48	Other spondylopathies
M49	Spondylopathies in diseases classified elsewhere
M50	Cervical disc disorders
M51	Other intervertebral disc disorders
M53	Other dorsopathies, not elsewhere classified
Q675	Congenital deformity of spine
Q761	Klippel-Feil syndrome
Q764	Other congenital malformations of spine, not associated with scoliosis
DYSLIPIDEMIA	
Included ICD-10 codes and labels	
E78	Disorders of lipoprotein metabolism and other lipidaemias
EAR, NOSE, THROAT DISEASES	
Included ICD-10 codes and labels	
H604	Cholesteatoma of external ear
H661	Chronic tubotympanic suppurative otitis media
H662	Chronic atticoantral suppurative otitis media
H663	Other chronic suppurative otitis media
H701	Chronic mastoiditis
H71	Cholesteatoma of middle ear
H731	Chronic myringitis
H741	Adhesive middle ear disease

H810	MÚniÞre disease
H831	Labyrinthine fistula
H832	Labyrinthine dysfunction
H95	Postprocedural disorders of ear and mastoid process, not elsewhere classified
J300	Vasomotor rhinitis
J31	Chronic rhinitis, nasopharyngitis and pharyngitis
J32	Chronic sinusitis
J33	Nasal polyp
J341	Cyst and mucocele of nose and nasal sinus
J342	Deviated nasal septum
J343	Hypertrophy of nasal turbinates
J35	Chronic diseases of tonsils and adenoids
J37	Chronic laryngitis and laryngotracheitis
J380	Paralysis of vocal cords and larynx
J386	Stenosis of larynx
K051	Chronic gingivitis
K053	Chronic periodontitis
K07	Dentofacial anomalies [including malocclusion]
K110	Atrophy of salivary gland
K117	Disturbances of salivary secretion
Q30	Congenital malformations of nose
Q31	Congenital malformations of larynx
Q32	Congenital malformations of trachea and bronchus
Q35	Cleft palate
Q36	Cleft lip
Q37	Cleft palate with cleft lip
Q38	Other congenital malformations of tongue, mouth and pharynx
EPILEPSY	
Included ICD-10 codes and labels	
G40	Epilepsy
Excluded ICD-10 codes and labels	
G405	Special epileptic syndromes
ESOPHAGUS, STOMACH AND DUODENUM DISEASES	
Included ICD-10 codes and labels	
I85	Oesophageal varices
I864	Gastric varices
I982	Oesophageal varices without bleeding in diseases classified elsewhere
I983	Oesophageal varices with bleeding in diseases classified elsewhere
K21	Gastro-oesophageal reflux disease
K220	Achalasia of cardia
K222	Oesophageal obstruction
K224	Dyskinesia of oesophagus
K225	Diverticulum of oesophagus, acquired

K227	Barrett oesophagus
K230	Tuberculous oesophagitis
K231	Megaoesophagus in Chagas disease
K254	Gastric ulcer: Chronic or unspecified with haemorrhage
K255	Gastric ulcer: Chronic or unspecified with perforation
K256	Gastric ulcer: Chronic or unspecified with both haemorrhage and perforation
K257	Gastric ulcer: Chronic without haemorrhage or perforation
K264	Duodenal ulcer: Chronic or unspecified with haemorrhage
K265	Duodenal ulcer: Chronic or unspecified with perforation
K266	Duodenal ulcer: Chronic or unspecified with both haemorrhage and perforation
K267	Duodenal ulcer: Chronic without haemorrhage or perforation
K274	Peptic ulcer, site unspecified: Chronic or unspecified with haemorrhage
K275	Peptic ulcer, site unspecified: Chronic or unspecified with perforation
K276	Peptic ulcer, site unspecified: Chronic or unspecified with both haemorrhage and perforation
K277	Peptic ulcer, site unspecified: Chronic without haemorrhage or perforation
K284	Gastrojejunal ulcer: Chronic or unspecified with haemorrhage
K285	Gastrojejunal ulcer: Chronic or unspecified with perforation
K286	Gastrojejunal ulcer: Chronic or unspecified with both haemorrhage and perforation
K287	Gastrojejunal ulcer: Chronic without haemorrhage or perforation
K293	Chronic superficial gastritis
K294	Chronic atrophic gastritis
K295	Chronic gastritis, unspecified
K296	Other gastritis
K297	Gastritis, unspecified
K298	Duodenitis
K299	Gastroduodenitis, unspecified
K311	Adult hypertrophic pyloric stenosis
K312	Hourglass stricture and stenosis of stomach
K313	Pylorospasm, not elsewhere classified
K314	Gastric diverticulum
K315	Obstruction of duodenum
Q39	Congenital malformations of oesophagus
Q40	Other congenital malformations of upper alimentary tract
Z903	Acquired absence of part of stomach
GLAUCOMA	
Included ICD-10 codes and labels	
H401	Primary open-angle glaucoma
H402	Primary angle-closure glaucoma
H403	Glaucoma secondary to eye trauma
H404	Glaucoma secondary to eye inflammation
H405	Glaucoma secondary to other eye disorders
H406	Glaucoma secondary to drugs
H408	Other glaucoma

H409	Glaucoma, unspecified
HEART FAILURE	
Included ICD-10 codes and labels	
I110	Hypertensive heart disease with (congestive) heart failure
I130	Hypertensive heart and renal disease with (congestive) heart failure
I132	Hypertensive heart and renal disease with both (congestive) heart failure and renal failure
I27	Other pulmonary heart diseases
I280	Arteriovenous fistula of pulmonary vessels
I42	Cardiomyopathy
I43	Cardiomyopathy in diseases classified elsewhere
I50	Heart failure
I515	Myocardial degeneration
I517	Cardiomegaly
I528	Other heart disorders in other diseases classified elsewhere
Z941	Heart transplant status
Z943	Heart and lungs transplant status
HEMATOLOGICAL NEOPLASMS	
Included ICD-10 codes and labels	
C81	Hodgkin lymphoma
C82	Follicular lymphoma
C83	Non-follicular lymphoma
C84	Mature T/NK-cell lymphomas
C85	Other and unspecified types of non-Hodgkin lymphoma
C86	Other specified types of T/NK-cell lymphoma
C88	Malignant immunoproliferative diseases
C90	Multiple myeloma and malignant plasma cell neoplasms
C91	Lymphoid leukaemia
C92	Myeloid leukaemia
C93	Monocytic leukaemia
C94	Other leukaemias of specified cell type
C95	Leukaemia of unspecified cell type
C96	Other and unspecified malignant neoplasms of lymphoid, haematopoietic and related tissue
HYPERTENSION	
Included ICD-10 codes and labels	
I10	Essential (primary) hypertension
I11	Hypertensive heart disease
I12	Hypertensive renal disease
I13	Hypertensive heart and renal disease
I15	Secondary hypertension
INFLAMMATORY ARTHROPATHIES	
Included ICD-10 codes and labels	
M023	Reiter disease
M05	Seropositive rheumatoid arthritis

M06	Other rheumatoid arthritis
M07	Psoriatic and enteropathic arthropathies
M08	Juvenile arthritis
M09	Juvenile arthritis in diseases classified elsewhere
M10	Gout
M11	Other crystal arthropathies
M12	Other specific arthropathies
M13	Other arthritis
M14	Arthropathies in other diseases classified elsewhere
M45	Ankylosing spondylitis
M460	Spinal enthesopathy
M461	Sacroiliitis, not elsewhere classified
M468	Other specified inflammatory spondylopathies
M469	Inflammatory spondylopathy, unspecified
INFLAMMATORY BOWEL DISEASES	
Included ICD-10 codes and labels	
K50	Crohn disease [regional enteritis]
K51	Ulcerative colitis
ISCHEMIC HEART DISEASE	
Included ICD-10 codes and labels	
I20	Angina pectoris
I21	Acute myocardial infarction
I22	Subsequent myocardial infarction
I24	Other acute ischaemic heart diseases
I25	Chronic ischaemic heart disease
Z951	Presence of aortocoronary bypass graft
Z955	Presence of coronary angioplasty implant and graft
MIGRAINE AND FACIAL PAIN SYNDROMES	
Included ICD-10 codes and labels	
G43	Migraine
G440	Cluster headache syndrome
G441	Vascular headache, not elsewhere classified
G442	Tension-type headache
G443	Chronic post-traumatic headache
G448	Other specified headache syndromes
G50	Disorders of trigeminal nerve
MULTIPLE SCLEROSIS	
Included ICD-10 codes and labels	
G35	Multiple sclerosis
NEUROTIC, STRESS-RELATED AND SOMATOFORM DISEASES	
Included ICD-10 codes and labels	
F40	Phobic anxiety disorders
F41	Other anxiety disorders

F42	Obsessive-compulsive disorder
F43	Reaction to severe stress, and adjustment disorders
F44	Dissociative [conversion] disorders
F45	Somatoform disorders
F48	Other neurotic disorders
Excluded ICD-10 codes and labels	
F430	Acute stress reaction
F432	Adjustment disorders
OBESITY	
Included ICD-10 codes and labels	
E66	Obesity
OSTEOARTHRITIS AND OTHER DEGENERATIVE JOINT DISEASES	
Included ICD-10 codes and labels	
M15	Polyarthrosis
M16	Coxarthrosis [arthrosis of hip]
M17	Gonarthrosis [arthrosis of knee]
M18	Arthrosis of first carpometacarpal joint
M19	Other arthrosis
M362	Haemophilic arthropathy
M363	Arthropathy in other blood disorders
OSTEOPOROSIS	
Included ICD-10 codes and labels	
M80	Osteoporosis with pathological fracture
M81	Osteoporosis without pathological fracture
M82	Osteoporosis in diseases classified elsewhere
OTHER CARDIOVASCULAR DISEASES	
Included ICD-10 codes and labels	
I09	Other rheumatic heart diseases
I281	Aneurysm of pulmonary artery
I310	Chronic adhesive pericarditis
I311	Chronic constrictive pericarditis
I456	Pre-excitation syndrome
I495	Sick sinus syndrome
I498	Other specified cardiac arrhythmias
I70	Atherosclerosis
I71	Aortic aneurysm and dissection
I72	Other aneurysm and dissection
I790	Aneurysm of aorta in diseases classified elsewhere
I791	Aortitis in diseases classified elsewhere
I950	Idiopathic hypotension
I951	Orthostatic hypotension
I958	Other hypotension
Q20	Congenital malformations of cardiac chambers and connections

Q21	Congenital malformations of cardiac septa
Q24	Other congenital malformations of heart
Q25	Congenital malformations of great arteries
Q26	Congenital malformations of great veins
Q27	Other congenital malformations of peripheral vascular system
Q28	Other congenital malformations of circulatory system
Z958	Presence of other cardiac and vascular implants and grafts
Z959	Presence of cardiac and vascular implant and graft, unspecified
Excluded ICD-10 codes and labels	
I091	Rheumatic diseases of endocardium, valve unspecified
I098	Other specified rheumatic heart diseases
I702	Atherosclerosis of arteries of extremities
OTHER DIGESTIVE DISEASES	
Included ICD-10 codes and labels	
K660	Peritoneal adhesions
K900	Coeliac disease
K901	Tropical sprue
K902	Blind loop syndrome, not elsewhere classified
K911	Postgastric surgery syndromes
K93	Disorders of other digestive organs in diseases classified elsewhere
Q41	Congenital absence, atresia and stenosis of small intestine
Q42	Congenital absence, atresia and stenosis of large intestine
Q43	Other congenital malformations of intestine
R15	Faecal incontinence
Z904	Acquired absence of other parts of digestive tract
Z980	Intestinal bypass and anastomosis status
OTHER EYE DISEASES	
Included ICD-10 codes and labels	
H022	Lagophthalmos
H023	Blepharochalasis
H024	Ptosis of eyelid
H025	Other disorders affecting eyelid function
H04	Disorders of lacrimal system
H05	Disorders of orbit
H104	Chronic conjunctivitis
H17	Corneal scars and opacities
H184	Corneal degeneration
H185	Hereditary corneal dystrophies
H186	Keratoconus
H187	Other corneal deformities
H188	Other specified disorders of cornea
H189	Disorder of cornea, unspecified
H193	Keratitis and keratoconjunctivitis in other diseases classified elsewhere

H198	Other disorders of sclera and cornea in diseases classified elsewhere
H201	Chronic iridocyclitis
H21	Other disorders of iris and ciliary body
H310	Chorioretinal scars
H311	Choroidal degeneration
H312	Hereditary choroidal dystrophy
H318	Other specified disorders of choroid
H319	Disorder of choroid, unspecified
H33	Retinal detachments and breaks
H352	Other proliferative retinopathy
H353	Degeneration of macula and posterior pole
H354	Peripheral retinal degeneration
H355	Hereditary retinal dystrophy
H357	Separation of retinal layers
H358	Other specified retinal disorders
H359	Retinal disorder, unspecified
H36	Retinal disorders in diseases classified elsewhere
H47	Other disorders of optic [2nd] nerve and visual pathways
H48	Disorders of optic [2nd] nerve and visual pathways in diseases classified elsewhere
H49	Paralytic strabismus
H51	Other disorders of binocular movement
Q10	Congenital malformations of eyelid, lacrimal apparatus and orbit
Q11	Anophthalmos, microphthalmos and macrophthalmos
Q13	Congenital malformations of anterior segment of eye
Q14	Congenital malformations of posterior segment of eye
Q15	Other congenital malformations of eye
Z947	Corneal transplant status
Excluded ICD-10 codes and labels	
H043	Acute and unspecified inflammation of lacrimal passages
H050	Acute inflammation of orbit
H470	Disorders of optic nerve, not elsewhere classified
H471	Papilloedema, unspecified
H481	Retrobulbar neuritis in diseases classified elsewhere
OTHER GENITOURINARY DISEASES	
Included ICD-10 codes and labels	
B901	Sequelae of genitourinary tuberculosis
N200	Calculus of kidney
N202	Calculus of kidney with calculus of ureter
N209	Urinary calculus, unspecified
N210	Calculus in bladder
N218	Other lower urinary tract calculus
N219	Calculus of lower urinary tract, unspecified
N22	Calculus of urinary tract in diseases classified elsewhere

N301	Interstitial cystitis (chronic)
N302	Other chronic cystitis
N303	Trigonitis
N304	Irradiation cystitis
N31	Neuromuscular dysfunction of bladder, not elsewhere classified
N320	Bladder-neck obstruction
N323	Diverticulum of bladder
N328	Other specified disorders of bladder
N329	Bladder disorder, unspecified
N33	Bladder disorders in diseases classified elsewhere
N35	Urethral stricture
N393	Stress incontinence
N394	Other specified urinary incontinence
N480	Leukoplakia of penis
N484	Impotence of organic origin
N489	Disorder of penis, unspecified
N701	Chronic salpingitis and oophoritis
N711	Chronic inflammatory disease of uterus
N731	Chronic parametritis and pelvic cellulitis
N734	Female chronic pelvic peritonitis
N736	Female pelvic peritoneal adhesions
N761	Subacute and chronic vaginitis
N763	Subacute and chronic vulvitis
N81	Female genital prolapse
N88	Other noninflammatory disorders of cervix uteri
N895	Stricture and atresia of vagina
N905	Atrophy of vulva
N952	Postmenopausal atrophic vaginitis
Q54	Hypospadias
Q620	Congenital hydronephrosis
Q621	Atresia and stenosis of ureter
Q622	Congenital megaloureter
Q623	Other obstructive defects of renal pelvis and ureter
Q624	Agenesis of ureter
Q627	Congenital vesico-uretero-renal reflux
Q628	Other congenital malformations of ureter
Q638	Other specified congenital malformations of kidney
Q639	Congenital malformation of kidney, unspecified
Q640	Epispadias
Q641	Exstrophy of urinary bladder
Q643	Other atresia and stenosis of urethra and bladder neck
Q644	Malformation of urachus
Q645	Congenital absence of bladder and urethra

