

BMJ Open

BMJ Open is committed to open peer review. As part of this commitment we make the peer review history of every article we publish publicly available.

When an article is published we post the peer reviewers' comments and the authors' responses online. We also post the versions of the paper that were used during peer review. These are the versions that the peer review comments apply to.

The versions of the paper that follow are the versions that were submitted during the peer review process. They are not the versions of record or the final published versions. They should not be cited or distributed as the published version of this manuscript.

BMJ Open is an open access journal and the full, final, typeset and author-corrected version of record of the manuscript is available on our site with no access controls, subscription charges or pay-per-view fees (<http://bmjopen.bmj.com>).

If you have any questions on BMJ Open's open peer review process please email info.bmjopen@bmj.com

BMJ Open

Multimorbidity patterns in low-middle and high income regions: A multi-region latent class analysis using ATHLOS harmonized cohorts.

Journal:	<i>BMJ Open</i>
Manuscript ID	bmjopen-2019-034441
Article Type:	Original research
Date Submitted by the Author:	19-Sep-2019
Complete List of Authors:	<p>Bayes-Marin, Iveta; Parc Sanitari Sant Joan de Déu, Research, Innovation and Teaching Unit; Centro de Investigación Biomédica en Red de Salud Mental, CIBERSAM</p> <p>Sanchez-Niubo, Albert ; Parc Sanitari Sant Joan de Déu, Research, Innovation and Teaching Unit; Centro de Investigación Biomédica en Red de Salud Mental, CIBERSAM</p> <p>Egea-Cortés, Laia; Parc Sanitari Sant Joan de Déu, Research, Innovation and Teaching Unit</p> <p>Nguyen, Hai; King's College London, Health Service and Population Research Department</p> <p>Prina, A. Matthew; King's College London,</p> <p>Haro, Josep; Parc Sanitari Sant Joan de Deu, CIBERSAM, University of Barcelona; King Sand University</p> <p>Olaya, Beatriz; Parc Sanitari Sant Joan de Deu, Research and Development Unit</p>
Keywords:	multimorbidity, non-communicable diseases (NCDs), low- and middle-income countries (LMICs), high-income country (HICs), Latent class analysis (LCA)

SCHOLARONE™
Manuscripts



I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our [licence](#).

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which [Creative Commons](#) licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

1
2
3 **Original Research Article**
4
5

6
7 **Manuscript title:**
8

9
10 Multimorbidity patterns in low-middle and high income regions: A multi-region latent class analysis
11
12 using ATHLOS harmonized cohorts.
13
14

15
16 **Author's full names and affiliations:**
17

18 Ivet Bayes-Marin, ^{a,b} Albert Sanchez-Niubo, ^{a,b} Laia Egea-Cortés, ^a Hai Nguyen, ^c Matthew Prina, ^c Josep
19
20 Maria Haro, ^{a,b} and Beatriz Olaya ^{a,b}
21
22

23
24
25 From the ^aResearch, Innovation and Teaching Unit, Parc Sanitari Sant Joan de Déu, Sant Boi de
26
27 Llobregat, Spain, ^bCentro de Investigación Biomédica en Red de Salud Mental, CIBERSAM, Madrid,
28
29 Spain, ^cHealth Service and Population Research Department, Institute of Psychiatry, Psychology and
30
31 Neuroscience, King's College London, London, United Kingdom
32
33

34
35
36 **Corresponding author details:**
37

38
39 Correspondence: Albert Sanchez- Niubo. Research, Innovation and Teaching Unit, Parc Sanitari Sant
40
41 Joan de Déu, Dr Antoni Pujadas 42, Sant Boi de Llobregat, Barcelona 08830, Spain. E-mail:
42
43 albert.sanchez@pssjd.org
44
45

46
47 **ORCID of the authors:**
48

49 Ivet Bayes-Marin - 0000-0002-3816-5244
50

51 Albert Sanchez-Niubo - 0000-0003-0309-181X
52

53 Laia Egea-Cortés - 0000-0003-4002-8254
54

55 Hai Nguyen - 0000-0003-2171-1955
56

57 Matthew Prina - 0000-0001-6698-3263
58
59
60

1
2
3 Josep Maria Haro - 0000-0002-3984-277X
4

5 Beatriz Olaya - 0000-0003-2046-3929
6
7
8
9

10 **Word count (excluding title page, abstract, references, figures and tables): 4108**
11
12
13

14 **Abstract**

15
16 **Objectives:** Our aim was to determine clusters of non-communicable diseases (NCDs) in a very large,
17 population-based sample of middle-aged and older adults from low- and middle-income and high-
18 income regions. Additionally, we explored the associations with several covariates.
19
20
21

22
23 **Design:** The total sample was 72 140 people aged 50+ from three population-based studies (ELSA,
24 SHARE and SAGE) included in the ATHLOS project and representing eight regions with high and low
25 income countries. Variables were previously harmonized using an ex-post strategy. Eight NCDs were
26 used in latent class analysis. Multinomial models were made to calculate associations with
27 covariates. All the analyses were stratified by age (50-64 and 65+ years old).
28
29
30

31
32 **Results:** Three clusters were identified: “cardio-metabolic” (8.93% in participants aged 50-64 and
33 27.22% in those aged 65+), “respiratory-mental-articular” (3.91% and 5.27%) and “healthy” (87.16%
34 and 67.51%). In the youngest group, Russia presented the highest prevalence of the “cardio-
35 metabolic” group (18.8%) and England the “respiratory-mental-articular” (5.1%). In the older group,
36 Russia had the highest proportion of both classes (48.3% and 9%). Both the youngest and older
37 African participants presented the highest proportion of the “healthy” class. Older age, being female,
38 widowed, and with low levels of education and income were related to an increased risk of
39 multimorbidity. Physical activity was a protective factor in both age groups and smoking a risk factor
40 for the “respiratory-mental-articular”.
41
42
43
44
45
46
47
48
49
50
51
52
53

54 **Conclusions:** Multimorbidity is common worldwide, especially in HICs and Russia. Health policies in
55 each country addressing coordination and support are needed to face the complexity of a pattern of
56 growing multimorbidity.
57
58
59
60

1
2
3
4
5 **Keywords:** multimorbidity, non-communicable diseases (NCDs), low- and middle- income countries
6
7 (LMICs), high-income countries (HICs), latent class analysis (LCA).
8
9

10 11 12 **Strengths and limitations of this study**

- 13
14
15 • This study used a large, harmonized, multi-regional database, which allowed us to compare
16
17 two age groups (50-64 and 65 or older) as well as disease prevalence and clusters of
18
19 conditions in regions with differing incomes.
- 20
21
22 • The presence or absence of the non-communicable diseases was based on self-reported
23
24 measures, and thus might be affected by measurement errors or lack of accuracy.
- 25
26
27 • Only diseases that were common across studies were included in the analyses. This might
28
29 have led to a smaller number of latent classes, or to different patterns of multimorbidity.
- 30
31
32 • When performing Latent Class Analysis, the 3-class solution was forced as we aimed to do
33
34 comparisons among age subsamples and regions in terms of disease prevalence as well as
35
36 protective and risk factors. Invariance analysis should have been performed.
- 37
38 • The use of multiple imputations for missing data in the covariates could carry some bias.
39

40 41 **Background**

42
43 By 2050, the population aged 60 years and older is expected to reach 2 billion worldwide, compared
44
45 to 900 million in 2015 [1]. Along with this rapid increase, the incidence of chronic conditions or non-
46
47 communicable diseases (NCDs) is also on the rise, having become the leading cause of morbidity and
48
49 disability worldwide [2].

50
51
52 Multimorbidity, defined as the co-existence of two or more chronic conditions, is more common in
53
54 older adults and is often more prevalent in people of lower socioeconomic status [3]. Multimorbidity
55
56 is thought to account for 65% of total health care expenses in high-income countries (HICs) because
57
58 of the huge associated healthcare utilization [4]. Due to the increasing prevalence of multimorbidity,
59
60

1
2
3 the managing of multiple conditions has become an unavoidable international research priority,
4 because of the high impact on the quality of life of patients and caregivers and on healthcare systems
5
6
7 [3].
8

9
10 Most studies on the prevalence of multimorbidity in older people come from HICs, while data from
11 middle-aged adults and LMICs are much more limited [5–8]. LMICs are experiencing an increase in
12 life expectancy that, together with changes in lifestyle and environment exposures, are triggering
13 changes in their disease burden profile [3,9]. Few studies have compared patterns of multimorbidity
14 between HICs and LMICs. Afshar et al. [10] used population-based chronic disease data from the
15 World Health Survey to compare multimorbidity prevalence across 27 LMICs and one HIC, and used
16 gross domestic product (GDP) to study inter-country socioeconomic differences. They found high
17 multimorbidity prevalence in all the studied countries, and a positive but non-linear relationship
18 between country GDP and multimorbidity prevalence, suggesting the influence of other factors, such
19 as lifestyles, social conditions, and differences among health systems.
20
21
22
23
24
25
26
27
28
29
30
31

32 The lack of study of differences in multimorbidity between HICs and LMICs may be due to the use of
33 different methodologies, which might hinder comparisons of prevalence and multimorbidity patterns
34 across countries. The integration of data from different studies would allow us to determine
35 differences across regions and cohorts, as well as to explore risk and protective factors involved in
36 the clustering of chronic conditions, thereby improving our understanding of the problem and the
37 creation of adapted medical guidelines.
38
39
40
41
42
43
44

45 This study aimed to: a) identify multimorbidity clusters in middle-aged (50-64) and older adults (+65)
46 from different regions, classified as LMICs and HICs; b) investigate the associations between
47 multimorbidity clusters and sociodemographic, economic, lifestyles and health status variables; and
48
49
50
51
52
53 c) explore differences across regions.
54
55
56
57
58
59
60

Methods

Study design and data extraction

The present study used data from the Ageing Trajectories of Health: Longitudinal Opportunities and Synergies (ATHLOS) project [11]. Longitudinal data from 17 international cohort studies related to health and ageing were harmonized with the aim of obtaining an integrated dataset and achieving a better understanding of ageing and health processes.

We selected three studies due to their inclusion of the variables of interest and the possibility of comparing HICs and LMICs. Baseline samples of the following studies were included in the analyses: the World Health Organization's Study on Global Ageing and Adult Health (SAGE) [12], the English Longitudinal Study of Ageing (ELSA)[13], and the Survey of Health, Ageing and Retirement in Europe Study (SHARE) [14]. These panel studies included non-institutionalized people aged 50 years and older. SAGE comprises six LMICs according to The World Bank Classification [15], namely Ghana, South Africa, Mexico, India, China and Russia; ELSA includes the English population and SHARE covers eleven countries of the European Union and Israel at baseline, considered as HICs [15].

The analyses presented focused on people aged 50 years of age and older who were part of the core sample of each study and who completed a non-proxy interview at baseline. Mexico was excluded from the analyses due the high percentage of missingness (62.08%) in the variables of interest.

Variables

The following variables were the result of a stringent, ex-post harmonization process using systematic harmonization methodology and tools from Maelstrom Research [16].

Eight NCDs were used to conduct the analysis, including those that were available in the three studies: diabetes, hypertension, asthma, chronic lung disease, joint disorders (arthritis, rheumatism or osteoarthritis), angina or myocardial infarction, stroke, and depression. The presence or absence of these conditions was self-reported and based on a medical diagnosis. Depression was assessed with standardized tools, such as the Composite International Diagnostic Interview (CIDI) in the SAGE

1
2
3 study, the Center for Epidemiologic Studies Depression Scale (CES-D) in ELSA, and the EURO-D in
4 SHARE [17–19]. A dichotomous variable (yes/no) was created using the proper cut-off score for each
5
6 tool and population.
7
8

9
10 Self-reported demographic variables included age, gender, level of education (primary or less,
11 secondary, and tertiary), marital status (single, married or currently cohabiting, separated or
12 divorced, and widowed) and quintiles of household wealth (first quintile indicating lowest level). Life-
13
14 styles and health behaviors were ‘ever smoked’ any type of tobacco and physical activity referring to
15 the practice of vigorous exercise during the last two weeks, both coded as *yes* or *no*. Other health
16
17 related variables were self-rated health (good, moderate, or poor), presence or absence of loneliness
18
19 feelings in the last week, difficulties in activities of daily living (ADL), cognitive performance, and
20
21 number of diseases.
22
23

24
25 To assess difficulties in ADL, we used a set of daily self-care activities such as problems in using the
26
27 toilet, bathing or showering, getting dressed, eating, moving, or getting in or out of bed. Each of the
28
29 ADL difficulties was coded into a *yes/no* if the person answered ‘severe’ or ‘extreme/cannot do it’. To
30
31 build the set of ADL difficulties, we coded *yes* if the person reported at least one difficulty in any of
32
33 the five items.
34
35

36
37 Immediate and delayed recall was assessed using the 10-word learning list task, and verbal fluency
38
39 utilizing the animal naming test [20]. Continuous total scores were used to perform the analyses.
40
41

42
43 Number of diseases was built by adding up the occurrences of all the above-mentioned NCDs.
44

45
46 Finally, a 7-level regional membership variable was created in order to analyze regional differences,
47
48 based on the World Health Organization (WHO) and the United Nations Statistical Division (UNSD)
49
50 regional classification [21,22]. Moreover, the World Bank Classification was used to classify these
51
52 regions into HICs or LMICs [15]. SAGE includes Africa (Ghana and South Africa), China and India, all of
53
54 them considered as LMICs. SHARE countries were grouped into three regions: Northern Europe
55
56 (Denmark, Sweden), Southern Europe (Greece, Italy, and Spain) and Western Europe (Austria,
57
58 Belgium, France, Germany, Israel, Netherlands, and Switzerland). ELSA and SHARE regions were
59
60

1
2
3 considered as HICs. Ghana and South Africa were grouped together and named as Africa for practical
4 purposes as well as due to their smaller sample size. These countries are not necessarily
5 representative of the whole continent.
6
7
8
9

10 11 12 *Statistical analysis*

13
14 Descriptive statistics were used to summarize information regarding sociodemographic economic
15 variables and disease prevalence among regions. Confidence intervals (IC95%) were calculated for
16 categorical variables in order to make comparisons across regions.
17
18

19
20 Latent Class Analysis (LCA) was conducted stratified by age (50-64, +65). Eight NCDs (diabetes,
21 hypertension, asthma, chronic lung disease, joint disorders, angina-myocardial infarction, stroke, and
22 depression) were used as observed indicators. Region was used as cluster when conducting LCA in
23 order to accurately describe disease proportions, indicating that the subjects were not independent
24 random draws, but rather were nested within clusters [23].
25
26
27
28
29

30
31 The optimal number of latent classes was determined using the adjusted Bayesian Information
32 Criterion (aBIC), the consistent Akaike Information Criterion (CAIC), and the Entropy Index. Lower
33 values of aBIC and CAIC indicate better fit, whereas entropy index values higher than 0.80 indicate
34 that the latent classes are highly discriminating [24]. The average posterior probability indicates how
35 well a model classifies individuals into their most likely class. Values higher than 0.70 indicate well-
36 identified classes [25]. Additionally, interpretability and clinical judgment were used.
37
38
39
40
41
42
43
44

45
46 Missing data in one of the indicators was handled with the full-information maximum likelihood
47 (FIML) technique, assuming missing-at-random (MAR) [26]. Missing data in the covariates was
48 handled using multiple imputation by chained equations (MICE) assuming missing-at-random (MAR)
49 [26]. The imputation model included the outcome (group membership in one of the latent classes)
50 and all the variables used in the regression models (See Additional file 1: Table S1).
51
52
53
54

55
56 Multinomial logistic regression models were used to assess the association of each multimorbidity
57 class with several outcomes adjusted for gender, age, marital status, education level, wealth, and
58
59
60

1
2
3 region. Regression models were conducted separately in one hundred imputed datasets and results
4
5 combined using Rubin's rules [27].
6

7 All analyses were conducted with Stata SE version 13.1 (College Station, TX). LCA analyses were
8
9 performed using a Stata plugin [23].
10
11
12
13

14 **Results**

15 *Descriptive analysis*

16
17 People with missing values in sex and age were excluded, resulting in a final sample of 72 140
18
19 individuals aged 50 years old (Table 1). The mean age ranged from 62 (SD=9.02) in Southern Asia to
20
21 65 years (10.18) in Russia and (10.26) in England. Some 54% were females, 72% were married or
22
23 cohabitating, and 39% had secondary education. Russia presented the highest number of conditions
24
25 (mean 1.66) compared to Africa (0.64), China (0.80), and India (0.72).
26
27

28
29 The most prevalent conditions in the total sample were hypertension (31.2%, 95%CI = 30.9-31.6) and
30
31 joint disorders (22.4%, 95%CI = 22.0-22.7).
32
33

34
35 Hypertension was particularly high in Russia (56.5%, 95%CI = 54.9-58.1) compared to the other
36
37 regions. Diabetes prevalence was greater in Southern (11.9%, 95%CI = 11.2-12.7) and Western
38
39 Europe (10.4%, 95%CI = 9.9-10.9), whereas Africa and China presented the lowest proportions.
40
41 Similarly, myocardial infarction-angina was highly prevalent in Russia (33.1%, 95%CI = 31.6-34.6),
42
43 followed by countries of Northern (13.8%, 95%CI = 12.8-14.8), Southern (11.7%, 95%CI = 11.0-12.4)
44
45 and Western Europe (13.1%, 95%CI =12.6-13.6); England, Africa, and India presented the lowest
46
47 proportion of diabetes.
48
49
50
51
52
53
54
55
56
57
58
59
60

Table 1 Main characteristics of the total sample and by regions

Region	Total	Africa ^a	China ^b	India ^b	Russia ^b	England ^c	Northern Europe ^d	Southern Europe ^e	Western Europe ^f
	N=72140	n=7950	n=12840	n=6558	n=3887	n=11 517	n=4573	n=7465	n=17 350
Age, mean (SD)	64.05 (9.96)	63.60 (10.26)	63.07 (9.31)	61.80 (9.02)	65.06 (10.18)	65.06 (10.26)	64.78 (10.34)	65.02 (10.17)	64.34 (9.96)
Female, % (IC 95%)	54.0 (53.6-54.4)	52.3 (51.2-53.4)	52.9 (52.1-54.8)	49.6 (48.4-50.8)	64.6 (63.0-66.1)	54.6 (53.7-55.5)	53.9 (51.8-54.7)	55.4 (54.2-56.5)	54.1 (53.4-54.9)
Marital status, % (IC 95%)									
Single	4.1 (4.0-4.3)	6.7 (6.2-7.3)	0.9 (0.7-1.6)	1.0 (0.8-1.2)	2.8 (2.3-3.4)	5.0 (4.6-5.4)	5.0 (4.8-6.2)	6.3 (5.8-6.9)	4.9 (4.6-5.3)
Married	71.5 (71.1-72.0)	54.5 (53.4-55.6)	83.5 (82.8-84.1)	74.1 (73.1-75.2)	56.1 (54.5-57.7)	69.1 (68.2-69.9)	72.0 (71.0-73.7)	73.3 (72.3-74.3)	73.3 (72.7-74.0)
Divorced	5.9 (5.7-6.2)	10.4 (9.7-11.1)	1.7 (1.5-2.0)	0.6 (0.5-0.9)	8.3 (7.4-9.2)	9.0 (8.5-9.5)	9.0 (8.8-10.5)	2.6 (2.3-3.0)	6.8 (6.5-7.2)
Widowed	18.4 (18.1-18.6)	27.3 (26.3-28.3)	13.9 (13.3-14.6)	24.3 (23.2-25.3)	32.7 (31.2-34.2)	16.9 (16.3-17.6)	12.0 (11.6-13.5)	17.8 (16.9-18.6)	14.9 (14.4-15.4)
Education level % (IC 95%)									
Primary or less	33.4 (33.0-33.7)	30.5 (29.5-31.5)	37.9 (37.1-38.8)	25.5 (24.5-26.6)	9.9 (9.0-10.9)	42.4 (41.5-43.4)	29.0 (28.3-30.9)	59.4 (58.3-60.5)	23.2 (22.6-23.8)
Secondary	38.7 (38.4-39.1)	19.0 (18.2-19.9)	33.2 (32.4-34.0)	18.2 (17.3-19.2)	69.3 (67.9-70.8)	37.5 (36.7-38.4)	45.0 (43.6-46.5)	31.3 (30.2-32.3)	55.1 (54.3-55.8)
Tertiary	12.0 (11.9-12.2)	3.9 (3.5-4.3)	4.7 (4.3-5.0)	5.0 (4.4-5.5)	19.7 (18.5-21.0)	11.1 (10.6-11.7)	24.0 (23.0-25.5)	9.1 (8.4-9.7)	20.8 (20.2-21.4)
Wealth (quintiles% (IC 95%))									
1 st (worse)	19.3 (19.1-19.6)	19.3 (18.4-20.2)	19.8 (19.1-20.5)	16.2 (15.3-17.1)	18.2 (17.0-19.5)	19.0 (18.3-19.7)	20.0 (19.6-21.9)	20.3 (19.4-21.2)	19.9 (19.3-20.5)
2 nd	19.7 (19.4-20.0)	19.7 (18.4-20.2)	19.7 (19.0-20.4)	18.6 (17.6-19.5)	19.8 (18.5-21.1)	19.3 (18.6-20.1)	20.0 (19.3-21.6)	20.4 (19.5-21.3)	19.9 (19.4-20.6)
3 rd	19.7 (19.4-19.9)	19.8 (18.9-20.7)	20.1 (19.4-20.8)	18.4 (17.5-19.4)	20.3 (19.0-21.6)	19.7 (18.9-20.4)	20.0 (19.0-21.3)	19.8 (18.9-20.7)	19.5 (18.9-20.1)
4 th	19.9 (19.7-20.3)	20.5 (19.6-21.4)	20.5 (19.8-21.2)	21.5 (20.5-22.5)	20.0 (18.8-21.3)	19.6 (18.9-20.3)	19.0 (18.5-20.8)	19.4 (18.6-20.4)	19.3 (18.8-19.9)
5 th (best)	20.1 (19.8-20.4)	20.4 (19.5-21.3)	19.7 (19.0-20.4)	24.8 (23.8-25.9)	21.6 (20.3-22.9)	19.6 (18.9-20.3)	18.0 (17.8-20.1)	18.9 (18.0-19.8)	19.4 (18.8-20.0)
N^o diseases, mean (SD)	1.02 (1.14)	0.64 (0.94)	0.80 (0.99)	0.72 (0.97)	1.66 (1.38)	1.19 (1.13)	1.02 (1.10)	1.28 (1.25)	1.10 (1.16)
Diseases, % (IC 95%)									
Diabetes	8.5 (8.3-8.7)	6.6 (6.1-7.2)	6.5 (6.1-7.0)	7.3 (6.7-7.9)	9.0 (8.1-10.0)	7.4 (6.9-7.9)	8.0 (7.4-9.0)	11.9 (11.2-12.7)	10.4 (9.9-10.9)
Hypertension	31.2 (30.9-31.6)	21.5 (20.6-22.4)	27.4 (26.6-28.2)	17.5 (16.6-18.5)	56.5 (54.9-58.1)	37.8 (36.9-38.7)	29.0 (28.0-30.7)	35.6 (34.5-36.7)	32.3 (31.6-33.0)
Joint disorders	22.4 (22.0-22.7)	17.8 (16.9-18.6)	22.1 (21.4-22.8)	17.9 (17.0-18.9)	35.2 (33.7-36.7)	32.5 (31.6-33.3)	15.7 (14.6-16.8)	26.3 (25.3-27.3)	16.8 (16.3-17.4)
Asthma	5.5 (5.3-5.6)	4.1 (3.7-4.5)	2.4 (2.2-2.7)	6.9 (6.3-7.6)	3.4 (2.9-4.0)	11.7 (11.1-12.3)	7.8 (6.8-8.4)	4.1 (3.7-4.6)	4.1 (3.8-4.4)
Chronic lung disease	6.1 (5.9-6.3)	1.5 (1.2-1.7)	8.6 (8.1-9.1)	4.1 (3.6-4.6)	17.9 (16.8-19.2)	6.5 (6.1-7.0)	4.0 (3.9-5.2)	5.4 (4.9-6.0)	4.9 (4.6-5.3)
MI - Angina	10.0 (9.8-10.3)	4.5 (4.1-5.0)	8.8 (8.3-9.3)	4.9 (4.4-5.5)	33.1 (31.6-34.6)	3.3 (3.0-3.6)	13.4 (12.8-14.8)	11.7 (11.0-12.4)	13.1 (12.6-13.6)
Stroke	3.8 (3.7-3.9)	3.2 (2.8-3.6)	3.5 (3.2-3.8)	2.2 (1.9-2.6)	6.0 (5.3-6.8)	4.5 (4.1-4.9)	5.0 (4.4-5.7)	3.1 (2.8-3.6)	3.9 (3.7-4.2)
Depression	15.3 (15.0-15.5)	5.8 (5.3-6.3)	1.2 (1.1-1.4)	12.1 (11.3-12.9)	5.2 (4.5-5.9)	16.5 (15.8-17.2)	18.0 (17.7-20.0)	31.7 (30.6-32.7)	25.0 (24.3-25.6)

Note. ^aSAGE study - Africa: Ghana, South Africa; ^bSAGE study; ^cELSA study- England; ^dSHARE study - Northern Europe: Denmark, Sweden; ^eSHARE study - Southern Europe: Greece, Italy, Spain; ^fSHARE study - Western Europe: Austria, Belgium, France, Germany, Israel, Netherlands, Switzerland.
The analyses were performed before multiple imputation procedure

http://bmjopen.bmj.com/ on April 17, 2024 by guest Protected by copyright.

1
2
3 Joint disorders were more prevalent in Russia (35.2%, 95%CI = 33.7-36.7) and England (32.5%, 95%CI
4 =31.6-33.3). The prevalence of asthma was greater in England than other regions (11.7%, 95%CI =
5 11.1-12.3) and chronic lung disease was greater in Russia (17.9%, 95%CI = 16.8-19.2).
6
7

8
9
10 As for the prevalence of depression, European countries presented the highest values, especially in
11 Southern (31.7%, 95%CI = 30.6-32.7) and Western Europe (25.0%, 95%CI = 24.3-25.6), whereas LMICs
12 showed very low proportions, especially in China, where only 1.2% of people aged 50+ presented
13 depression.
14
15
16
17

18 19 20 *Multimorbidity patterns* 21

22
23 Table 2 displays the aBIC, CAIC, and entropy values, proportions, and average posterior probability of
24 each latent class, for a two- to five-class model in both age subsamples. In the younger subsample
25 (50-64), the five-class solution yielded the lowest aBIC and CAIC values and the highest entropy value
26 (0.67). However, it was dismissed because one of the latent classes was very infrequent and the
27 posterior probabilities were far below .70. Similarly, the four-class model was rejected for an
28 inadequate posterior probability value in one of the classes (0.52). The model finally selected was the
29 3-class model. The three-class solution was also chosen for the older age group because of lower
30 posterior probability values in the four- and five-class models in spite of lower aBIC and CAIC values.
31
32
33
34
35
36
37
38
39
40
41
42

43 We named each latent class according to the most prevalent diseases within each latent class. Figure
44 1 shows the distribution of each condition across latent classes in the total sample and by regions.
45 The “cardio-metabolic” class presented excess prevalence of diabetes, hypertension, myocardial
46 infarction or angina, and stroke, comprising 8.93% of the total sample in the younger group and
47 27.22% in the older group. The “respiratory-mental-articular” class, which comprised 3.91% and
48 5.27% of each sample, respectively, showed greater prevalence of joint disorders, asthma, chronic
49 lung diseases, and depression. Finally, the “healthy” class presented low prevalence of conditions,
50 comprising 87.16% of the sample in the first age group and 67.51% in the second group.
51
52
53
54
55
56
57
58
59
60

Table 2 Comparison between models in individuals aged 50-64 and +65

No. of latent classes	Aged 50-64					Aged ≥65				
	Information Criteria Indices		Classification Quality	Latent classes, n (%)	Average posterior probability	Information Criteria Indices		Classification Quality	Latent classes, n (%)	Average posterior probability
	aBIC	CAIC	Entropy			aBIC	CAIC	Entropy		
2	1512.91	1583.94	0.51	33023 (82.15)	.88	1603.30	1674.33	0.33	21113 (66.10)	.83
				7177 (17.85)	.75				10827 (33.90)	.75
3	875.23	983.86	0.63	3589 (8.93)	.76	1032.83	1141.46	0.59	8693 (27.22)	.71
				1571 (3.91)	.67				1684 (5.27)	.68
				35040 (87.16)	.87				21563 (67.51)	.81
4	777.14	923.37	0.43	25701 (63.93)	.52	817.60	963.83	0.66	9557 (29.92)	.65
				626 (1.56)	.72				1474 (4.61)	.78
				7754 (19.29)	.80				17220 (53.91)	.86
				6119 (15.22)	.68				3689 (11.55)	.77
5	661.04	844.87	0.67	4578 (11.39)	.64	689.22	873.05	0.54	11094 (34.73)	.77
				247 (0.61)	.67				1094 (3.43)	.71
				32423 (80.65)	.85				14155 (44.32)	.76
				1359 (3.38)	.48				1148 (3.59)	.63
				1593 (3.96)	.64				4449 (13.93)	.72

Note: Boldface indicates the final selected model. *aBIC* adjusted Bayesian Information Criterion, *CAIC* Consistent Akaike Information Criterion

6/bmjopen-2019-034441 on 19 April 2024 by guest. Protected by copyright. http://bmjopen.bmj.com/

1
2
3 Differences in the proportions of multimorbidity classes were found across regions (Figure 1). The
4
5 “cardio-metabolic” class (18.8%, 95%CI =17.1-20.6) was significantly greater in Russia than in other
6
7 regions, and England (5.1%, 95%CI =4.5-5.7) showed a higher proportion of individuals classified into
8
9 the “respiratory-mental-articular” class. The “healthy” class was higher in Africa (91.5%, 95%CI =
10
11 90.7-92.3), China (90.8%, 95%CI = 90.1-91.4), and India (89.5%, 95%CI = 88.5-90.4), and remarkably
12
13 lower in Russia (71.6%, 95%CI = 69.5-73.6) compared to other regions.
14
15
16
17
18
19

20 ----- FIGURE 1 HERE -----
21
22
23

24 Similar results were found for the older group (Figure 2). In Russia, the “cardio-metabolic” class was
25
26 significantly higher than in other regions (48.3%, 95%CI = 46.1-50.6) whereas the “healthy” class was
27
28 the least frequent class compared with the rest of regions (38.4%, 95%CI = 36.2-40.6), followed by
29
30 Southern Europe (52.6%, 95%CI = 50.9-54.2). Africa and India showed lower proportions of
31
32 individuals classified into the “cardio-metabolic” class (12.9%, 95%CI = 11.8-14.1 and 11.2%, 95%CI =
33
34 10.0-12.6, respectively).
35
36
37
38

39 ----- FIGURE 2 HERE -----
40
41
42

43 *Association between multimorbidity classes and covariates*

44 Table 3 shows the unadjusted and adjusted relative risk ratios (RRR) for the younger sample. The
45
46 “healthy” class was used as the reference group. Both multimorbidity classes (cardio-metabolic and
47
48 respiratory-mental-articular) were associated with all the covariates, except for smoking status in the
49
50 “cardio-metabolic” class. Compared with the “healthy” class, individuals classified into the “cardio-
51
52 metabolic” and “respiratory-mental-articular” classes were more likely to be older (RRR = 1.09,
53
54 95%CI = 1.08-1.10; RRR = 1.06, 95%CI = 1.04-1.07, respectively) and being widowed (RRR = 1.4, 95%CI
55
56 = 1.1-1.7) and divorced in the “respiratory-mental-articular” class (RRR = 1.7, 95%CI = 1.2-2.4). Being
57
58
59
60

1
2
3 a male, having tertiary education, and high levels of wealth had a protective effect for being in both
4
5 multimorbidity groups, compared with the “healthy” class.
6

7 Taking Africa as the reference category, participants from Russia were more likely to be classified
8
9 into the “cardio-metabolic” class (RRR = 3.6, 95%CI = 3.0-4.2), whereas individuals from England (RRR
10
11 = 5.6, 95%CI = 4.1-7.7), Northern Europe (RRR = 2.8, 95%CI = 1.9-4.1) and India (RRR = 2.2, 95%CI =
12
13 1.5-3.2) showed higher risk of being in the “respiratory-mental-articular” class.
14
15

16 After adjusting for covariates, both latent classes were more likely to be associated with the
17
18 presence of feelings of loneliness (RRR = 1.8, 95%CI = 1.7-2.0; RRR = 2.5, 95%CI = 2.0-3.0),
19
20 limitations in ADL (RRR = 3.2, 95%CI = 2.9-3.6; RRR = 3.9, 95%CI = 3.3-4.7) and worse health status
21
22 (RRR = 12.8, 95%CI = 11.3-14.4; RRR = 12.9, 95%CI = 10.5-16.0). Physical activity had a protective
23
24 effect for being in these classes and having smoked was a risk factor only for being classified into the
25
26 “respiratory-mental-articular” class (RRR = 1.5, 95%CI = 1.2-1.7). Better performance in verbal
27
28 memory was significantly associated with less risk of being classified into the two multimorbidity
29
30 classes. Similarly, higher scores in verbal fluency (RRR = 0.96, 95%CI = 0.96-0.97; RRR = 0.89, 95%CI =
31
32 0.97-0.99) were a protective factor for multimorbidity, compared with the healthy individuals group.
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Table 3 Association between latent multimorbidity membership and outcomes in individuals aged 50-64

Outcomes	Aged 50-64			
	Unadjusted		Adjusted ^a	
	“Cardio-metabolic” class	“Respiratory-mental-articular” class	“Cardio-metabolic” class	“Respiratory-mental-articular” class
Gender				
Female	1.0	1.0		
Male	0.9 (0.8-1.0)	0.6 (0.5-0.7)		
Age (years)	1.1 (1.1-1.1)	1.1 (1.0-1.1)		
Marital status				
Single	1.0	1.0		
Married	0.9 (0.8-1.1)	0.7 (0.5-1.0)		
Divorced	1.1 (0.9-1.4)	1.7 (1.2-2.4)		
Widowed	1.4 (1.1-1.7)	1.1 (0.8-1.6)		
Education level				
Primary or less	1.0	1.0		
Secondary	0.9 (0.9-1.0)	0.7 (0.6-0.8)		
Tertiary	0.8 (0.7-0.9)	0.6 (0.5-0.7)		
Wealth				
1 st (worse)	1.0	1.0		
2 nd	1.0 (0.9-1.1)	0.8 (0.6-1.0)		
3 rd	0.9 (0.9-1.1)	0.7 (0.5-0.8)		
4 th	0.9 (0.8-1.0)	0.6 (0.5-0.8)		
5 th (best)	0.8 (0.7-0.9)	0.5 (0.4-0.6)		
Region				
Africa	1.0	1.0		
China	0.9 (0.8-1.0)	0.3 (0.2-0.5)		
India	0.9 (0.8-1.1)	2.2 (1.5-3.2)		
Russia	3.6 (3.0-4.2)	2.2 (1.4-3.5)		
England	1.1 (0.9-1.3)	5.6 (4.1-7.7)		
Northern Europe	1.3 (1.1-1.5)	2.8 (1.9-4.1)		
Southern Europe	1.5 (1.3-1.8)	1.7 (1.2-2.5)		
Western Europe	1.5 (1.3-1.7)	1.9 (1.3-2.6)		
Loneliness (yes/no)	1.9 (1.8-2.1)	2.5 (2.1-3.0)	1.8 (1.7-2.0)	2.5 (2.0-3.0)
Ever smoked (yes/no)	1.0 (0.9-1.1)	1.7 (1.5-2.0)	1.1 (1.0-1.2)	1.5 (1.2-1.7)
Physical activity (yes/no)	0.5 (0.5-0.5)	0.6 (0.5-0.7)	0.4 (0.4-0.5)	0.5 (0.5-0.6)
Limitations in ADL (yes/no)	2.5 (2.3-2.7)	3.2 (2.7-3.7)	3.2 (2.9-3.6)	3.9 (3.3-4.7)
Self-rated health				
Good	1.0	1.0	1.0	1.0
Moderate	3.7 (3.3-3.9)	2.5 (2.1-2.9)	4.8 (4.3-5.2)	3.9 (3.2-4.7)
Poor	8.5 (7.7-9.4)	7.4 (6.1-8.8)	12.8 (11.3-14.4)	12.9 (10.5-16.0)
Memory: Immediate recall	0.9 (0.9-0.9)	0.9 (0.9-1.0)	0.9 (0.9-0.9)	0.9 (0.9-1.0)
Memory: Delayed recall	0.9 (0.9-0.9)	0.9 (0.8-0.9)	0.9 (0.9-0.9)	0.9 (0.9-1.0)
Verbal fluency	1.0 (1.0-1.0)	1.0 (1.0-1.0)	1.0 (1.0-1.0)	1.0 (1.0-1.0)

Note: The reference group for the multimorbidity group variable was the “Healthy” class
 Relative Risk Ratios (95% confidence interval) from multinomial logistic regression models
 Models were run in 100 imputed datasets and results combined using Rubin’s rules
 ADL Activities of Daily Living

^aAdjusted for gender, age, marital status, education level, wealth and region

1
2
3 Table 4 shows the RRR for the older subsample. Individuals in lower wealth quintiles had greater risk
4 of being classified into the “cardio-metabolic” class (RRR = 1.1, 95%CI = 1.0-1.2), and individuals who
5
6
7 were in the fourth and fifth quintile (RRR = 0.8, 95%CI = 0.7-1.0) were less likely to be classified into
8
9
10 the “respiratory-mental-articular” class, compared with the “healthy” group. Compared with African
11
12 participants, all regions had greater risk of being classified into the “respiratory-mental-articular”
13
14 class, compared to the “healthy” class, especially participants from Russia (RRR = 14.5, 95%CI = 10.3-
15
16 20.3), compared to Africa.

17
18 Likewise, individuals who had ever smoked had a higher risk of being in the “respiratory-mental-
19
20 articular” group (RRR = 1.8, 95%CI = 1.5-2.0) and a lower risk of being classified into the “cardio-
21
22 metabolic” class (RRR = 1.0, 95%CI = 0.9-1.0). Better scores in verbal memory and verbal fluency
23
24 were significantly associated with less risk of being classified into the multimorbidity groups
25
26 compared with the class of the “healthy” individuals.
27
28
29
30
31

32 Discussion

33
34 To the best of our knowledge, this is the first multi-region study to use harmonized data to compare
35
36 multimorbidity patterns across different regions from three distinct population-based cohorts. We
37
38 identified three latent classes of multimorbidity based on the presence of eight NCDs: the “cardio-
39
40 metabolic”, the “respiratory-mental-articular” and the “healthy” class. The same clusters were
41
42 identified in another study using SAGE original data, applying exploratory factor analysis (EFA) in a
43
44 sample of 41 909 individuals aged 50 years or older [28]. Similarly, a study of a representative
45
46 sample of Spanish community-dwelling adults over 50 years old also found three latent classes using
47
48 eleven chronic conditions, showing similar diseases distributions among the multimorbidity clusters
49
50
51 [29].
52
53
54
55
56
57
58
59
60

Table 4 Association between latent multimorbidity membership and outcomes in individuals aged ≥ 65

Outcomes	Aged ≥ 65			
	Unadjusted		Adjusted ^a	
	"Cardio-metabolic" class	"Respiratory-mental-articular" class	"Cardio-metabolic" class	"Respiratory-mental-articular" class
Gender				
Female	1.0	1.0		
Male	0.7 (0.7-0.8)	0.9 (0.8-1.0)		
Age (years)	1.0 (1.0-1.0)	1.0 (1.0-1.0)		
Marital status				
Single	1.0	1.0		
Married	1.2 (1.00-1.4)	1.2 (0.9-1.6)		
Divorced	1.1 (1.0-1.4)	1.8 (1.2-2.6)		
Widowed	1.5 (1.3-1.8)	1.6 (1.2-2.3)		
Education level				
Primary or less	1.0	1.0		
Secondary	1.0 (0.9-1.0)	0.7 (0.6-0.8)		
Tertiary	0.9 (0.8-1.0)	0.6 (0.5-0.7)		
Wealth				
1 st (worse)	1.0	1.0		
2 nd	1.1 (1.0-1.2)	1.1 (0.9-1.2)		
3 rd	1.1 (1.0-1.2)	1.0 (0.9-1.2)		
4 th	1.1 (1.0-1.2)	0.8 (0.7-1.0)		
5 th (best)	1.2 (1.0-1.2)	0.7 (0.6-0.9)		
Region				
Africa	1.0	1.0		
China	1.8 (1.6-2.1)	4.1 (3.0-5.7)		
India	0.9 (0.8-1.0)	2.9 (2.0-4.2)		
Russia	8.0 (7.0-9.2)	14.5 (10.3-20.3)		
England	2.2 (2.0-2.5)	6.2 (4.5-8.5)		
Northern Europe	1.9 (1.6-2.2)	4.2 (2.9-6.1)		
Southern Europe	2.8 (2.5-3.2)	6.5 (4.7-9.0)		
Western Europe	2.1 (1.9-2.3)	3.6 (2.6-5.0)		
Loneliness (yes/no)	1.4 (1.4-1.6)	1.9 (1.7-2.2)	1.3 (1.2-1.5)	1.9 (1.7-2.3)
Ever smoked (yes/no)	0.8 (0.8-0.8)	1.4 (1.3-1.6)	1.0 (0.9-1.0)	1.8 (1.5-2.0)
Physical activity (yes/no)	0.5 (0.5-0.6)	0.5 (0.4-0.6)	0.5 (0.5-0.6)	0.4 (0.4-0.5)
Limitations in ADL (yes/no)	1.8 (1.7-1.9)	2.8 (2.5-3.1)	2.3 (2.1-2.5)	4.0 (3.5-4.6)
Self-rated health				
Good	1.0	1.0	1.0	1.0
Moderate	2.5 (2.3-2.7)	3.9 (3.3-4.6)	3.1 (2.9-3.4)	5.7 (4.8-6.8)
Poor	4.3 (4.6)	10.9 (9.3-12.8)	6.2 (5.6-6.9)	19.4 (16.2-23.4)
Memory: Immediate recall	0.9 (0.9-0.9)	0.9 (0.9-0.9)	0.9 (0.9-0.9)	0.9 (0.9-1.0)
Memory: Delayed recall	0.9 (0.9-0.9)	0.9 (0.9-0.9)	0.9 (0.9-0.9)	0.9 (0.9-1.0)
Verbal fluency	1.0 (0.9-1.00)	1.0 (0.9-1.0)	0.9 (0.9-1.0)	1.0 (1.0-1.0)

Note: The reference group for the multimorbidity group variable was the "Healthy" class
 Relative Risk Ratios (95% confidence interval) from multinomial logistic regression models
 Models were run in 100 imputed datasets and results combined using Rubin's rules
 ADL Activities of Daily Living

^aAdjusted for gender, age, marital status, education level and wealth

1
2
3 In our study, for both age groups the majority of the sample was classified into the “healthy” class,
4 87.16% and 67.51%, respectively. This latent group has previously been described in studies which
5 applied LCA [29–32]. Likewise, the other two identified classes are similar to those reported in a
6 systematic review based on 14 studies of multimorbidity patterns [33]. In this review, the most
7 prevalent diseases in the “cardio-metabolic” group were diabetes, hypertension, heart diseases,
8 hyperlipidemia, and obesity; and in the second group conditions such as mental disorders, thyroid
9 disease, neurological disease, pain, asthma or chronic lung diseases, musculoskeletal disorders,
10 obesity, and gastroesophageal reflux disease were included. Despite the fact that we included a
11 smaller number of diseases, we found analogous patterns. In our study, 8.93% of the younger group
12 (50-64) and 27.22% of the older were classified into the “cardio-metabolic” class, including
13 individuals with higher prevalence of diabetes, hypertension, myocardial infarction or angina, and
14 stroke. This clustering of diseases is similar to the metabolic syndrome, which has metabolically-
15 related cardiovascular risk factors and greater risk of stroke and diabetes [34]. Lastly, the least
16 prevalent group was the “respiratory-mental-articular” class, consisting of greater prevalence of joint
17 disorders, asthma, chronic lung diseases, and depression. Association between depression and
18 arthritis has commonly been reported, with socioeconomic and disease factors reported as being
19 involved in its association, as well as systemic inflammation mechanisms [3]. Nevertheless, the links
20 between depression and chronic lung diseases, and chronic lung diseases and arthritis, despite
21 having been studied, remain unclear [34,35].

22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47 Analogous latent multimorbidity classes have been found among both age groups. Despite this,
48 certain aspects should be pointed out. As expected, the proportion of participants classified into the
49 “healthy” class was greater in participants aged 50-64 (87.16%) compared to those aged +65 years
50 old (67.51%), illustrating higher multimorbidity in elderly individuals. The distribution of chronic
51 conditions was also less clear in the older subsample. For example, both joint disorders and angina-
52 myocardial infarction were similarly present in the “cardio-metabolic” and “respiratory-mental-
53
54
55
56
57
58
59
60

1
2
3 articular” categories, whereas in the youngest participant (50-64) subsample we observed a more
4
5 differentiated profile of those chronic conditions that cluster into one latent class. For example,
6
7 respiratory-related diseases (asthma, chronic lung diseases) are highly presented in the “respiratory-
8
9 mental-articular” class, while very infrequent among middle-aged people classified into the “cardio-
10
11 metabolic” group. It is worth mentioning that although depression is frequently observed among
12
13 participants classified into the “respiratory-mental-articular” class, it is not infrequent among people
14
15 within the “cardio-vascular” class. This may be due to the relationship between mental and medical
16
17 disorders, which has frequently been reported, suggesting a bidirectional association between them
18
19 [36]. On the one hand, medical conditions could be accompanied by a high symptom burden, leading
20
21 to depression, and, on the other, depression could be a risk factor for medical conditions, since
22
23 depressive symptoms could increase the incidence of behaviors, such as smoking, alcohol intake,
24
25 poor diet, or physical inactivity, which are risk factors for NCDs [3,36].
26
27
28
29

30 One important implication of our findings is the relatively high proportion of people aged 50-64 with
31
32 multimorbidity. Thus, preventive and intervention programs are also needed for this population to
33
34 mitigate the multimorbidity burden.
35

36 Our results show that these multimorbidity patterns are qualitatively different, but only if compared
37
38 to the “healthy” class in terms of sociodemographic and economic characteristics, lifestyles, and
39
40 health status variables. As has been reported in the literature, being older, female, widowed, with a
41
42 lower level of education and lower socioeconomic status are related to an increased risk of
43
44 multimorbidity [3]. In addition, those individuals with multiple chronic conditions were more likely
45
46 to have limitations in ADL, especially those classified into the “respiratory-mental-articular” group,
47
48 similar to what was found in another study of multimorbidity [29]. Physical activity seems to be a
49
50 protective factor for being classified into the “respiratory-mental-articular” class, whereas smokers
51
52 were more likely to be classified into the “respiratory-mental-articular” class, but not the “cardio-
53
54 metabolic” class. This is inconsistent with the literature, since cigarette smoking is considered a
55
56
57
58
59
60

1
2
3 major cause of cardiovascular diseases (CVDs). However, smoking is probably the most complex and
4
5 least understood risk factor for CVDs [37].
6
7
8
9

10 One interesting finding is the association between cognition outcomes and multimorbidity in both
11
12 age subsamples. Better performance in verbal memory and fluency was related to less risk of being
13
14 classified into the multimorbidity groups, with similar results among latent classes. Impaired
15
16 cognition has been associated with conditions such as arthritis [38], depression [39], and respiratory
17
18 diseases [40], cardiovascular conditions, diabetes [41], hypertension [42], and coronary heart
19
20 diseases [43].
21
22

23 Concerning the regional distribution of multimorbidity, Russia accounted for the highest burden as
24
25 opposed to Africa, China, and India. The “cardio-metabolic” class is especially common in this
26
27 country, with a prevalence of 18.82% in the youngest and 48.34% in the older subsample. Prevalence
28
29 of CVDs, such as hypertension, myocardial infarction or angina, and stroke, was also higher in Russia.
30
31 This high proportion could be related to the high rate of alcohol consumption and rapid societal
32
33 changes experienced in this country which might account for increased risk of circulatory diseases
34
35 [44,45]. Followed by Russia, European regions showed higher rates of multimorbidity. NCDs such as
36
37 hypertension, joint disorders, respiratory diseases, and depression were highly prevalent, especially
38
39 in England and Southern Europe, where the “respiratory-mental-articular” class was highly prevalent
40
41 in both age subsamples. The relationship between mood disorders such as depression and joint
42
43 disorders has been previously reported in other studies, though the underlying cause remains
44
45 unclear [28,29,33]. Notwithstanding, previous studies suggested that the emotional burden of joint
46
47 disorders may contribute to the onset of psychiatric disorders [28,46].
48
49
50

51 LMICs such as Africa, China, and India showed lower rates of multimorbidity compared to Russia and
52
53 other HICs. However, there was a wide variation in terms of some diseases, such as respiratory
54
55 diseases and depression. Asthma and chronic lung diseases were highly prevalent in India and China,
56
57 influenced by factors such as increasing smoking rates, air pollution, and occupational lung diseases
58
59
60

1
2
3 in these countries [47]. As reported in previous studies [48], depression was remarkably prevalent in
4
5 India, whereas the lowest prevalence was observed in China. This is in line with previous
6
7 epidemiological studies on the prevalence of depression in Chinese older people, suggesting
8
9 differences in diagnostic criteria that make depression less diagnosed; somatic symptoms are more
10
11 prevalent in this population instead of sadness, and lack of interest and energy. Moreover, stigma
12
13 and prejudice in Chinese population might also contribute to under-reporting depressive symptoms
14
15
16 [49,50].
17
18
19
20

21 The highest burden of multimorbidity in HICs could be explained by an increased level of
22
23 development in the HICs. Notwithstanding, LMICs are experiencing a change in lifestyle and
24
25 environmental exposures which contributes, as in HICs, to multimorbidity. Thus, the increased
26
27 burden of NCDs, in addition to the existing burden of infectious diseases such as HIV/AIDS, worsens
28
29 multimorbidity management [3] .
30
31
32
33

34 *Strengths and limitations*

35
36 A major strength of this study is the use of a large, harmonized, multi-regional database. Research on
37
38 multimorbidity has typically been hampered by several factors, such as the exclusion of patients with
39
40 multimorbidity from participation, targeting of research mostly on elderly individuals, and a shortage
41
42 of studies focusing on LMICs. The ATHLOS study allowed us to compare two age groups (50-64 and
43
44 65 or older) as well as disease prevalence and clusters of conditions in regions with differing incomes
45
46 in a very large, diverse population-based study of middle-aged and older adults.
47
48

49 Some limitations should be considered when interpreting our findings. First, the presence or absence
50
51 of the NCDs was based on self-reported measures, and thus might be affected by measurement
52
53 errors or lack of accuracy. Nevertheless, some authors sustain self-reported diagnostics as a well-
54
55 established method for the measurement of multimorbidity in population-based studies [51–54].
56
57
58 Second, we could only focus on those diseases that were common across studies. Conditions such as
59
60

1
2
3 obesity, cancer, kidney disease, and neurological illness were not evaluated. This might have led to a
4 smaller number of latent classes, or to different patterns of multimorbidity.
5
6

7 Third, when performing LCA, the 3-class solution was forced. In order to determine whether the
8 latent classes were equivalent, invariance analysis should have been performed [55]. Nevertheless,
9 this solution was forced as we aimed to do comparisons among age subsamples and regions in terms
10 of disease prevalence as well as protective and risk factors.
11
12
13
14

15 Finally, the use of multiple imputations could carry some bias. Despite this, the use of multiple
16 imputation procedures is widely advocated when missing data occur in one or more covariates in a
17 regression model and under a MAR assumption [56,57].
18
19
20
21
22
23
24

25 **Acknowledgments**

26
27
28
29 - The English Longitudinal Study of Ageing (ELSA):

30
31 ELSA is supported by the U.S. National Institute of Aging, the National Centre for Social Research, the
32 University College London (UCL), and the Institute for Fiscal Studies. The authors gratefully
33 acknowledge the UK Data Service and UCL who provided data for this paper.
34
35
36
37
38
39

40 - The Study on Global Ageing and Adult Health (SAGE):

41
42 The SAGE study is funded by the U.S. National Institute on Aging and has received financial support
43 through Interagency Agreements (OGHA 04034785; YA1323-08-CN-0020; Y1-AG-1005-01) and Grants
44 (R01-AG034479; IR21-AG034263-0182). The authors gratefully acknowledge the World Health
45 Organization who provided data for this paper.
46
47
48
49

50
51
52
53 - The Survey of Health, Ageing and Retirement in Europe (SHARE):

54
55 The SHARE study is funded by the European Commission through FP5 (QLK6-CT-2001-00360), FP6
56 (SHARE-I3: RII-CT-2006-062193, COMPARE: CIT5-CT-2005-028857, SHARELIFE: CIT4-CT-2006-028812)
57 and FP7 (SHARE-PREP: N°211909, SHARE-LEAP: N°227822, SHARE M4: N°261982). Additional funding
58
59
60

1
2
3 from the German Ministry of Education and Research, the Max Planck Society for the Advancement
4 of Science, the U.S. National Institute on Aging (U01_AG09740-13S2, P01_AG005842, P01_AG08291,
5 P30_AG12815, R21_AG025169, Y1-AG-4553-01, IAG_BSR06-11, OGHA_04-064,
6 HHSN271201300071C) and from various national funding sources is gratefully acknowledged (see
7 www.share-project.org).
8
9
10
11
12
13
14
15
16
17

18 **Funding:** This work was supported by the five-year Ageing Trajectories of Health: Longitudinal
19 Opportunities and Synergies (ATHLOS) project. The ATHLOS project has received funding from the
20 European Union's Horizon 2020 research and innovation program under grant agreement No
21 635316.
22
23
24
25
26
27

28 **Competing interests:** The authors report no conflicts of interest.
29

30 **Patient consent for publication:** not required
31

32 **Ethics approval:** The study protocol was approved by the Committee on the Ethics of Clinical
33 Research, CEIC Fundació Sant Joan de Déu (Protocol No: PIC-22-15). All data were anonymized and
34 EHR confidentially was respected in accordance with national and international law.
35
36
37
38
39
40
41

42 **Data sharing statement:** The original studies data are available on their respective websites: the
43 Study on Global Ageing and Adult Health - SAGE (<https://www.who.int/healthinfo/sage/en/>), the
44 English Longitudinal Study of Ageing – ELSA (<https://www.elsa-project.ac.uk/>), and the Survey of
45 Health, Ageing and Retirement in Europe – SHARE (<http://www.share-project.org/home0.html>). R
46 codes for harmonizing the included variables, as well as the STATA codes for the performed analysis
47 are available on <https://github.com/athlosproject/athlos-project.github.io>
48
49
50
51
52
53
54
55
56

57 **Additional files:** Additional file 1: Table S1 Proportion of missingness for indicators of latent classes
58 and variables included in the imputation model.
59
60

Author contributions:

IB: Participated in the database management, drafted the paper, carried out the statistical analyses and worked on the interpretation of data. She also gave final approval of the version to be published and agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved; **AS:** Participated in the study design, database management, statistical support and critical revision of the paper. He also gave final approval of the version to be published and agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved; **LE:** Participated in the interpretation of data and critical revision of the paper; **HN:** She participated in critical revision of the paper and gave final approval of the version to be published and agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved; **MP:** Participated in the study design, and critical revision of the paper. He also gave final approval of the version to be published and agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved; **JMH:** Participated in the study design, acquisition of data, interpretation of data and critical revision of the paper. He also gave final approval of the version to be published and agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved; **BO:** Participated in the acquisition of data, study design, database management, and critical revision of the paper. She also gave final approval of the version to be published and agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

References

- 1 WHO. Ageing and Health. <https://www.who.int/news-room/fact-sheets/detail/ageing-and-health> (11 December 2018, date last accessed)
- 2 Wang H, Naghavi M, Allen C, *et al*. Global, regional, and national life expectancy, all-cause mortality, and cause-specific mortality for 249 causes of death, 1980–2015: a systematic analysis for the Global Burden of Disease Study 2015. *Lancet* 2016;**388**:1459–544. doi:10.1016/S0140-6736(16)31012-1
- 3 Multimorbidity: a priority for global health research. 2018.
- 4 Parekh AK, Barton MB. The challenge of multiple comorbidity for the us health care system. *JAMA - J Am Med Assoc* 2010;**303**:1303–4. doi:10.1001/jama.2010.381
- 5 Mugisha J, Rosenbaum S, Ward PB, *et al*. Chronic physical conditions, multimorbidity and physical activity across 46 low- and middle-income countries. *Int J Behav Nutr Phys Act* 2017;**14**. doi:10.1186/s12966-017-0463-5
- 6 Chatterji S, Byles J, Cutler D, *et al*. Health, functioning, and disability in older adults - Present status and future implications. *Lancet* 2015;**385**:563–75. doi:10.1016/S0140-6736(14)61462-8
- 7 Arokiasamy P, Uttamacharya U, Jain K, *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* 2015;**13**:1–16. doi:10.1186/s12916-015-0402-8
- 8 Bao J, Chua KC, Prina M, *et al*. Multimorbidity and care dependence in older adults: A longitudinal analysis of findings from the 10/66 study. *BMC Public Health* 2019;**19**:1–10. doi:10.1186/s12889-019-6961-4
- 9 Maximova T, Pati S, Biritwum RB, *et al*. The impact of multimorbidity on adult physical and