Q646	Congenital diverticulum of bladder
Q647	Other congenital malformations of bladder and urethra
Q648	Other specified congenital malformations of urinary system
Q649	Congenital malformation of urinary system, unspecified
Z906	Acquired absence of other organs of urinary tract
Z907	Acquired absence of genital organ(s)
Z960	Presence of urogenital implants
OTHER METABOLIC DISEASES	
Included ICD-10 codes and labels	
E20	Hypoparathyroidism
E21	Hyperparathyroidism and other disorders of parathyroid gland
E22	Hyperfunction of pituitary gland
E23	Hypofunction and other disorders of pituitary gland
E24	Cushing syndrome
E25	Adrenogenital disorders
E26	Hyperaldosteronism
E27	Other disorders of adrenal gland
E28	Ovarian dysfunction
E29	Testicular dysfunction
E31	Polyglandular dysfunction
E34	Other endocrine disorders
E35	Disorders of endocrine glands in diseases classified elsewhere
E40	Kwashiorkor
E41	Nutritional marasmus
E42	Marasmic kwashiorkor
E43	Unspecified severe protein-energy malnutrition
E44	Protein-energy malnutrition of moderate and mild degree
E45	Retarded development following protein-energy malnutrition
E46	Unspecified protein-energy malnutrition
E64	Sequelae of malnutrition and other nutritional deficiencies
E70	Disorders of aromatic amino-acid metabolism
E71	Disorders of branched-chain amino-acid metabolism and fatty-acid metabolism
E72	Other disorders of amino-acid metabolism
E74	Other disorders of carbohydrate metabolism
E75	Disorders of sphingolipid metabolism and other lipid storage disorders
E76	Disorders of glycosaminoglycan metabolism
E77	Disorders of glycoprotein metabolism
E79	Disorders of purine and pyrimidine metabolism
E80	Disorders of porphyrin and bilirubin metabolism
E83	Disorders of mineral metabolism
E84	Cystic fibrosis
E85	Amyloidosis
E88	Other metabolic disorders

E89	Postprocedural endocrine and metabolic disorders, not elsewhere classified
K903	Pancreatic steatorrhoea
K904	Malabsorption due to intolerance, not elsewhere classified
K908	Other intestinal malabsorption
K909	Intestinal malabsorption, unspecified
K912	Postsurgical malabsorption, not elsewhere classified
M83	Adult osteomalacia
M88	Paget disease of bone [osteitis deformans]
N25	Disorders resulting from impaired renal tubular function
Excluded ICD-10 codes and labels	
E231	Drug-induced hypopituitarism
E242	Drug-induced Cushing syndrome
E244	Alcohol-induced pseudo-Cushing syndrome
E273	Drug-induced adrenocortical insufficiency
E343	Short stature, not elsewhere classified
E344	Constitutional tall stature
E350	Disorders of thyroid gland in diseases classified elsewhere
E441	Mild protein-energy malnutrition
E790	Hyperuricaemia without signs of inflammatory arthritis and tophaceous disease
E804	Gilbert syndrome
E883	Tumour lysis syndrome
E890	Postprocedural hypothyroidism
E892	Postprocedural hypoparathyroidism
OTHER MUSCULOSKELETAL AND JOINT DISEASES	
Included ICD-10 codes and labels	
B902	Sequelae of tuberculosis of bones and joints
M212	Flexion deformity
M213	Wrist or foot drop (acquired)
M214	Flat foot [pes planus] (acquired)
M215	Acquired clawhand, clubhand, clawfoot and clubfoot
M216	Other acquired deformities of ankle and foot
M217	Unequal limb length (acquired)
M218	Other specified acquired deformities of limbs
M219	Acquired deformity of limb, unspecified
M22	Disorders of patella
M23	Internal derangement of knee
M24	Other specific joint derangements
M252	Flail joint
M253	Other instability of joint
M357	Hypermobility syndrome
M61	Calcification and ossification of muscle
M652	Calcific tendinitis
M653	Trigger finger

M654	Radial styloid tenosynovitis [de Quervain]
M700	Chronic crepitant synovitis of hand and wrist
M720	Palmar fascial fibromatosis [Dupuytren]
M722	Plantar fascial fibromatosis
M724	Pseudosarcomatous fibromatosis
M750	Adhesive capsulitis of shoulder
M751	Rotator cuff syndrome
M753	Calcific tendinitis of shoulder
M754	Impingement syndrome of shoulder
M797	Fibromyalgia
M841	Nonunion of fracture [pseudarthrosis]
M89	Other disorders of bone
M91	Juvenile osteochondrosis of hip and pelvis
M93	Other osteochondropathies
M94	Other disorders of cartilage
M96	Postprocedural musculoskeletal disorders, not elsewhere classified
M99	Biomechanical lesions, not elsewhere classified
Q65	Congenital deformities of hip
Q66	Congenital deformities of feet
Q68	Other congenital musculoskeletal deformities
Q71	Reduction defects of upper limb
Q72	Reduction defects of lower limb
Q73	Reduction defects of unspecified limb
Q74	Other congenital malformations of limb(s)
Q77	Osteochondrodysplasia with defects of growth of tubular bones and spine
Q78	Other osteochondrodysplasias
Q796	Ehlers-Danlos syndrome
Q798	Other congenital malformations of musculoskeletal system
Q87	Other specified congenital malformation syndromes affecting multiple systems
S382	Traumatic amputation of external genital organs
S48	Traumatic amputation of shoulder and upper arm
S58	Traumatic amputation of forearm
S68	Traumatic amputation of wrist and hand
S78	Traumatic amputation of hip and thigh
S88	Traumatic amputation of lower leg
S98	Traumatic amputation of ankle and foot
T05	Traumatic amputations involving multiple body regions
T096	Traumatic amputation of trunk, level unspecified
T116	Traumatic amputation of upper limb, level unspecified
T136	Traumatic amputation of lower limb, level unspecified
T147	Crushing injury and traumatic amputation of unspecified body region
T90	Sequelae of injuries of head
T91	Sequelae of injuries of neck and trunk

T92	Sequelae of injuries of upper limb
T93	Sequelae of injuries of lower limb
T94	Sequelae of injuries involving multiple and unspecified body regions
T95	Sequelae of burns, corrosions and frostbite
T96	Sequelae of poisoning by drugs, medicaments and biological substances
T97	Sequelae of toxic effects of substances chiefly nonmedicinal as to source
T98	Sequelae of other and unspecified effects of external causes
Z440	Fitting and adjustment of artificial arm (complete)(partial)
Z441	Fitting and adjustment of artificial leg (complete)(partial)
Z891	Acquired absence of hand and wrist
Z892	Acquired absence of upper limb above wrist
Z893	Acquired absence of both upper limbs [any level]
Z894	Acquired absence of foot and ankle
Z895	Acquired absence of leg at or below knee
Z896	Acquired absence of leg above knee
Z897	Acquired absence of both lower limbs [any level, except toes alone]
Z898	Acquired absence of upper and lower limbs [any level]
Z899	Acquired absence of limb, unspecified
Z946	Bone transplant status
Z966	Presence of orthopaedic joint implants
Z971	Presence of artificial limb (complete)(partial)
OTHER NEUROLOGICAL DISEASES	
Included ICD-10 codes and labels	
B900	Sequelae of central nervous system tuberculosis
D482	Neoplasm of uncertain or unknown behaviour: Peripheral nerves and autonomic nervous system
G041	Tropical spastic paraplegia
G09	Sequelae of inflammatory diseases of central nervous system
G10	Huntington disease
G11	Hereditary ataxia
G12	Spinal muscular atrophy and related syndromes
G13	Systemic atrophies primarily affecting central nervous system in diseases classified elsewhere
G24	Dystonia
G25	Other extrapyramidal and movement disorders
G26	Extrapyramidal and movement disorders in diseases classified elsewhere
G32	Other degenerative disorders of nervous system in diseases classified elsewhere
G37	Other demyelinating diseases of central nervous system
G51	Facial nerve disorders
G52	Disorders of other cranial nerves
G53	Cranial nerve disorders in diseases classified elsewhere
G70	Myasthenia gravis and other myoneural disorders
G71	Primary disorders of muscles

G723	Periodic paralysis
G724	Inflammatory myopathy, not elsewhere classified
G728	Other specified myopathies
G729	Myopathy, unspecified
G73	Disorders of myoneural junction and muscle in diseases classified elsewhere
G80	Cerebral palsy
G81	Hemiplegia
G82	Paraplegia and tetraplegia
G83	Other paralytic syndromes
G90	Disorders of autonomic nervous system
G91	Hydrocephalus
G938	Other specified disorders of brain
G939	Disorder of brain, unspecified
G95	Other diseases of spinal cord
G99	Other disorders of nervous system in diseases classified elsewhere
M471	Other spondylosis with myelopathy
Q00	Anencephaly and similar malformations
Q01	Encephalocele
Q02	Microcephaly
Q03	Congenital hydrocephalus
Q04	Other congenital malformations of brain
Q05	Spina bifida
Q06	Other congenital malformations of spinal cord
Q07	Other congenital malformations of nervous system
Q760	Spina bifida occulta
Excluded ICD-10 codes and labels	
G130	Paraneoplastic neuromyopathy and neuropathy
G131	Other systemic atrophy primarily affecting central nervous system in neoplastic disease
G251	Drug-induced tremor
G254	Drug-induced chorea
G256	Drug-induced tics and other tics of organic origin
G510	Bell palsy
G732	Other myasthenic syndromes in neoplastic disease
G733	Myasthenic syndromes in other diseases classified elsewhere
G734	Myopathy in infectious and parasitic diseases classified elsewhere
G838	Other specified paralytic syndromes
OTHER PSYCHIATRIC AND BEHAVIORAL DISEASES	
Included ICD-10 codes and labels	
F04	Organic amnesic syndrome, not induced by alcohol and other psychoactive substances
F06	Other mental disorders due to brain damage and dysfunction and to physical disease
F07	Personality and behavioural disorders due to brain disease, damage and dysfunction
F09	Unspecified organic or symptomatic mental disorder
F102	Mental and behavioural disorders due to use of alcohol: Dependence syndrome

F106	Mental and behavioural disorders due to use of alcohol: Amnesic syndrome
F107	Mental and behavioural disorders due to use of alcohol: Residual and late-onset psychotic disorder
F112	Mental and behavioural disorders due to use of opioids: Dependence syndrome
F116	Mental and behavioural disorders due to use of opioids: Amnesic syndrome
F117	Mental and behavioural disorders due to use of opioids: Residual and late-onset psychotic disorder
F122	Mental and behavioural disorders due to use of cannabinoids: Dependence syndrome
F126	Mental and behavioural disorders due to use of cannabinoids: Amnesic syndrome
F127	Mental and behavioural disorders due to use of cannabinoids: Residual and late-onset psychotic disorder
F132	Mental and behavioural disorders due to use of sedatives or hypnotics: Dependence syndrome
F136	Mental and behavioural disorders due to use of sedatives or hypnotics: Amnesic syndrome
F137	Mental and behavioural disorders due to use of sedatives or hypnotics: Residual and late-onset psychotic disorder
F142	Mental and behavioural disorders due to use of cocaine: Dependence syndrome
F146	Mental and behavioural disorders due to use of cocaine: Amnesic syndrome
F147	Mental and behavioural disorders due to use of cocaine: Residual and late-onset psychotic disorder
F152	Mental and behavioural disorders due to use of other stimulants, including caffeine: Dependence syndrome
F156	Mental and behavioural disorders due to use of other stimulants, including caffeine: Amnesic syndrome
F157	Mental and behavioural disorders due to use of other stimulants, including caffeine: Residual and late-onset psychotic disorder
F162	Mental and behavioural disorders due to use of hallucinogens: Dependence syndrome
F166	Mental and behavioural disorders due to use of hallucinogens: Amnesic syndrome
F167	Mental and behavioural disorders due to use of hallucinogens: Residual and late-onset psychotic disorder
F172	Mental and behavioural disorders due to use of tobacco: Dependence syndrome
F176	Mental and behavioural disorders due to use of tobacco: Amnesic syndrome
F177	Mental and behavioural disorders due to use of tobacco: Residual and late-onset psychotic disorder
F182	Mental and behavioural disorders due to use of volatile solvents: Dependence syndrome
F186	Mental and behavioural disorders due to use of volatile solvents: Amnesic syndrome
F187	Mental and behavioural disorders due to use of volatile solvents: Residual and late-onset psychotic disorder
F192	Mental and behavioural disorders due to multiple drug use and use of other psychoactive substances: Dependence syndrome
F196	Mental and behavioural disorders due to multiple drug use and use of other psychoactive substances: Amnesic syndrome
F197	Mental and behavioural disorders due to multiple drug use and use of other psychoactive substances: Residual and late-onset psychotic disorder
F50	Eating disorders
F52	Sexual dysfunction, not caused by organic disorder or disease
F60	Specific personality disorders

F61	Mixed and other personality disorders
F62	Enduring personality changes, not attributable to brain damage and disease
F63	Habit and impulse disorders
F68	Other disorders of adult personality and behaviour
F70	Mild mental retardation
F71	Moderate mental retardation
F72	Severe mental retardation
F73	Profound mental retardation
F78	Other mental retardation
F79	Unspecified mental retardation
F80	Specific developmental disorders of speech and language
F81	Specific developmental disorders of scholastic skills
F82	Specific developmental disorder of motor function
F83	Mixed specific developmental disorders
F84	Pervasive developmental disorders
F88	Other disorders of psychological development
F89	Unspecified disorder of psychological development
F95	Tic disorders
F99	Mental disorder, not otherwise specified
OTHER RESPIRATORY DISEASES	
Included ICD-10 codes and labels	
B909	Sequelae of respiratory and unspecified tuberculosis
E662	Extreme obesity with alveolar hypoventilation
J60	Coalworker pneumoconiosis
J61	Pneumoconiosis due to asbestos and other mineral fibres
J62	Pneumoconiosis due to dust containing silica
J63	Pneumoconiosis due to other inorganic dusts
J64	Unspecified pneumoconiosis
J65	Pneumoconiosis associated with tuberculosis
J66	Airway disease due to specific organic dust
J67	Hypersensitivity pneumonitis due to organic dust
J684	Chronic respiratory conditions due to chemicals, gases, fumes and vapours
J701	Chronic and other pulmonary manifestations due to radiation
J703	Chronic drug-induced interstitial lung disorders
J704	Drug-induced interstitial lung disorders, unspecified
J84	Other interstitial pulmonary diseases
J92	Pleural plaque
J941	Fibrothorax
J953	Chronic pulmonary insufficiency following surgery
J955	Postprocedural subglottic stenosis
J961	Chronic respiratory failure
J98	Other respiratory disorders
Q33	Congenital malformations of lung