- 1
2
3 mental health in low- and middle-income countries: what does the study on global ageing and
4 adult health (SAGE) reveal? *BMC Med* 2015;**13**:1–16. doi:10.1186/s12916-015-0402-8
5
6
7
8
9 10 Afshar S, Roderick PJ, Kowal P, *et al.* *Applied demography and public health in the 21st*
11 *century*. Cham, Switzerland: : Springer International Publishing AG 2016.
12
13 doi:10.1249/mss.0000000000001652
14
15
16 11 Sanchez-Niubo A, Egea-Cortés L, Olaya B *et al.* Cohort profile: the ageing trajectories of health
17 - longitudinal opportunities and synergies (ATHLOS) project. *Int J Epidemiol* 2019;**0**:1–11.
18
19 doi:org/10.1093/ije/dyz077
20
21
22
23 12 Kowal P, Chatterji S, Naidoo N, *et al.* Data resource profile: The World Health Organization
24 study on global AGEing and adult health (SAGE). *Int J Epidemiol* 2013;**41**:1639–49.
25
26
27 doi:10.1093/ije/dys210
28
29
30
31 13 Steptoe A, Breeze E, Banks J, *et al.* Cohort profile: The English Longitudinal Study of Ageing. *Int*
32 *J Epidemiol* 2013;**42**:1640–8. doi:10.1093/ije/dys168
33
34
35
36 14 Börsch-Supan A, Brandt M, Hunkler C, *et al.* Data resource profile: The survey of health,
37 ageing and retirement in europe (SHARE). *Int J Epidemiol* 2013;**42**:992–1001.
38
39
40
41
42 doi:10.1093/ije/dyt088
43
44
45 15 IDA TWBI. World Bank Open Data. <https://data.worldbank.org/> (5 March 2019, date last
46 accessed)
47
48
49 16 Fortier I, Raina P, Van den Heuvel ER, Griffith LE, Craig C, Saliba M, Doiron D, Stolk RP,
50 Knoppers BM, Ferretti V, Granda P BP. Maelstrom Research guidelines for rigorous
51 retrospective data harmonization. 2017;:103–5. doi:10.1093/ije/dyw075
52
53
54
55
56 17 Ronald C. Kessler TBÜ. The World Mental Health (WMH) survey initiative version of the World
57 Health Organization (WHO) Composite Diagnostic Interview (CIDI). *Int J Methods Psychiatr*
58
59
60

- 1
2
3 Res;**13**:93–121. doi:org/10.1002/mpr.168
4
5
6 18 Radloff L S. The CES-D scale: A self report depression scale for research in the general
7
8 population. *Appl Psychol Meas* 1977;**1**:385–401. doi:10.1002/ardp.19552881111
9
10
11 19 Reischies F, Lobo A, Turrina C, *et al.* Development of the EURO–D scale – a European Union
12
13 initiative to compare symptoms of depression in 14 European centres. *Br J Psychiatry*
14
15 2008;**174**:330–8. doi:10.1192/bjp.174.4.330
16
17
18 20 Morris JC, Heyman A, Mohs RC, Hughes JP, van Belle G, Fillenbaum G *et al.* The consortium to
19
20 establish a registry for Alzheimer’s disease (CERAD). Part I. Clinical and neuropsychological
21
22 assesment of Alzheimer’s disease. *Neurology* 1989;**39**:1159–65. doi:10.1212/wnl.39.9.1159
23
24
25
26 21 WHO. Country groupings.
27
28 https://www.who.int/quantifying_ehimpacts/global/ebdcountgroup/en/ (7 February 2019,
29
30 date last accessed)
31
32
33
34 22 United Nations Statistical Division (UNSD). Countries or areas/geographical regions.
35
36 <https://unstats.un.org/unsd/methodology/m49/> (7 February 2019, date last accessed)
37
38
39 23 Lanza ST, Dziak JJ, Huang L WA, Collins L. *LCA Stata plugin users ’ guide (version 1.2.1)*. Penn
40
41 State: : University Park: The Methodology Center 2018.
42
43
44 24 Tein JY, Coxe S, Cham H. Statistical power to detect the correct number of classes in latent
45
46 profile analysis. *Struct Equ Model* 2013;**20**:640–57. doi:10.1080/10705511.2013.824781
47
48
49 25 Nylund-Gibson K, Choi AY. Ten frequently asked questions about latent class analysis. *Transl*
50
51 *Issues Psychol Sci* 2018;**4**:440–61. doi:10.1037/tps0000176
52
53
54
55 26 StataCorp. *Stata: Release 13. Statistical Software*. College St. 2013.
56
57
58 27 Rubin DB. *Multiple Imputation for nonresponse in surveys*. United States of America: : John
59
60

- 1
2
3 Wiley & Sons 1987. doi:10.2307/3172772
4
5
6 28 Garin N, Olaya B, Perales J, *et al.* Multimorbidity patterns in a national representative sample
7
8 of the Spanish adult population. *PLoS One* 2014;**9**:e84794. doi:10.1371/journal.pone.0084794
9
10
11 29 Olaya B, Moneta MV, Caballero FF, *et al.* Latent class analysis of multimorbidity patterns and
12
13 associated outcomes in Spanish older adults: A prospective cohort study. *BMC Geriatr*
14
15 2017;**17**. doi:10.1186/s12877-017-0586-1
16
17
18
19 30 Yen L, Valderas JM, McRae IS, *et al.* Multimorbidity and comorbidity of chronic diseases
20
21 among the senior Australians: prevalence and patterns. *PLoS One* 2014;**9**:e83783.
22
23 doi:10.1371/journal.pone.0083783
24
25
26
27 31 Larsen FB, Pedersen MH, Friis K, *et al.* A Latent class analysis of multimorbidity and the
28
29 relationship to socio-demographic factors and health-related quality of life. A national
30
31 population-based study of 162,283 Danish Adults. *PLoS One* 2017;**12**:1–17.
32
33 doi:10.1371/journal.pone.0169426
34
35
36 32 Whitson HE, Johnson KS, Sloane R, *et al.* Identifying patterns of multimorbidity in older
37
38 americans: application of latent class analysis. *J Am Geriatr Soc* 2016;**64**:1668–73.
39
40 doi:10.1111/jgs.14201. Identifying
41
42
43
44 33 Prados-Torres A, Calderón-Larrañaga A, Hancoco-Saavedra J, *et al.* Multimorbidity patterns: A
45
46 systematic review. *J Clin Epidemiol* 2014;**67**:254–66. doi:10.1016/j.jclinepi.2013.09.021
47
48
49
50 34 Alberti KGM, Zimmet P, Shaw J. Metabolic syndrome -a new world-wide definition. A
51
52 consensus statement from the International Diabetes Federation. *Abstr Appl Anal*
53
54 2006;**23**:469–80. doi:10.1155/2010/307409
55
56
57 35 Margaretten M, Julian L, Katz P, *et al.* Depression in patients with rheumatoid arthritis:
58
59 description, causes and mechanisms. *Int J Clin Rheumtol* 2011;**6**:617–23.
60

1
2
3 doi:10.2217/IJR.11.6.Depression
4
5

- 6 36 Druss BG, Walker ER. Mental disorders and medical comorbidity. 2011.
7
8
9 37 Barbara Messner DB. Smoking and cardiovascular disease: Mechanisms of endothelial
10 dysfunction and early atherogenesis. 2014;**34**:509–15. doi:10.1161/ATVBAHA.113.300156
11
12
13
14 38 Meade T, Manolios N, Cumming SR, *et al.* Cognitive impairment in rheumatoid arthritis: A
15 systematic review. *Arthritis Care Res* 2018;**70**:39–52. doi:10.1002/acr.23243
16
17
18
19 39 Lam RW, Kennedy SH, McIntyre RS, *et al.* Cognitive dysfunction in major depressive disorder:
20 effects on psychosocial functioning and implications for treatment. *Can J Psychiatry*
21
22 2014;**59**:649–54. doi:10.1177/070674371405901206
23
24
25
26
27 40 Dodd JW. Lung disease as a determinant of cognitive decline and dementia. *Alzheimer's Res*
28
29 *Ther* 2015;**7**:1–8. doi:10.1186/s13195-015-0116-3
30
31
32
33 41 Rose Mayeda E, Whitmer RA, Yaffe K. Diabetes and cognition. *Clin Geriatr Med* 2015;**31**:101–
34
35 ix. doi:10.1016/j.cger.2014.08.021
36
37
38 42 Walker KA, Power MC, Gottesman RF. Defining the relationship between hypertension,
39
40 cognitive decline, and dementia: a review. 2018;**19**:24. doi:10.1007/s11906-017-0724-3
41
42
43 43 Deckers K, Schievink SHJ, Rodriguez MMF, *et al.* Coronary heart disease and risk for cognitive
44
45 impairment or dementia: Systematic review and meta-analysis. *PLoS One* 2017;**12**:1–15.
46
47 doi:10.1371/journal.pone.0184244
48
49
50
51 44 Leon DA, Shkolnikov VM, McKee M, *et al.* Alcohol increases circulatory disease mortality in
52
53 Russia: Acute and chronic effects or misattribution of cause? *Int J Epidemiol* 2010;**39**:1279–90.
54
55 doi:10.1093/ije/dyq102
56
57
58 45 Petrukhin IS, Lunina EY. Cardiovascular disease risk factors and mortality in Russia: Challenges
59
60

- 1
2
3 and barriers. *Public Health Rev* 2012;**33**:436–49. doi:10.1007/BF03391645
4
5
6 46 Y. He, M. Zhang, E. H. B. Lin et al. Mental disorders among persons with arthritis: results from
7
8 the World Mental Health Surveys. 2008;**38**:1639–50. doi:10.1017/S0033291707002474
9
10
11 47 Jamrozik E, Musk AW. Respiratory health issues in the Asia-Pacific region: An overview.
12
13 *Respirology* 2011;**16**:3–12. doi:10.1111/j.1440-1843.2010.01844.x
14
15
16 48 Sandeep Grover, Alakananda Dutt AA. An overview of Indian research in depression. *Indian J*
17
18 *Psychiatry* 2010;**52**:178–88. doi:10.4103/0019-5545.69231
19
20
21
22 49 Gu L, Xie J, Long J, et al. Epidemiology of major depressive disorder in Mainland China: A
23
24 Systematic review. *PLoS One* 2013;**8**:e65356. doi:10.1371/journal.pone.0065356
25
26
27 50 Hou Z, Jiang W, Yin Y, et al. The current situation on major depressive disorder in China:
28
29 Research on mechanisms and clinical practice. *Neurosci Bull* 2016;**32**:389–97.
30
31 doi:10.1007/s12264-016-0037-6
32
33
34
35 51 Huntley AL, Johnson R, Purdy S, et al. Measures of multimorbidity and morbidity burden for
36
37 use in primary care and community settings: A systematic review and guide. *Ann Fam Med*
38
39 2012;**10**:134–41. doi:10.1370/afm.1363
40
41
42
43 52 Dal Grande E, Fullerton S, Taylor AW. Reliability of self-reported health risk factors and
44
45 chronic conditions questions collected using the telephone in South Australia, Australia. *BMC*
46
47 *Med Res Methodol* 2012;**12**. doi:10.1186/1471-2288-12-108
48
49
50 53 Ortolá-Devesa JB, Egea-Caparrós JM, Huerta JM, et al. Accuracy of self-reported diabetes,
51
52 hypertension, and hyperlipidemia in the adult spanish population. DINO study findings. *Rev*
53
54 *Española Cardiol (English Ed)* 2009;**62**:143–52. doi:10.1016/s1885-5857(09)71532-4
55
56
57 54 Martin LM, Leff M, Calonge N, et al. Validation of self-reported chronic conditions and health
58
59 services in a managed care population. *Am J Prev Med* 2000;**18**:215–8. doi:10.1016/S0749-
60

1
2
3 3797(99)00158-0
4
5

6 55 Finch H. A comparison of statistics for assessing model invariance in latent class analysis. *Open*
7
8 *J Stat* 2015;**5**:191–210. doi:10.4236/ojs.2015.53022
9

10
11 56 White IR, Carlin JB. Bias and efficiency of multiple imputation compared with complete-case
12
13 analysis for missing covariate values. *Stat Med* 2010;**29**:2920–31. doi:10.1002/sim.3944
14
15

16 57 Sterne JAC, White IR, Carlin JB, *et al*. Multiple imputation for missing data in epidemiological
17
18 and clinical research: Potential and pitfalls. *BMJ* 2009;**339**:157–60. doi:10.1136/bmj.b2393
19
20
21
22
23
24

25 **Figure 1** Prevalence of diseases in the three latent classes in the total sample and by regions (subsample 50-64
26
27 years)
28
29

30 **Figure 2** Prevalence of diseases in the three latent classes in the total sample and by regions (subsample +65
31
32 years)
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

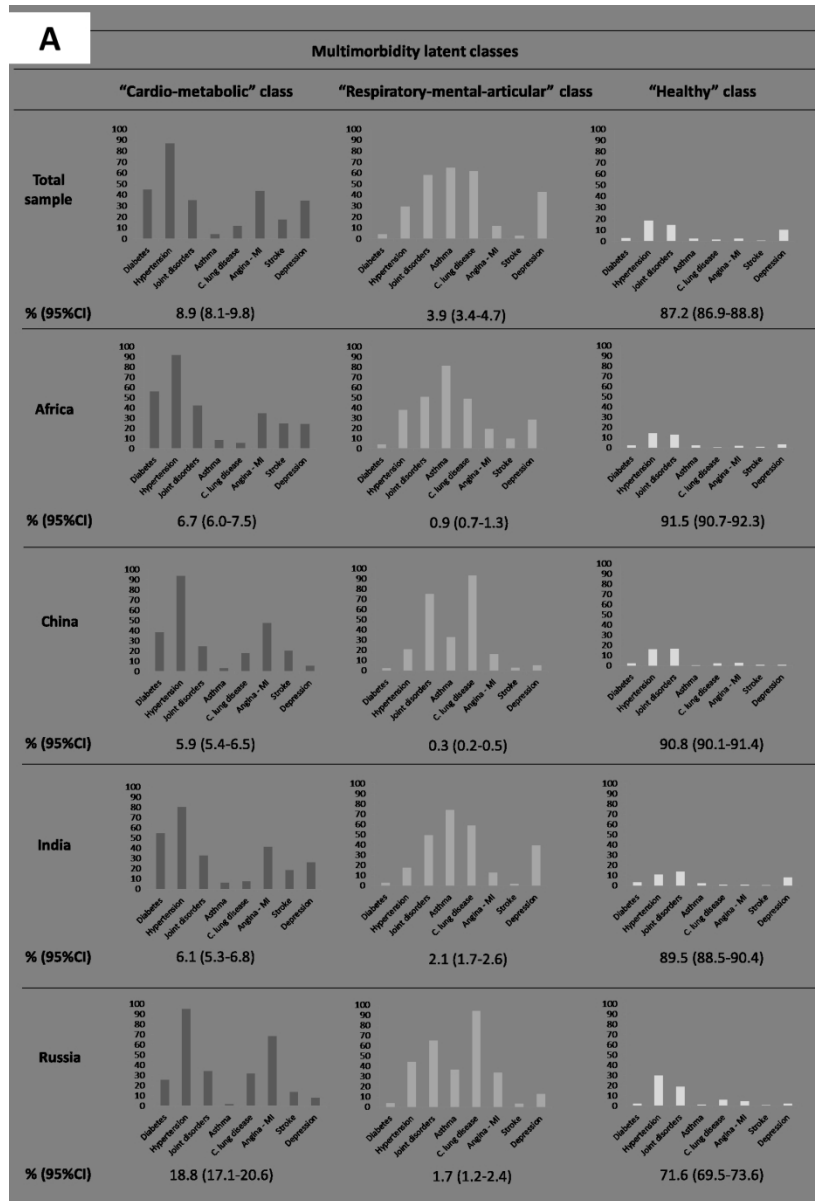


Figure 1 Prevalence of diseases in the three latent classes in the total sample and by regions (subsample 50-64 years). Part a

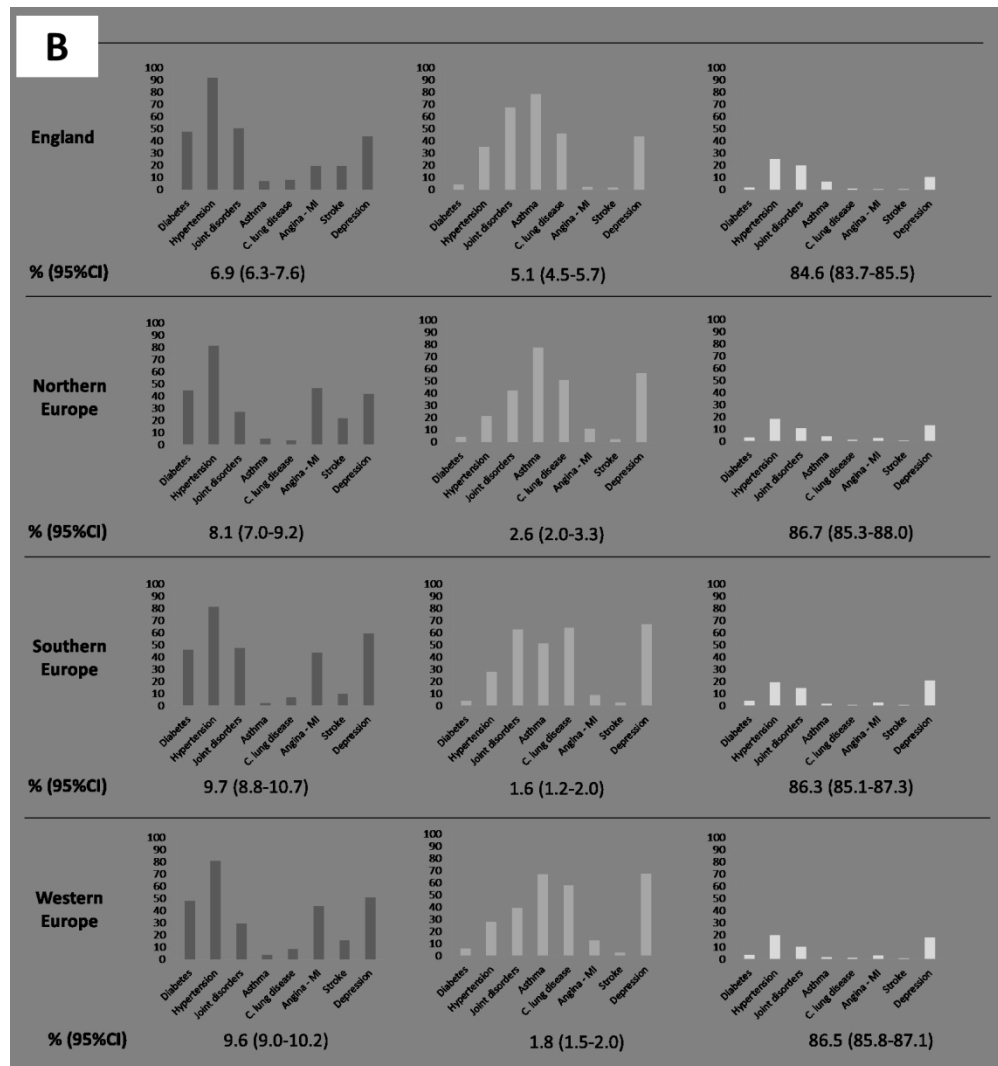


Figure 1 Prevalence of diseases in the three latent classes in the total sample and by regions (subsample 50-64 years). Part b

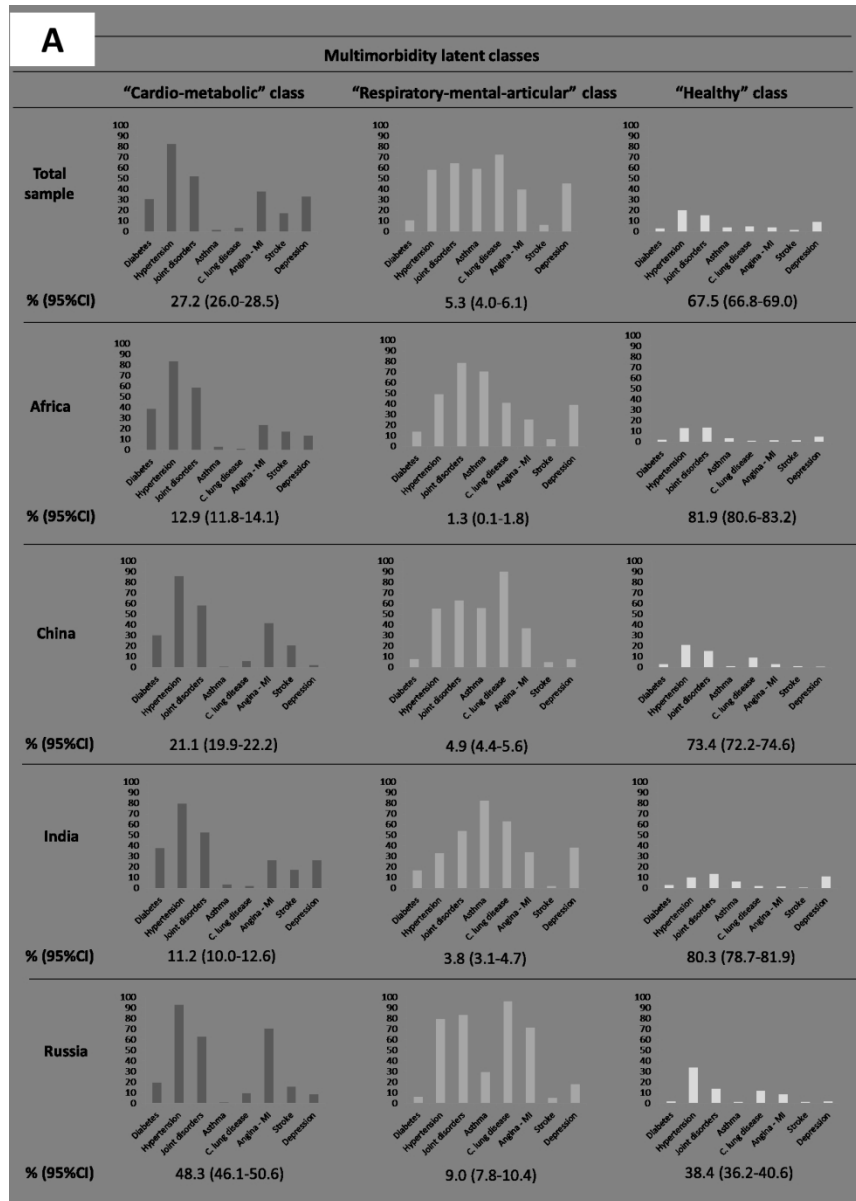


Figure 2 Prevalence of diseases in the three latent classes in the total sample and by regions (subsample +65 years). Part a

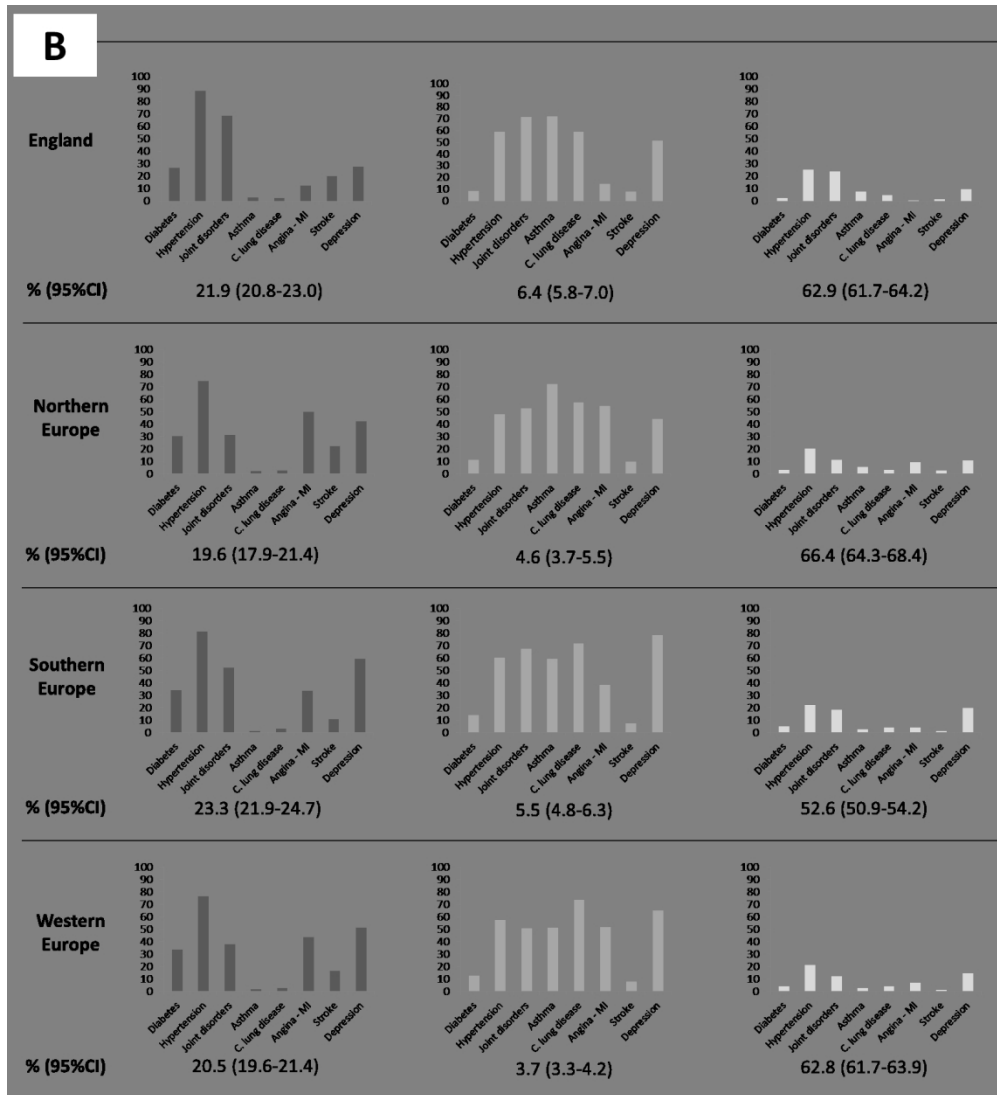


Figure 2 Prevalence of diseases in the three latent classes in the total sample and by regions (subsample +65 years). Part b

Table S1. Proportion of missingness for indicators of latent classes and variables included in the imputation model

(a) Indicators of multimorbidity classes	
	n (%) missingness
Diabetes	62 (0.08)
Hypertension	154 (0.21)
Joint disorders	31 (0.04)
Asthma	82 (0.11)
Chronic lung disease	51 (0.07)
MI - Angina	63 (0.08)
Stroke	41 (0.05)
Depression	1233 (1.70)
(b) Variables used in the regression model	
Sex	0
Age	0
Country	0
Study	0
Marital status	106 (0.14)
Education	11475 (15.91)
Wealth	844 (1.17)
Loneliness	10985 (15.23)
Ever smoked	201 (0.27)
Physical activity	258 (0.35)
Self-rated health	218 (0.30)
ADL – Using the toilet	355 (0.49)
ADL – Bathing or showering	340 (0.47)
ADL – Getting dressed	304 (0.42)
ADL - Eating	400 (0.55)
ADL – Getting in or out of bed	310 (0.42)
ADL – Moving around the house	345 (0.47)
Memory: Immediate recall	1442 (1.99)
Memory: Delayed recall	1448 (2.00)
Verbal fluency	1641 (2.27)

BMJ Open

Multimorbidity patterns in low-middle and high income regions: A multi-region latent class analysis using ATHLOS harmonized cohorts.

Journal:	<i>BMJ Open</i>
Manuscript ID	bmjopen-2019-034441.R1
Article Type:	Original research
Date Submitted by the Author:	17-Apr-2020
Complete List of Authors:	Bayes-Marin, Iveta; Parc Sanitari Sant Joan de Déu, Research, Innovation and Teaching Unit; Centro de Investigación Biomédica en Red de Salud Mental, CIBERSAM Sanchez-Niubo, Albert ; Parc Sanitari Sant Joan de Déu, Research, Innovation and Teaching Unit; Centro de Investigación Biomédica en Red de Salud Mental, CIBERSAM Egea-Cortés, Laia; Parc Sanitari Sant Joan de Déu, Research, Innovation and Teaching Unit Nguyen, Hai; King's College London, Health Service and Population Research Department Prina, A.Matthew; King's College London, Fernández, Daniel; Parc Sanitari Sant Joan de Déu, Research, Innovation and Teaching Unit; CIBERSAM Haro, Josep; Parc Sanitari Sant Joan de Deu, CIBERSAM, University of Barcelona; King Sand University Olaya, Beatriz; Parc Sanitari Sant Joan de Deu, Research and Development Unit
Primary Subject Heading:	Epidemiology
Secondary Subject Heading:	Global health
Keywords:	multimorbidity, non-communicable diseases (NCDs), low- and middle-income countries (LMICs), high-income country (HICs), Latent class analysis (LCA)

SCHOLARONE™
Manuscripts



I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our [licence](#).

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which [Creative Commons](#) licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

Original Research Article**Manuscript title:**

Multimorbidity patterns in low-middle and high income regions: A multi-region latent class analysis using ATHLOS harmonized cohorts.