Q34	Other congenital malformations of respiratory system
Z902	Acquired absence of lung [part of]
Z942	Lung transplant status
Z943	Heart and lungs transplant status
Z963	Presence of artificial larynx
Excluded ICD-10 codes and labels	
J981	Pulmonary collapse
OTHER SKIN DISEASES	
Included ICD-10 codes and labels	
L13	Other bullous disorders
L28	Lichen simplex chronicus and prurigo
L301	Dyshidrosis [pompholyx]
L43	Lichen planus
L508	Other urticaria
L581	Chronic radiodermatitis
L85	Other epidermal thickening
Q80	Congenital ichthyosis
Q81	Epidermolysis bullosa
Q821	Xeroderma pigmentosum
Q822	Mastocytosis
Q829	Congenital malformation of skin, unspecified
Excluded ICD-10 codes and labels	
L432	Lichenoid drug reaction
PARKINSON AND PARKINSONISM	
Included ICD-10 codes and labels	
G20	Parkinson disease
G21	Secondary parkinsonism
G22	Parkinsonism in diseases classified elsewhere
G23	Other degenerative diseases of basal ganglia
Excluded ICD-10 codes and labels	
G210	Malignant neuroleptic syndrome
PERIPHERAL NEUROPATHY	
Included ICD-10 codes and labels	
B91	Sequelae of poliomyelitis
G14	Postpolio syndrome
G54	Nerve root and plexus disorders
G55	Nerve root and plexus compressions in diseases classified elsewhere
G56	Mononeuropathies of upper limb
G57	Mononeuropathies of lower limb
G58	Other mononeuropathies
G59	Mononeuropathy in diseases classified elsewhere
G60	Hereditary and idiopathic neuropathy
G628	Other specified polyneuropathies

G629	Polyneuropathy, unspecified
G63	Polyneuropathy in diseases classified elsewhere
M472	Other spondylosis with radiculopathy
M531	Cervicobrachial syndrome
M541	Radiculopathy
Excluded ICD-10 codes and labels	
G631	Polyneuropathy in neoplastic disease
PERIPHERAL VASCULAR DISEASE	
Included ICD-10 codes and labels	
I702	Atherosclerosis of arteries of extremities
I73	Other peripheral vascular diseases
I792	Peripheral angiopathy in diseases classified elsewhere
I798	Other disorders of arteries, arterioles and capillaries in diseases classified elsewhere
Excluded ICD-10 codes and labels	
I731	Thromboangiitis obliterans [Buerger]
I738	Other specified peripheral vascular diseases
PROSTATE DISEASES	
Included ICD-10 codes and labels	
N40	Hyperplasia of prostate
N411	Chronic prostatitis
N418	Other inflammatory diseases of prostate
SCHIZOPHRENIA AND DELUSIONAL DISEASES	
Included ICD-10 codes and labels	
F20	Schizophrenia
F22	Persistent delusional disorders
F24	Induced delusional disorder
F25	Schizoaffective disorders
F28	Other nonorganic psychotic disorders
SLEEP DISORDERS	
Included ICD-10 codes and labels	
F510	Nonorganic insomnia
F511	Nonorganic hypersomnia
F512	Nonorganic disorder of the sleep-wake schedule
F513	Sleepwalking [somnambulism]
G47	Sleep disorders
SOLID NEOPLASMS	
Included ICD-10 codes and labels	
C	Malignant neoplasms
D00	Carcinoma in situ of oral cavity, oesophagus and stomach
D01	Carcinoma in situ of other and unspecified digestive organs
D02	Carcinoma in situ of middle ear and respiratory system
D03	Melanoma in situ
D04	Carcinoma in situ of skin

D05	Carcinoma in situ of breast
D06	Carcinoma in situ of cervix uteri
D07	Carcinoma in situ of other and unspecified genital organs
D09	Carcinoma in situ of other and unspecified sites
D320	Benign neoplasm: Cerebral meninges
D321	Benign neoplasm: Spinal meninges
D329	Benign neoplasm: Meninges, unspecified
D330	Benign neoplasm: Brain, supratentorial
D331	Benign neoplasm: Brain, infratentorial
D332	Benign neoplasm: Brain, unspecified
D333	Benign neoplasm: Cranial nerves
D334	Benign neoplasm: Spinal cord
Q85	Phakomatoses, not elsewhere classified
Excluded ICD-10 codes and labels	
C81	Hodgkin lymphoma
C82	Follicular lymphoma
C83	Non-follicular lymphoma
C84	Mature T/NK-cell lymphomas
C85	Other and unspecified types of non-Hodgkin lymphoma
C86	Other specified types of T/NK-cell lymphoma
C88	Malignant immunoproliferative diseases
C90	Multiple myeloma and malignant plasma cell neoplasms
C91	Lymphoid leukaemia
C92	Myeloid leukaemia
C93	Monocytic leukaemia
C94	Other leukaemias of specified cell type
C95	Leukaemia of unspecified cell type
C96	Other and unspecified malignant neoplasms of lymphoid, haematopoietic and related tissue
THYROID DISEASES	
Included ICD-10 codes and labels	
E00	Congenital iodine-deficiency syndrome
E01	Iodine-deficiency-related thyroid disorders and allied conditions
E02	Subclinical iodine-deficiency hypothyroidism
E03	Other hypothyroidism
E05	Thyrotoxicosis [hyperthyroidism]
E062	Chronic thyroiditis with transient thyrotoxicosis
E063	Autoimmune thyroiditis
E065	Other chronic thyroiditis
E07	Other disorders of thyroid
E350	Disorders of thyroid gland in diseases classified elsewhere
E890	Postprocedural hypothyroidism
Excluded ICD-10 codes and labels	
E035	Myxoedema coma

VENOUS AND LYMPHATIC DISEASES	
Included ICD-10 codes and labels	
I780	Hereditary haemorrhagic telangiectasia
I83	Varicose veins of lower extremities
I87	Other disorders of veins
I89	Other noninfective disorders of lymphatic vessels and lymph nodes
I972	Postmastectomy lymphoedema syndrome
Q820	Hereditary lymphoedema

Supplementary table 4. Percentage (**row percentages**) of participants moving from one cluster to another between baseline and 6 years.

					ROW PERCENTAGE				
	6-YEAR CLUSTERS								
	Heart and vasc.	Heart and cogn.	Neurops. and resp.	Eye	MSK, resp. and immun.	Unspecific	Death	Dropout	Total
BASELINE CLUSTERS									
Psychiatric and respiratory, n	2	7	51	10	11	18	35	25	159
%	1,26	4,4	32,08	6,29	6,92	11,32	22,01	15,72	100
Heart, n	15	63	7	4	3	4	146	35	277
%	5,42	22,74	2,53	1,44	1,08	1,44	52,71	12,64	100
Eye and Cancer, n	7	22	27	78	6	17	105	43	305
%	2,3	7,21	8,85	25,57	1,97	5,57	34,43	14,1	100
Cognitive and sensory, n	3	3	33	9	2	4	233	19	306
%	0,98	0,98	10,78	2,94	0,65	1,31	76,14	6,21	100
RESP-MSKRespiratory and MSK	12	13	60	38	109	68	91	65	456
%	2,63	2,85	13,16	8,33	23,9	14,91	19,96	14,25	100
Unspecific, n	37	44	87	88	107	647	203	215	1.428
%	2,59	3,08	6,09	6,16	7,49	45,31	14,22	15,06	100
Total, n	76	152	265	227	238	758	813	402	2.931
%	2,59	5,19	9,04	7,74	8,12	25,86	27,74	13,72	100

Supplementary table 5. Percentage (**column percentages**) of participants moving from one cluster to another between baseline and 6 years.

					COLUMN PERCENTAGE				
	6-YEAR CLUSTERS								
	Heart and vasc.	Heart and cogn.	Neurops. and resp.	Eye	MSK, resp. and immun.	Unspecific	Death	Dropout	Total
BASELINE CLUSTERS									
Psychiatric and respiratory, n	2	7	51	10	11	18	35	25	159
%	2,63	4,61	19,25	4,41	4,62	2,37	4,31	6,22	5,42
Heart, n	15	63	7	4	3	4	146	35	277
%	19,74	41,45	2,64	1,76	1,26	0,53	17,96	8,71	9,45
Eye and Cancer, n	7	22	27	78	6	17	105	43	305
%	9,21	14,47	10,19	34,36	2,52	2,24	12,92	10,7	10,41
Cognitive and sensory, n	3	3	33	9	2	4	233	19	306
%	3,95	1,97	12,45	3,96	0,84	0,53	28,66	4,73	10,44
RESP-MSKRespiratory and MSK	12	13	60	38	109	68	91	65	456
%	15,79	8,55	22,64	16,74	45,8	8,97	11,19	16,17	15,56
Unspecific, n	37	44	87	88	107	647	203	215	1.428
%	48,68	28,95	32,83	38,77	44,96	85,36	24,97	53,48	48,72
Total, n	76	152	265	227	238	758	813	402	2.931
%	100	100	100	100	100	100	100	100	100

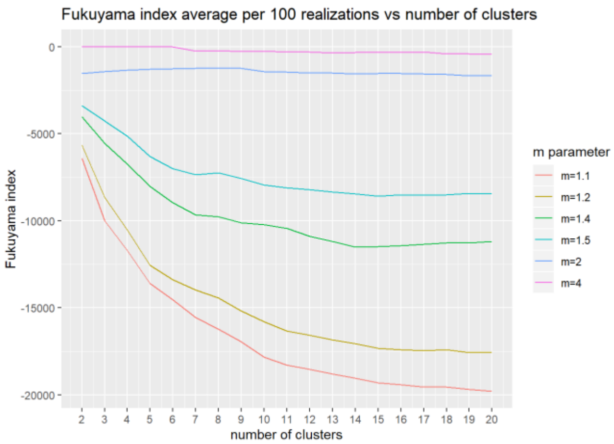
Supplementary table 6. Percentage (**row percentages**) of participants moving from one cluster to another between 6 and 12 years.

						ROW PERCENTAGE			
	12-YEAR CLUSTERS								
	Vascular	Cardiometabolic	Respiratory	Neuropsyc	Eye and MSK	Unspecific	Death	Dropout	Total
6-YEAR CLUSTERS									
Heart and vascular, n	19	4	0	1	1	2	37	12	76
%	25	5,26	0	1,32	1,32	2,63	48,68	15,79	100
Heart and cogn., n	5	31	2	4	3	0	93	14	152
%	3,29	20,39	1,32	2,63	1,97	0	61,18	9,21	100
Neurops. and resp, n	4	5	19	31	12	22	143	29	265
%	1,51	1,89	7,17	11,7	4,53	8,3	53,96	10,94	100
Eye, n	10	12	5	42	52	4	79	23	227
%	4,41	5,29	2,2	18,5	22,91	1,76	34,8	10,13	100
MSK, resp and imm, n	12	7	73	18	33	34	32	29	238
%	5,04	2,94	30,67	7,56	13,87	14,29	13,45	12,18	100
Unspecific, n	13	57	19	24	99	337	93	116	758
%	1,72	7,52	2,51	3,17	13,06	44,46	12,27	15,3	100
Total, n	63	116	118	120	200	399	477	223	1.716
%	3,67	6,76	6,88	6,99	11,66	23,25	27,8	13	100

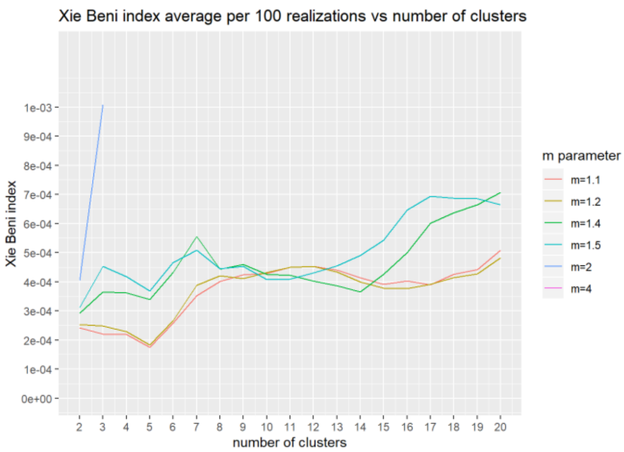
Supplementary table 7. Percentage (**column percentages**) of participants moving from one cluster to another between 6 and 12 years.

						COLUMN PERCENTAGE			
	12-YEAR CLUSTERS								
	Vascular	Cardiometabolic	Respiratory	Neuropsyc	Eye and MSK	Unspecific	Death	Dropout	Total
6-YEAR CLUSTERS									
Heart and vascular, n	19	4	0	1	1	2	37	12	76
%	30,16	3,45	0	0,83	0,5	0,5	7,76	5,38	4,43
Heart and cogn., n	5	31	2	4	3	0	93	14	152
%	7,94	26,72	1,69	3,33	1,5	0	19,5	6,28	8,86
Neurops. and resp, n	4	5	19	31	12	22	143	29	265
%	6,35	4,31	16,1	25,83	6	5,51	29,98	13	15,44
Eye, n	10	12	5	42	52	4	79	23	227
%	15,87	10,34	4,24	35	26	1	16,56	10,31	13,23
MSK, resp and imm, n	12	7	73	18	33	34	32	29	238
%	19,05	6,03	61,86	15	16,5	8,52	6,71	13	13,87
Unspecific, n	13	57	19	24	99	337	93	116	758
%	20,63	49,14	16,1	20	49,5	84,46	19,5	52,02	44,17
Total, n	63	116	118	120	200	399	477	223	1.716
%	100	100	100	100	100	100	100	100	100

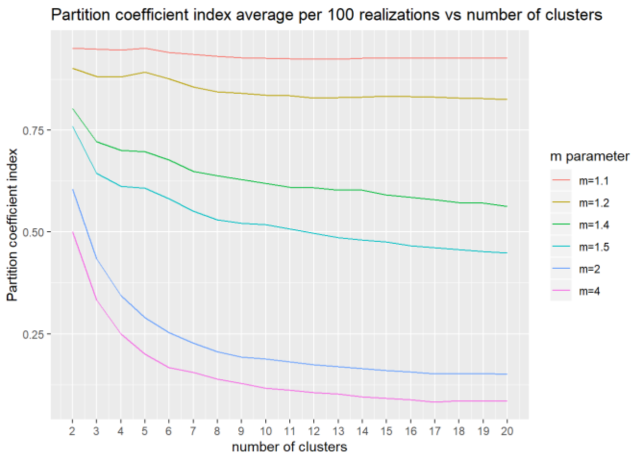
Validation indices at baseline



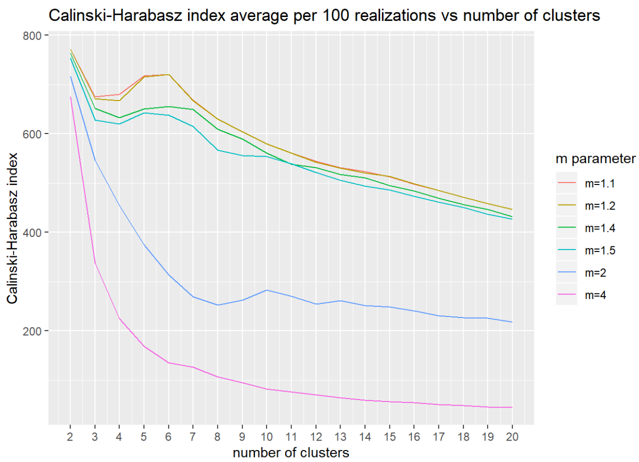
Supplementary figure 1. Fukuyama index across increasing number of clusters at baseline.



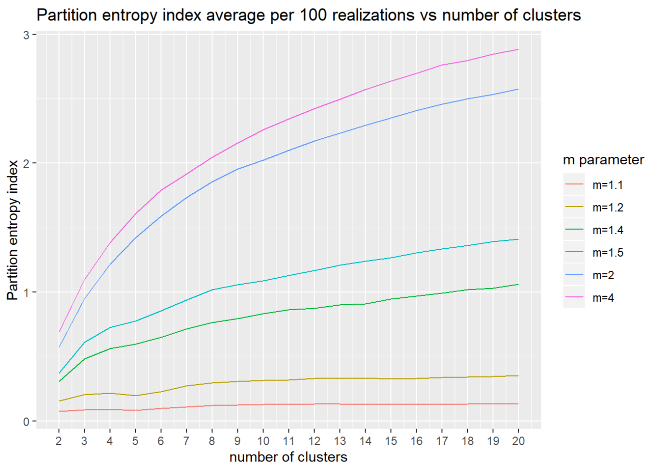
Supplementary figure 2. Xie Beni index across increasing number of clusters at baseline.



Supplementary figure 3. Partition coefficient index across increasing number of clusters at baseline.

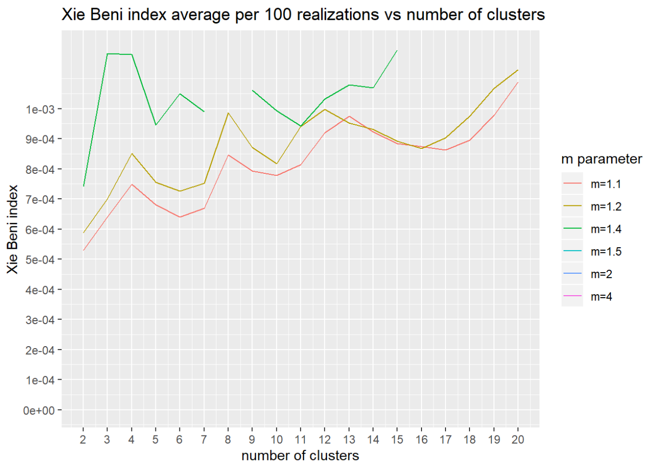


Supplementary figure 4. Calinski-Harabasz index across increasing number of clusters at baseline.

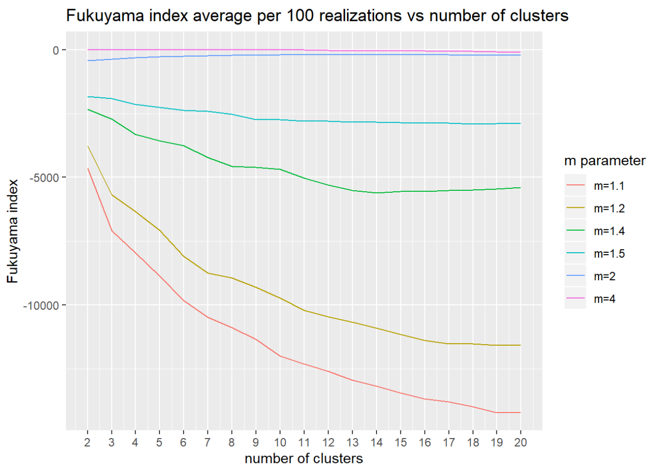


Supplementary figure 5. Partition entropy index across increasing number of clusters at baseline.

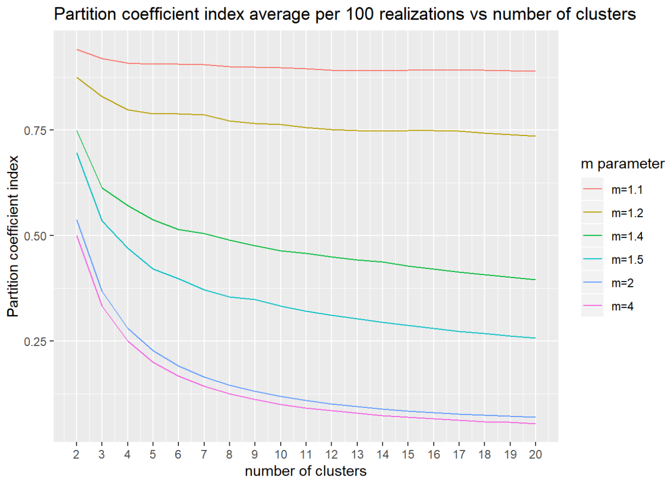
Validation indices at 6 years



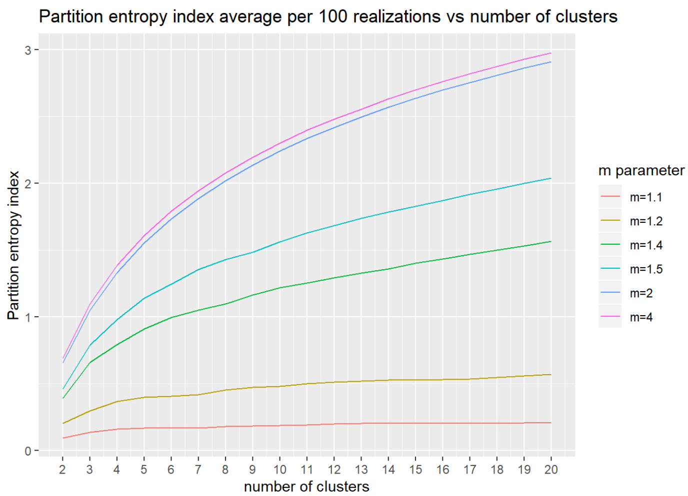
Supplementary figure 6. Xie Beni index across increasing number of clusters at 6 years.



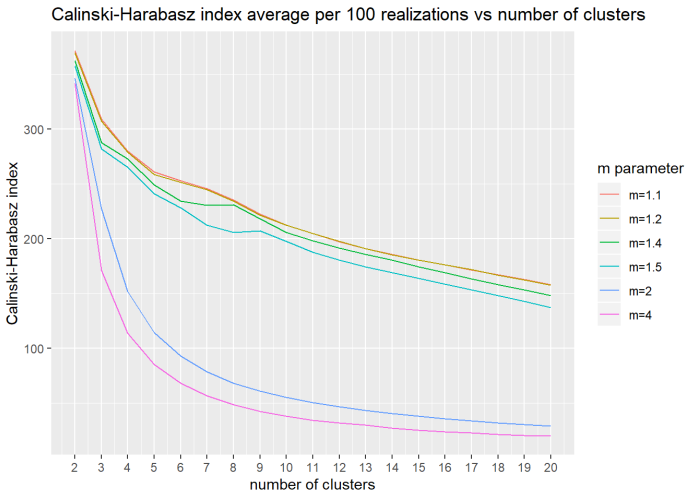
Supplementary figure 7. Fukuyama index across increasing number of clusters at 6 years.



Supplementary figure 8. Partition coefficient index across increasing number of clusters at 6 years.

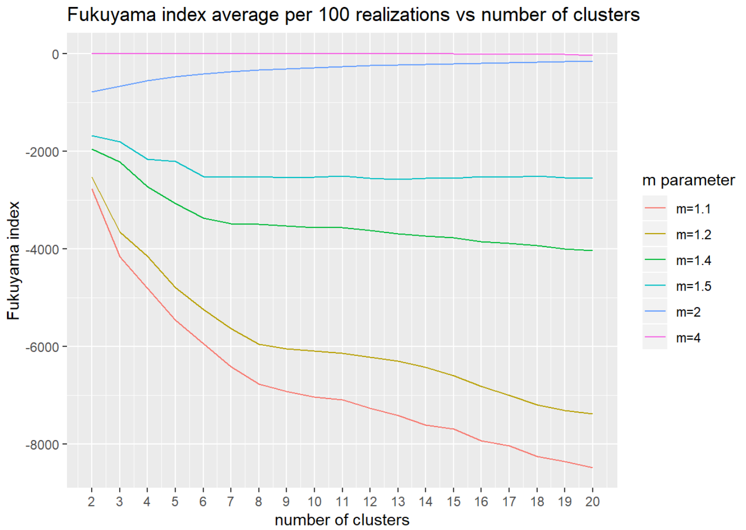


Supplementary figure 9. Partition entropy index across increasing number of clusters at 6 years.

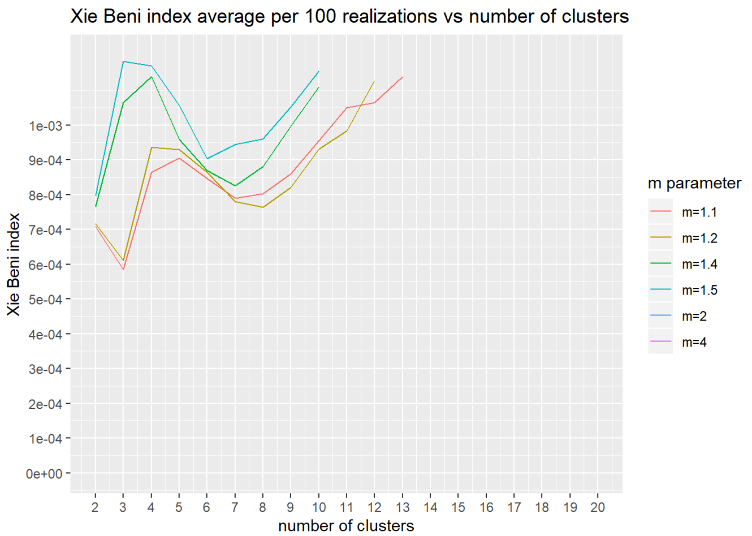


Supplementary figure 10. Calinski-Harabasz index across increasing number of clusters at 6 years.

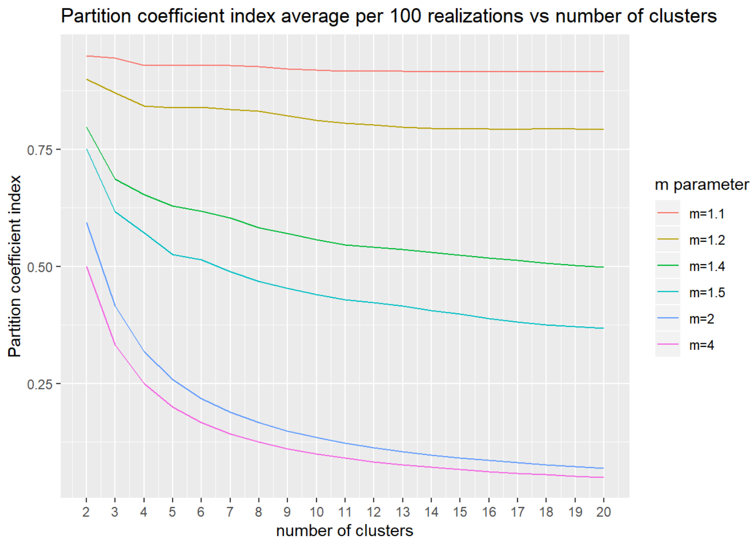
Validation indices at 12 years



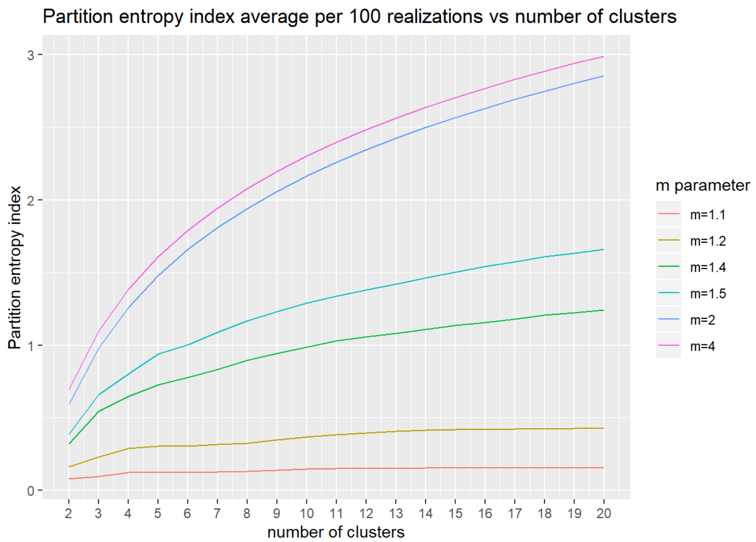
Supplementary figure 11. Fukuyama index across increasing number of clusters at 12 years.



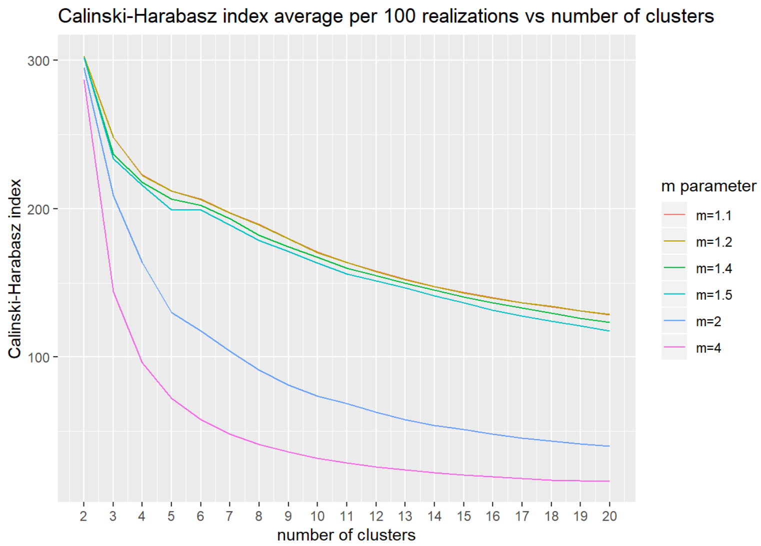
Supplementary figure 12. Xie Beni index across increasing number of clusters at 12 years.



Supplementary figure 13. Partition coefficient index across increasing number of clusters at 12 years.



Supplementary figure 14. Partition entropy index across increasing number of clusters at 12 years.



Supplementary figure 15. Calinski-Harabasz index across increasing number of clusters at 12 years.

10.3 Supplementary files Study 3

Supplementary Table 1. Disease prevalence by age group and follow-up wave.