Author's full names and affiliations:

Ivet Bayes-Marin, ^{a,b} Albert Sanchez-Niubo, ^{a,b} Laia Egea-Cortés, ^a Hai Nguyen, ^c Matthew Prina, ^c Daniel Fernández, ^{a,b,d} Josep Maria Haro, ^{a,b} and Beatriz Olaya ^{a,b}

From the ^aResearch, Innovation and Teaching Unit, Parc Sanitari Sant Joan de Déu, Sant Boi de Llobregat, Spain, ^bCentro de Investigación Biomédica en Red de Salud Mental, CIBERSAM, Madrid, Spain, ^cHealth Service and Population Research Department, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, United Kingdom, ^dSerra Hünter fellow. Department of Statistics and Operations Research, Polytechnic University of Catalonia, Barcelona, Spain.

Corresponding author details:

Correspondence: Albert Sanchez- Niubo. Research, Innovation and Teaching Unit, Parc Sanitari Sant Joan de Déu, Dr Antoni Pujadas 42, Sant Boi de Llobregat, Barcelona 08830, Spain. E-mail: albert.sanchez@pssjd.org

ORCID of the authors:

Ivet Bayes-Marin - 0000-0002-3816-5244

Albert Sanchez-Niubo - 0000-0003-0309-181X

Laia Egea-Cortés - 0000-0003-4002-8254

Hai Nguyen - 0000-0003-2171-1955

1
2
3 Matthew Prina - 0000-0001-6698-3263
4

5 Daniel Fernández - 0000-0003-0012-2094
6

7 Josep Maria Haro - 0000-0002-3984-277X
8

9
10 Beatriz Olaya - 0000-0003-2046-3929
11

12
13
14 **Word count (excluding title page, abstract, references, figures and tables): 5081**
15
16

17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For peer review only

Abstract

Objectives: Our aim was to determine clusters of non-communicable diseases (NCDs) in a very large, population-based sample of middle-aged and older adults from low- and middle-income (LMICs) and high-income (HICs) regions. Additionally, we explored the associations with several covariates.

Design: The total sample was 72 140 people aged 50+ from three population-based studies (ELSA, SHARE and SAGE) included in the ATHLOS project and representing eight regions with LMICs and HICs. Variables were previously harmonized using an ex-post strategy. Eight NCDs were used in latent class analysis. Multinomial models were made to calculate associations with covariates. All the analyses were stratified by age (50-64 and 65+ years old).

Results: Three clusters were identified: “cardio-metabolic” (8.93% in participants aged 50-64 and 27.22% in those aged 65+), “respiratory-mental-articular” (3.91% and 5.27%) and “healthy” (87.16% and 67.51%). In the younger group, Russia presented the highest prevalence of the “cardio-metabolic” group (18.8%) and England the “respiratory-mental-articular” (5.1%). In the older group, Russia had the highest proportion of both classes (48.3% and 9%). Both the younger and older African participants presented the highest proportion of the “healthy” class. Older age, being woman, widowed, and with low levels of education and income were related to an increased risk of multimorbidity. Physical activity was a protective factor in both age groups and smoking a risk factor for the “respiratory-mental-articular”.

Conclusions: Multimorbidity is common worldwide, especially in HICs and Russia. Health policies in each country addressing coordination and support are needed to face the complexity of a pattern of growing multimorbidity.

Keywords: multimorbidity, non-communicable diseases (NCDs), low- and middle- income countries (LMICs), high-income countries (HICs), latent class analysis (LCA).

Strengths and limitations of this study

- This study used a large, harmonized, multi-regional database, which allowed us to compare two age groups as well as disease prevalence in regions with differing incomes.
- The presence or absence of the non-communicable diseases was based on self-reported measures, and thus might be affected by measurement errors or lack of accuracy.
- Only common diseases across studies were included in the analyses, so this might have led to a smaller number of latent classes, or to different clusters.
- When performing Latent Class Analysis, we forced the solution as we aimed to do comparisons among age subsamples and regions in terms of disease prevalence as well as protective and risk factors.
- The use of multiple imputations for missing data in the covariates could carry some bias.

Background

By 2050, the population aged 60 years and older is expected to reach 2 billion worldwide, compared to 900 million in 2015 [1]. Along with this rapid increase, the incidence of chronic conditions (CC) or non-communicable diseases (NCDs) is also on the rise, having become the leading cause of morbidity and disability worldwide [2].

Multimorbidity, defined as the co-existence of two or more CC, is more common in older adults and is often more prevalent in people of lower socioeconomic status [3]. Multimorbidity is thought to account for 65% of total health care expenses in high-income countries (HICs) because of the huge associated healthcare utilization [4]. Due to the increasing prevalence of multimorbidity, the managing of multiple conditions has become an unavoidable international research priority, because of the high impact on the quality of life of patients and caregivers and on healthcare systems [3].

Most studies on the prevalence of multimorbidity in older people come from HICs, while data from middle-aged adults and low- and middle- income countries (LMICs) are much more limited [5–8]. LMICs are experiencing an increase in life expectancy that, together with changes in lifestyle and environment exposures, are triggering changes in their disease burden profile [3,9]. Few studies have compared patterns of multimorbidity between HICs and LMICs. Afshar et al. [10] used population-based chronic disease data from the World Health Survey to compare multimorbidity prevalence across 27 LMICs and one HIC, and used gross domestic product (GDP) to study inter-country socioeconomic differences. They found high multimorbidity prevalence in all countries, and a positive but non-linear relationship between country GDP and multimorbidity prevalence, suggesting the influence of other factors, such as lifestyles, social conditions, and differences across health systems. Four latent classes were identified in a cross-sectional sample of Australian seniors aged 50 years and over, using self-reported diagnosis of eleven conditions, including cancer and Parkinson's disease [11]. Another study, focusing on complex health care needs of Italian elderly people, found five clusters using 15 diseases [12]. A study conducted in a sample of 162,283 people from a survey of Danish population identified seven latent classes considering 15 chronic diseases and seven age

1
2
3 groups, ranging from 16 to 104 years [13]. These differences could be explained in light of variations
4
5 in collection methods, data sources, populations, diseases included, and the analysis performed
6
7 [11,14,15].
8

9
10 Similarly, the lack of study of differences in multimorbidity between HICs and LMICs may be due to
11
12 the use of different methodologies, which might hinder comparisons of prevalence and
13
14 multimorbidity patterns across countries. The integration of data from different studies would allow
15
16 us to determine differences across regions and cohorts, as well as to explore risk and protective
17
18 factors involved in the clustering of CC, thereby improving our understanding of the problem and the
19
20 creation of adapted medical guidelines.
21
22

23
24
25 This study aimed to: a) identify multimorbidity clusters in middle-aged (50-64) and older adults (+65)
26
27 from different regions, classified as LMICs and HICs; b) investigate the associations between
28
29 multimorbidity clusters and sociodemographic, economic, lifestyles and health status variables; and
30
31 c) explore differences across regions.
32
33
34
35

36 **Methods**

37 *Study design and data extraction*

38
39 The present study used data from the Ageing Trajectories of Health: Longitudinal Opportunities and
40
41 Synergies (ATHLOS) project [16]. Longitudinal data from 17 international cohort studies related to
42
43 health and ageing were harmonized with the aim of obtaining an integrated dataset and achieving a
44
45 better understanding of ageing and health processes.
46
47

48
49 We selected three studies due to their inclusion of the variables of interest and the possibility of
50
51 comparing HICs and LMICs. Baseline samples of the following studies were included in the analyses:
52
53 the World Health Organization's Study on Global Ageing and Adult Health (SAGE) [17], the English
54
55 Longitudinal Study of Ageing (ELSA)[18], and the Survey of Health, Ageing and Retirement in Europe
56
57 Study (SHARE) [19]. These panel studies included non-institutionalized people aged 50 years and
58
59
60

1
2
3 older. SAGE comprises six LMICs according to The World Bank Classification [20], namely Ghana,
4 South Africa, Mexico, India, China and Russia; ELSA includes the English population and SHARE covers
5 eleven countries of the European Union and Israel at baseline, considered as HICs [20].
6
7

8
9 The analyses presented focused on people aged 50 years of age and older who were part of the core
10 sample of each study and who completed a non-proxy interview at baseline. We excluded from the
11 analyses those participants who participated via proxy due to cognitive problems or severe physical
12 limitations. Moreover, people with missing values in sex and age were excluded, resulting in a final
13 sample of 72 140 individuals. Mexico was excluded from the analyses due the high percentage of
14 missingness in the variables of interest (See Additional file 1: Table S1).
15
16
17
18
19
20
21
22

23 24 25 *Patient and public involvement*

26
27 No patient involved.
28
29
30
31

32 33 *Variables*

34 The following variables were the result of a stringent, ex-post harmonization process using
35 systematic harmonization methodology and tools from Maelstrom Research [21].
36
37

38 Eight NCDs were used to conduct the analysis, including those that were available in the three
39 studies: diabetes, hypertension, asthma, chronic lung disease, joint disorders (arthritis, rheumatism
40 or osteoarthritis), angina or myocardial infarction, stroke, and depression. The presence or absence
41 of these conditions was self-reported and based on a medical diagnosis. Depression was assessed
42 with standardized tools, such as the Composite International Diagnostic Interview (CIDI) in the SAGE
43 study, the Center for Epidemiologic Studies Depression Scale (CES-D) in ELSA, and the EURO-D in
44 SHARE [22–24]. A dichotomous variable (yes/no) was created using the indicated cut-off score for
45 each tool and population based on previous studies [22,25,26].
46
47
48
49
50
51
52
53
54
55

56 Self-reported demographic variables included age, sex, level of education (primary or less, secondary,
57 and tertiary), marital status (single, married or currently cohabiting, separated or divorced, and
58
59
60

1
2
3 widowed) and quintiles of household wealth (first quintile indicating lowest level). Life-styles and
4 health behaviors were 'ever smoked' any type of tobacco and physical activity referring to the
5 practice of vigorous exercise during the last two weeks, both coded as *yes* or *no*. Other health-
6 related variables were self-rated health (good, moderate, or poor), presence or absence of loneliness
7 feelings in the last week, difficulties in activities of daily living (ADL), cognitive performance, and
8 number of diseases.
9

10
11 To assess difficulties in ADL, we used a set of daily self-care activities, which were common across
12 studies, such as problems in using the toilet, bathing or showering, getting dressed, eating, moving,
13 or getting in or out of bed. Each of the ADL difficulties was coded into a *yes/no* if the person
14 answered 'severe' or 'extreme/cannot do it'. To build the set of ADL difficulties, we coded *yes* if the
15 person reported at least one difficulty in any of the six items.
16

17 Immediate and delayed recall was assessed using the 10-word learning list task, and verbal fluency
18 utilizing the animal naming test [27]. Continuous total scores were used to perform the analyses.
19

20 Number of diseases was built by adding up the occurrences of all the above-mentioned NCDs.
21

22 Finally, a 7-level regional membership variable was created in order to analyze regional differences,
23 based on the World Health Organization (WHO) and the United Nations Statistical Division (UNSD)
24 regional classification [28,29]. Moreover, the World Bank Classification was used to classify these
25 regions into HICs or LMICs [20]. SAGE includes Africa (Ghana and South Africa), China and India, all of
26 them considered as LMICs. SHARE countries were grouped into three regions: Northern Europe
27 (Denmark, Sweden), Southern Europe (Greece, Italy, and Spain) and Western Europe (Austria,
28 Belgium, France, Germany, Israel, Netherlands, and Switzerland). ELSA and SHARE regions were
29 considered as HICs. Ghana and South Africa were grouped together and named as Africa for practical
30 purposes as well as due to their smaller sample size. These countries are not necessarily
31 representative of the whole continent.
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Statistical analysis

All the analyses were performed using data from the baseline. Descriptive statistics were used to summarize information regarding sociodemographic economic variables and disease prevalence among regions. Confidence intervals (95%IC) were calculated for categorical variables in order to make comparisons across regions.

Latent Class Analysis (LCA) was conducted stratified by age (50-64, +65). Eight NCDs (diabetes, hypertension, asthma, chronic lung disease, joint disorders, angina-myocardial infarction, stroke, and depression) were used as observed indicators without using covariates since we aimed to identify latent classes only based on disease variables. Region was used as cluster when conducting LCA in order to accurately describe disease proportions, indicating that the subjects were not independent random draws, but rather were nested within clusters [30] .

The optimal number of latent classes was determined using the adjusted Bayesian Information Criterion (aBIC), the consistent Akaike Information Criterion (CAIC), and the Entropy Index. Lower values of aBIC and CAIC indicate better fit, whereas entropy index values higher than 0.80 indicate that the latent classes are highly discriminating [31]. The average posterior probability indicates how well a model classifies individuals into their most likely class. Values higher than 0.70 indicate well-identified classes [32]. Additionally, interpretability and clinical judgment were used [32,33].

Missing data in one of the indicators was handled with the full-information maximum likelihood (FIML) technique, assuming missing-at-random (MAR) [34]. Missing data in the covariates was handled using multiple imputation by chained equations (MICE) assuming MAR [34]. The imputation model included the outcome (group membership in one of the latent classes) and all the variables used in the regression models. In the Additional file 1: Table S2-10 there is a report of those variables and the percentage of missingness of each region in the variables of interest.

Multinomial logistic regression models were used to assess the association of each multimorbidity class with several outcomes adjusted for sex, age, marital status, education level, wealth, and region at baseline. Due to potential collinearity between income and education, we checked the significance

1
2
3 and magnitude of the correlation between both variables. The association was small, and thus both
4
5 covariates were included as separate variables in the models. Regression models were conducted
6
7 separately in one hundred imputed datasets and results combined using Rubin's rules [35].
8

9
10 All analyses were conducted with Stata SE version 13.1 (College Station, TX). LCA analyses were
11
12 performed using a Stata plugin [30].
13
14
15

16 **Results**

17 *Descriptive analysis*

18
19 In Table 1 are presented the main characteristics of the sample by region. The mean age ranged from
20
21 62 (SD=9.02) in Southern Asia to 65 years (10.18) in Russia and (10.26) in England. Some 54% were
22
23 women, 72% were married or cohabitating, and 39% had secondary education. Russia presented the
24
25 highest number of conditions (mean 1.66) compared to Africa (0.64), China (0.80), and India (0.72).
26
27

28
29 The most prevalent conditions in the total sample were hypertension (31.2%, 95%CI = 30.9-31.6) and
30
31 joint disorders (22.4%, 95%CI = 22.0-22.7). Hypertension was particularly high in Russia (56.5%,
32
33 95%CI = 54.9-58.1) compared to the other regions. Diabetes prevalence was greater in Southern
34
35 (11.9%, 95%CI = 11.2-12.7) and Western Europe (10.4%, 95%CI = 9.9-10.9), whereas Africa and China
36
37 presented the lowest proportions. Similarly, myocardial infarction-angina was highly prevalent in
38
39 Russia (33.1%, 95%CI = 31.6-34.6), followed by countries of Northern (13.8%, 95%CI = 12.8-14.8),
40
41 Southern (11.7%, 95%CI = 11.0-12.4) and Western Europe (13.1%, 95%CI =12.6-13.6).
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Table 1 Main characteristics of the total sample and by regions

Region	Total	Africa ^a	China ^b	India ^b	Russia ^b	England ^c	Northern Europe ^d	Southern Europe ^e	Western Europe ^f
	N=72140	n=7950	n=12840	n=6558	n=3887	n=11 517	n=4573	n=7465	n=17 350
Age, mean (SD)	64.05 (9.96)	63.60 (10.26)	63.07 (9.31)	61.80 (9.02)	65.06 (10.18)	65.06 (10.26)	64.78 (10.34)	65.02 (10.17)	64.34 (9.96)
Female, % (95% CI)	54.0 (53.6-54.4)	52.3 (51.2-53.4)	52.9 (52.1-54.8)	49.6 (48.4-50.8)	64.6 (63.0-66.1)	54.6 (53.7-55.5)	53.9 (51.8-54.7)	55.4 (54.2-56.5)	54.1 (53.4-54.9)
Marital status, % (95% CI)									
Single	4.1 (4.0-4.3)	6.7 (6.2-7.3)	0.9 (0.7-1.6)	1.0 (0.8-1.2)	2.8 (2.3-3.4)	5.0 (4.6-5.4)	5.0 (4.8-6.2)	6.3 (5.8-6.9)	4.9 (4.6-5.3)
Married	71.5 (71.1-72.0)	54.5 (53.4-55.6)	83.5 (82.8-84.1)	74.1 (73.1-75.2)	56.1 (54.5-57.7)	69.1 (68.2-69.9)	72.0 (71.0-73.7)	73.3 (72.3-74.3)	73.3 (72.7-74.0)
Divorced	5.9 (5.7-6.2)	10.4 (9.7-11.1)	1.7 (1.5-2.0)	0.6 (0.5-0.9)	8.3 (7.4-9.2)	9.0 (8.5-9.5)	9.0 (8.8-10.5)	2.6 (2.3-3.0)	6.8 (6.5-7.2)
Widowed	18.4 (18.1-18.6)	27.3 (26.3-28.3)	13.9 (13.3-14.6)	24.3 (23.2-25.3)	32.7 (31.2-34.2)	16.9 (16.3-17.6)	12.1 (11.6-13.5)	17.8 (16.9-18.6)	14.9 (14.4-15.4)
Education level % (95% CI)									
Primary or less	33.4 (33.0-33.7)	30.5 (29.5-31.5)	37.9 (37.1-38.8)	25.5 (24.5-26.6)	9.9 (9.0-10.9)	42.4 (41.5-43.4)	29.9 (28.3-30.9)	59.4 (58.3-60.5)	23.2 (22.6-23.8)
Secondary	38.7 (38.4-39.1)	19.0 (18.2-19.9)	33.2 (32.4-34.0)	18.2 (17.3-19.2)	69.3 (67.9-70.8)	37.5 (36.7-38.4)	45.0 (43.6-46.5)	31.3 (30.2-32.3)	55.1 (54.3-55.8)
Tertiary	12.0 (11.9-12.2)	3.9 (3.5-4.3)	4.7 (4.3-5.0)	5.0 (4.4-5.5)	19.7 (18.5-21.0)	11.1 (10.6-11.7)	24.1 (23.0-25.5)	9.1 (8.4-9.7)	20.8 (20.2-21.4)
Wealth (quintiles% (95% CI)									
1 st (worse)	19.3 (19.1-19.6)	19.3 (18.4-20.2)	19.8 (19.1-20.5)	16.2 (15.3-17.1)	18.2 (17.0-19.5)	19.0 (18.3-19.7)	20.0 (19.6-21.9)	20.3 (19.4-21.2)	19.9 (19.3-20.5)
2 nd	19.7 (19.4-20.0)	19.7 (18.4-20.2)	19.7 (19.0-20.4)	18.6 (17.6-19.5)	19.8 (18.5-21.1)	19.3 (18.6-20.1)	20.0 (19.3-21.6)	20.4 (19.5-21.3)	19.9 (19.4-20.6)
3 rd	19.7 (19.4-19.9)	19.8 (18.9-20.7)	20.1 (19.4-20.8)	18.4 (17.5-19.4)	20.3 (19.0-21.6)	19.7 (18.9-20.4)	20.0 (19.0-21.3)	19.8 (18.9-20.7)	19.5 (18.9-20.1)
4 th	19.9 (19.7-20.3)	20.5 (19.6-21.4)	20.5 (19.8-21.2)	21.5 (20.5-22.5)	20.0 (18.8-21.3)	19.6 (18.9-20.3)	19.9 (18.5-20.8)	19.4 (18.6-20.4)	19.3 (18.8-19.9)
5 th (best)	20.1 (19.8-20.4)	20.4 (19.5-21.3)	19.7 (19.0-20.4)	24.8 (23.8-25.9)	21.6 (20.3-22.9)	19.6 (18.9-20.3)	18.9 (17.8-20.1)	18.9 (18.0-19.8)	19.4 (18.8-20.0)
N^o diseases, mean (SD)	1.02 (1.14)	0.64 (0.94)	0.80 (0.99)	0.72 (0.97)	1.66 (1.38)	1.19 (1.13)	1.02 (1.10)	1.28 (1.25)	1.10 (1.16)
Diseases, % (95% CI)									
Diabetes	8.5 (8.3-8.7)	6.6 (6.1-7.2)	6.5 (6.1-7.0)	7.3 (6.7-7.9)	9.0 (8.1-10.0)	7.4 (6.9-7.9)	8.0 (7.4-9.0)	11.9 (11.2-12.7)	10.4 (9.9-10.9)
Hypertension	31.2 (30.9-31.6)	21.5 (20.6-22.4)	27.4 (26.6-28.2)	17.5 (16.6-18.5)	56.5 (54.9-58.1)	37.8 (36.9-38.7)	29.9 (28.0-30.7)	35.6 (34.5-36.7)	32.3 (31.6-33.0)
Joint disorders	22.4 (22.0-22.7)	17.8 (16.9-18.6)	22.1 (21.4-22.8)	17.9 (17.0-18.9)	35.2 (33.7-36.7)	32.5 (31.6-33.3)	15.7 (14.6-16.8)	26.3 (25.3-27.3)	16.8 (16.3-17.4)
Asthma	5.5 (5.3-5.6)	4.1 (3.7-4.5)	2.4 (2.2-2.7)	6.9 (6.3-7.6)	3.4 (2.9-4.0)	11.7 (11.1-12.3)	7.8 (6.8-8.4)	4.1 (3.7-4.6)	4.1 (3.8-4.4)
Chronic lung disease	6.1 (5.9-6.3)	1.5 (1.2-1.7)	8.6 (8.1-9.1)	4.1 (3.6-4.6)	17.9 (16.8-19.2)	6.5 (6.1-7.0)	4.0 (3.9-5.2)	5.4 (4.9-6.0)	4.9 (4.6-5.3)
MI - Angina	10.0 (9.8-10.3)	4.5 (4.1-5.0)	8.8 (8.3-9.3)	4.9 (4.4-5.5)	33.1 (31.6-34.6)	3.3 (3.0-3.6)	13.4 (12.8-14.8)	11.7 (11.0-12.4)	13.1 (12.6-13.6)
Stroke	3.8 (3.7-3.9)	3.2 (2.8-3.6)	3.5 (3.2-3.8)	2.2 (1.9-2.6)	6.0 (5.3-6.8)	4.5 (4.1-4.9)	5.0 (4.4-5.7)	3.1 (2.8-3.6)	3.9 (3.7-4.2)
Depression	15.3 (15.0-15.5)	5.8 (5.3-6.3)	1.2 (1.1-1.4)	12.1 (11.3-12.9)	5.2 (4.5-5.9)	16.5 (15.8-17.2)	18.0 (17.7-20.0)	31.7 (30.6-32.7)	25.0 (24.3-25.6)

Note. ^aSAGE study - Africa: Ghana, South Africa; ^bSAGE study; ^cELSA study- England; ^dSHARE study - Northern Europe: Denmark, Sweden; ^eSHARE study - Southern Europe: Greece, Italy, Spain; ^fSHARE study - Western Europe: Austria, Belgium, France, Germany, Israel, Netherlands, Switzerland.
The analyses were performed before multiple imputation procedure

http://bmjopen.bmj.com/ or http://open.bmj.com/ April 17, 2024 by guest Protected by copyright.

1
2
3 Joint disorders were more prevalent in Russia (35.2%, 95%CI = 33.7-36.7) and England (32.5%, 95%CI
4 =31.6-33.3). The prevalence of asthma was greater in England than other regions (11.7%, 95%CI =
5 11.1-12.3) and chronic lung disease was greater in Russia (17.9%, 95%CI = 16.8-19.2).
6
7
8

9
10 As for the prevalence of depression, European countries presented the highest values, especially in
11 Southern (31.7%, 95%CI = 30.6-32.7) and Western Europe (25.0%, 95%CI = 24.3-25.6), whereas LMICs
12 showed very low proportions, especially in China, where only 1.2% of people aged 50+ presented
13 depression.
14
15
16
17
18

19 *Multimorbidity patterns*

20
21 Table 2 displays the aBIC, CAIC, and entropy values, proportions, and average posterior probability of
22 each latent class, for a two- to five-class model in both age subsamples. In the younger subsample
23 (50-64), the five-class solution yielded the lowest aBIC and CAIC values and the highest entropy value
24 (0.67). However, it was dismissed because one of the latent classes was very infrequent and the
25 posterior probabilities were far below 0.70. Similarly, the four-class model was rejected for an
26 inadequate posterior probability value in one of the classes (0.52). The model finally selected was the
27 three-class model. The three-class solution was also chosen for the older age group because of lower
28 posterior probability values in the four- and five-class models in spite of lower aBIC and CAIC values.
29
30
31
32
33
34
35
36
37
38
39
40
41
42

43 We named each latent class according to the most prevalent diseases within each latent class. Figure
44 1 shows the distribution of each condition across the three latent classes (“cardio-metabolic”,
45 “respiratory-mental-articular”, and “healthy” class) in the total sample and by regions. The “cardio-
46 metabolic” class presented excess prevalence of diabetes, hypertension, myocardial infarction or
47 angina, and stroke, comprising 8.93% of the total sample in the younger group and 27.22% in the
48 older group. The “respiratory-mental-articular” class, which comprised 3.91% and 5.27% of each
49 sample, respectively, showed greater prevalence of joint disorders, asthma, chronic lung diseases,
50
51
52
53
54
55
56
57
58
59
60

1
2
3 and depression. Finally, the “healthy” class presented low prevalence of conditions, comprising
4
5 87.16% of the sample in the first age group and 67.51% in the second group.
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For peer review only

Table 2 Comparison between models in individuals aged 50-64 and +65

No. of latent classes	Aged 50-64					Aged ≥65				
	Information Criteria Indices		Classification Quality	Latent classes, n (%)	Average posterior probability	Information Criteria Indices		Classification Quality	Latent classes, n (%)	Average posterior probability
	aBIC	CAIC	Entropy			aBIC	CAIC	Entropy		
2	1512.91	1583.94	0.51	33023 (82.15)	.88	1603.30	1674.33	0.33	21113 (66.10)	.83
				7177 (17.85)	.75				10827 (33.90)	.75
3	875.23	983.86	0.63	3589 (8.93)	.76	1032.83	1141.46	0.59	8693 (27.22)	.71
				1571 (3.91)	.67				1684 (5.27)	.68
				35040 (87.16)	.87				21563 (67.51)	.81
4	777.14	923.37	0.43	25701 (63.93)	.52	817.60	963.83	0.66	9557 (29.92)	.65
				626 (1.56)	.72				1474 (4.61)	.78
				7754 (19.29)	.80				17220 (53.91)	.86
				6119 (15.22)	.68				3689 (11.55)	.77
5	661.04	844.87	0.67	4578 (11.39)	.64	689.22	873.05	0.54	11094 (34.73)	.77
				247 (0.61)	.67				1094 (3.43)	.71
				32423 (80.65)	.85				14155 (44.32)	.76
				1359 (3.38)	.48				1148 (3.59)	.63
				1593 (3.96)	.64			4449 (13.93)	.72	

Note: Boldface indicates the final selected model. *aBIC* adjusted Bayesian Information Criterion, *CAIC* Consistent Akaike Information Criterion

1
2
3 Differences in the proportions of multimorbidity classes were found across regions (Figure 1). The
4
5 “cardio-metabolic” class (18.8%, 95%CI =17.1-20.6) was significantly greater in Russia than in other
6
7 regions, and England (5.1%, 95%CI =4.5-5.7) showed a higher proportion of individuals classified into
8
9 the “respiratory-mental-articular” class. The “healthy” class was higher in Africa (91.5%, 95%CI =
10
11 90.7-92.3), China (90.8%, 95%CI = 90.1-91.4), and India (89.5%, 95%CI = 88.5-90.4), and remarkably
12
13 lower in Russia (71.6%, 95%CI = 69.5-73.6) compared to other regions.
14
15
16
17
18
19

20 ----- FIGURE 1 HERE -----
21
22
23

24 Similar results were found for the older group (Figure 2). In Russia, the “cardio-metabolic” class was
25
26 significantly higher than in other regions (48.3%, 95%CI = 46.1-50.6) whereas the “healthy” class was
27
28 the least frequent class compared with the rest of regions (38.4%, 95%CI = 36.2-40.6), followed by
29
30 Southern Europe (52.6%, 95%CI = 50.9-54.2). Africa and India showed lower proportions of
31
32 individuals classified into the “cardio-metabolic” class (12.9%, 95%CI = 11.8-14.1 and 11.2%, 95%CI =
33
34 10.0-12.6, respectively).
35
36
37
38

39 ----- FIGURE 2 HERE -----
40
41
42

43 *Association between multimorbidity classes and covariates*

44 In the Additional file 2: Table S1 are presented the unadjusted relative risk ratios (RRR) for both age
45
46 subsamples. The “healthy” class was used as the reference group. In the case of the younger
47
48 subsample, and compared with the “healthy” class, individuals classified into the “cardio-metabolic”
49
50 and “respiratory-mental-articular” classes were more likely to be older (RRR = 1.09, 95%CI = 1.08-
51
52 1.10; RRR = 1.06, 95%CI = 1.04-1.07, respectively) and being widowed (RRR = 1.4, 95%CI = 1.1-1.7)
53
54 and divorced in the “respiratory-mental-articular” class (RRR = 1.7, 95%CI = 1.2-2.4). Being a man,
55
56 having tertiary education, and high levels of wealth had a protective effect for being in both
57
58 multimorbidity groups, compared with the “healthy” class. Similarly, those individuals from the older
59
60

1
2
3 subsample who were in the fourth and fifth quintile (RRR = 0.8, 95%CI = 0.7-1.0) were less likely to be
4
5 classified into the “respiratory-mental-articular” class, compared with the “healthy” group.