Sexagenarians

	N_Baseline	Prev_Baseline	N_6 years	Prev_6 years	N_12 years	Prev_12 years	median
Hypertension	797	61.12	765	73.21	672	79.43	**73.21**
Dyslipidemia	678	51.99	673	64.4	600	70.92	**64.4**
Osteoarthritis and other degenerative joint diseases	129	9.89	297	28.42	419	49.53	**28.42**
Obesity	202	15.49	209	20	189	22.34	**20**
Other musculoskeletal and joint diseases	52	3.99	167	15.98	260	30.73	**15.98**
Solid neoplasms	76	5.83	157	15.02	225	26.6	**15.02**
Colitis and related diseases	76	5.83	146	13.97	200	23.64	**13.97**
Cataract and other lens diseases	17	1.3	142	13.59	342	40.43	**13.59**
Chronic kidney diseases	130	9.97	131	12.54	140	16.55	**12.54**
Depression and mood diseases	113	8.67	122	11.67	126	14.89	**11.67**
Diabetes	91	6.98	120	11.48	112	13.24	**11.48**
Other eye diseases	23	1.76	116	11.1	222	26.24	**11.1**
Dorsopathies	82	6.29	115	11	148	17.49	**11**
Thyroid diseases	103	7.9	115	11	109	12.88	**11**
Neurotic, stress-related and somatoform diseases	40	3.07	94	9	122	14.42	**9**
Ischemic heart disease	73	5.6	92	8.8	108	12.77	**8.8**
Asthma	82	6.29	90	8.61	95	11.23	**8.61**
Other genitourinary diseases	17	1.3	88	8.42	228	26.95	**8.42**
Esophagus, stomach and duodenum diseases	52	3.99	85	8.13	132	15.6	**8.13**
Anemia	45	3.45	77	7.37	107	12.65	**7.37**
Prostate diseases	32	2.45	72	6.89	112	13.24	**6.89**
COPD, emphysema, chronic bronchitis	41	3.14	67	6.41	80	9.46	**6.41**
Osteoporosis	30	2.3	66	6.32	97	11.47	**6.32**

Autoimmune diseases	34	2.61	65	6.22	80	9.46	**6.22**
Deafness, hearing impairment	24	1.84	64	6.12	143	16.9	**6.12**
Allergy	26	1.99	62	5.93	91	10.76	**5.93**
Inflammatory arthropathies	41	3.14	59	5.65	82	9.69	**5.65**
Atrial fibrillation	36	2.76	57	5.45	95	11.23	**5.45**
Sleep disorders	32	2.45	55	5.26	71	8.39	**5.26**
Cerebrovascular disease	33	2.53	54	5.17	94	11.11	**5.17**
Migraine and facial pain syndromes	38	2.91	48	4.59	59	6.97	**4.59**
Other cardiovascular diseases	23	1.76	48	4.59	62	7.33	**4.59**
Chronic pancreas, biliary tract and gallbladder diseases	23	1.76	42	4.02	45	5.32	**4.02**
Ear, nose, throat diseases	15	1.15	41	3.92	100	11.82	**3.92**
Glaucoma	18	1.38	38	3.64	56	6.62	**3.64**
Other psychiatric and behavioral diseases	20	1.53	36	3.44	50	5.91	**3.44**
Heart failure	17	1.3	34	3.25	61	7.21	**3.25**
Other metabolic diseases	18	1.38	34	3.25	46	5.44	**3.25**
Peripheral neuropathy	11	0.84	33	3.16	76	8.98	**3.16**
Venous and lymphatic diseases	11	0.84	30	2.87	62	7.33	**2.87**
Other neurological diseases	18	1.38	29	2.78	42	4.96	**2.78**
Blindness, visual impairment	10	0.77	26	2.49	41	4.85	**2.49**
Peripheral vascular disease	12	0.92	26	2.49	30	3.55	**2.49**
Cardiac valve diseases	9	0.69	25	2.39	59	6.97	**2.39**
Inflammatory bowel diseases	14	1.07	18	1.72	18	2.13	1.72
Chronic infectious diseases	6	0.46	17	1.63	19	2.25	1.63
Other digestive diseases	9	0.69	17	1.63	24	2.84	1.63
Blood and blood forming organ diseases	8	0.61	16	1.53	20	2.36	1.53
Other respiratory diseases	9	0.69	15	1.44	15	1.77	1.44
Parkinson and parkinsonism	5	0.38	15	1.44	21	2.48	1.44

Dementia	6	0.46	12	1.15	28	3.31	1.15
Hematological neoplasms	6	0.46	12	1.15	9	1.06	1.06
Chronic ulcer of the skin	5	0.38	11	1.05	16	1.89	1.05
Chronic liver diseases	4	0.31	11	1.05	8	0.95	0.95
Bradycardias and conduction diseases	5	0.38	9	0.86	12	1.42	0.86
Other skin diseases	1	0.08	8	0.77	21	2.48	0.77
Epilepsy	6	0.46	7	0.67	9	1.06	0.67
Schizophrenia and delusional diseases	6	0.46	6	0.57	2	0.24	0.46
Multiple sclerosis	2	0.15	2	0.19	1	0.12	0.15
Chromosomal abnormalities	0	0	0	0	0	0	0

Septuagenarians

	N_Baseline	Prev_Baseline	N_6 years	Prev_6 years	N_12 years	Prev_12 years	median
Hypertension	710	75.61	557	87.17	331	92.46	**87.17**
Dyslipidemia	477	50.8	401	62.75	256	71.51	**62.75**
Chronic kidney diseases	335	35.68	263	41.16	186	51.96	**41.16**
Osteoarthritis and other degenerative joint diseases	142	15.12	251	39.28	201	56.15	**39.28**
Cataract and other lens diseases	60	6.39	208	32.55	222	62.01	**32.55**
Colitis and related diseases	111	11.82	179	28.01	167	46.65	**28.01**
Solid neoplasms	115	12.25	165	25.82	156	43.58	**25.82**
Ischemic heart disease	160	17.04	148	23.16	93	25.98	**23.16**
Other eye diseases	47	5.01	139	21.75	169	47.21	**21.75**
Anemia	87	9.27	131	20.5	127	35.47	**20.5**
Other musculoskeletal and joint diseases	48	5.11	128	20.03	141	39.39	**20.03**
Deafness, hearing impairment	75	7.99	121	18.94	165	46.09	**18.94**
Atrial fibrillation	101	10.76	119	18.62	96	26.82	**18.62**

Obesity	124	13.21	115	18	71	19.83	**18**
Heart failure	83	8.84	109	17.06	99	27.65	**17.06**
Thyroid diseases	94	10.01	104	16.28	74	20.67	**16.28**
Depression and mood diseases	80	8.52	102	15.96	82	22.91	**15.96**
Osteoporosis	75	7.99	101	15.81	95	26.54	**15.81**
Cerebrovascular disease	82	8.73	99	15.49	90	25.14	**15.49**
Dementia	39	4.15	94	14.71	81	22.63	**14.71**
Diabetes	103	10.97	89	13.93	66	18.44	**13.93**
Dorsopathies	49	5.22	85	13.3	88	24.58	**13.3**
Other genitourinary diseases	31	3.3	79	12.36	109	30.45	**12.36**
Prostate diseases	59	6.28	73	11.42	47	13.13	**11.42**
Glaucoma	42	4.47	70	10.95	63	17.6	**10.95**
COPD, emphysema, chronic bronchitis	58	6.18	69	10.8	50	13.97	**10.8**
Asthma	67	7.14	64	10.02	47	13.13	**10.02**
Neurotic, stress-related and somatoform diseases	30	3.19	64	10.02	78	21.79	**10.02**
Autoimmune diseases	54	5.75	62	9.7	55	15.36	**9.7**
Esophagus, stomach and duodenum diseases	38	4.05	62	9.7	64	17.88	**9.7**
Cardiac valve diseases	35	3.73	52	8.14	58	16.2	**8.14**
Other psychiatric and behavioral diseases	18	1.92	52	8.14	40	11.17	**8.14**
Inflammatory arthropathies	40	4.26	51	7.98	44	12.29	**7.98**
Other cardiovascular diseases	33	3.51	51	7.98	55	15.36	**7.98**
Blindness, visual impairment	16	1.7	45	7.04	86	24.02	**7.04**
Peripheral neuropathy	18	1.92	36	5.63	45	12.57	**5.63**
Bradycardias and conduction diseases	21	2.24	35	5.48	25	6.98	**5.48**
Other neurological diseases	27	2.88	34	5.32	38	10.61	**5.32**
Ear, nose, throat diseases	9	0.96	31	4.85	51	14.25	**4.85**
Peripheral vascular disease	16	1.7	31	4.85	21	5.87	**4.85**

Sleep disorders	19	2.02	30	4.69	19	5.31	**4.69**
Other metabolic diseases	12	1.28	28	4.38	44	12.29	**4.38**
Allergy	19	2.02	23	3.6	27	7.54	**3.6**
Venous and lymphatic diseases	6	0.64	23	3.6	25	6.98	**3.6**
Chronic pancreas, biliary tract and gallbladder diseases	15	1.6	22	3.44	22	6.15	**3.44**
Parkinson and parkinsonism	17	1.81	21	3.29	19	5.31	**3.29**
Other respiratory diseases	16	1.7	18	2.82	13	3.63	**2.82**
Migraine and facial pain syndromes	13	1.38	14	2.19	19	5.31	**2.19**
Other digestive diseases	3	0.32	13	2.03	20	5.59	**2.03**
Chronic infectious diseases	3	0.32	12	1.88	8	2.23	1.88
Chronic ulcer of the skin	6	0.64	11	1.72	20	5.59	1.72
Inflammatory bowel diseases	9	0.96	11	1.72	10	2.79	1.72
Epilepsy	14	1.49	5	0.78	5	1.4	1.4
Blood and blood forming organ diseases	6	0.64	8	1.25	11	3.07	1.25
Hematological neoplasms	6	0.64	6	0.94	7	1.96	0.94
Other skin diseases	3	0.32	5	0.78	7	1.96	0.78
Chronic liver diseases	2	0.21	3	0.47	2	0.56	0.47
Schizophrenia and delusional diseases	5	0.53	2	0.31	1	0.28	0.31
Chromosomal abnormalities	0	0	1	0.16	1	0.28	0.16
Multiple sclerosis	2	0.21	1	0.16	0	0	0.16

Octogenarians and beyond

	N_ Baseline	Prev_ Baseline	N_ 3 years	Prev_ 3 years	N_ 6 years	Prev_ 6 years	N_ 9 years	Prev_ 9 years	N_ 12 years	Prev_ 12 years	median
Hypertension	770	68.75	545	85.16	339	90.64	198	94.29	90	95.74	**90.64**
Chronic kidney diseases	652	58.21	426	66.56	264	70.59	156	74.29	76	80.85	**70.59**

Dyslipidemia	403	35.98	307	47.97	202	54.01	125	59.52	62	65.96	**54.01**
Deafness, hearing impairment	287	25.62	260	40.62	185	49.47	130	61.9	80	85.11	**49.47**
Colitis and related diseases	238	21.25	231	36.09	178	47.59	117	55.71	61	64.89	**47.59**
Anemia	273	24.38	210	32.81	155	41.44	97	46.19	47	50	**41.44**
Cataract and other lens diseases	107	9.55	163	25.47	143	38.24	111	52.86	67	71.28	**38.24**
Heart failure	253	22.59	205	32.03	136	36.36	82	39.05	38	40.43	**36.36**
Osteoarthritis and other degenerative joint diseases	154	13.75	157	24.53	129	34.49	102	48.57	55	58.51	**34.49**
Dementia	277	24.73	186	29.06	129	34.49	71	33.81	37	39.36	**33.81**
Ischemic heart disease	281	25.09	186	29.06	125	33.42	71	33.81	30	31.91	**31.91**
Other eye diseases	97	8.66	136	21.25	115	30.75	76	36.19	48	51.06	**30.75**
Solid neoplasms	108	9.64	106	16.56	93	24.87	62	29.52	35	37.23	**24.87**
Other musculoskeletal and joint diseases	122	10.89	113	17.66	91	24.33	74	35.24	40	42.55	**24.33**
Blindness, visual impairment	118	10.54	120	18.75	83	22.19	64	30.48	38	40.43	**22.19**
Cerebrovascular disease	150	13.39	126	19.69	77	20.59	46	21.9	26	27.66	**20.59**
Atrial fibrillation	187	16.7	131	20.47	98	26.2	49	23.33	19	20.21	**20.47**
Osteoporosis	123	10.98	103	16.09	75	20.05	50	23.81	32	34.04	**20.05**
Glaucoma	129	11.52	108	16.88	71	18.98	45	21.43	23	24.47	**18.98**
Depression and mood diseases	117	10.45	110	17.19	70	18.72	48	22.86	30	31.91	**18.72**
Dorsopathies	85	7.59	74	11.56	67	17.91	52	24.76	30	31.91	**17.91**
Thyroid diseases	155	13.84	97	15.16	56	14.97	40	19.05	18	19.15	**15.16**
Diabetes	102	9.11	72	11.25	53	14.17	30	14.29	17	18.09	**14.17**
Esophagus, stomach and duodenum diseases	56	5	53	8.28	47	12.57	37	17.62	21	22.34	**12.57**
Inflammatory arthropathies	55	4.91	49	7.66	47	12.57	25	11.9	20	21.28	**11.9**
Obesity	70	6.25	66	10.31	44	11.76	32	15.24	16	17.02	**11.76**
Other genitourinary diseases	37	3.3	39	6.09	40	10.7	31	14.76	27	28.72	**10.7**
Prostate diseases	45	4.02	48	7.5	40	10.7	29	13.81	11	11.7	**10.7**
Autoimmune diseases	64	5.71	54	8.44	39	10.43	25	11.9	19	20.21	**10.43**

Neurotic, stress-related and somatoform diseases	35	3.12	44	6.88	39	10.43	39	18.57	26	27.66	**10.43**
Other cardiovascular diseases	60	5.36	50	7.81	35	9.36	25	11.9	18	19.15	**9.36**
COPD, emphysema, chronic bronchitis	68	6.07	52	8.12	36	9.63	21	10	8	8.51	**8.51**
Cardiac valve diseases	39	3.48	39	6.09	30	8.02	20	9.52	14	14.89	**8.02**
Other psychiatric and behavioral diseases	36	3.21	32	5	34	9.09	18	8.57	7	7.45	**7.45**
Asthma	56	5	39	6.09	24	6.42	13	6.19	7	7.45	**6.19**
Other neurological diseases	20	1.79	20	3.12	22	5.88	18	8.57	17	18.09	**5.88**
Chronic ulcer of the skin	19	1.7	20	3.12	20	5.35	19	9.05	12	12.77	**5.35**
Parkinson and parkinsonism	18	1.61	20	3.12	20	5.35	15	7.14	10	10.64	**5.35**
Peripheral neuropathy	20	1.79	23	3.59	19	5.08	20	9.52	13	13.83	**5.08**
Sleep disorders	19	1.7	21	3.28	17	4.55	12	5.71	5	5.32	**4.55**
Other metabolic diseases	21	1.88	21	3.28	16	4.28	10	4.76	18	19.15	**4.28**
Bradycardias and conduction diseases	36	3.21	16	2.5	14	3.74	10	4.76	5	5.32	**3.74**
Peripheral vascular disease	26	2.32	19	2.97	14	3.74	13	6.19	11	11.7	**3.74**
Venous and lymphatic diseases	9	0.8	16	2.5	14	3.74	14	6.67	8	8.51	**3.74**
Migraine and facial pain syndromes	31	2.77	16	2.5	13	3.48	9	4.29	6	6.38	**3.48**
Ear, nose, throat diseases	7	0.62	12	1.88	9	2.41	12	5.71	10	10.64	**2.41**
Other respiratory diseases	12	1.07	11	1.72	10	2.67	6	2.86	2	2.13	**2.13**
Hematological neoplasms	12	1.07	12	1.88	7	1.87	5	2.38	2	2.13	1.88
Allergy	9	0.8	7	1.09	7	1.87	8	3.81	2	2.13	1.87
Chronic pancreas, biliary tract and gallbladder diseases	15	1.34	8	1.25	6	1.6	3	1.43	2	2.13	1.43
Inflammatory bowel diseases	7	0.62	4	0.62	5	1.34	4	1.9	3	3.19	1.34
Other digestive diseases	5	0.45	8	1.25	3	0.8	4	1.9	6	6.38	1.25
Blood and blood forming organ diseases	4	0.36	5	0.78	4	1.07	6	2.86	4	4.26	1.07
Chronic infectious diseases	3	0.27	4	0.62	4	1.07	3	1.43	3	3.19	1.07
Epilepsy	10	0.89	8	1.25	6	1.6	2	0.95	1	1.06	1.06
Schizophrenia and delusional diseases	10	0.89	6	0.94	5	1.34	2	0.95	0	0	0.94

Other skin diseases	0	0	3	0.47	1	0.27	3	1.43	1	1.06	0.47
Chronic liver diseases	1	0.09	3	0.47	2	0.53	0	0	0	0	0.09
Chromosomal abnormalities	0	0	0	0	0	0	0	0	0	0	0
Multiple sclerosis	0	0	0	0	0	0	0	0	0	0	0

Supplementary Table 2. Description of multimorbidity patterns in terms of the top 10 diseases characterizing them by age group and follow-up wave.

Sexagenarians

Baseline	**Unspecific**					**Cardiovascular and anemia**					**Cardio-metabolic**					**Psychiatric-endocrine and sensorial**		
Problem	Prev	OE	Exc		Problem	Prev	OE	Exc		Problem	Prev	OE	Exc		Problem	Prev	OE	Exc
Dyslipidemia	52.42	1.01	84.81		Peripheral vascular disease	40.00	43.47	16.67		Cardiac valve diseases	10.23	14.82	100.00		Neurotic, stress-related and somatoform diseases	22.81	7.44	65.00
Hypertension	60.16	0.98	82.81		Heart failure	40.00	30.68	11.76		Other cardiovascular diseases	23.86	13.53	91.30		Blindness, visual impairment	5.26	6.86	60.00
Autoimmune diseases	2.55	0.98	82.35		Other metabolic diseases	40.00	28.98	11.11		Peripheral vascular disease	11.36	12.35	83.33		Glaucoma	8.77	6.35	55.56
Venous and lymphatic diseases	0.82	0.97	81.82		Other cardiovascular diseases	40.00	22.68	8.70		Heart failure	15.91	12.20	82.35		Other metabolic diseases	7.89	5.72	50.00
Deafness, hearing impairment	1.64	0.89	75.00		COPD, emphysema, chronic bronchitis	60.00	19.08	7.32		Atrial fibrillation	18.18	6.59	44.44		Other neurological diseases	7.89	5.72	50.00
Obesity	13.58	0.88	73.76		Inflammatory arthropathies	40.00	12.72	4.88		Ischemic heart disease	36.36	6.50	43.84		Peripheral neuropathy	4.39	5.20	45.45
Ear, nose, throat diseases	1.00	0.87	73.33		Anemia	40.00	11.59	4.44		Diabetes	37.50	5.37	36.26		Depression and mood diseases	41.23	4.76	41.59

Solid neoplasms	5.01	0.86	72.37		Colitis and related diseases	60.00	10.29	3.95		COPD, emphysema, chronic bronchitis	13.64	4.34	29.27		Other psychiatric and behavioral diseases	7.02	4.58	40.00
Other genitourinary diseases	1.09	0.84	70.59		Osteoporosis	20.00	8.69	3.33		Sleep disorders	10.23	4.17	28.13		Thyroid diseases	30.70	3.89	33.98
Osteoarthritis and other degenerative joint diseases	8.02	0.81	68.22		Chronic kidney diseases	80.00	8.02	3.08		Inflammatory arthropathies	12.50	3.98	26.83		Osteoporosis	8.77	3.81	33.33
6 years	**Unspecific**					**Cardiovascular and anemia**					**Cardio-metabolic**					**Psychiatric-endocrine and sensorial**		
Problem	Prev	OE	Exc		Problem	Prev	OE	Exc		Problem	Prev	OE	Exc		Problem	Prev	OE	Exc
Dyslipidemia	65.15	1.01	53.05		Peripheral vascular disease	34.62	13.91	69.23		Other cardiovascular diseases	22.03	4.80	54.17		Glaucoma	8.87	2.44	76.32
Hypertension	67.52	0.92	48.37		Heart failure	30.77	9.46	47.06		Heart failure	14.41	4.43	50.00		Blindness, visual impairment	5.81	2.34	73.08
Obesity	14.78	0.74	38.76		Other cardiovascular diseases	42.31	9.21	45.83		Ischemic heart disease	38.14	4.33	48.91		Other neurological diseases	6.42	2.31	72.41
Solid neoplasms	10.58	0.70	36.94		Cardiac valve diseases	19.23	8.04	40.00		Cardiac valve diseases	10.17	4.25	48.00		Ear, nose, throat diseases	8.56	2.18	68.29
Venous and lymphatic diseases	2.01	0.70	36.67		Other psychiatric and	23.08	6.70	33.33		Diabetes	43.22	3.76	42.50		Depression and mood diseases	24.46	2.10	65.57

					behavioral diseases													
Chronic pancreas, biliary tract and gallbladder diseases	2.74	0.68	35.71		Anemia	40.38	5.48	27.27		Atrial fibrillation	15.25	2.80	31.58		Osteoporosis	12.84	2.03	63.64
Chronic kidney diseases	8.03	0.64	33.59		Atrial fibrillation	28.85	5.29	26.32		Prostate diseases	18.64	2.71	30.56		Thyroid diseases	22.02	2.00	62.61
Osteoarthritis and other degenerative joint diseases	17.52	0.62	32.32		Other metabolic diseases	15.38	4.73	23.53		Sleep disorders	11.86	2.25	25.45		Other metabolic diseases	6.12	1.88	58.82
Prostate diseases	4.20	0.61	31.94		Peripheral neuropathy	13.46	4.26	21.21		Peripheral vascular disease	5.08	2.04	23.08		Allergy	11.01	1.86	58.06
Esophagus, stomach and duodenum diseases	4.93	0.61	31.76		Blindness, visual impairment	9.62	3.86	19.23		COPD, emphysema, chronic bronchitis	12.71	1.98	22.39		Neurotic, stress-related and somatoform diseases	16.51	1.84	57.45
12 years	**Unspecific**					**Cardiovascular and anemia**					**Cardio-metabolic**					**Psychiatric-endocrine and sensorial**		
Problem	Prev	OE	Exc		Problem	Prev	OE	Exc		Problem	Prev	OE	Exc		Problem	Prev	OE	Exc
Chronic pancreas, biliary tract and	5.24	0.98	22.22		Peripheral vascular disease	18.18	5.13	86.67		Ischemic heart disease	42.45	3.33	41.67		Depression and mood diseases	19.95	1.34	64.29

gallbladder diseases																		
Dyslipidemia	68.06	0.96	21.67		Other cardiovascular diseases	29.37	4.01	67.74		Heart failure	21.70	3.01	37.70		Neurotic, stress-related and somatoform diseases	18.23	1.26	60.66
Hypertension	75.92	0.96	21.58		Heart failure	25.17	3.49	59.02		Cardiac valve diseases	20.75	2.98	37.29		Thyroid diseases	16.26	1.26	60.55
Venous and lymphatic diseases	6.28	0.86	19.35		Cardiac valve diseases	20.28	2.91	49.15		Diabetes	38.68	2.92	36.61		Osteoarthritis and other degenerative joint diseases	62.07	1.25	60.14
Solid neoplasms	19.90	0.75	16.89		Glaucoma	18.88	2.85	48.21		Prostate diseases	35.85	2.71	33.93		Deafness, hearing impairment	20.44	1.21	58.04
Obesity	12.04	0.54	12.17		Other psychiatric and behavioral diseases	16.78	2.84	48.00		Other cardiovascular diseases	16.98	2.32	29.03		Migraine and facial pain syndromes	8.37	1.20	57.63
Other genitourinary diseases	14.14	0.52	11.84		Blindness, visual impairment	13.29	2.74	46.34		Atrial fibrillation	22.64	2.02	25.26		Other musculoskeletal and joint diseases	36.21	1.18	56.54
Allergy	5.24	0.49	10.99		Atrial fibrillation	30.07	2.68	45.26		Obesity	37.74	1.69	21.16		Other genitourinary diseases	31.03	1.15	55.26
Diabetes	6.28	0.47	10.71		COPD, emphysema, chronic bronchitis	23.08	2.44	41.25		Sleep disorders	13.21	1.57	19.72		Cataract and other lens diseases	45.81	1.13	54.39
Esophagus, stomach and	7.33	0.47	10.61		Anemia	30.77	2.43	41.12		COPD, emphysema,	13.21	1.40	17.50		Osteoporosis	12.56	1.10	52.58

duodenum diseases										chronic bronchitis								

Septuagenarians

Baseline	**Unspecific**					**Cardiovascular and diabetes**					**Neuro-vascular and skin-sensorial**					**Neuro-psychiatric and sensorial**		
Problem	Prev	OE	Exc		Problem	Prev	OE	Exc		Problem	Prev	OE	Exc		Problem	Prev	OE	Exc
Dyslipidemia	54.59	1.07	74.84		Other respiratory diseases	13.16	7.72	62.50		Venous and lymphatic diseases	25.00	39.13	16.67		Parkinson and parkinsonism	7.32	4.04	88.24
Sleep disorders	2.14	1.06	73.68		Peripheral vascular disease	13.16	7.72	62.50		Chronic ulcer of the skin	25.00	39.13	16.67		Allergy	7.80	3.86	84.21
Hypertension	74.92	0.99	69.01		Bradycardias and conduction diseases	17.11	7.65	61.90		Other neurological diseases	50.00	17.39	7.41		Other psychiatric and behavioral diseases	6.34	3.31	72.22
Thyroid diseases	9.79	0.98	68.09		Other cardiovascular diseases	26.32	7.49	60.61		Blindness, visual impairment	25.00	14.67	6.25		Peripheral neuropathy	6.34	3.31	72.22
Other digestive diseases	0.31	0.96	66.67		Venous and lymphatic diseases	3.95	6.18	50.00		Peripheral vascular disease	25.00	14.67	6.25		Neurotic, stress-related and somatoform diseases	9.76	3.05	66.67
Deafness, hearing impairment	7.34	0.92	64.00		Heart failure	51.32	5.81	46.99		Parkinson and parkinsonism	25.00	13.81	5.88		Depression and mood diseases	24.39	2.86	62.50
Glaucoma	3.98	0.89	61.90		Cardiac valve diseases	21.05	5.65	45.71		Allergy	25.00	12.36	5.26		Blindness, visual impairment	4.88	2.86	62.50

Chronic kidney diseases	31.35	0.88	61.19		Diabetes	48.68	4.44	35.92		Dementia	50.00	12.04	5.13		Migraine and facial pain syndromes	3.90	2.82	61.54
Solid neoplasms	10.70	0.87	60.87		Chronic ulcer of the skin	2.63	4.12	33.33		Other cardiovascular diseases	25.00	7.11	3.03		Dementia	11.71	2.82	61.54
Chronic pancreas, biliary tract and gallbladder diseases	1.38	0.86	60.00		Other digestive diseases	1.32	4.12	33.33		Asthma	50.00	7.01	2.99		Other neurological diseases	7.80	2.71	59.26
6 years	**Unspecific**					**Cardiovascular and diabetes**					**Neuro-vascular and skin-sensorial**					**Neuro-psychiatric and sensorial**		
Problem	Prev	OE	Exc		Problem	Prev	OE	Exc		Problem	Prev	OE	Exc		Problem	Prev	OE	Exc
Dyslipidemia	68.39	1.09	26.43		Bradycardias and conduction diseases	35.38	6.46	65.71		Chronic ulcer of the skin	15.09	8.77	72.73		Dementia	19.95	1.36	77.66
Hypertension	85.16	0.98	23.70		Other respiratory diseases	15.38	5.46	55.56		Parkinson and parkinsonism	20.75	6.32	52.38		Deafness, hearing impairment	24.59	1.30	74.38
Thyroid diseases	12.26	0.75	18.27		Peripheral vascular disease	20.00	4.12	41.94		Peripheral neuropathy	26.42	4.69	38.89		Ear, nose, throat diseases	6.28	1.30	74.19
Chronic pancreas, biliary tract and gallbladder diseases	2.58	0.75	18.18		Other cardiovascular diseases	32.31	4.05	41.18		Allergy	15.09	4.19	34.78		Other psychiatric and behavioral diseases	10.38	1.28	73.08

Solid neoplasms	18.71	0.72	17.58		Heart failure	64.62	3.79	38.53		Peripheral vascular disease	18.87	3.89	32.26		Migraine and facial pain syndromes	2.73	1.25	71.43
Obesity	12.90	0.72	17.39		Cardiac valve diseases	27.69	3.40	34.62		Other metabolic diseases	16.98	3.88	32.14		Glaucoma	13.66	1.25	71.43
Chronic kidney diseases	29.03	0.71	17.11		Other digestive diseases	6.15	3.02	30.77		Other cardiovascular diseases	30.19	3.78	31.37		Other genitourinary diseases	15.30	1.24	70.89
Prostate diseases	7.74	0.68	16.44		Atrial fibrillation	55.38	2.97	30.25		Venous and lymphatic diseases	13.21	3.67	30.43		Other eye diseases	26.50	1.22	69.78
Osteoarthritis and other degenerative joint diseases	25.81	0.66	15.94		Diabetes	40.00	2.87	29.21		Other neurological diseases	18.87	3.55	29.41		Neurotic, stress-related and somatoform diseases	12.02	1.20	68.75
Other genitourinary diseases	7.74	0.63	15.19		Ischemic heart disease	64.62	2.79	28.38		Blindness, visual impairment	20.75	2.95	24.44		Other musculoskeletal and joint diseases	23.77	1.19	67.97
12 years	**Unspecific**					**Cardiovascular and diabetes**					**Neuro-vascular and skin-sensorial**					**Neuro-psychiatric and sensorial**		
Problem	Prev	OE	Exc		Problem	Prev	OE	Exc		Problem	Prev	OE	Exc		Problem	Prev	OE	Exc
Thyroid diseases	23.81	1.15	6.76		Bradycardias and conduction diseases	33.33	4.77	36.00		Chronic ulcer of the skin	13.74	2.46	90.00		Thyroid diseases	21.79	1.05	52.70
Hypertension	95.24	1.03	6.04		Venous and lymphatic diseases	29.63	4.24	32.00		Parkinson and parkinsonism	12.21	2.30	84.21		Sleep disorders	5.59	1.05	52.63

Dyslipidemia	66.67	0.93	5.47		Other respiratory diseases	14.81	4.08	30.77		Peripheral vascular disease	12.21	2.08	76.19		Dyslipidemia	73.18	1.02	51.17
Osteoarthritis and other degenerative joint diseases	42.86	0.76	4.48		Peripheral vascular disease	18.52	3.16	23.81		Other cardiovascular diseases	29.77	1.94	70.91		Cataract and other lens diseases	62.57	1.01	50.45
Asthma	9.52	0.73	4.26		Diabetes	48.15	2.61	19.70		Other psychiatric and behavioral diseases	21.37	1.91	70.00		Esophagus, stomach and duodenum diseases	17.88	1.00	50.00
COPD, emphysema, chronic bronchitis	9.52	0.68	4.00		Other cardiovascular diseases	37.04	2.41	18.18		Chronic pancreas, biliary tract and gallbladder diseases	11.45	1.86	68.18		Deafness, hearing impairment	45.25	0.98	49.09
Solid neoplasms	28.57	0.66	3.85		Cardiac valve diseases	37.04	2.29	17.24		Allergy	13.74	1.82	66.67		Chronic kidney diseases	50.28	0.97	48.39
Dorsopathies	14.29	0.58	3.41		COPD, emphysema, chronic bronchitis	29.63	2.12	16.00		Other neurological diseases	19.08	1.80	65.79		Osteoarthritis and other degenerative joint diseases	53.63	0.96	47.76
Obesity	9.52	0.48	2.82		Migraine and facial pain syndromes	11.11	2.09	15.79		Other digestive diseases	9.92	1.78	65.00		Hypertension	88.27	0.95	47.73
Chronic kidney diseases	23.81	0.46	2.69		Prostate diseases	25.93	1.97	14.89		Other metabolic diseases	21.37	1.74	63.64		Glaucoma	16.76	0.95	47.62