6
7 Regarding the association of regions and multimorbidity groups, some differences were found in the
8
9 younger subsample. Taking Africa as the reference category, participants from Russia were more
10
11 likely to be classified into the “cardio-metabolic” class (RRR = 3.6, 95%CI = 3.0-4.2), whereas
12
13 individuals from England (RRR = 5.6, 95%CI = 4.1-7.7), Northern Europe (RRR = 2.8, 95%CI = 1.9-4.1)
14
15 and India (RRR = 2.2, 95%CI = 1.5-3.2) showed higher risk of being in the “respiratory-mental-
16
17 articular” class. In the case of the older subsample, all regions had greater risk of being classified into
18
19 the “respiratory-mental-articular” class, compared to the “healthy” class, especially participants from
20
21 Russia (RRR = 14.5, 95%CI = 10.3-20.3), compared to Africa.
22
23
24
25
26
27

28 Table 3 shows the adjusted relative risk ratios (RRR) for both age subsamples, taking the “healthy”
29
30 class as the reference group. Both multimorbidity classes (cardio-metabolic and respiratory-mental-
31
32 articular) were associated with all the covariates in the younger group, except for smoking status in
33
34 the “cardio-metabolic” class.
35
36
37
38

39 In the younger individuals subsample, both latent classes were more likely to be associated with the
40
41 presence of feelings of loneliness (RRR = 1.8, 95%CI = 1.7-2.0; RRR = 2.5, 95%CI = 2.0-3.0), limitations
42
43 in ADL (RRR = 3.2, 95%CI = 2.9-3.6; RRR = 3.9, 95%CI = 3.3-4.7) and worse health status (RRR = 12.8,
44
45 95%CI = 11.3-14.4; RRR = 12.9, 95%CI = 10.5-16.0). Physical activity had a protective effect for being
46
47 in these classes and having smoked was a risk factor only for being classified into the “respiratory-
48
49 mental-articular” class (RRR = 1.5, 95%CI = 1.2-1.7). Conversely, those older individuals who had ever
50
51 smoked had a higher risk of being in the “respiratory-mental-articular” group (RRR = 1.8, 95%CI = 1.5-
52
53 2.0) and a lower risk of being classified into the “cardio-metabolic” class (RRR = 1.0, 95%CI = 0.9-1.0).
54
55 For both age subsamples, better performance in verbal memory was significantly associated with less
56
57
58
59
60

risk of being classified into the two multimorbidity classes. Similarly, higher scores in verbal fluency were a protective factor for multimorbidity, compared with the healthy individuals group.

Table 3 Association between latent multimorbidity membership and outcomes in individuals aged 50-64 and ≥65

Outcomes ^a	Aged 50-64		Aged ≥65	
	“Cardio-metabolic” class	“Respiratory-mental-articular” class	“Cardio-metabolic” class	“Respiratory-mental-articular” class
Loneliness (yes/no)	1.8 (1.7-2.0)	2.5 (2.0-3.0)	1.3 (1.2-1.5)	1.9 (1.7-2.3)
Ever smoked (yes/no)	1.1 (1.0-1.2)	1.5 (1.2-1.7)	1.0 (0.9-1.0)	1.8 (1.5-2.0)
Physical activity (yes/no)	0.4 (0.4-0.5)	0.5 (0.5-0.6)	0.5 (0.5-0.6)	0.4 (0.4-0.5)
Limitations in ADL (yes/no)	3.2 (2.9-3.6)	3.9 (3.3-4.7)	2.3 (2.1-2.5)	4.0 (3.5-4.6)
Self-rated health				
Good	1.0	1.0	1.0	1.0
Moderate	4.8 (4.3-5.2)	3.9 (3.2-4.7)	3.1 (2.9-3.4)	5.7 (4.8-6.8)
Poor	12.8 (11.3-14.4)	12.9 (10.5-16.0)	6.2 (5.6-6.9)	19.4 (16.2-23.4)
Memory: Immediate recall	0.9 (0.9-0.9)	0.9 (0.9-1.0)	0.9 (0.9-0.9)	0.9 (0.9-1.0)
Memory: Delayed recall	0.9 (0.9-0.9)	0.9 (0.9-1.0)	0.9 (0.9-0.9)	0.9 (0.9-1.0)
Verbal fluency	1.0 (1.0-1.0)	1.0 (1.0-1.0)	0.9 (0.9-1.0)	1.0 (1.0-1.0)

Note: The reference group for the multimorbidity group variable was the “Healthy” class
 Relative Risk Ratios (95% confidence interval) from multinomial logistic regression models
 Models were run in 100 imputed datasets and results combined using Rubin’s rules
 ADL Activities of Daily Living
^aAdjusted for sex, age, marital status, education level, wealth and region

Discussion

To the best of our knowledge, this is the first multi-region study to use harmonized data to compare multimorbidity patterns across different regions from three distinct population-based cohorts. We identified three latent classes of multimorbidity based on the presence of eight NCDs: the “cardio-metabolic”, the “respiratory-mental-articular” and the “healthy” class. The same clusters were identified in another study using SAGE original data, applying exploratory factor analysis (EFA) in a sample of 41 909 individuals aged 50 years or older [36]. Similarly, a study of a representative sample of Spanish community-dwelling adults over 50 years old also found three latent classes using eleven chronic conditions, showing similar diseases distributions among the multimorbidity clusters [37].

1
2
3 In our study, for both age groups the majority of the sample was classified into the “healthy” class,
4 87.16% and 67.51%, respectively. This latent group has previously been described in studies which
5 applied LCA [11,13,37,38]. Likewise, the other two identified classes are similar to those reported in a
6 systematic review based on 14 studies of multimorbidity patterns [39]. In this review, the most
7 prevalent diseases in the “cardio-metabolic” group were diabetes, hypertension, heart diseases,
8 hyperlipidemia, and obesity; and in the second group conditions such as mental disorders, thyroid
9 disease, neurological disease, pain, asthma or chronic lung diseases, musculoskeletal disorders,
10 obesity, and gastroesophageal reflux disease were included. Despite the fact that we included a
11 smaller number of diseases, we found analogous patterns. In our study, 8.93% of the younger group
12 (50-64) and 27.22% of the older were classified into the “cardio-metabolic” class, including
13 individuals with higher prevalence of diabetes, hypertension, myocardial infarction or angina, and
14 stroke. This clustering of diseases is similar to the metabolic syndrome, which has metabolically-
15 related cardiovascular risk factors and greater risk of stroke and diabetes [40]. Lastly, the least
16 prevalent group was the “respiratory-mental-articular” class, consisting of greater prevalence of joint
17 disorders, asthma, chronic lung diseases, and depression. Association between depression and
18 arthritis has commonly been reported, with socioeconomic and disease factors reported as being
19 involved in its association, as well as systemic inflammation mechanisms [3]. Nevertheless, the links
20 between depression and chronic lung diseases, and chronic lung diseases and arthritis, despite
21 having been studied, remain unclear [40,41].

22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47 Analogous latent multimorbidity classes have been found among both age groups. Despite this,
48 certain aspects should be pointed out. As expected, the proportion of participants classified into the
49 “healthy” class was greater in participants aged 50-64 (87.16%) compared to those aged +65 years
50 old (67.51%), illustrating higher multimorbidity in elderly individuals. The distribution of CC was also
51 less clear in the older subsample. For example, both joint disorders and angina–myocardial infarction
52 were similarly present in the “cardio-metabolic” and “respiratory-mental-articular” categories,
53
54
55
56
57
58
59
60

1
2
3 whereas in the younger participants (50-64) subsample we observed a more differentiated profile of
4 those chronic conditions that cluster into one latent class. For example, respiratory-related diseases
5 (asthma, chronic lung diseases) are highly presented in the “respiratory-mental-articular” class, while
6
7 very infrequent among middle-aged people classified into the “cardio-metabolic” group. It is worth
8
9 mentioning that although depression is frequently observed among participants classified into the
10
11 “respiratory-mental-articular” class, it is not infrequent among people within the “cardio-vascular”
12
13 class. This may be due to the relationship between mental and physical disorders, which has
14
15 frequently been reported, suggesting a bidirectional association between them [42]. On the one
16
17 hand, medical conditions could be accompanied by a high symptom burden, leading to depression,
18
19 and, on the other, depression could be a risk factor for medical conditions, since depressive
20
21 symptoms could increase the incidence of behaviors, such as smoking, alcohol intake, poor diet, or
22
23 physical inactivity, which are risk factors for NCDs [3,42].
24
25
26
27
28

29
30 One important implication of our findings is the relatively high proportion of people aged 50-64 with
31
32 multimorbidity. Thus, preventive and intervention programs are also needed for this population to
33
34 mitigate the multimorbidity burden.
35

36
37 Our results show that these multimorbidity patterns are qualitatively different, but only if compared
38
39 to the “healthy” class in terms of sociodemographic and economic characteristics, lifestyles, and
40
41 health status variables. As has been reported in the literature, being older, woman, widowed, with a
42
43 lower level of education and lower socioeconomic status are related to an increased risk of
44
45 multimorbidity [3]. In addition, those individuals with multiple chronic conditions were more likely
46
47 to have limitations in ADL, especially those classified into the “respiratory-mental-articular” group,
48
49 similar to what was found in another study of multimorbidity [37]. Physical activity seems to be a
50
51 protective factor for being classified into the “respiratory-mental-articular” class, whereas smokers
52
53 were more likely to be classified into the “respiratory-mental-articular” class, but not the “cardio-
54
55 metabolic” class. This is inconsistent with the literature, since cigarette smoking is considered a
56
57
58
59
60

1
2
3 major cause of cardiovascular diseases (CVDs). However, smoking is probably the most complex and
4
5 least understood risk factor for CVDs [43].
6
7
8
9

10 One interesting finding is the association between cognition outcomes and multimorbidity in both
11
12 age subsamples. Better performance in verbal memory and fluency was related to less risk of being
13
14 classified into the multimorbidity groups, with similar results among latent classes. Impaired
15
16 cognition has been associated with conditions such as arthritis [44], depression [45], and respiratory
17
18 diseases [46], cardiovascular conditions, diabetes [47], hypertension [48], and coronary heart
19
20 diseases [49].
21
22
23
24

25 Concerning the regional distribution of multimorbidity, Russia accounted for the highest burden as
26
27 opposed to Africa, China, and India. The “cardio-metabolic” class is especially common in this
28
29 country, with a prevalence of 18.82% in the younger and 48.34% in the older subsample. Prevalence
30
31 of CVDs, such as hypertension, myocardial infarction or angina, and stroke, was also higher in Russia.
32
33 This high proportion could be related to the high rate of alcohol consumption and rapid societal
34
35 changes experienced in this country which might account for increased risk of circulatory diseases
36
37 [50,51]. Followed by Russia, European regions showed higher rates of multimorbidity. NCDs such as
38
39 hypertension, joint disorders, respiratory diseases, and depression were highly prevalent, especially
40
41 in England and Southern Europe, where the “respiratory-mental-articular” class was highly prevalent
42
43 in both age subsamples. The relationship between mood disorders such as depression and joint
44
45 disorders has been previously reported in other studies, though the underlying cause remains
46
47 unclear [36,37,39]. Notwithstanding, previous studies suggested that the emotional burden of joint
48
49 disorders may contribute to the onset of psychiatric disorders [36,52].
50
51
52
53

54 LMICs such as Africa, China, and India showed lower rates of multimorbidity compared to Russia and
55
56 other HICs. However, there was a wide variation in terms of some diseases, such as respiratory
57
58 diseases and depression. Asthma and chronic lung diseases were highly prevalent in India and China,
59
60

1
2
3 influenced by factors such as increasing smoking rates, air pollution, and occupational lung diseases
4
5 in these countries [53]. As reported in previous studies [54], depression was remarkably prevalent in
6
7 India, whereas the lowest prevalence was observed in China. This is in line with previous
8
9 epidemiological studies on the prevalence of depression in Chinese older people, suggesting
10
11 differences in diagnostic criteria that make depression less diagnosed; somatic symptoms are more
12
13 prevalent in this population instead of sadness, and lack of interest and energy. Moreover, stigma
14
15 and prejudice in Chinese population might also contribute to under-reporting depressive symptoms
16
17 [55,56]. Furthermore, the variation found across regions in terms of depression prevalence could be
18
19 due to cultural differences in expressions or expectations of mood disorders or mental health[57].
20
21
22

23
24
25 The highest burden of multimorbidity in HICs could be explained by an increased level of
26
27 development in the HICs. Notwithstanding, LMICs are experiencing a change in lifestyle and
28
29 environmental exposures which contributes, as in HICs, to multimorbidity. Thus, the increased
30
31 burden of NCDs, in addition to the existing burden of infectious diseases such as HIV/AIDS, worsens
32
33 multimorbidity management [3] . Moreover, the differences found in the regional distribution of
34
35 multimorbidity might be linked to different stages of development of their health systems, since
36
37 there are differences between LMICs and HICs in terms of opportunities and barriers to improving
38
39 the organization, integration, and delivery of multimorbidity care [3].
40
41
42
43
44

45 *Strengths and limitations*

46
47 A major strength of this study is the use of a large, harmonized, multi-regional database. Research on
48
49 multimorbidity has typically been hampered by several factors, such as the exclusion of patients with
50
51 multimorbidity from participation, targeting of research mostly on elderly individuals, and a shortage
52
53 of studies focusing on LMICs. The ATHLOS study allowed us to compare two age groups (50-64 and
54
55 65 or older) as well as disease prevalence and clusters of conditions in regions with differing incomes
56
57 in a very large, diverse population-based study of middle-aged and older adults.
58
59
60

1
2
3 Some limitations should be considered when interpreting our findings. First, the presence or absence
4 of the NCDs was based on self-reported measures, and thus might be affected by measurement
5 errors or lack of accuracy. Nevertheless, some authors sustain self-reported diagnostics as a well-
6 established method for the measurement of multimorbidity in population-based studies [58–61].
7
8 Second, we could only focus on those diseases that were common across studies. Conditions such as
9 obesity, cancer, kidney disease, and neurological illness were not evaluated. This might have led to a
10 smaller number of latent classes, or to different patterns of multimorbidity.
11
12 Third, when performing LCA, the three-class solution was forced. In order to determine whether the
13 latent classes were equivalent, invariance analysis should have been performed [62]. Nevertheless,
14 this solution was forced as we aimed to do comparisons among age subsamples and regions in terms
15 of disease prevalence as well as protective and risk factors.
16
17 Finally, the use of multiple imputations could carry some bias. Despite this, the use of multiple
18 imputation procedures is widely advocated when missing data occur in one or more covariates in a
19 regression model and under a MAR assumption [63,64].
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36

37 The results of this study suggest that NCDs cluster together in non-random associations across
38 several regions worldwide. The three qualitatively distinct entities are also linked to several
39 sociodemographic and economic characteristics, lifestyles, and health status variables. A deeper
40 understanding of the interactions across regions and the studied variables is needed. Knowledge
41 regarding broad patterns of conditions may contribute to the creation and implementation of
42 guidelines that consider clusters of conditions instead of single diseases, since multimorbidity has
43 become an unavoidable reality. Future efforts should focus on the underlying mechanisms of these
44 clusters as well as their stability over time using longitudinal data. Moreover, cohort and age effects
45 should be explored as might influence the likelihood of reporting some diagnosis and hence result in
46 different multimorbidity patterns.
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Acknowledgments

- The English Longitudinal Study of Ageing (ELSA):

ELSA is supported by the U.S. National Institute of Aging, the National Centre for Social Research, the University College London (UCL), and the Institute for Fiscal Studies. The authors gratefully acknowledge the UK Data Service and UCL who provided data for this paper.

- The Study on Global Ageing and Adult Health (SAGE):

The SAGE study is funded by the U.S. National Institute on Aging and has received financial support through Interagency Agreements (OGHA 04034785; YA1323-08-CN-0020; Y1-AG-1005-01) and Grants (R01-AG034479; IR21-AG034263-0182). The authors gratefully acknowledge the World Health Organization who provided data for this paper.

- The Survey of Health, Ageing and Retirement in Europe (SHARE):

The SHARE study is funded by the European Commission through FP5 (QLK6-CT-2001-00360), FP6 (SHARE-I3: RII-CT-2006-062193, COMPARE: CIT5-CT-2005-028857, SHARELIFE: CIT4-CT-2006-028812) and FP7 (SHARE-PREP: N°211909, SHARE-LEAP: N°227822, SHARE M4: N°261982). Additional funding from the German Ministry of Education and Research, the Max Planck Society for the Advancement of Science, the U.S. National Institute on Aging (U01_AG09740-13S2, P01_AG005842, P01_AG08291, P30_AG12815, R21_AG025169, Y1-AG-4553-01, IAG_BSR06-11, OGHA_04-064, HHSN271201300071C) and from various national funding sources is gratefully acknowledged (see www.share-project.org).

Funding: This work was supported by the five-year Ageing Trajectories of Health: Longitudinal Opportunities and Synergies (ATHLOS) project. The ATHLOS project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No

635316. DF's work has been supported by grant RTI2018-100927-J-I00 Retos Investigación from Ministerio de Ciencia e Innovación (MCI), by Marsden grant E2987-3648 (Royal Society of New Zealand), and by grant 2017 SGR 622 (GRBIO) from the Departament d'Economia i Coneixement de la Generalitat de Catalunya (Spain). This work, grant number RTI2018-100927-J-I00, is supported by the Ministerio de Ciencia e Innovación (MCI, Spain), by the Agencia Estatal de Investigación (AEI, Spain) and by the European Regional Development Fund FEDER (FEDER, UE).

Competing interests: The authors report no conflicts of interest.

Patient consent for publication: not required

Ethics approval: The study protocol was approved by the Committee on the Ethics of Clinical Research, CEIC Fundació Sant Joan de Déu (Protocol No: PIC-22-15). All data were anonymized and EHR confidentially was respected in accordance with national and international law.

Data sharing statement: The original studies data are available on their respective websites: the Study on Global Ageing and Adult Health - SAGE (<https://www.who.int/healthinfo/sage/en/>), the English Longitudinal Study of Ageing – ELSA (<https://www.elsa-project.ac.uk/>), and the Survey of Health, Ageing and Retirement in Europe – SHARE (<http://www.share-project.org/home0.html>). R codes for harmonizing the included variables, as well as the STATA codes for the performed analysis are available on <https://github.com/athlosproject/athlos-project.github.io>

Additional files:

Additional file 1: **Table S1.** Proportion of missingness in the variables of interest in Mexico; **Table S2.** Overall proportion of missingness for indicators of latent classes and variables included in the imputation model; **Table S3.** Proportion of missingness for indicators of latent classes and variables included in the imputation model for region 1 (SAGE study - Africa: Ghana, South Africa); **Table S4.** Proportion of missingness for indicators of latent classes and variables included in the imputation

1
2
3 model for region 2 (SAGE study - Eastern Asia: China); **Table S5.** Proportion of missingness for
4 indicators of latent classes and variables included in the imputation model for region 2 (SAGE study -
5 Southern Asia: India); **Table S6.** Proportion of missingness for indicators of latent classes and
6 variables included in the imputation model for region 2 (SAGE study - Eastern Europe: the Russian
7 Federation); **Table S7.** Proportion of missingness for indicators of latent classes and variables
8 included in the imputation model for region 2 (ELSA study- England); **Table S8.** Proportion of
9 missingness for indicators of latent classes and variables included in the imputation model for region
10 2 (SHARE study - Northern Europe: Denmark, Sweden); **Table S9.** Proportion of missingness for
11 indicators of latent classes and variables included in the imputation model for region 2 (SHARE study
12 - Southern Europe: Greece, Italy, Spain); **Table S10.** Proportion of missingness for indicators of latent
13 classes and variables included in the imputation model for region 2 (SHARE study - Western Europe:
14 Austria, Belgium, France, Germany, Israel, Netherlands, Switzerland).

15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31 Additional file 2: **Table S1.** Association between latent multimorbidity membership and outcomes in
32 individuals aged 50-64 and ≥ 65
33
34
35
36

37 **Author contributions:**

38
39 **IB:** Participated in the database management, drafted the paper, carried out the statistical analyses
40 and worked on the interpretation of data. She also gave final approval of the version to be published
41 and agreed to be accountable for all aspects of the work in ensuring that questions related to the
42 accuracy or integrity of any part of the work are appropriately investigated and resolved; **AS:**
43 Participated in the study design, database management, statistical support and critical revision of the
44 paper. He also gave final approval of the version to be published and agreed to be accountable for all
45 aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the
46 work are appropriately investigated and resolved; **LE:** Participated in the interpretation of data and
47 critical revision of the paper; **HN:** She participated in critical revision of the paper and gave final
48 approval of the version to be published and agreed to be accountable for all aspects of the work in
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 ensuring that questions related to the accuracy or integrity of any part of the work are appropriately
4 investigated and resolved; **MP**: Participated in the study design, and critical revision of the paper. He
5 also gave final approval of the version to be published and agreed to be accountable for all aspects of
6 the work in ensuring that questions related to the accuracy or integrity of any part of the work are
7 appropriately investigated and resolved; **DF**: Participated in the study design, database management,
8 statistical support and critical revision of the paper. He also gave final approval of the version to be
9 published and agreed to be accountable for all aspects of the work in ensuring that questions related
10 to the accuracy or integrity of any part of the work are appropriately investigated and resolved; **JMH**:
11 Participated in the study design, acquisition of data, interpretation of data and critical revision of the
12 paper. He also gave final approval of the version to be published and agreed to be accountable for all
13 aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the
14 work are appropriately investigated and resolved; **BO**: Participated in the acquisition of data, study
15 design, database management, and critical revision of the paper. She also gave final approval of the
16 version to be published and agreed to be accountable for all aspects of the work in ensuring that
17 questions related to the accuracy or integrity of any part of the work are appropriately investigated
18 and resolved.

References

- 1 WHO. Ageing and Health. <https://www.who.int/news-room/fact-sheets/detail/ageing-and-health> (11 December 2018, date last accessed)
- 2 Wang H, Naghavi M, Allen C, *et al*. Global, regional, and national life expectancy, all-cause mortality, and cause-specific mortality for 249 causes of death, 1980–2015: a systematic analysis for the Global Burden of Disease Study 2015. *Lancet* 2016;**388**:1459–544.
doi:10.1016/S0140-6736(16)31012-1

- 1
2
3 3 Multimorbidity: a priority for global health research. 2018.
4
5
6 4 Parekh AK, Barton MB. The challenge of multiple comorbidity for the us health care system.
7
8 *JAMA - J Am Med Assoc* 2010;**303**:1303–4. doi:10.1001/jama.2010.381
9
10
11 5 Mugisha J, Rosenbaum S, Ward PB, *et al.* Chronic physical conditions, multimorbidity and
12
13 physical activity across 46 low- and middle-income countries. *Int J Behav Nutr Phys Act*
14
15 2017;**14**. doi:10.1186/s12966-017-0463-5
16
17
18
19 6 Chatterji S, Byles J, Cutler D, *et al.* Health, functioning, and disability in older adults - Present
20
21 status and future implications. *Lancet* 2015;**385**:563–75. doi:10.1016/S0140-6736(14)61462-8
22
23
24 7 Arokiasamy P, Uttamacharya U, Jain K, *et al.* The impact of multimorbidity on adult physical
25
26 and mental health in low- and middle-income countries: What does the study on global
27
28 ageing and adult health (SAGE) reveal? *BMC Med* 2015;**13**:1–16. doi:10.1186/s12916-015-
29
30 0402-8
31
32
33
34 8 Bao J, Chua KC, Prina M, *et al.* Multimorbidity and care dependence in older adults: A
35
36 longitudinal analysis of findings from the 10/66 study. *BMC Public Health* 2019;**19**:1–10.
37
38 doi:10.1186/s12889-019-6961-4
39
40
41
42 9 Maximova T, Pati S, Biritwum RB, *et al.* The impact of multimorbidity on adult physical and
43
44 mental health in low- and middle-income countries: what does the study on global ageing and
45
46 adult health (SAGE) reveal? *BMC Med* 2015;**13**:1–16. doi:10.1186/s12916-015-0402-8
47
48
49 10 Afshar S, Roderick PJ, Kowal P, *et al.* *Applied demography and public health in the 21st*
50
51 *century*. Cham, Switzerland: : Springer International Publishing AG 2016.
52
53 doi:10.1249/mss.0000000000001652
54
55
56 11 Yen L, Valderas JM, McRae IS, *et al.* Multimorbidity and comorbidity of chronic diseases
57
58 among the senior Australians: prevalence and patterns. *PLoS One* 2014;**9**:e83783.
59
60

- 1
2
3 doi:10.1371/journal.pone.0083783
4
5
6 12 Baldo V, Corti MC, Boccuzzo G, *et al.* Multimorbidity patterns in high-need, high-cost elderly
7
8 patients. *PLoS One* 2018;**13**:e0208875. doi:10.1371/journal.pone.0208875
9
10
11 13 Larsen FB, Pedersen MH, Friis K, *et al.* A Latent class analysis of multimorbidity and the
12
13 relationship to socio-demographic factors and health-related quality of life. A national
14
15 population-based study of 162,283 Danish Adults. *PLoS One* 2017;**12**:1–17.
16
17 doi:10.1371/journal.pone.0169426
18
19
20
21 14 Alcalá-Nalvaiz JT, Sicras-Mainar A, Calderón-Larrañaga A, *et al.* Multimorbidity Patterns in
22
23 Primary Care: Interactions among Chronic Diseases Using Factor Analysis. *PLoS One*
24
25 2012;**7**:e32190. doi:10.1371/journal.pone.0032190
26
27
28
29 15 Guisado-Clavero M, Roso-Llorach A, López-Jimenez T, *et al.* Multimorbidity patterns in the
30
31 elderly: A prospective cohort study with cluster analysis. *BMC Geriatr* 2018;**18**:1–11.
32
33 doi:10.1186/s12877-018-0705-7
34
35
36 16 Sanchez-Niubo A, Egea-Cortés L, Olaya B *et al.* Cohort profile: the ageing trajectories of health
37
38 - longitudinal opportunities and synergies (ATHLOS) project. *Int J Epidemiol* 2019;**0**:1–11.
39
40 doi:org/10.1093/ije/dyz077
41
42
43
44 17 Kowal P, Chatterji S, Naidoo N, *et al.* Data resource profile: The World Health Organization
45
46 study on global AGEing and adult health (SAGE). *Int J Epidemiol* 2013;**41**:1639–49.
47
48 doi:10.1093/ije/dys210
49
50
51 18 Steptoe A, Breeze E, Banks J, *et al.* Cohort profile: The English Longitudinal Study of Ageing. *Int*
52
53 *J Epidemiol* 2013;**42**:1640–8. doi:10.1093/ije/dys168
54
55
56 19 Börsch-Supan A, Brandt M, Hunkler C, *et al.* Data resource profile: The survey of health,
57
58 ageing and retirement in europe (SHARE). *Int J Epidemiol* 2013;**42**:992–1001.
59
60

- 1
2
3 doi:10.1093/ije/dyt088
4
5
6 20 IDA TWBI. World Bank Open Data. <https://data.worldbank.org/> (5 March 2019, date last
7
8 accessed)
9
10
11 21 Fortier I, Raina P, Van den Heuvel ER, Griffith LE, Craig C, Saliba M, Doiron D, Stolk RP,
12
13 Knoppers BM, Ferretti V, Granda P BP. Maelstrom Research guidelines for rigorous
14
15 retrospective data harmonization. 2017;;103–5. doi:10.1093/ije/dyw075
16
17
18 22 Ronald C. Kessler TBÜ. The World Mental Health (WMH) survey initiative version of the World
19
20 Health Organization (WHO) Composite Diagnostic Interview (CIDI). *Int J Methods Psychiatr Res*
21
22 2004;**13**:93–121. doi:org/10.1002/mpr.168
23
24
25
26 23 Radloff L S. The CES-D scale: A self report depression scale for research in the general
27
28 population. *Appl Psychol Meas* 1977;**1**:385–401. doi:10.1002/ardp.19552881111
29
30
31
32 24 Reischies F, Lobo A, Turrina C, *et al.* Development of the EURO–D scale – a European Union
33
34 initiative to compare symptoms of depression in 14 European centres. *Br J Psychiatry*
35
36 2008;**174**:330–8. doi:10.1192/bjp.174.4.330
37
38
39 25 Zivin K, Llewellyn DJ, Lang IA, *et al.* Depression among older adults in the United States and
40
41 England. *Am J Geriatr Psychiatry* 2010;**18**:1036–44. doi:10.1097/JGP.0b013e3181dba6d2
42
43
44 26 Prince MJ, Beekman ATF, Deeg DJH, *et al.* Depression symptoms in late life assessed during
45
46 the EURO-D scale. Effect of age, gender and marital status in 14 European centres. *Br J*
47
48 *Psychiatry* 1999;**174**:339–45. doi:10.1192/bjp.174.4.339
49
50
51
52 27 Morris JC, Heyman A, Mohs RC, Hughes JP, van Belle G, Fillenbaum G *et al.* The consortium to
53
54 establish a registry for Alzheimer’s disease (CERAD). Part I. Clinical and neuropsychological
55
56 assesment of Alzheimer’s disease. *Neurology* 1989;**39**:1159–65. doi:10.1212/wnl.39.9.1159
57
58
59
60 28 WHO. Country groupings.

- 1
2
3 https://www.who.int/quantifying_ehimpacts/global/ebdcountgroup/en/ (7 February 2019,
4
5 date last accessed)
6
7
8
9 29 United Nations Statistical Division (UNSD). Countries or areas/geographical regions.
10
11 <https://unstats.un.org/unsd/methodology/m49/> (7 February 2019, date last accessed)
12
13
14 30 Lanza ST, Dziak JJ, Huang L WA, Collins L. *LCA Stata plugin users' guide (version 1.2.1)*. Penn
15
16 State: : University Park: The Methodology Center 2018.
17
18
19 31 Tein JY, Coxe S, Cham H. Statistical power to detect the correct number of classes in latent
20
21 profile analysis. *Struct Equ Model* 2013;**20**:640–57. doi:10.1080/10705511.2013.824781
22
23
24 32 Nylund-Gibson K, Choi AY. Ten frequently asked questions about latent class analysis. *Transl*
25
26 *Issues Psychol Sci* 2018;**4**:440–61. doi:10.1037/tps0000176
27
28
29 33 Masyn KE. Latent Class Analysis and Finite Mixture Modeling. Published Online First: 21 March
30
31 2013. doi:10.1093/OXFORDHB/9780199934898.013.0025
32
33
34 34 StataCorp. *Stata: Release 13. Statistical Software*. College St. 2013.
35
36
37
38 35 Rubin DB. *Multiple Imputation for nonresponse in surveys*. United States of America: : John
39
40 Wiley & Sons 1987. doi:10.2307/3172772
41
42
43 36 Garin N, Olaya B, Perales J, *et al*. Multimorbidity patterns in a national representative sample
44
45 of the Spanish adult population. *PLoS One* 2014;**9**:e84794. doi:10.1371/journal.pone.0084794
46
47
48 37 Olaya B, Moneta MV, Caballero FF, *et al*. Latent class analysis of multimorbidity patterns and
49
50 associated outcomes in Spanish older adults: A prospective cohort study. *BMC Geriatr*
51
52 2017;**17**. doi:10.1186/s12877-017-0586-1
53
54
55
56 38 Whitson HE, Johnson KS, Sloane R, *et al*. Identifying patterns of multimorbidity in older
57
58 americans: application of latent class analysis. *J Am Geriatr Soc* 2016;**64**:1668–73.
59
60

- 1
2
3 doi:10.1111/jgs.14201. Identifying
4
5
6 39 Prados-Torres A, Calderón-Larrañaga A, Hanco-Saavedra J, *et al.* Multimorbidity patterns: A
7
8 systematic review. *J Clin Epidemiol* 2014;**67**:254–66. doi:10.1016/j.jclinepi.2013.09.021
9
10
11 40 Alberti KGM, Zimmet P, Shaw J. Metabolic syndrome -a new world-wide definition. A
12
13 consensus statement from the International Diabetes Federation. *Abstr Appl Anal*
14
15 2006;**23**:469–80. doi:10.1155/2010/307409
16
17
18
19 41 Margaretten M, Julian L, Katz P, *et al.* Depression in patients with rheumatoid arthritis:
20
21 description, causes and mechanisms. *Int J Clin Rheumtol* 2011;**6**:617–23.
22
23 doi:10.2217/IJR.11.6. Depression
24
25
26 42 Druss BG, Walker ER. Mental disorders and medical comorbidity. 2011.
27
28 [https://www.rwjf.org/content/dam/farm/reports/issue_briefs/2011/rwjf69438/subassets/rw](https://www.rwjf.org/content/dam/farm/reports/issue_briefs/2011/rwjf69438/subassets/rwjf69438_1)
29
30 [jf69438_1](https://www.rwjf.org/content/dam/farm/reports/issue_briefs/2011/rwjf69438/subassets/rwjf69438_1)
31
32
33
34 43 Barbara Messner DB. Smoking and cardiovascular disease: Mechanisms of endothelial
35
36 dysfunction and early atherogenesis. 2014;**34**:509–15. doi:10.1161/ATVBAHA.113.300156
37
38
39 44 Meade T, Manolios N, Cumming SR, *et al.* Cognitive impairment in rheumatoid arthritis: A
40
41 systematic review. *Arthritis Care Res* 2018;**70**:39–52. doi:10.1002/acr.23243
42
43
44 45 Lam RW, Kennedy SH, McIntyre RS, *et al.* Cognitive dysfunction in major depressive disorder:
45
46 effects on psychosocial functioning and implications for treatment. *Can J Psychiatry*
47
48 2014;**59**:649–54. doi:10.1177/070674371405901206
49
50
51
52 46 Dodd JW. Lung disease as a determinant of cognitive decline and dementia. *Alzheimer's Res*
53
54 *Ther* 2015;**7**:1–8. doi:10.1186/s13195-015-0116-3
55
56
57 47 Rose Mayeda E, Whitmer RA, Yaffe K. Diabetes and cognition. *Clin Geriatr Med* 2015;**31**:101–
58
59 ix. doi:10.1016/j.cger.2014.08.021
60

- 1
2
3 48 Walker KA, Power MC, Gottesman RF. Defining the relationship between hypertension,
4 cognitive decline, and dementia: a review. 2018;**19**:24. doi:10.1007/s11906-017-0724-3
5
6
7
8
9 49 Deckers K, Schievink SHJ, Rodriquez MMF, *et al.* Coronary heart disease and risk for cognitive
10 impairment or dementia: Systematic review and meta-analysis. *PLoS One* 2017;**12**:1–15.
11
12 doi:10.1371/journal.pone.0184244
13
14
15
16 50 Leon DA, Shkolnikov VM, McKee M, *et al.* Alcohol increases circulatory disease mortality in
17 Russia: Acute and chronic effects or misattribution of cause? *Int J Epidemiol* 2010;**39**:1279–90.
18
19 doi:10.1093/ije/dyq102
20
21
22
23 51 Petrukhin IS, Lunina EY. Cardiovascular disease risk factors and mortality in Russia: Challenges
24 and barriers. *Public Health Rev* 2012;**33**:436–49. doi:10.1007/BF03391645
25
26
27
28
29 52 Y. He, M. Zhang, E. H. B. Lin *et al.* Mental disorders among persons with arthritis: results from
30 the World Mental Health Surveys. 2008;**38**:1639–50. doi:10.1017/S0033291707002474
31
32
33
34 53 Jamrozik E, Musk AW. Respiratory health issues in the Asia-Pacific region: An overview.
35
36 *Respirology* 2011;**16**:3–12. doi:10.1111/j.1440-1843.2010.01844.x
37
38
39
40 54 Sandeep Grover, Alakananda Dutt AA. An overview of Indian research in depression. *Indian J*
41
42 *Psychiatry* 2010;**52**:178–88. doi:10.4103/0019-5545.69231
43
44
45 55 Gu L, Xie J, Long J, *et al.* Epidemiology of major depressive disorder in Mainland China: A
46
47 Systematic review. *PLoS One* 2013;**8**:e65356. doi:10.1371/journal.pone.0065356
48
49
50 56 Hou Z, Jiang W, Yin Y, *et al.* The current situation on major depressive disorder in China:
51
52 Research on mechanisms and clinical practice. *Neurosci Bull* 2016;**32**:389–97.
53
54 doi:10.1007/s12264-016-0037-6
55
56
57 57 Castro-Costa E, Dewey M, Stewart R, *et al.* Ascertaining late-life depressive symptoms in
58
59 Europe: An evaluation of the survey version of the EURO-D scale in 10 nations. The SHARE
60

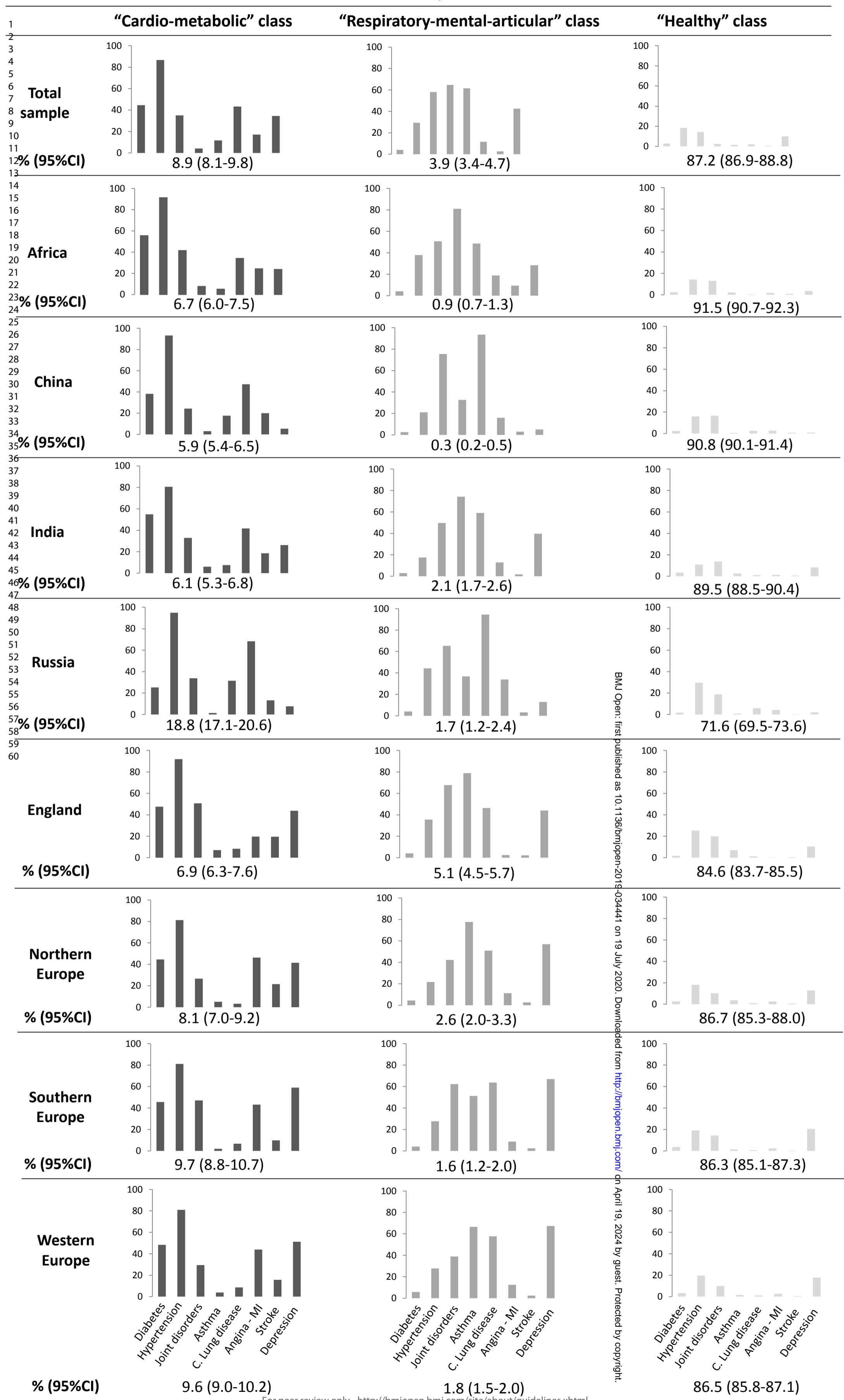
- 1
2
3 project. *Int J Methods Psychiatr Res* 2008;**17**:12–29. doi:10.1002/mpr.236
4
5
6 58 Huntley AL, Johnson R, Purdy S, *et al*. Measures of multimorbidity and morbidity burden for
7 use in primary care and community settings: A systematic review and guide. *Ann Fam Med*
8 2012;**10**:134–41. doi:10.1370/afm.1363
9
10
11
12
13 59 Dal Grande E, Fullerton S, Taylor AW. Reliability of self-reported health risk factors and
14 chronic conditions questions collected using the telephone in South Australia, Australia. *BMC*
15 *Med Res Methodol* 2012;**12**. doi:10.1186/1471-2288-12-108
16
17
18
19
20
21 60 Ortolá-Devesa JB, Egea-Caparrós JM, Huerta JM, *et al*. Accuracy of self-reported diabetes,
22 hypertension, and hyperlipidemia in the adult spanish population. DINO study findings. *Rev*
23 *Española Cardiol (English Ed)* 2009;**62**:143–52. doi:10.1016/s1885-5857(09)71532-4
24
25
26
27
28
29 61 Martin LM, Leff M, Calonge N, *et al*. Validation of self-reported chronic conditions and health
30 services in a managed care population. *Am J Prev Med* 2000;**18**:215–8. doi:10.1016/S0749-
31 3797(99)00158-0
32
33
34
35
36 62 Finch H. A comparison of statistics for assessing model invariance in latent class analysis. *Open*
37 *J Stat* 2015;**5**:191–210. doi:10.4236/ojs.2015.53022
38
39
40
41
42 63 White IR, Carlin JB. Bias and efficiency of multiple imputation compared with complete-case
43 analysis for missing covariate values. *Stat Med* 2010;**29**:2920–31. doi:10.1002/sim.3944
44
45
46
47 64 Sterne JAC, White IR, Carlin JB, *et al*. Multiple imputation for missing data in epidemiological
48 and clinical research: Potential and pitfalls. *BMJ* 2009;**339**:157–60. doi:10.1136/bmj.b2393
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 **Figure 1** Prevalence of diseases in the three latent classes in the total sample and by regions (subsample 50-64
4
5 years)

6
7
8 **Figure 2** Prevalence of diseases in the three latent classes in the total sample and by regions (subsample +65
9
10 years)

11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For peer review only



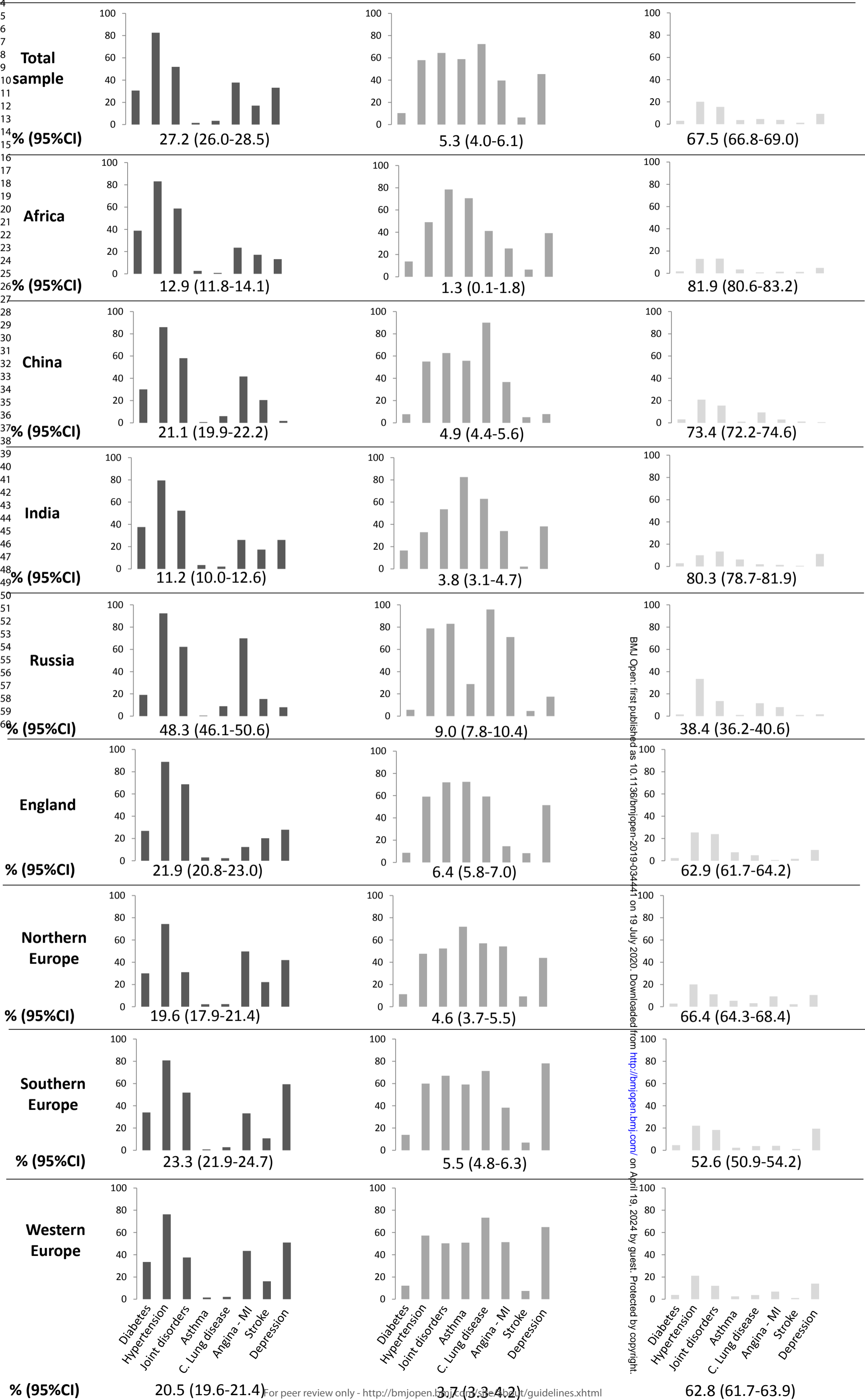
BMJ Open: first published as 10.1136/bmjopen-2019-034441 on 19 July 2020. Downloaded from <http://bmjopen.bmj.com/> on April 19, 2024 by guest. Protected by copyright.

Multimorbidity latent classes

“Cardio-metabolic” class

“Respiratory-mental-articular” class

“Healthy” class



BMJ Open: first published as 10.1136/bmjopen-2019-034441 on 19 July 2020. Downloaded from <http://bmjopen.bmj.com/> on April 19, 2024 by guest. Protected by copyright.

Table S1. Proportion of missingness in the variables of interest in Mexico

(a) Indicators of multimorbidity classes	
	n (%) missingness
Diabetes	2814 (51.65)
Hypertension	2815 (51.67)
Joint disorders	2814 (51.65)
Asthma	2814 (51.65)
Chronic lung disease	2814 (51.65)
Myocardial infarction – Angina	2814 (51.65)
Stroke	2814 (51.65)
Depression	2815 (51.67)
(b) Variables used in the regression model	
Sex	2706 (49.67)
Age	2707 (49.69)
Country	0 (0.00)
Study	0 (0.00)
Marital status	2811 (51.60)
Education	3264 (59.91)
Wealth	1922 (35.28)
Loneliness	2819 (51.74)
Ever smoked	2814 (51.65)
Vigorous exercise in last 2 weeks	2814 (51.65)
Self-rated health	2811 (51.6)
ADL – Using the toilet	2814 (51.65)
ADL – Bathing or showering	2815 (51.67)
ADL – Getting dressed	0 (0.00)
ADL - Eating	2814 (51.65)
ADL – Getting in or out of bed	2813 (51.63)
ADL – Moving around the house	2817 (51.71)
Memory: Immediate recall	2854 (52.39)
Memory: Delayed recall	2854 (52.39)
Verbal fluency	2852 (52.35)

Table S2. Overall proportion of missingness for indicators of latent classes and variables included in the imputation model

(a) Indicators of multimorbidity classes	
	n (%) missingness
Diabetes	62 (0.08)
Hypertension	154 (0.21)
Joint disorders	31 (0.04)
Asthma	82 (0.11)
Chronic lung disease	51 (0.07)
Myocardial infarction – Angina	63 (0.08)
Stroke	41 (0.05)
Depression	1233 (1.70)
(b) Variables used in the regression model	
Sex	0 (0.00)
Age	0 (0.00)
Country	0 (0.00)
Study	0 (0.00)
Marital status	106 (0.14)
Education	11475 (15.91)
Wealth	844 (1.17)
Loneliness	10985 (15.23)
Ever smoked	201 (0.27)
Vigorous exercise in last 2 weeks	258 (0.35)
Self-rated health	218 (0.30)
ADL – Using the toilet	355 (0.49)
ADL – Bathing or showering	340 (0.47)
ADL – Getting dressed	304 (0.42)
ADL - Eating	400 (0.55)
ADL – Getting in or out of bed	310 (0.42)
ADL – Moving around the house	345 (0.47)
Memory: Immediate recall	1442 (1.99)
Memory: Delayed recall	1448 (2.00)
Verbal fluency	1641 (2.27)

Table S3. Proportion of missingness for indicators of latent classes and variables included in the imputation model for region 1 (SAGE study - Africa: Ghana, South Africa)

(a) Indicators of multimorbidity classes	
	n (%) missingness
Diabetes	1 (0.01)
Hypertension	1 (0.01)
Joint disorders	3 (0.03)
Asthma	1 (0.01)
Chronic lung disease	1 (0.01)
Myocardial infarction – Angina	3 (0.03)
Stroke	1 (0.01)
Depression	17 (0.21)
(b) Variables used in the regression model	
Sex	0 (0.00)
Age	0 (0.00)
Country	0 (0.00)
Study	0 (0.00)
Marital status	89 (1.11)
Education	3705 (46.6)
Wealth	23 (0.28)
Loneliness	48 (0.60)
Ever smoked	7 (0.08)
Vigorous exercise in last 2 weeks	12 (0.15)
Self-rated health	15 (0.18)
ADL – Using the toilet	40 (0.50)
ADL – Bathing or showering	32 (0.40)
ADL – Getting dressed	27 (0.33)
ADL - Eating	31 (0.38)
ADL – Getting in or out of bed	32 (0.40)
ADL – Moving around the house	35 (0.44)
Memory: Immediate recall	36 (0.45)
Memory: Delayed recall	31 (0.38)
Verbal fluency	9 (0.11)

Table S4. Proportion of missingness for indicators of latent classes and variables included in the imputation model for region 2 (SAGE study - Eastern Asia: China)

(a) Indicators of multimorbidity classes	
	n (%) missingness
Diabetes	35 (0.27)
Hypertension	127 (0.98)
Joint disorders	2 (0.01)
Asthma	53 (0.41)
Chronic lung disease	23 (0.17)
Myocardial infarction – Angina	34 (0.26)
Stroke	13 (0.10)
Depression	146 (1.13)
(b) Variables used in the regression model	
Sex	0 (0.00)
Age	0 (0.00)
Country	0 (0.00)
Study	0 (0.00)
Marital status	8 (0.06)
Education	3105 (24.18)
Wealth	18 (0.14)
Loneliness	74 (0.57)
Ever smoked	2 (0.01)
Vigorous exercise in last 2 weeks	45 (0.35)
Self-rated health	16 (0.12)
ADL – Using the toilet	51 (0.39)
ADL – Bathing or showering	52 (0.40)
ADL – Getting dressed	41 (0.31)
ADL - Eating	46 (0.35)
ADL – Getting in or out of bed	36 (0.28)
ADL – Moving around the house	48 (0.37)
Memory: Immediate recall	317 (2.46)
Memory: Delayed recall	399 (3.10)
Verbal fluency	335 (2.60)

Table S5. Proportion of missingness for indicators of latent classes and variables included in the imputation model for region 2 (SAGE study - Southern Asia: India)

(a) Indicators of multimorbidity classes	
	n (%) missingness
Diabetes	0 (0.00)
Hypertension	1 (0.01)
Joint disorders	1 (0.01)
Asthma	0 (0.00)
Chronic lung disease	0 (0.00)
Myocardial infarction – Angina	0 (0.00)
Stroke	1 (0.01)
Depression	1 (0.01)
(b) Variables used in the regression model	
Sex	0 (0.00)
Age	0 (0.00)
Country	0 (0.00)
Study	0 (0.00)
Marital status	0 (0.00)
Education	0 (0.00)
Wealth	38 (0.57)
Loneliness	10 (0.15)
Ever smoked	1 (0.01)
Vigorous exercise in last 2 weeks	0 (0.00)
Self-rated health	0 (0.00)
ADL – Using the toilet	19 (0.28)
ADL – Bathing or showering	12 (0.18)
ADL – Getting dressed	11 (0.16)
ADL - Eating	73 (1.11)
ADL – Getting in or out of bed	7 (0.10)
ADL – Moving around the house	24 (0.36)
Memory: Immediate recall	88 (1.34)
Memory: Delayed recall	88 (1.34)
Verbal fluency	71 (1.08)

Table S6. Proportion of missingness for indicators of latent classes and variables included in the imputation model for region 2 (SAGE study - Eastern Europe: the Russian Federation)

(a) Indicators of multimorbidity classes	
	n (%) missingness
Diabetes	5 (0.12)
Hypertension	4 (0.10)
Joint disorders	3 (0.07)
Asthma	6 (0.15)
Chronic lung disease	5 (0.12)
Myocardial infarction – Angina	5 (0.12)
Stroke	5 (0.12)
Depression	14 (0.36)
(b) Variables used in the regression model	
Sex	0 (0.00)
Age	0 (0.00)
Country	0 (0.00)
Study	0 (0.00)
Marital status	6 (0.15)
Education	41 (1.05)
Wealth	5 (0.12)
Loneliness	79 (2.03)
Ever smoked	1 (0.02)
Vigorous exercise in last 2 weeks	0 (0.00)
Self-rated health	5 (0.12)
ADL – Using the toilet	47 (1.20)
ADL – Bathing or showering	46 (1.18)
ADL – Getting dressed	27 (0.69)
ADL - Eating	52 (1.33)
ADL – Getting in or out of bed	37 (0.95)
ADL – Moving around the house	40 (1.02)
Memory: Immediate recall	153 (3.93)
Memory: Delayed recall	81 (2.08)
Verbal fluency	208 (5.35)

Table S7. Proportion of missingness for indicators of latent classes and variables included in the imputation model for region 2 (ELSA study- England)

(a) Indicators of multimorbidity classes	
	n (%) missingness
Diabetes	2 (0.01)
Hypertension	2 (0.01)
Joint disorders	3 (0.02)
Asthma	3 (0.02)
Chronic lung disease	3 (0.02)
Myocardial infarction – Angina	2 (0.01)
Stroke	2 (0.01)
Depression	372 (3.23)
(b) Variables used in the regression model	
Sex	0 (0.00)
Age	0 (0.00)
Country	0 (0.00)
Study	0 (0.00)
Marital status	0 (0.00)
Education	1023 (8.88)
Wealth	330 (2.86)
Loneliness	356 (3.09)
Ever smoked	173 (1.50)
Vigorous exercise in last 2 weeks	176 (1.52)
Self-rated health	175 (1.51)
ADL – Using the toilet	173 (1.50)
ADL – Bathing or showering	173 (1.50)
ADL – Getting dressed	173 (1.50)
ADL - Eating	173 (1.50)
ADL – Getting in or out of bed	173 (1.50)
ADL – Moving around the house	173 (1.50)
Memory: Immediate recall	360 (3.12)
Memory: Delayed recall	373 (3.23)
Verbal fluency	362 (3.14)

Table S8. Proportion of missingness for indicators of latent classes and variables included in the imputation model for region 2 (SHARE study - Northern Europe: Denmark, Sweden)

(a) Indicators of multimorbidity classes	
	n (%) missingness
Diabetes	2 (0.04)
Hypertension	2 (0.04)
Joint disorders	2 (0.04)
Asthma	2 (0.04)
Chronic lung disease	2 (0.04)
Myocardial infarction – Angina	2 (0.04)
Stroke	2 (0.04)
Depression	78 (1.70)
(b) Variables used in the regression model	
Sex	0 (0.00)
Age	0 (0.00)
Country	0 (0.00)
Study	0 (0.00)
Marital status	0 (0.00)
Education	52 (1.13)
Wealth	6 (0.13)
Loneliness	1384 (30.26)
Ever smoked	4 (0.08)
Vigorous exercise in last 2 weeks	6 (0.13)
Self-rated health	2 (0.04)
ADL – Using the toilet	3 (0.06)
ADL – Bathing or showering	3 (0.06)
ADL – Getting dressed	3 (0.06)
ADL - Eating	3 (0.06)
ADL – Getting in or out of bed	3 (0.06)
ADL – Moving around the house	3 (0.06)
Memory: Immediate recall	68 (1.48)
Memory: Delayed recall	66 (1.44)
Verbal fluency	83 (1.81)

Table S9. Proportion of missingness for indicators of latent classes and variables included in the imputation model for region 2 (SHARE study - Southern Europe: Greece, Italy, Spain)