Baseline	**Unspecific**					**Respiratory-circulatory and skin**					**Cardio-respiratory and Neurological**					**Neuro-sensorial**		
Problem	Prev	OE	Exc		Problem	Prev	OE	Exc		Problem	Prev	OE	Exc		Problem	Prev	OE	Exc
Hypertension	72.28	1.05	76.88		Venous and lymphatic diseases	42.86	53.33	66.67		Bradycardias and conduction diseases	9.80	3.05	69.44		Other digestive diseases	9.38	21.00	60.00
Dyslipidemia	37.73	1.05	76.67		Chronic ulcer of the skin	57.14	33.68	42.11		Asthma	14.90	2.98	67.86		Other neurological diseases	21.88	12.25	35.00
Obesity	6.11	0.98	71.43		Other respiratory diseases	28.57	26.67	33.33		COPD, emphysema, chronic bronchitis	18.04	2.97	67.65		Ear, nose, throat diseases	6.25	10.00	28.57
Chronic kidney diseases	56.78	0.98	71.32		Peripheral vascular disease	28.57	12.31	15.38		Other metabolic diseases	5.49	2.93	66.67		Parkinson and parkinsonism	15.63	9.72	27.78
Thyroid diseases	13.43	0.97	70.97		Other metabolic diseases	21.43	11.43	14.29		Migraine and facial pain syndromes	7.45	2.69	61.29		Peripheral vascular disease	21.88	9.42	26.92
Solid neoplasms	9.28	0.96	70.37		Inflammatory arthropathies	35.71	7.27	9.09		Other respiratory diseases	2.75	2.56	58.33		Other cardiovascular diseases	46.88	8.75	25.00
Dementia	23.57	0.95	69.68		Other cardiovascular diseases	35.71	6.67	8.33		Sleep disorders	4.31	2.54	57.89		Neurotic, stress-related and somatoform diseases	25.00	8.00	22.86
Cataract and other lens diseases	8.79	0.92	67.29		COPD, emphysema,	28.57	4.71	5.88		Chronic ulcer of the skin	3.92	2.31	52.63		Migraine and facial pain syndromes	18.75	6.77	19.35

					chronic bronchitis													
Deafness, hearing impairment	23.57	0.92	67.25		Parkinson and parkinsonism	7.14	4.44	5.56		Neurotic, stress-related and somatoform diseases	7.06	2.26	51.43		Sleep disorders	9.38	5.53	15.79
Osteoarthritis and other degenerative joint diseases	12.58	0.91	66.88		Asthma	21.43	4.29	5.36		Parkinson and parkinsonism	3.53	2.20	50.00		Peripheral neuropathy	9.38	5.25	15.00
3 years	**Unspecific**					**Respiratory-circulatory and skin**					**Cardio-respiratory and Neurological**					**Neuro-sensorial**		
Problem	Prev	OE	Exc		Problem	Prev	OE	Exc		Problem	Prev	OE	Exc		Problem	Prev	OE	Exc
Dyslipidemia	51.34	1.07	49.84		Venous and lymphatic diseases	41.67	16.67	62.50		Other psychiatric and behavioral diseases	8.21	1.64	71.88		Other neurological diseases	26.32	8.42	50.00
Hypertension	87.25	1.02	47.71		Other respiratory diseases	25.00	14.55	54.55		Blindness, visual impairment	30.36	1.62	70.83		Peripheral vascular disease	21.05	7.09	42.11
Dorsopathies	10.74	0.93	43.24		Chronic ulcer of the skin	37.50	12.00	45.00		COPD, emphysema, chronic bronchitis	12.86	1.58	69.23		Sleep disorders	21.05	6.42	38.10
Osteoarthritis and other degenerative joint diseases	22.48	0.92	42.68		Peripheral vascular disease	25.00	8.42	31.58		Depression and mood diseases	26.43	1.54	67.27		Other digestive diseases	7.89	6.32	37.50
Thyroid diseases	13.76	0.91	42.27		Asthma	41.67	6.84	25.64		Asthma	9.29	1.52	66.67		Parkinson and parkinsonism	18.42	5.89	35.00

Chronic kidney diseases	59.73	0.90	41.78		COPD, emphysema, chronic bronchitis	45.83	5.64	21.15		Other genitourinary diseases	8.93	1.47	64.10		Peripheral neuropathy	21.05	5.86	34.78
Cataract and other lens diseases	22.82	0.90	41.72		Other cardiovascular diseases	37.50	4.80	18.00		Cerebrovascular disease	28.57	1.45	63.49		Ear, nose, throat diseases	10.53	5.61	33.33
Deafness, hearing impairment	36.24	0.89	41.54		Other metabolic diseases	12.50	3.81	14.29		Bradycardias and conduction diseases	3.57	1.43	62.50		Other cardiovascular diseases	39.47	5.05	30.00
Obesity	9.06	0.88	40.91		Other genitourinary diseases	20.83	3.42	12.82		Heart failure	45.36	1.42	61.95		Migraine and facial pain syndromes	10.53	4.21	25.00
Dementia	25.50	0.88	40.86		Other digestive diseases	4.17	3.33	12.50		Other eye diseases	29.64	1.39	61.03		Other metabolic diseases	13.16	4.01	23.81
6 years	**Unspecific**					**Respiratory-circulatory and skin**					**Cardio-respiratory and Neurological**					**Neuro-sensorial**		
Problem	Prev	OE	Exc		Problem	Prev	OE	Exc		Problem	Prev	OE	Exc		Problem	Prev	OE	Exc
Dyslipidemia	60.55	1.12	32.67		Venous and lymphatic diseases	41.94	11.20	92.86		Blindness, visual impairment	31.28	1.41	73.49		Parkinson and parkinsonism	28.21	5.27	55.00
Hypertension	91.74	1.01	29.50		Peripheral vascular disease	29.03	7.76	64.29		Glaucoma	26.15	1.38	71.83		Bradycardias and conduction diseases	17.95	4.79	50.00
Osteoarthritis and other degenerative joint diseases	34.86	1.01	29.46		Chronic ulcer of the skin	35.48	6.64	55.00		Other psychiatric and behavioral diseases	11.79	1.30	67.65		Peripheral neuropathy	23.08	4.54	47.37

Deafness, hearing impairment	48.62	0.98	28.65		Other respiratory diseases	16.13	6.03	50.00		Atrial fibrillation	32.31	1.23	64.29		Sleep disorders	20.51	4.51	47.06
Chronic kidney diseases	63.30	0.90	26.14		Asthma	32.26	5.03	41.67		Depression and mood diseases	23.08	1.23	64.29		Other neurological diseases	23.08	3.92	40.91
Dementia	30.28	0.88	25.58		Other cardiovascular diseases	41.94	4.48	37.14		Cerebrovascular disease	25.13	1.22	63.64		Peripheral vascular disease	12.82	3.42	35.71
Dorsopathies	14.68	0.82	23.88		Other digestive diseases	3.23	4.02	33.33		Other eye diseases	36.41	1.18	61.74		Other digestive diseases	2.56	3.20	33.33
Obesity	9.17	0.78	22.73		Other metabolic diseases	16.13	3.77	31.25		Other musculoskeletal and joint diseases	28.72	1.18	61.54		Ear, nose, throat diseases	7.69	3.20	33.33
Solid neoplasms	19.27	0.77	22.58		COPD, emphysema, chronic bronchitis	35.48	3.69	30.56		Migraine and facial pain syndromes	4.10	1.18	61.54		Other cardiovascular diseases	28.21	3.01	31.43
Cataract and other lens diseases	28.44	0.74	21.68		Other genitourinary diseases	38.71	3.62	30.00		Colitis and related diseases	55.38	1.16	60.67		Neurotic, stress-related and somatoform diseases	28.21	2.70	28.21
9 years	**Unspecific**					**Respiratory-circulatory and skin**					**Cardio-respiratory and Neurological**					**Neuro-sensorial**		
Problem	Prev	OE	Exc		Problem	Prev	OE	Exc		Problem	Prev	OE	Exc		Problem	Prev	OE	Exc
Diabetes	15.15	1.06	16.67		Venous and lymphatic diseases	54.17	8.13	92.86		Other psychiatric and behavioral diseases	13.45	1.57	88.89		Parkinson and parkinsonism	32.35	4.53	73.33

Esophagus, stomach and duodenum diseases	18.18	1.03	16.22		Peripheral vascular disease	33.33	5.38	61.54		Glaucoma	26.89	1.25	71.11		Other neurological diseases	29.41	3.43	55.56
Dyslipidemia	60.61	1.02	16.00		Chronic ulcer of the skin	41.67	4.61	52.63		Blindness, visual impairment	36.13	1.19	67.19		Bradycardias and conduction diseases	14.71	3.09	50.00
Dementia	33.33	0.99	15.49		Other digestive diseases	8.33	4.38	50.00		Other eye diseases	42.86	1.18	67.11		Peripheral neuropathy	29.41	3.09	50.00
Hypertension	90.91	0.96	15.15		Other respiratory diseases	12.50	4.38	50.00		Migraine and facial pain syndromes	5.04	1.18	66.67		Inflammatory arthropathies	32.35	2.72	44.00
Chronic kidney diseases	69.70	0.94	14.74		Other cardiovascular diseases	45.83	3.85	44.00		Obesity	17.65	1.16	65.63		Other cardiovascular diseases	32.35	2.72	44.00
Anemia	42.42	0.92	14.43		Asthma	20.83	3.37	38.46		Cardiac valve diseases	10.92	1.15	65.00		Sleep disorders	14.71	2.57	41.67
Deafness, hearing impairment	54.55	0.88	13.85		COPD, emphysema, chronic bronchitis	29.17	2.92	33.33		Depression and mood diseases	26.05	1.14	64.58		Peripheral vascular disease	14.71	2.38	38.46
Thyroid diseases	15.15	0.80	12.50		Other neurological diseases	20.83	2.43	27.78		Solid neoplasms	33.61	1.14	64.52		Migraine and facial pain syndromes	8.82	2.06	33.33
Osteoarthritis and other degenerative joint diseases	36.36	0.75	11.76		Other genitourinary diseases	33.33	2.26	25.81		Cataract and other lens diseases	59.66	1.13	63.96		Ear, nose, throat diseases	11.76	2.06	33.33
12 years	**Unspecific**					**Respiratory-circulatory and skin**					**Cardio-respiratory and Neurological**					**Neuro-sensorial**		

Problem	Prev	OE	Exc		Problem	Prev	OE	Exc		Problem	Prev	OE	Exc		Problem	Prev	OE	Exc
Cerebrovascular disease	33.33	1.21	7.69		Venous and lymphatic diseases	53.33	6.27	100.00		Other psychiatric and behavioral diseases	12.24	1.64	85.71		Ear, nose, throat diseases	25.00	2.35	60.00
Hypertension	100.00	1.04	6.67		Other respiratory diseases	13.33	6.27	100.00		Diabetes	24.49	1.35	70.59		Parkinson and parkinsonism	25.00	2.35	60.00
Deafness, hearing impairment	83.33	0.98	6.25		Chronic ulcer of the skin	60.00	4.70	75.00		Cardiac valve diseases	18.37	1.23	64.29		Bradycardias and conduction diseases	12.50	2.35	60.00
Solid neoplasms	33.33	0.90	5.71		Peripheral vascular disease	53.33	4.56	72.73		Thyroid diseases	22.45	1.17	61.11		Peripheral neuropathy	29.17	2.11	53.85
Thyroid diseases	16.67	0.87	5.56		Other cardiovascular diseases	60.00	3.13	50.00		Glaucoma	28.57	1.17	60.87		Other neurological diseases	37.50	2.07	52.94
Dementia	33.33	0.85	5.41		Asthma	20.00	2.69	42.86		Cataract and other lens diseases	81.63	1.15	59.70		Migraine and facial pain syndromes	12.50	1.96	50.00
Autoimmune diseases	16.67	0.82	5.26		COPD, emphysema, chronic bronchitis	20.00	2.35	37.50		Other eye diseases	57.14	1.12	58.33		Other digestive diseases	12.50	1.96	50.00
Chronic kidney diseases	66.67	0.82	5.26		Heart failure	86.67	2.14	34.21		Blindness, visual impairment	44.90	1.11	57.89		Neurotic, stress-related and somatoform diseases	50.00	1.81	46.15
Other musculoskeletal and joint diseases	33.33	0.78	5.00		Atrial fibrillation	40.00	1.98	31.58		Atrial fibrillation	22.45	1.11	57.89		Prostate diseases	20.83	1.78	45.45

Dyslipidemia	50.00	0.76	4.84		Parkinson and parkinsonism	20.00	1.88	30.00		Obesity	18.37	1.08	56.25		Inflammatory arthropathies	37.50	1.76	45.00

Supplementary Table 3. Description of multimorbidity patterns in terms of sociodemographic, clinical and functional characteristics by age group and follow-up wave.

Sexagenarians

Baseline								
	UNSP N=1097	**CV & ANEMIA N=5**	**CARDIO-META N=88**	**PSY-ENDOC & SENS N=114**	p.overall	N		
Age	62.9 (2.88)	65.3 (2.75)	64.4 (2.79)	63.7 (2.90)	<0.001	1303		
Sex:					0.003	1304		
Men	477 (43.5%)	2 (40.0%)	52 (59.1%)	38 (33.3%)				
Women	620 (56.5%)	3 (60.0%)	36 (40.9%)	76 (66.7%)				
Education:					0.065	1302		
Elementary	75 (6.84%)	1 (20.0%)	9 (10.3%)	8 (7.02%)				
High school	465 (42.4%)	4 (80.0%)	46 (52.9%)	46 (40.4%)				
University	556 (50.7%)	0 (0.00%)	32 (36.8%)	60 (52.6%)				
# chronic diseases	2.23 (1.28)	9.00 (2.35)	5.52 (1.89)	5.00 (1.58)	<0.001	1304		
# drugs	2.11 (2.18)	10.0 (4.00)	6.25 (3.61)	4.94 (3.25)	<0.001	1300		
Walking speed	1.29 (0.29)	0.82 (0.58)	1.06 (0.40)	1.15 (0.32)	<0.001	1290		
MMSE	29.3 (1.43)	29.0 (1.73)	28.9 (2.09)	29.3 (0.98)	0.100	1280		
6 years								
	UNSP N=548	**CV & ANEMIA N=52**	**CARDIO-META N=118**	**PSY-ENDOC & SENS N=327**	**DROPOUT N=181**	**DEATH N=78**	p.overall	N
Age	62.5 (2.80)	64.5 (2.78)	63.8 (2.92)	63.1 (2.91)	63.1 (2.93)	64.3 (2.86)	<0.001	1303
Sex:							<0.001	1304
Men	244 (44.5%)	22 (42.3%)	76 (64.4%)	100 (30.6%)	77 (42.5%)	50 (64.1%)		
Women	304 (55.5%)	30 (57.7%)	42 (35.6%)	227 (69.4%)	104 (57.5%)	28 (35.9%)		

Education:							0.001	1302
Elementary	33 (6.02%)	1 (1.92%)	7 (5.93%)	20 (6.12%)	18 (10.1%)	14 (17.9%)		
High school	230 (42.0%)	32 (61.5%)	50 (42.4%)	133 (40.7%)	78 (43.6%)	38 (48.7%)		
University	285 (52.0%)	19 (36.5%)	61 (51.7%)	174 (53.2%)	83 (46.4%)	26 (33.3%)		
# chronic diseases	3.02 (1.35)	11.3 (2.68)	6.36 (1.95)	6.39 (1.94)	. (.)	. (.)	<0.001	1045
# drugs	2.62 (2.22)	8.87 (3.89)	6.18 (3.73)	5.10 (3.26)	. (.)	. (.)	<0.001	1028
Walking speed	1.28 (0.30)	0.93 (0.42)	1.16 (0.37)	1.12 (0.38)	. (.)	. (.)	<0.001	1040
MMSE	28.8 (1.76)	28.1 (1.84)	28.7 (1.25)	28.8 (1.31)	. (.)	. (.)	0.045	1030

12 years

	UNSP N=191	**CV & ANEMIA N=143**	**CARDIO-META N=106**	**PSY-ENDOC & SENS N=406**	**DROPOUT N=304**	**DEATH N=154**	p.overall	N
Age	62.3 (2.72)	63.8 (2.90)	63.4 (2.96)	62.7 (2.83)	63.0 (2.90)	64.2 (2.88)	<0.001	1303
Sex:							<0.001	1304
Men	85 (44.5%)	53 (37.1%)	75 (70.8%)	130 (32.0%)	137 (45.1%)	89 (57.8%)		
Women	106 (55.5%)	90 (62.9%)	31 (29.2%)	276 (68.0%)	167 (54.9%)	65 (42.2%)		
Education:							0.005	1302
Elementary	12 (6.28%)	10 (6.99%)	7 (6.60%)	16 (3.94%)	31 (10.3%)	17 (11.0%)		
High school	80 (41.9%)	61 (42.7%)	39 (36.8%)	166 (40.9%)	143 (47.4%)	72 (46.8%)		
University	99 (51.8%)	72 (50.3%)	60 (56.6%)	224 (55.2%)	128 (42.4%)	65 (42.2%)		
# chronic diseases	3.77 (1.29)	13.0 (2.83)	8.20 (2.06)	7.53 (2.10)	. (.)	. (.)	<0.001	846
# drugs	2.83 (2.07)	9.23 (5.04)	6.12 (3.43)	4.62 (3.04)	. (.)	. (.)	<0.001	844
Walking speed	1.24 (0.30)	0.89 (0.36)	1.07 (0.33)	1.08 (0.34)	. (.)	. (.)	<0.001	845
MMSE	28.7 (1.59)	28.4 (2.04)	28.0 (3.03)	28.5 (2.37)	. (.)	. (.)	0.139	840

Unspecific (USP); Cardiovascular and anemia (CV & ANEMIA); Cardio-metabolic (CARDIO-META) and Psychiatric-endocrine and sensorial (PSY-ENDOC & SENS).