(a) Indicators of multimorbidity classes	
	n (%) missingness
Diabetes	12 (0.16)
Hypertension	12 (0.16)
Joint disorders	12 (0.16)
Asthma	12 (0.16)
Chronic lung disease	12 (0.16)
Myocardial infarction – Angina	12 (0.16)
Stroke	12 (0.16)
Depression	194 (2.59)
(b) Variables used in the regression model	
Sex	0 (0.00)
Age	0 (0.00)
Country	0 (0.00)
Study	0 (0.00)
Marital status	1 (0.01)
Education	20 (0.26)
Wealth	95 (1.27)
Loneliness	2715 (36.37)
Ever smoked	3 (0.04)
Vigorous exercise in last 2 weeks	4 (0.05)
Self-rated health	0 (0.00)
ADL – Using the toilet	3 (0.04)
ADL – Bathing or showering	3 (0.04)
ADL – Getting dressed	3 (0.04)
ADL - Eating	3 (0.04)
ADL – Getting in or out of bed	3 (0.04)
ADL – Moving around the house	3 (0.04)
Memory: Immediate recall	125 (1.67)
Memory: Delayed recall	125 (1.67)
Verbal fluency	176 (2.35)

Table S10. Proportion of missingness for indicators of latent classes and variables included in the imputation model for region 2 (SHARE study - Western Europe: Austria, Belgium, France, Germany, Israel, Netherlands, Switzerland)

(a) Indicators of multimorbidity classes	
	n (%) missingness
Diabetes	5 (0.02)
Hypertension	5 (0.02)
Joint disorders	5 (0.02)
Asthma	5 (0.02)
Chronic lung disease	5 (0.02)
Myocardial infarction – Angina	5 (0.02)
Stroke	5 (0.02)
Depression	411 (2.36)
(b) Variables used in the regression model	
Sex	0 (0.00)
Age	0 (0.00)
Country	0 (0.00)
Study	0 (0.00)
Marital status	2 (0.01)
Education	165 (0.95)
Wealth	329 (1.89)
Loneliness	6319 (36.42)
Ever smoked	10 (0.05)
Vigorous exercise in last 2 weeks	15 (0.08)
Self-rated health	5 (0.02)
ADL – Using the toilet	19 (0.10)
ADL – Bathing or showering	19 (0.10)
ADL – Getting dressed	19 (0.10)
ADL - Eating	19 (0.10)
ADL – Getting in or out of bed	19 (0.10)
ADL – Moving around the house	19 (0.10)
Memory: Immediate recall	295 (1.7)
Memory: Delayed recall	285 (1.64)
Verbal fluency	397 (2.28)

Table S1. Association between latent multimorbidity membership and outcomes in individuals aged 50-64 and ≥65

Outcomes ^a	Aged 50-64		Aged ≥65	
	“Cardio-metabolic” class	“Respiratory-mental-articular” class	“Cardio-metabolic” class	“Respiratory-mental-articular” class
Sex				
Woman	1.0	1.0	1.0	1.0
Male	0.9 (0.8-1.0)	0.6 (0.5-0.7)	0.7 (0.7-0.8)	0.9 (0.8-1.0)
Age (years)	1.1 (1.1-1.1)	1.1 (1.0-1.1)	1.0 (1.0-1.0)	1.0 (1.0-1.0)
Marital status				
Single	1.0	1.0	1.0	1.0
Married	0.9 (0.8-1.1)	0.7 (0.5-1.0)	1.2 (1.00-1.4)	1.2 (0.9-1.6)
Divorced	1.1 (0.9-1.4)	1.7 (1.2-2.4)	1.1 (1.0-1.4)	1.8 (1.2-2.6)
Widowed	1.4 (1.1-1.7)	1.1 (0.8-1.6)	1.5 (1.3-1.8)	1.6 (1.2-2.3)
Education level				
Primary or less	1.0	1.0	1.0	1.0
Secondary	0.9 (0.9-1.0)	0.7 (0.6-0.8)	1.0 (0.9-1.0)	0.7 (0.6-0.8)
Tertiary	0.8 (0.7-0.9)	0.6 (0.5-0.7)	0.9 (0.8-1.0)	0.6 (0.5-0.7)
Wealth				
1 st (worse)	1.0	1.0	1.0	1.0
2 nd	1.0 (0.9-1.1)	0.8 (0.6-1.0)	1.1 (1.0-1.2)	1.1 (0.9-1.2)
3 rd	0.9 (0.9-1.1)	0.7 (0.5-0.8)	1.1 (1.0-1.2)	1.0 (0.9-1.2)
4 th	0.9 (0.8-1.0)	0.6 (0.5-0.8)	1.1 (1.0-1.2)	0.8 (0.7-1.0)
5 th (best)	0.8 (0.7-0.9)	0.5 (0.4-0.6)	1.2 (1.0-1.2)	0.7 (0.6-0.9)
Region				
Africa	1.0	1.0	1.0	1.0
China	0.9 (0.8-1.0)	0.3 (0.2-0.5)	1.8 (1.6-2.1)	4.1 (3.0-5.7)
India	0.9 (0.8-1.1)	2.2 (1.5-3.2)	0.9 (0.8-1.0)	2.9 (2.0-4.2)
Russia	3.6 (3.0-4.2)	2.2 (1.4-3.5)	8.0 (7.0-9.2)	14.5 (10.3-20.3)
England	1.1 (0.9-1.3)	5.6 (4.1-7.7)	2.2 (2.0-2.5)	6.2 (4.5-8.5)
Northern Europe	1.3 (1.1-1.5)	2.8 (1.9-4.1)	1.9 (1.6-2.2)	4.2 (2.9-6.1)
Southern Europe	1.5 (1.3-1.8)	1.7 (1.2-2.5)	2.8 (2.5-3.2)	6.5 (4.7-9.0)
Western Europe	1.5 (1.3-1.7)	1.9 (1.3-2.6)	2.1 (1.9-2.3)	3.6 (2.6-5.0)
Loneliness (yes/no)	1.9 (1.8-2.1)	2.5 (2.1-3.0)	1.4 (1.4-1.6)	1.9 (1.7-2.2)
Ever smoked (yes/no)	1.0 (0.9-1.1)	1.7 (1.5-2.0)	0.8 (0.8-0.8)	1.4 (1.3-1.6)
Physical activity (yes/no)	0.5 (0.5-0.5)	0.6 (0.5-0.7)	0.5 (0.5-0.6)	0.5 (0.4-0.6)
Limitations in ADL (yes/no)	2.5 (2.3-2.7)	3.2 (2.7-3.7)	1.8 (1.7-1.9)	2.8 (2.5-3.1)
Self-rated health				
Good	1.0	1.0	1.0	1.0
Moderate	3.7 (3.3-3.9)	2.5 (2.1-2.9)	2.5 (2.3-2.7)	3.9 (3.3-4.6)
Poor	8.5 (7.7-9.4)	7.4 (6.1-8.8)	4.3 (4.6)	10.9 (9.3-12.8)
Memory: Immediate recall	0.9 (0.9-0.9)	0.9 (0.9-1.0)	0.9 (0.9-0.9)	0.9 (0.9-0.9)
Memory: Delayed recall	0.9 (0.9-0.9)	0.9 (0.8-0.9)	0.9 (0.9-0.9)	0.9 (0.9-0.9)
Verbal fluency	1.0 (1.0-1.0)	1.0 (1.0-1.0)	1.0 (0.9-1.00)	1.0 (0.9-1.0)

Note: The reference group for the multimorbidity group variable was the “Healthy” class
 Relative Risk Ratios (95% confidence interval) from multinomial logistic regression models
 Models were run in 100 imputed datasets and results combined using Rubin’s rules
 ADL Activities of Daily Living
^aUnadjusted analyses.

1
2
3
4
5
6
7
8
9

STROBE CHECKLIST: “MULTIMORBIDITY PATTERNS IN LOW-MIDDLE AND HIGH INCOME REGIONS: A MULTI-REGION LATENT CLASS ANALYSIS USING ATHLOS HARMONIZED COHORTS”

10

1. Title and abstract

- 11 (a) Indicate the study’s design with a commonly used term in the title or the abstract
12 **(page 1 and 3)**
13 (b) Provide in the abstract an informative and balanced summary of what was done and
14 what was found **(page 3)**
15

16

Introduction

- 17
18 2. Background/rationale: Explain the scientific background and rationale for the
19 investigation being reported **(page 5-6)**
20 3. Objectives: State specific objectives, including any prespecified hypotheses **(page 6,**
21 **lines 10-13)**
22

23

Methods

- 24 4. Study design: Present key elements of study design early in the paper **(page 6, lines**
25 **16-19)**
26 5. Setting: Describe the setting, locations, and relevant dates, including periods of
27 recruitment, exposure, follow-up, and data collection **(page 6, lines 20-24, page 7, lines**
28 **1-3).**
29 6. Participants: Give the eligibility criteria, and the sources and methods of selection of
30 Participants **(page 7, lines 4-9).**
31 7. Variables: Clearly define all outcomes, exposures, predictors, potential confounders,
32 and effect modifiers. Give diagnostic criteria, if applicable. **(page 7, lines 13- page 8).**
33 8. Data sources/measurement: For each variable of interest, give sources of data and
34 details of methods of assessment (measurement). Describe comparability of assessment
35 methods if there is more than one group **(page 7, lines 13- page 8).**
36 9. Bias: Describe any efforts to address potential sources of bias **(page 9, lines 19-23).**
37 10. Study size: Explain how the study size was arrived at **(page 7, lines 4-9).**
38 11. Quantitative variables: Explain how quantitative variables were handled in the
39 analyses. If applicable, describe which groupings were chosen and why. **(page 7, lines**
40 **13- page 8).**
41 12. Statistical methods **(page 9-10):**
42
43 (a) Describe all statistical methods, including those used to control for
44 confounding
45 (b) Describe any methods used to examine subgroups and interactions
46 (c) Explain how missing data were addressed
47 (d) If applicable, describe analytical methods taking account of sampling
48 strategy
49 (e) Describe any sensitivity analyses
50
51
52
53
54

55

Results

- 56
57 13. Participants **(page 7, lines 4-9):**
58
59
60

- 1
- 2
- 3 (a) Report numbers of individuals at each stage of study—eg numbers potentially eligible,
- 4 examined for eligibility, confirmed eligible, included in the study, completing follow-up,
- 5 and analysed
- 6 (b) Give reasons for non-participation at each stage
- 7 (c) Consider use of a flow diagram
- 8
- 9

10 14. Descriptive data (**page 10, lines 8-18; page 11, lines 1-7**):

- 11 (a) Give characteristics of study participants (eg demographic, clinical, social) and
- 12 information on exposures and potential confounders
- 13 (b) Indicate number of participants with missing data for each variable of interest
- 14 (**Supplementary file 1**)
- 15
- 16

17 15. Outcome data: Report numbers of outcome events or summary measures (**pages 12-15**).

18 16. Main results (**page 15, lines 14-21; to page 17**):

- 19
- 20
- 21 (a) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their
- 22 precision (eg, 95% confidence interval). Make clear which confounders were adjusted
- 23 for and why they were included
- 24 (b) Report category boundaries when continuous variables were categorized
- 25 (c) If relevant, consider translating estimates of relative risk into absolute risk for a
- 26 meaningful time period
- 27

28 17. Other analyses: Report other analyses done—eg analyses of subgroups and interactions, and

29 sensitivity analyses. **Not applicable.**

30

31 **Discussion**

32 18. Key results: Summarise key results with reference to study objectives (**page 17, lines 5-8**).

33 19. Limitations: Discuss limitations of the study, taking into account sources of potential bias or

34 imprecision. Discuss both direction and magnitude of any potential bias (**page 21, lines 19-24,**

35 **page 22, 1-14**).

36 20. Interpretation: Give a cautious overall interpretation of results considering objectives,

37 limitations, multiplicity of analyses, results from similar studies, and other relevant evidence

38 (**pages 18-21**).

39 21. Generalisability: Discuss the generalisability (external validity) of the study results (**page 19,**

40 **13-15**).

41

42 **Other information**

43 22. Funding: Give the source of funding and the role of the funders for the present study and, if

44 applicable, for the original study on which the present article is based (**page 23, lines 20-22,**

45 **page 24, lines 1-6**).

46

47

48

49

50

51

52

53

54

55

56

57

58

59

60

BMJ Open

Multimorbidity patterns in low-middle and high income regions: A multi-region latent class analysis using ATHLOS harmonized cohorts.

Journal:	<i>BMJ Open</i>
Manuscript ID	bmjopen-2019-034441.R2
Article Type:	Original research
Date Submitted by the Author:	17-May-2020
Complete List of Authors:	Bayes-Marin, Iveta; Parc Sanitari Sant Joan de Déu, Research, Innovation and Teaching Unit; Centro de Investigación Biomédica en Red de Salud Mental, CIBERSAM Sanchez-Niubo, Albert ; Parc Sanitari Sant Joan de Déu, Research, Innovation and Teaching Unit; Centro de Investigación Biomédica en Red de Salud Mental, CIBERSAM Egea-Cortés, Laia; Parc Sanitari Sant Joan de Déu, Research, Innovation and Teaching Unit Nguyen, Hai; King's College London, Health Service and Population Research Department Prina, A. Matthew; King's College London, Fernández, Daniel; Parc Sanitari Sant Joan de Déu, Research, Innovation and Teaching Unit; CIBERSAM Haro, Josep; Parc Sanitari Sant Joan de Deu, CIBERSAM, University of Barcelona; King Sand University Olaya, Beatriz; Parc Sanitari Sant Joan de Deu, Research and Development Unit
Primary Subject Heading:	Epidemiology
Secondary Subject Heading:	Global health
Keywords:	multimorbidity, non-communicable diseases (NCDs), low- and middle-income countries (LMICs), high-income country (HICs), Latent class analysis (LCA)

SCHOLARONE™
Manuscripts



I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our [licence](#).

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which [Creative Commons](#) licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

Original Research Article**Manuscript title:**

Multimorbidity patterns in low-middle and high income regions: A multi-region latent class analysis using ATHLOS harmonized cohorts.

Author's full names and affiliations:

Ivet Bayes-Marin, ^{a,b,c} Albert Sanchez-Niubo, ^{a,b} Laia Egea-Cortés, ^d Hai Nguyen, ^e Matthew Prina, ^e Daniel Fernández, ^{a,b,f} Josep Maria Haro, ^{a,b,c} and Beatriz Olaya ^{a,b}

From the ^aResearch, Innovation and Teaching Unit, Parc Sanitari Sant Joan de Déu, Sant Boi de Llobregat, Spain, ^bCentro de Investigación Biomédica en Red de Salud Mental, CIBERSAM, Madrid, Spain, ^cDepartment of Medicine, Universitat de Barcelona, Barcelona, Spain, ^dCenter for Epidemiological Studies on Sexually Transmitted Infections and HIV/AIDS of Catalonia (CEEISCAT), Catalonia Public Health Agency (ASPCAT), Badalona, Spain, ^eHealth Service and Population Research Department, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, United Kingdom, ^fSerra Hünter fellow. Department of Statistics and Operations Research, Polytechnic University of Catalonia, Barcelona, Spain.

Corresponding author details:

Correspondence: Albert Sanchez- Niubo. Research, Innovation and Teaching Unit, Parc Sanitari Sant Joan de Déu, Dr Antoni Pujadas 42, Sant Boi de Llobregat, Barcelona 08830, Spain. E-mail:

albert.sanchez@pssjd.org

ORCID of the authors:

Ivet Bayes-Marin - 0000-0002-3816-5244

1
2
3 Albert Sanchez-Niubo - 0000-0003-0309-181X
4

5 Laia Egea-Cortés - 0000-0003-4002-8254
6

7 Hai Nguyen - 0000-0003-2171-1955
8

9 Matthew Prina - 0000-0001-6698-3263
10

11 Daniel Fernández - 0000-0003-0012-2094
12

13 Josep Maria Haro - 0000-0002-3984-277X
14

15 Beatriz Olaya - 0000-0003-2046-3929
16
17
18
19
20

21 **Word count (excluding title page, abstract, references, figures and tables): 5183**
22
23

24 25 **Abstract**

26
27 **Objectives:** Our aim was to determine clusters of non-communicable diseases (NCDs) in a very large,
28 population-based sample of middle-aged and older adults from low- and middle-income (LMICs) and
29 high-income (HICs) regions. Additionally, we explored the associations with several covariates.
30
31

32
33 **Design:** The total sample was 72 140 people aged 50+ from three population-based studies (ELSA,
34 SHARE and SAGE) included in the ATHLOS project and representing eight regions with LMICs and
35 HICs. Variables were previously harmonized using an ex-post strategy. Eight NCDs were used in latent
36 class analysis. Multinomial models were made to calculate associations with covariates. All the
37 analyses were stratified by age (50-64 and 65+ years old).
38
39

40
41 **Results:** Three clusters were identified: “cardio-metabolic” (8.93% in participants aged 50-64 and
42 27.22% in those aged 65+), “respiratory-mental-articular” (3.91% and 5.27%) and “healthy” (87.16%
43 and 67.51%). In the younger group, Russia presented the highest prevalence of the “cardio-
44 metabolic” group (18.8%) and England the “respiratory-mental-articular” (5.1%). In the older group,
45 Russia had the highest proportion of both classes (48.3% and 9%). Both the younger and older
46 African participants presented the highest proportion of the “healthy” class. Older age, being
47 woman, widowed, and with low levels of education and income were related to an increased risk of
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 multimorbidity. Physical activity was a protective factor in both age groups and smoking a risk factor
4
5 for the “respiratory-mental-articular”.

6
7 **Conclusions:** Multimorbidity is common worldwide, especially in HICs and Russia. Health policies in
8
9 each country addressing coordination and support are needed to face the complexity of a pattern of
10
11 growing multimorbidity.
12
13

14
15
16 **Keywords:** multimorbidity, non-communicable diseases (NCDs), low- and middle- income countries
17
18 (LMICs), high-income countries (HICs), latent class analysis (LCA).
19
20
21

22 23 **Strengths and limitations of this study**

- 24
25
26 • This study used a large, harmonized, multi-regional database, which allowed us to compare
27
28 two age groups as well as disease prevalence in regions with differing incomes.
- 29
30
31 • The presence or absence of the non-communicable diseases was based on self-reported
32
33 measures, and thus might be affected by measurement errors or lack of accuracy.
- 34
35
36 • Only common diseases across studies were included in the analyses, so this might have led to
37
38 a smaller number of latent classes, or to different clusters.
- 39
40
41 • When performing Latent Class Analysis, we forced the solution as we aimed to do comparisons
42
43 among age subsamples and regions in terms of disease prevalence as well as protective and
44
45 risk factors.
- 46
47
48 • The use of multiple imputations for missing data in the covariates could carry some bias.

49 50 **Background**

51
52 By 2050, the population aged 60 years and older is expected to reach 2 billion worldwide, compared
53
54 to 900 million in 2015 [1]. Along with this rapid increase, the incidence of chronic conditions (CC) or
55
56 non-communicable diseases (NCDs) is also on the rise, having become the leading cause of morbidity
57
58 and disability worldwide [2].
59
60

1
2
3 Multimorbidity, defined as the co-existence of two or more CC, is more common in older adults and is
4 often more prevalent in people of lower socioeconomic status [3]. Multimorbidity is thought to
5 account for 65% of total health care expenses in high-income countries (HICs) because of the huge
6 associated healthcare utilization [4]. Due to the increasing prevalence of multimorbidity, the managing
7 of multiple conditions has become an unavoidable international research priority, because of the high
8 impact on the quality of life of patients and caregivers and on healthcare systems [3].
9

10
11 Most studies on the prevalence of multimorbidity in older people come from HICs, while data from
12 middle-aged adults and low- and middle- income countries (LMICs) are much more limited [5–8]. LMICs
13 are experiencing an increase in life expectancy that, together with changes in lifestyle and environment
14 exposures, are triggering changes in their disease burden profile [3,9]. Few studies have compared
15 patterns of multimorbidity between HICs and LMICs. Afshar et al. [10] used population-based chronic
16 disease data from the World Health Survey to compare multimorbidity prevalence across 27 LMICs
17 and one HIC, and used gross domestic product (GDP) to study inter-country socioeconomic differences.
18 They found high multimorbidity prevalence in all countries, and a positive but non-linear relationship
19 between country GDP and multimorbidity prevalence, suggesting the influence of other factors, such
20 as lifestyles, social conditions, and differences across health systems. Four latent classes were
21 identified in a cross-sectional sample of Australian seniors aged 50 years and over, using self-reported
22 diagnosis of eleven conditions, including cancer and Parkinson's disease [11]. Another study, focusing
23 on complex health care needs of Italian elderly people, found five clusters using 15 diseases [12]. A
24 study conducted in a sample of 162,283 people from a survey of Danish population identified seven
25 latent classes considering 15 chronic diseases and seven age groups, ranging from 16 to 104 years [13].
26 These differences could be explained in light of variations in collection methods, data sources,
27 populations, diseases included, and the analysis performed [11,14,15].
28

29
30 Similarly, the lack of study of differences in multimorbidity between HICs and LMICs may be due to the
31 use of different methodologies, which might hinder comparisons of prevalence and multimorbidity
32 patterns across countries. The integration of data from different studies would allow us to determine
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 differences across regions and cohorts, as well as to explore risk and protective factors involved in the
4 clustering of CC, thereby improving our understanding of the problem and the creation of adapted
5 medical guidelines.
6
7
8
9

10
11
12 This study aimed to: a) identify multimorbidity clusters in middle-aged (50-64) and older adults (+65)
13 from different regions, classified as LMICs and HICs; b) investigate the associations between
14 multimorbidity clusters and sociodemographic, economic, lifestyles and health status variables; and c)
15 explore differences across regions.
16
17
18
19
20
21
22

23 **Methods**

24 *Study design and data extraction*

25
26 The present study used data from the Ageing Trajectories of Health: Longitudinal Opportunities and
27 Synergies (ATHLOS) project [16]. Longitudinal data from 17 international cohort studies related to
28 health and ageing were harmonized with the aim of obtaining an integrated dataset and achieving a
29 better understanding of ageing and health processes.
30
31
32
33
34
35

36 We selected three studies due to their inclusion of the variables of interest and the possibility of
37 comparing HICs and LMICs. Baseline samples of the following studies were included in the analyses:
38 the World Health Organization's Study on Global Ageing and Adult Health (SAGE) [17], the English
39 Longitudinal Study of Ageing (ELSA)[18], and the Survey of Health, Ageing and Retirement in Europe
40 Study (SHARE) [19]. These panel studies included non-institutionalized people aged 50 years and older.
41
42 SAGE comprises six LMICs according to The World Bank Classification [20], namely Ghana, South Africa,
43 Mexico, India, China and Russia; ELSA includes the English population and SHARE covers eleven
44 countries of the European Union and Israel at baseline, considered as HICs [20].
45
46
47
48
49
50
51
52
53

54 The analyses presented focused on people aged 50 years of age and older who were part of the core
55 sample of each study and who completed a non-proxy interview at baseline. We excluded from the
56 analyses those participants who participated via proxy due to cognitive problems or severe physical
57
58
59
60

1
2
3 limitations. Moreover, people with missing values in sex and age were excluded, resulting in a final
4
5 sample of 72 140 individuals. Mexico was excluded from the analyses due the high percentage of
6
7 missingness in the variables of interest (See Additional file 1: Table S1).
8
9

10 11 12 *Patient and public involvement*

13
14 No patient involved.
15
16

17 18 19 *Variables*

20
21 The following variables were the result of a stringent, ex-post harmonization process using systematic
22
23 harmonization methodology and tools from Maelstrom Research [21].

24
25 Eight NCDs were used to conduct the analysis, including those that were available in the three studies:
26
27 diabetes, hypertension, asthma, chronic lung disease, joint disorders (arthritis, rheumatism or
28
29 osteoarthritis), angina or myocardial infarction, stroke, and depression. The presence or absence of
30
31 these conditions was self-reported and based on a medical diagnosis. Depression was assessed with
32
33 standardized tools, such as the Composite International Diagnostic Interview (CIDI) in the SAGE study,
34
35 the Center for Epidemiologic Studies Depression Scale (CES-D) in ELSA, and the EURO-D in SHARE [22–
36
37 24]. A dichotomous variable (yes/no) was created using the indicated cut-off score for each tool and
38
39 population based on previous studies [22,25,26].
40
41
42

43
44 Self-reported demographic variables included age, sex, level of education (primary or less, secondary,
45
46 and tertiary), marital status (single, married or currently cohabiting, separated or divorced, and
47
48 widowed) and quintiles of household wealth (first quintile indicating lowest level). Life-styles and
49
50 health behaviors were 'ever smoked' any type of tobacco and physical activity referring to the practice
51
52 of vigorous exercise during the last two weeks, both coded as *yes* or *no*. Other health-related variables
53
54 were self-rated health (good, moderate, or poor), presence or absence of loneliness feelings in the last
55
56 week, difficulties in activities of daily living (ADL), cognitive performance, and number of diseases.
57
58
59
60

1
2
3 To assess difficulties in ADL, we used a set of daily self-care activities, which were common across
4 studies, such as problems in using the toilet, bathing or showering, getting dressed, eating, moving, or
5 getting in or out of bed. Each of the ADL difficulties was coded into a *yes/no* if the person answered
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

To assess difficulties in ADL, we used a set of daily self-care activities, which were common across studies, such as problems in using the toilet, bathing or showering, getting dressed, eating, moving, or getting in or out of bed. Each of the ADL difficulties was coded into a *yes/no* if the person answered 'severe' or 'extreme/cannot do it'. To build the set of ADL difficulties, we coded *yes* if the person reported at least one difficulty in any of the six items.

Immediate and delayed recall was assessed using the 10-word learning list task, and verbal fluency utilizing the animal naming test [27]. Continuous total scores were used to perform the analyses.

Number of diseases was built by adding up the occurrences of all the above-mentioned NCDs.

Finally, a 7-level regional membership variable was created in order to analyze regional differences, based on the World Health Organization (WHO) and the United Nations Statistical Division (UNSD) regional classification [28,29]. Moreover, the World Bank Classification was used to classify these regions into HICs or LMICs [20]. SAGE includes Africa (Ghana and South Africa), China and India, all of them considered as LMICs. SHARE countries were grouped into three regions: Northern Europe (Denmark, Sweden), Southern Europe (Greece, Italy, and Spain) and Western Europe (Austria, Belgium, France, Germany, Israel, Netherlands, and Switzerland). ELSA and SHARE regions were considered as HICs. Ghana and South Africa were grouped together and named as Africa for practical purposes as well as due to their smaller sample size. These countries are not necessarily representative of the whole continent.

Statistical analysis

All the analyses were performed using data from the baseline. Descriptive statistics were used to summarize information regarding sociodemographic economic variables and disease prevalence among regions. Confidence intervals (95%IC) were calculated for categorical variables in order to make comparisons across regions.

1
2
3 Latent Class Analysis (LCA) was conducted stratified by age (50-64, +65). Eight NCDs (diabetes,
4 hypertension, asthma, chronic lung disease, joint disorders, angina-myocardial infarction, stroke, and
5 depression) were used as observed indicators without using covariates since we aimed to identify
6 latent classes only based on disease variables. Region was used as cluster when conducting LCA in
7 order to accurately describe disease proportions, indicating that the subjects were not independent
8 random draws, but rather were nested within clusters [30].

9
10
11
12
13
14
15
16 The optimal number of latent classes was determined using the adjusted Bayesian Information
17 Criterion (aBIC), the consistent Akaike Information Criterion (CAIC), and the Entropy Index. Lower
18 values of aBIC and CAIC indicate better fit, whereas entropy index values higher than 0.80 indicate that
19 the latent classes are highly discriminating [31]. The average posterior probability indicates how well
20 a model classifies individuals into their most likely class. Values higher than 0.70 indicate well-
21 identified classes [32]. Additionally, interpretability and clinical judgment were used [32,33].

22
23
24
25
26
27
28
29
30 Missing data in one of the indicators was handled with the full-information maximum likelihood (FIML)
31 technique, assuming missing-at-random (MAR) [34]. Missing data in the covariates was handled using
32 multiple imputation by chained equations (MICE) assuming MAR [34]. The imputation model included
33 the outcome (group membership in one of the latent classes) and all the variables used in the
34 regression models. In the Additional file 1: Table S2-10 there is a report of those variables and the
35 percentage of missingness of each region in the variables of interest.