Septuagenarians

Baseline								
	UNSP N=654	**CV & DIAB N=76**	**NEUROVASC & SKIN N=4**	**NEUROPSY & SENS N=205**	p.overall	N		
Age	75.1 (2.99)	75.6 (2.98)	77.0 (3.10)	75.8 (2.96)	0.025	937		
Sex:					0.007	939		
Men	231 (35.3%)	41 (53.9%)	1 (25.0%)	68 (33.2%)				
Women	423 (64.7%)	35 (46.1%)	3 (75.0%)	137 (66.8%)				
Education:					0.650	933		
Elementary	108 (16.6%)	16 (21.1%)	1 (25.0%)	25 (12.4%)				
High school	357 (54.8%)	42 (55.3%)	2 (50.0%)	113 (56.2%)				
University	187 (28.7%)	18 (23.7%)	1 (25.0%)	63 (31.3%)				
# chronic diseases	3.24 (1.44)	7.42 (2.65)	10.5 (2.65)	6.16 (1.73)	<0.001	939		
# drugs	3.30 (2.55)	7.29 (4.02)	10.0 (2.31)	6.69 (3.76)	<0.001	938		
Walking speed	1.08 (0.34)	0.78 (0.40)	0.29 (0.49)	0.83 (0.40)	<0.001	916		
MMSE	28.7 (2.02)	28.8 (1.30)	19.7 (17.0)	27.3 (5.68)	<0.001	907		
6 years								
	UNSP N=155	**CV & DIAB N=65**	**NEUROVASC & SKIN N=53**	**NEUROPSY & SENS N=366**	**DROPOUT N=124**	**DEATH N=176**	p.overall	N
age	74.3 (2.84)	75.2 (3.01)	76.0 (2.96)	75.2 (2.97)	76.1 (2.93)	75.7 (2.99)	<0.001	937
sex:							0.001	939
Men	53 (34.2%)	36 (55.4%)	21 (39.6%)	110 (30.1%)	43 (34.7%)	78 (44.3%)		
Women	102 (65.8%)	29 (44.6%)	32 (60.4%)	256 (69.9%)	81 (65.3%)	98 (55.7%)		
education:							0.084	933
Elementary	17 (11.0%)	15 (23.1%)	4 (7.55%)	59 (16.1%)	22 (18.0%)	33 (19.2%)		
High school	87 (56.1%)	28 (43.1%)	31 (58.5%)	197 (53.8%)	74 (60.7%)	97 (56.4%)		
University	51 (32.9%)	22 (33.8%)	18 (34.0%)	110 (30.1%)	26 (21.3%)	42 (24.4%)		

# chronic diseases	3.95 (1.49)	10.5 (2.98)	13.2 (2.91)	8.00 (2.30)	. (.)	. (.)	<0.001	639
# drugs	3.01 (2.20)	8.77 (3.68)	9.68 (3.79)	6.41 (3.61)	. (.)	. (.)	<0.001	638
Walking speed	1.00 (0.33)	0.73 (0.43)	0.55 (0.35)	0.75 (0.42)	. (.)	. (.)	<0.001	632
MMSE	27.7 (2.83)	27.8 (2.65)	27.3 (2.11)	26.3 (5.06)	. (.)	. (.)	0.002	605
12 years								
	UNSP N=21	**CV & DIAB N=27**	**NEUROVASC & SKIN N=131**	**NEUROPSY & SENS N=179**	**DROPOUT N=224**	**DEATH N=357**	p.overall	N
age	73.2 (2.27)	74.4 (2.84)	75.1 (2.96)	74.4 (2.85)	75.5 (3.00)	75.9 (2.94)	<0.001	937
sex:							<0.001	939
Men	8 (38.1%)	15 (55.6%)	33 (25.2%)	57 (31.8%)	66 (29.5%)	162 (45.4%)		
Women	13 (61.9%)	12 (44.4%)	98 (74.8%)	122 (68.2%)	158 (70.5%)	195 (54.6%)		
education:							0.069	933
Elementary	2 (9.52%)	8 (29.6%)	20 (15.3%)	18 (10.1%)	37 (16.7%)	65 (18.4%)		
High school	10 (47.6%)	13 (48.1%)	67 (51.1%)	99 (55.3%)	126 (56.8%)	199 (56.4%)		
University	9 (42.9%)	6 (22.2%)	44 (33.6%)	62 (34.6%)	59 (26.6%)	89 (25.2%)		
# chronic diseases	4.67 (1.39)	12.2 (2.92)	16.2 (3.88)	9.88 (2.39)	. (.)	. (.)	<0.001	358
# drugs	3.57 (2.23)	8.30 (4.81)	9.65 (5.05)	6.14 (3.28)	. (.)	. (.)	<0.001	356
Walking speed	1.01 (0.31)	0.74 (0.40)	0.47 (0.37)	0.75 (0.39)	. (.)	. (.)	<0.001	355
MMSE	28.0 (1.45)	27.0 (2.62)	24.6 (6.02)	25.7 (6.05)	. (.)	. (.)	0.029	345

Unspecific (USP); Cardiovascular and diabetes (CV & DIAB); Neuro-vascular and skin-sensorial (NEUROVASC & SKIN); and Neuro-psychiatric and sensorial (NEUROPSY & SENS).

Octogenarians and beyond

Baseline								
	UNSP N=819	**RESP-CIRCULA & SKIN N=14**	**CARDIORESP & NEURO N=255**	**NEURO-SENS N=32**	p.overall	N		

Age	87.8 (5.17)	85.8 (4.57)	88.6 (4.94)	87.0 (4.52)	0.053	1114		
Sex:					0.496	1120		
Men	192 (23.4%)	2 (14.3%)	67 (26.3%)	10 (31.2%)				
Women	627 (76.6%)	12 (85.7%)	188 (73.7%)	22 (68.8%)				
Education:					0.356	1096		
Elementary	259 (32.3%)	2 (14.3%)	76 (30.6%)	10 (31.2%)				
High school	407 (50.7%)	10 (71.4%)	141 (56.9%)	18 (56.2%)				
University	136 (17.0%)	2 (14.3%)	31 (12.5%)	4 (12.5%)				
# chronic diseases	4.58 (1.86)	10.8 (2.97)	7.67 (2.31)	8.31 (2.72)	<0.001	1120		
# drugs	4.50 (2.99)	9.36 (5.29)	7.52 (3.47)	8.47 (4.08)	<0.001	1110		
Walking speed	0.58 (0.42)	0.51 (0.34)	0.42 (0.37)	0.49 (0.42)	<0.001	1035		
MMSE	25.0 (7.26)	28.0 (2.09)	23.9 (8.32)	27.5 (2.57)	0.019	963		
3 years								
	UNSP N=298	**RESP-CIRCULA & SKIN N=24**	**CARDIORESP & NEURO N=280**	**NEURO-SENS N=38**	**DROPOUT N=102**	**DEATH N=378**	p.overall	N
Age	86.5 (4.69)	84.7 (3.38)	86.9 (4.61)	85.9 (4.02)	86.2 (4.30)	90.7 (4.97)	<0.001	1114
Sex:							0.117	1120
Men	61 (20.5%)	8 (33.3%)	77 (27.5%)	14 (36.8%)	22 (21.6%)	89 (23.5%)		
Women	237 (79.5%)	16 (66.7%)	203 (72.5%)	24 (63.2%)	80 (78.4%)	289 (76.5%)		
Education:							0.011	1096
Elementary	80 (27.2%)	5 (20.8%)	79 (28.5%)	14 (36.8%)	32 (32.7%)	137 (37.5%)		
High school	149 (50.7%)	15 (62.5%)	153 (55.2%)	19 (50.0%)	50 (51.0%)	190 (52.1%)		
University	65 (22.1%)	4 (16.7%)	45 (16.2%)	5 (13.2%)	16 (16.3%)	38 (10.4%)		
# chronic diseases	5.73 (1.75)	12.2 (3.96)	9.36 (2.45)	10.5 (3.43)	. (.)	. (.)	<0.001	640
# drugs	3.94 (2.77)	7.62 (5.68)	5.93 (3.28)	6.84 (4.24)	4.77 (3.17)	5.95 (3.57)	<0.001	1110
Walking speed	0.56 (0.43)	0.42 (0.40)	0.43 (0.41)	0.46 (0.37)	. (.)	. (.)	0.001	627

MMSE	26.8 (4.81)	28.4 (1.70)	26.5 (5.09)	27.8 (1.66)	27.0 (3.70)	20.5 (10.1)	<0.001	963
6 years								
	UNSP N=109	**RESP-CIRCULA & SKIN N=31**	**CARDIORESP & NEURO N=195**	**NEURO-SENS N=39**	**DROPOUT N=157**	**DEATH N=589**	p.overall	N
Age	85.3 (4.27)	84.0 (3.79)	86.0 (4.08)	85.3 (3.78)	86.2 (4.25)	89.9 (5.06)	<0.001	1114
Sex:							0.037	1120
Men	22 (20.2%)	10 (32.3%)	49 (25.1%)	17 (43.6%)	31 (19.7%)	142 (24.1%)		
Women	87 (79.8%)	21 (67.7%)	146 (74.9%)	22 (56.4%)	126 (80.3%)	447 (75.9%)		
Education:							0.018	1096
Elementary	30 (27.8%)	8 (25.8%)	45 (23.6%)	12 (30.8%)	49 (32.0%)	203 (35.4%)		
High school	58 (53.7%)	19 (61.3%)	102 (53.4%)	18 (46.2%)	79 (51.6%)	300 (52.3%)		
University	20 (18.5%)	4 (12.9%)	44 (23.0%)	9 (23.1%)	25 (16.3%)	71 (12.4%)		
# chronic diseases	6.49 (1.72)	14.9 (4.21)	10.3 (2.70)	11.5 (3.05)	. (.)	. (.)	<0.001	374
# drugs	4.75 (2.86)	9.45 (4.38)	8.03 (3.66)	8.55 (4.57)	. (.)	. (.)	<0.001	372
Walking speed	0.55 (0.37)	0.38 (0.35)	0.38 (0.35)	0.41 (0.36)	. (.)	. (.)	0.001	368
MMSE	23.6 (7.57)	24.9 (6.78)	24.0 (7.11)	25.8 (3.93)	. (.)	. (.)	0.388	318
9 years								
	UNSP N=33	**RESP-CIRCULA & SKIN N=24**	**CARDIORESP & NEURO N=119**	**NEURO-SENS N=34**	**DROPOUT N=171**	**DEATH N=739**	p.overall	N
Age	84.2 (3.02)	83.9 (3.34)	85.0 (3.70)	84.4 (3.49)	86.2 (4.23)	89.3 (5.12)	<0.001	1114
Sex:							0.134	1120
Men	8 (24.2%)	9 (37.5%)	25 (21.0%)	13 (38.2%)	34 (19.9%)	182 (24.6%)		
Women	25 (75.8%)	15 (62.5%)	94 (79.0%)	21 (61.8%)	137 (80.1%)	557 (75.4%)		
Education:							0.112	1096
Elementary	11 (33.3%)	5 (20.8%)	29 (24.4%)	6 (17.6%)	51 (30.5%)	245 (34.1%)		

High school	16 (48.5%)	17 (70.8%)	66 (55.5%)	18 (52.9%)	89 (53.3%)	370 (51.5%)		
University	6 (18.2%)	2 (8.33%)	24 (20.2%)	10 (29.4%)	27 (16.2%)	104 (14.5%)		
# chronic diseases	7.21 (1.67)	15.1 (4.26)	11.4 (2.95)	13.4 (2.65)	. (.)	. (.)	<0.001	210
# drugs	4.58 (2.48)	7.54 (4.40)	6.47 (3.07)	7.26 (3.95)	8.64 (5.60)	8.28 (4.26)	<0.001	372
Walking speed	0.54 (0.33)	0.34 (0.30)	0.38 (0.32)	0.39 (0.37)	. (.)	. (.)	0.061	209
MMSE	23.9 (8.24)	27.4 (1.78)	25.9 (4.77)	26.7 (3.60)	26.2 (2.17)	20.8 (8.78)	<0.001	318
12 years								
	UNSP N=6	**RESP-CIRCULA & SKIN N=15**	**CARDIORESP & NEURO N=49**	**NEURO-SENS N=24**	**DROPOUT N=191**	**DEATH N=835**	p.overall	N
Age	83.3 (3.57)	83.6 (2.80)	83.9 (2.94)	82.9 (2.52)	85.9 (4.20)	88.9 (5.10)	<0.001	1114
Sex:							0.370	1120
Men	3 (50.0%)	3 (20.0%)	10 (20.4%)	7 (29.2%)	38 (19.9%)	210 (25.1%)		
Women	3 (50.0%)	12 (80.0%)	39 (79.6%)	17 (70.8%)	153 (80.1%)	625 (74.9%)		
Education:							0.212	1096
Elementary	3 (50.0%)	3 (20.0%)	13 (26.5%)	3 (12.5%)	57 (30.5%)	268 (32.9%)		
High school	2 (33.3%)	11 (73.3%)	27 (55.1%)	13 (54.2%)	98 (52.4%)	425 (52.1%)		
University	1 (16.7%)	1 (6.67%)	9 (18.4%)	8 (33.3%)	32 (17.1%)	122 (15.0%)		
# chronic diseases	6.67 (3.27)	18.4 (4.60)	13.1 (3.33)	15.6 (3.13)	. (.)	. (.)	<0.001	94
# drugs	5.00 (5.73)	11.0 (5.74)	7.86 (3.64)	9.00 (4.16)	. (.)	. (.)	0.019	94
Walking speed	0.38 (0.37)	0.35 (0.27)	0.38 (0.36)	0.35 (0.37)	. (.)	. (.)	0.984	92
MMSE	19.8 (11.2)	22.2 (10.7)	21.4 (8.50)	23.8 (7.15)	. (.)	. (.)	0.727	80

Unspecific (USP); Respiratory-circulatory and skin (RESP-CIRCULA & SKIN); Cardio-respiratory and Neurological (CARDIORESP & NEURO); and Neuro-sensorial (NEURO-SENS).

MMSE: Mini Mental State Examination