36
37
38
39
40
41
42
43 Adjusted multinomial logistic regression models were used to assess the association between the
44 outcome (multimorbidity classes, with the "Healthy" class as the reference category) and several
45 variables: loneliness, ever smoked, physical activity, limitations in ADL, self-rated health, immediate
46 recall, delayed recall and verbal fluency. The model was additionally adjusted for sex, age, marital
47 status, education level, wealth, and the region at baseline. Due to potential collinearity between
48 income and education, we checked the significance and magnitude of the correlation between both
49 variables. The association was small, and thus both covariates were included as separate variables in
50
51
52
53
54
55
56
57
58
59
60

1
2
3 the models. Regression models were conducted separately in one hundred imputed datasets and
4
5 results combined using Rubin's rules [35].
6

7 All analyses were conducted with Stata SE version 13.1 (College Station, TX). LCA analyses were
8
9 performed using a Stata plugin [30].
10
11
12
13

14 **Results**

15 *Descriptive analysis*

16
17 In Table 1 are presented the main characteristics of the sample by region. The mean age ranged from
18
19 62 (SD=9.02) in Southern Asia to 65 years (10.18) in Russia and (10.26) in England. Some 54% were
20
21 women, 72% were married or cohabitating, and 39% had secondary education. Russia presented the
22
23 highest number of conditions (mean 1.66) compared to Africa (0.64), China (0.80), and India (0.72).
24
25

26
27 The most prevalent conditions in the total sample were hypertension (31.2%, 95%CI = 30.9-31.6) and
28
29 joint disorders (22.4%, 95%CI = 22.0-22.7). Hypertension was particularly high in Russia (56.5%, 95%CI
30
31 = 54.9-58.1) compared to the other regions. Diabetes prevalence was greater in Southern (11.9%,
32
33 95%CI = 11.2-12.7) and Western Europe (10.4%, 95%CI = 9.9-10.9), whereas Africa and China presented
34
35 the lowest proportions. Similarly, myocardial infarction-angina was highly prevalent in Russia (33.1%,
36
37 95%CI = 31.6-34.6), followed by countries of Northern (13.8%, 95%CI = 12.8-14.8), Southern (11.7%,
38
39 95%CI = 11.0-12.4) and Western Europe (13.1%, 95%CI = 12.6-13.6).
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Table 1 Main characteristics of the total sample and by regions

Region	Total	Africa ^a	China ^b	India ^b	Russia ^b	England ^c	Northern Europe ^d	Southern Europe ^e	Western Europe ^f
	N=72140	n=7950	n=12840	n=6558	n=3887	n=11 517	n=4573	n=7465	n=17 350
Age, mean (SD)	64.05 (9.96)	63.60 (10.26)	63.07 (9.31)	61.80 (9.02)	65.06 (10.18)	65.06 (10.26)	64.78 (10.34)	65.02 (10.17)	64.34 (9.96)
Woman, % (95% CI)	54.0 (53.6-54.4)	52.3 (51.2-53.4)	52.9 (52.1-54.8)	49.6 (48.4-50.8)	64.6 (63.0-66.1)	54.6 (53.7-55.5)	53.7 (51.8-54.7)	55.4 (54.2-56.5)	54.1 (53.4-54.9)
Marital status, % (95% CI)									
Single	4.1 (4.0-4.3)	6.7 (6.2-7.3)	0.9 (0.7-1.6)	1.0 (0.8-1.2)	2.8 (2.3-3.4)	5.0 (4.6-5.4)	5.5 (4.8-6.2)	6.3 (5.8-6.9)	4.9 (4.6-5.3)
Married	71.5 (71.1-72.0)	54.5 (53.4-55.6)	83.5 (82.8-84.1)	74.1 (73.1-75.2)	56.1 (54.5-57.7)	69.1 (68.2-69.9)	72.0 (71.0-73.7)	73.3 (72.3-74.3)	73.3 (72.7-74.0)
Divorced	5.9 (5.7-6.2)	10.4 (9.7-11.1)	1.7 (1.5-2.0)	0.6 (0.5-0.9)	8.3 (7.4-9.2)	9.0 (8.5-9.5)	9.8 (8.8-10.5)	2.6 (2.3-3.0)	6.8 (6.5-7.2)
Widowed	18.4 (18.1-18.6)	27.3 (26.3-28.3)	13.9 (13.3-14.6)	24.3 (23.2-25.3)	32.7 (31.2-34.2)	16.9 (16.3-17.6)	12.7 (11.6-13.5)	17.8 (16.9-18.6)	14.9 (14.4-15.4)
Education level % (95% CI)									
Primary or less	33.4 (33.0-33.7)	30.5 (29.5-31.5)	37.9 (37.1-38.8)	25.5 (24.5-26.6)	9.9 (9.0-10.9)	42.4 (41.5-43.4)	29.9 (28.3-30.9)	59.4 (58.3-60.5)	23.2 (22.6-23.8)
Secondary	38.7 (38.4-39.1)	19.0 (18.2-19.9)	33.2 (32.4-34.0)	18.2 (17.3-19.2)	69.3 (67.9-70.8)	37.5 (36.7-38.4)	45.5 (43.6-46.5)	31.3 (30.2-32.3)	55.1 (54.3-55.8)
Tertiary	12.0 (11.9-12.2)	3.9 (3.5-4.3)	4.7 (4.3-5.0)	5.0 (4.4-5.5)	19.7 (18.5-21.0)	11.1 (10.6-11.7)	24.6 (23.0-25.5)	9.1 (8.4-9.7)	20.8 (20.2-21.4)
Wealth (quintiles% (95% CI))									
1 st (worse)	19.3 (19.1-19.6)	19.3 (18.4-20.2)	19.8 (19.1-20.5)	16.2 (15.3-17.1)	18.2 (17.0-19.5)	19.0 (18.3-19.7)	20.7 (19.6-21.9)	20.3 (19.4-21.2)	19.9 (19.3-20.5)
2 nd	19.7 (19.4-20.0)	19.7 (18.4-20.2)	19.7 (19.0-20.4)	18.6 (17.6-19.5)	19.8 (18.5-21.1)	19.3 (18.6-20.1)	20.0 (19.3-21.6)	20.4 (19.5-21.3)	19.9 (19.4-20.6)
3 rd	19.7 (19.4-19.9)	19.8 (18.9-20.7)	20.1 (19.4-20.8)	18.4 (17.5-19.4)	20.3 (19.0-21.6)	19.7 (18.9-20.4)	20.0 (19.0-21.3)	19.8 (18.9-20.7)	19.5 (18.9-20.1)
4 th	19.9 (19.7-20.3)	20.5 (19.6-21.4)	20.5 (19.8-21.2)	21.5 (20.5-22.5)	20.0 (18.8-21.3)	19.6 (18.9-20.3)	19.9 (18.5-20.8)	19.4 (18.6-20.4)	19.3 (18.8-19.9)
5 th (best)	20.1 (19.8-20.4)	20.4 (19.5-21.3)	19.7 (19.0-20.4)	24.8 (23.8-25.9)	21.6 (20.3-22.9)	19.6 (18.9-20.3)	18.8 (17.8-20.1)	18.9 (18.0-19.8)	19.4 (18.8-20.0)
N^o diseases, mean (SD)	1.02 (1.14)	0.64 (0.94)	0.80 (0.99)	0.72 (0.97)	1.66 (1.38)	1.19 (1.13)	1.02 (1.10)	1.28 (1.25)	1.10 (1.16)
Diseases, % (95% CI)									
Diabetes	8.5 (8.3-8.7)	6.6 (6.1-7.2)	6.5 (6.1-7.0)	7.3 (6.7-7.9)	9.0 (8.1-10.0)	7.4 (6.9-7.9)	8.7 (7.4-9.0)	11.9 (11.2-12.7)	10.4 (9.9-10.9)
Hypertension	31.2 (30.9-31.6)	21.5 (20.6-22.4)	27.4 (26.6-28.2)	17.5 (16.6-18.5)	56.5 (54.9-58.1)	37.8 (36.9-38.7)	29.5 (28.0-30.7)	35.6 (34.5-36.7)	32.3 (31.6-33.0)
Joint disorders	22.4 (22.0-22.7)	17.8 (16.9-18.6)	22.1 (21.4-22.8)	17.9 (17.0-18.9)	35.2 (33.7-36.7)	32.5 (31.6-33.3)	15.9 (14.6-16.8)	26.3 (25.3-27.3)	16.8 (16.3-17.4)
Asthma	5.5 (5.3-5.6)	4.1 (3.7-4.5)	2.4 (2.2-2.7)	6.9 (6.3-7.6)	3.4 (2.9-4.0)	11.7 (11.1-12.3)	7.9 (6.8-8.4)	4.1 (3.7-4.6)	4.1 (3.8-4.4)
Chronic lung disease	6.1 (5.9-6.3)	1.5 (1.2-1.7)	8.6 (8.1-9.1)	4.1 (3.6-4.6)	17.9 (16.8-19.2)	6.5 (6.1-7.0)	4.9 (3.9-5.2)	5.4 (4.9-6.0)	4.9 (4.6-5.3)
MI - Angina	10.0 (9.8-10.3)	4.5 (4.1-5.0)	8.8 (8.3-9.3)	4.9 (4.4-5.5)	33.1 (31.6-34.6)	3.3 (3.0-3.6)	13.9 (12.8-14.8)	11.7 (11.0-12.4)	13.1 (12.6-13.6)
Stroke	3.8 (3.7-3.9)	3.2 (2.8-3.6)	3.5 (3.2-3.8)	2.2 (1.9-2.6)	6.0 (5.3-6.8)	4.5 (4.1-4.9)	5.9 (4.4-5.7)	3.1 (2.8-3.6)	3.9 (3.7-4.2)
Depression	15.3 (15.0-15.5)	5.8 (5.3-6.3)	1.2 (1.1-1.4)	12.1 (11.3-12.9)	5.2 (4.5-5.9)	16.5 (15.8-17.2)	18.8 (17.7-20.0)	31.7 (30.6-32.7)	25.0 (24.3-25.6)

Note. ^aSAGE study - Africa: Ghana, South Africa; ^bSAGE study; ^cELSA study- England; ^dSHARE study - Northern Europe: Denmark, Sweden; ^eSHARE study - Southern Europe:

Greece, Italy, Spain; ^fSHARE study - Western Europe: Austria, Belgium, France, Germany, Israel, Netherlands, Switzerland.

The analyses were performed before multiple imputation procedure

1
2
3 Joint disorders were more prevalent in Russia (35.2%, 95%CI = 33.7-36.7) and England (32.5%, 95%CI
4 =31.6-33.3). The prevalence of asthma was greater in England than other regions (11.7%, 95%CI = 11.1-
5 12.3) and chronic lung disease was greater in Russia (17.9%, 95%CI = 16.8-19.2).
6
7

8
9
10 As for the prevalence of depression, European countries presented the highest values, especially in
11 Southern (31.7%, 95%CI = 30.6-32.7) and Western Europe (25.0%, 95%CI = 24.3-25.6), whereas LMICs
12 showed very low proportions, especially in China, where only 1.2% of people aged 50+ presented
13 depression.
14
15
16
17

18 19 20 *Multimorbidity patterns* 21

22
23 Table 2 displays the aBIC, CAIC, and entropy values, proportions, and average posterior probability of
24 each latent class, for a two- to five-class model in both age subsamples. In the younger subsample (50-
25 64), the five-class solution yielded the lowest aBIC and CAIC values and the highest entropy value
26 (0.67). However, it was dismissed because one of the latent classes was very infrequent and the
27 posterior probabilities were far below 0.70. Similarly, the four-class model was rejected for an
28 inadequate posterior probability value in one of the classes (0.52). The model finally selected was the
29 three-class model. The three-class solution was also chosen for the older age group because of lower
30 posterior probability values in the four- and five-class models in spite of lower aBIC and CAIC values.
31
32
33
34
35
36
37
38
39
40
41
42

43 We named each latent class according to the most prevalent diseases within each latent class. Figure
44 1 shows the distribution of each condition across the three latent classes (“cardio-metabolic”,
45 “respiratory-mental-articular”, and “healthy” class) in the total sample and by regions. The “cardio-
46 metabolic” class presented excess prevalence of diabetes, hypertension, myocardial infarction or
47 angina, and stroke, comprising 8.93% of the total sample in the younger group and 27.22% in the older
48 group. The “respiratory-mental-articular” class, which comprised 3.91% and 5.27% of each sample,
49 respectively, showed greater prevalence of joint disorders, asthma, chronic lung diseases, and
50
51
52
53
54
55
56
57
58
59
60

1
2
3 depression. Finally, the “healthy” class presented low prevalence of conditions, comprising 87.16% of
4
5 the sample in the first age group and 67.51% in the second group.
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For peer review only

Table 2 Comparison between models in individuals aged 50-64 and +65

No. of latent classes	Aged 50-64				Aged ≥65					
	Information Criteria Indices		Classification Quality	Latent classes, n (%)	Average posterior probability	Information Criteria Indices		Classification Quality	Latent classes, n (%)	Average posterior probability
	aBIC	CAIC				aBIC	CAIC			
2	1512.91	1583.94	0.51	33023 (82.15)	.88	1603.30	1674.33	0.69	21113 (66.10)	.83
				7177 (17.85)	.75				10827 (33.90)	.75
3	875.23	983.86	0.63	3589 (8.93)	.76	1032.83	1141.46	0.90	8693 (27.22)	.71
				1571 (3.91)	.67				1684 (5.27)	.68
				35040 (87.16)	.87				21563 (67.51)	.81
4	777.14	923.37	0.43	25701 (63.93)	.52	817.60	963.83	0.83	9557 (29.92)	.65
				626 (1.56)	.72				1474 (4.61)	.78
				7754 (19.29)	.80				17220 (53.91)	.86
				6119 (15.22)	.68				3689 (11.55)	.77
5	661.04	844.87	0.67	4578 (11.39)	.64	689.22	873.05	0.99	11094 (34.73)	.77
				247 (0.61)	.67				1094 (3.43)	.71
				32423 (80.65)	.85				14155 (44.32)	.76
				1359 (3.38)	.48				1148 (3.59)	.63
				1593 (3.96)	.64			4449 (13.93)	.72	

Note: Boldface indicates the final selected model. *aBIC* adjusted Bayesian Information Criterion, *CAIC* Consistent Akaike Information Criterion

http://bmjopen-2019-034441 on 19 April 2024 by guest. Protected by copyright.

1
2
3 Differences in the proportions of multimorbidity classes were found across regions (Figure 1). The
4
5 “cardio-metabolic” class (18.8%, 95%CI =17.1-20.6) was significantly greater in Russia than in other
6
7 regions, and England (5.1%, 95%CI =4.5-5.7) showed a higher proportion of individuals classified into
8
9 the “respiratory-mental-articular” class. The “healthy” class was higher in Africa (91.5%, 95%CI = 90.7-
10
11 92.3), China (90.8%, 95%CI = 90.1-91.4), and India (89.5%, 95%CI = 88.5-90.4), and remarkably lower
12
13 in Russia (71.6%, 95%CI = 69.5-73.6) compared to other regions.
14
15
16
17
18
19

20 ----- FIGURE 1 HERE -----
21
22
23

24 Similar results were found for the older group (Figure 2). In Russia, the “cardio-metabolic” class was
25
26 significantly higher than in other regions (48.3%, 95%CI = 46.1-50.6) whereas the “healthy” class was
27
28 the least frequent class compared with the rest of regions (38.4%, 95%CI = 36.2-40.6), followed by
29
30 Southern Europe (52.6%, 95%CI = 50.9-54.2). Africa and India showed lower proportions of individuals
31
32 classified into the “cardio-metabolic” class (12.9%, 95%CI = 11.8-14.1 and 11.2%, 95%CI = 10.0-12.6,
33
34 respectively).
35
36
37
38

39 ----- FIGURE 2 HERE -----
40
41
42

43 *Association between multimorbidity classes and covariates*

44 In the Additional file 2: Table S1 are presented the unadjusted relative risk ratios (RRR) for both age
45
46 subsamples. The “healthy” class was used as the reference group. In the case of the younger
47
48 subsample, and compared with the “healthy” class, individuals classified into the “cardio-metabolic”
49
50 and “respiratory-mental-articular” classes were more likely to be older (RRR = 1.09, 95%CI = 1.08-1.10;
51
52 RRR = 1.06, 95%CI = 1.04-1.07, respectively) and being widowed (RRR = 1.4, 95%CI = 1.1-1.7) and
53
54 divorced in the “respiratory-mental-articular” class (RRR = 1.7, 95%CI = 1.2-2.4). Being a man, having
55
56 tertiary education, and high levels of wealth had a protective effect for being in both multimorbidity
57
58 groups, compared with the “healthy” class. Similarly, those individuals from the older subsample who
59
60

1
2
3 were in the fourth and fifth quintile (RRR = 0.8, 95%CI = 0.7-1.0) were less likely to be classified into
4
5 the “respiratory-mental-articular” class, compared with the “healthy” group.
6

7 Regarding the association of regions and multimorbidity groups, some differences were found in the
8
9 younger subsample. Taking Africa as the reference category, participants from Russia were more likely
10
11 to be classified into the “cardio-metabolic” class (RRR = 3.6, 95%CI = 3.0-4.2), whereas individuals from
12
13 England (RRR = 5.6, 95%CI = 4.1-7.7), Northern Europe (RRR = 2.8, 95%CI = 1.9-4.1) and India (RRR =
14
15 2.2, 95%CI = 1.5-3.2) showed higher risk of being in the “respiratory-mental-articular” class. In the case
16
17 of the older subsample, all regions had greater risk of being classified into the “respiratory-mental-
18
19 articular” class, compared to the “healthy” class, especially participants from Russia (RRR = 14.5, 95%CI
20
21 = 10.3-20.3), compared to Africa.
22
23
24
25
26
27

28 Table 3 shows the adjusted relative risk ratios (RRR) for both age subsamples, taking the “healthy”
29
30 class as the reference group. Both multimorbidity classes (cardio-metabolic and respiratory-mental-
31
32 articular) were associated with all the covariates in the younger group, except for smoking status in
33
34 the “cardio-metabolic” class.
35
36
37
38

39 In the younger individuals subsample, both latent classes were more likely to be associated with the
40
41 presence of feelings of loneliness (RRR = 1.8, 95%CI = 1.7-2.0; RRR = 2.5, 95%CI = 2.0-3.0), limitations
42
43 in ADL (RRR = 3.2, 95%CI = 2.9-3.6; RRR = 3.9, 95%CI = 3.3-4.7) and worse health status (RRR = 12.8,
44
45 95%CI = 11.3-14.4; RRR = 12.9, 95%CI = 10.5-16.0). Physical activity had a protective effect for being in
46
47 these classes and having smoked was a risk factor only for being classified into the “respiratory-mental-
48
49 articular” class (RRR = 1.5, 95%CI = 1.2-1.7). Conversely, those older individuals who had ever smoked
50
51 had a higher risk of being in the “respiratory-mental-articular” group (RRR = 1.8, 95%CI = 1.5-2.0) and
52
53 a lower risk of being classified into the “cardio-metabolic” class (RRR = 1.0, 95%CI = 0.9-1.0). For both
54
55 age subsamples, better performance in verbal memory was significantly associated with less risk of
56
57
58
59
60

being classified into the two multimorbidity classes. Similarly, higher scores in verbal fluency were a protective factor for multimorbidity, compared with the healthy individuals group.

Table 3 Association between latent multimorbidity membership and outcomes in individuals aged 50-64 and ≥ 65

Variables ^a	Aged 50-64		Aged ≥ 65	
	“Cardio-metabolic” class	“Respiratory-mental-articular” class	“Cardio-metabolic” class	“Respiratory-mental-articular” class
Loneliness (yes/no)	1.8 (1.7-2.0)	2.5 (2.0-3.0)	1.3 (1.2-1.5)	1.9 (1.7-2.3)
Ever smoked (yes/no)	1.1 (1.0-1.2)	1.5 (1.2-1.7)	1.0 (0.9-1.0)	1.8 (1.5-2.0)
Physical activity (yes/no)	0.4 (0.4-0.5)	0.5 (0.5-0.6)	0.5 (0.5-0.6)	0.4 (0.4-0.5)
Limitations in ADL (yes/no)	3.2 (2.9-3.6)	3.9 (3.3-4.7)	2.3 (2.1-2.5)	4.0 (3.5-4.6)
Self-rated health				
Good	1.0	1.0	1.0	1.0
Moderate	4.8 (4.3-5.2)	3.9 (3.2-4.7)	3.1 (2.9-3.4)	5.7 (4.8-6.8)
Poor	12.8 (11.3-14.4)	12.9 (10.5-16.0)	6.2 (5.6-6.9)	19.4 (16.2-23.4)
Memory: Immediate recall	0.9 (0.9-0.9)	0.9 (0.9-1.0)	0.9 (0.9-0.9)	0.9 (0.9-1.0)
Memory: Delayed recall	0.9 (0.9-0.9)	0.9 (0.9-1.0)	0.9 (0.9-0.9)	0.9 (0.9-1.0)
Verbal fluency	1.0 (1.0-1.0)	1.0 (1.0-1.0)	0.9 (0.9-1.0)	1.0 (1.0-1.0)

Note: The reference group for the multimorbidity group variable was the “Healthy” class
 Relative Risk Ratios (95% confidence interval) from multinomial logistic regression models
 Models were run in 100 imputed datasets and results combined using Rubin’s rules
 ADL Activities of Daily Living
^aAdjusted for sex, age, marital status, education level, wealth and region

Discussion

To the best of our knowledge, this is the first multi-region study to use harmonized data to compare multimorbidity patterns across different regions from three distinct population-based cohorts. We identified three latent classes of multimorbidity based on the presence of eight NCDs: the “cardio-metabolic”, the “respiratory-mental-articular” and the “healthy” class. The same clusters were identified in another study using SAGE original data, applying exploratory factor analysis (EFA) in a sample of 41 909 individuals aged 50 years or older [36]. Similarly, a study of a representative sample of Spanish community-dwelling adults over 50 years old also found three latent classes using eleven chronic conditions, showing similar diseases distributions among the multimorbidity clusters [37].

In our study, for both age groups the majority of the sample was classified into the “healthy” class, 87.16% and 67.51%, respectively. This latent group has previously been described in studies which

1
2
3 applied LCA [11,13,37,38]. Likewise, the other two identified classes are similar to those reported in a
4
5 systematic review based on 14 studies of multimorbidity patterns [39]. In this review, the most
6
7 prevalent diseases in the “cardio-metabolic” group were diabetes, hypertension, heart diseases,
8
9 hyperlipidemia, and obesity; and in the second group conditions such as mental disorders, thyroid
10
11 disease, neurological disease, pain, asthma or chronic lung diseases, musculoskeletal disorders,
12
13 obesity, and gastroesophageal reflux disease were included. Despite the fact that we included a
14
15 smaller number of diseases, we found analogous patterns. In our study, 8.93% of the younger group
16
17 (50-64) and 27.22% of the older were classified into the “cardio-metabolic” class, including individuals
18
19 with higher prevalence of diabetes, hypertension, myocardial infarction or angina, and stroke. This
20
21 clustering of diseases is similar to the metabolic syndrome, which has metabolically-related
22
23 cardiovascular risk factors and greater risk of stroke and diabetes [40]. Lastly, the least prevalent group
24
25 was the “respiratory-mental-articular” class, consisting of greater prevalence of joint disorders,
26
27 asthma, chronic lung diseases, and depression. Association between depression and arthritis has
28
29 commonly been reported, with socioeconomic and disease factors reported as being involved in its
30
31 association, as well as systemic inflammation mechanisms [3]. Nevertheless, the links between
32
33 depression and chronic lung diseases, and chronic lung diseases and arthritis, despite having been
34
35 studied, remain unclear [40,41].
36
37
38
39
40
41
42

43 Analogous latent multimorbidity classes have been found among both age groups. Despite this, certain
44
45 aspects should be pointed out. As expected, the proportion of participants classified into the “healthy”
46
47 class was greater in participants aged 50-64 (87.16%) compared to those aged +65 years old (67.51%),
48
49 illustrating higher multimorbidity in elderly individuals. The distribution of CC was also less clear in the
50
51 older subsample. For example, both joint disorders and angina–myocardial infarction were similarly
52
53 present in the “cardio-metabolic” and “respiratory-mental-articular” categories, whereas in the
54
55 younger participants (50-64) subsample we observed a more differentiated profile of those chronic
56
57 conditions that cluster into one latent class. For example, respiratory-related diseases (asthma, chronic
58
59
60

1
2
3 lung diseases) are highly presented in the “respiratory-mental-articular” class, while very infrequent
4 among middle-aged people classified into the “cardio-metabolic” group. It is worth mentioning that
5
6 although depression is frequently observed among participants classified into the “respiratory-mental-
7
8 articular” class, it is not infrequent among people within the “cardio-vascular” class. This may be due
9
10 to the relationship between mental and physical disorders, which has frequently been reported,
11
12 suggesting a bidirectional association between them [42]. On the one hand, medical conditions could
13
14 be accompanied by a high symptom burden, leading to depression, and, on the other, depression could
15
16 be a risk factor for medical conditions, since depressive symptoms could increase the incidence of
17
18 behaviors, such as smoking, alcohol intake, poor diet, or physical inactivity, which are risk factors for
19
20 NCDs [3,42].
21
22
23
24

25 One important implication of our findings is the relatively high proportion of people aged 50-64 with
26
27 multimorbidity. Thus, preventive and intervention programs are also needed for this population to
28
29 mitigate the multimorbidity burden.
30
31

32 Our results show that these multimorbidity patterns are qualitatively different, but only if compared
33
34 to the “healthy” class in terms of sociodemographic and economic characteristics, lifestyles, and health
35
36 status variables. As has been reported in the literature, being older, woman, widowed, with a lower
37
38 level of education and lower socioeconomic status are related to an increased risk of multimorbidity
39
40 [3]. In addition, those individuals with multiple chronic conditions were more likely to have limitations
41
42 in ADL, especially those classified into the “respiratory-mental-articular” group, similar to what was
43
44 found in another study of multimorbidity [37]. Physical activity seems to be a protective factor for
45
46 being classified into the “respiratory-mental-articular” class, whereas smokers were more likely to be
47
48 classified into the “respiratory-mental-articular” class, but not the “cardio-metabolic” class. This is
49
50 inconsistent with the literature, since cigarette smoking is considered a major cause of cardiovascular
51
52 diseases (CVDs). However, smoking is probably the most complex and least understood risk factor for
53
54 CVDs [43].
55
56
57
58
59
60

1
2
3 One interesting finding is the association between cognition outcomes and multimorbidity in both age
4 subsamples. Better performance in verbal memory and fluency was related to less risk of being
5 classified into the multimorbidity groups, with similar results among latent classes. Impaired cognition
6 has been associated with conditions such as arthritis [44], depression [45], and respiratory diseases
7 [46], cardiovascular conditions, diabetes [47], hypertension [48], and coronary heart diseases [49].
8
9

10
11
12 Concerning the regional distribution of multimorbidity, Russia accounted for the highest burden as
13 opposed to Africa, China, and India. The “cardio-metabolic” class is especially common in this country,
14 with a prevalence of 18.82% in the younger and 48.34% in the older subsample. Prevalence of CVDs,
15 such as hypertension, myocardial infarction or angina, and stroke, was also higher in Russia. This high
16 proportion could be related to the high rate of alcohol consumption and rapid societal changes
17 experienced in this country which might account for increased risk of circulatory diseases [50,51].
18 Followed by Russia, European regions showed higher rates of multimorbidity. NCDs such as
19 hypertension, joint disorders, respiratory diseases, and depression were highly prevalent, especially in
20 England and Southern Europe, where the “respiratory-mental-articular” class was highly prevalent in
21 both age subsamples. The relationship between mood disorders such as depression and joint disorders
22 has been previously reported in other studies, though the underlying cause remains unclear [36,37,39].
23 Notwithstanding, previous studies suggested that the emotional burden of joint disorders may
24 contribute to the onset of psychiatric disorders [36,52].
25
26

27
28 LMICs such as Africa, China, and India showed lower rates of multimorbidity compared to Russia and
29 other HICs. However, there was a wide variation in terms of some diseases, such as respiratory diseases
30 and depression. Asthma and chronic lung diseases were highly prevalent in India and China, influenced
31 by factors such as increasing smoking rates, air pollution, and occupational lung diseases in these
32 countries [53]. As reported in previous studies [54], depression was remarkably prevalent in India,
33 whereas the lowest prevalence was observed in China. This is in line with previous epidemiological
34 studies on the prevalence of depression in Chinese older people, suggesting differences in diagnostic
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 criteria that make depression less diagnosed; somatic symptoms are more prevalent in this population
4
5 instead of sadness, and lack of interest and energy. Moreover, stigma and prejudice in Chinese
6
7 population might also contribute to under-reporting depressive symptoms [55,56]. Furthermore, the
8
9 variation found across regions in terms of depression prevalence could be due to cultural differences
10
11 in expressions or expectations of mood disorders or mental health [57].
12
13
14

15
16 The highest burden of multimorbidity in HICs could be explained by an increased level of development
17
18 in the HICs. Notwithstanding, LMICs are experiencing a change in lifestyle and environmental
19
20 exposures which contributes, as in HICs, to multimorbidity. Thus, the increased burden of NCDs, in
21
22 addition to the existing burden of infectious diseases such as HIV/AIDS, worsens multimorbidity
23
24 management [3]. Moreover, the differences found in the regional distribution of multimorbidity might
25
26 be linked to different stages of development of their health systems, since there are differences
27
28 between LMICs and HICs in terms of opportunities and barriers to improving the organization,
29
30 integration, and delivery of multimorbidity care [3].
31
32
33
34
35

36 *Strengths and limitations*

37
38 A major strength of this study is the use of a large, harmonized, multi-regional database. Research on
39
40 multimorbidity has typically been hampered by several factors, such as the exclusion of patients with
41
42 multimorbidity from participation, targeting of research mostly on elderly individuals, and a shortage
43
44 of studies focusing on LMICs. The ATHLOS study allowed us to compare two age groups (50-64 and 65
45
46 or older) as well as disease prevalence and clusters of conditions in regions with differing incomes in a
47
48 very large, diverse population-based study of middle-aged and older adults.
49
50

51
52 Some limitations should be considered when interpreting our findings. Firstly, the presence or absence
53
54 of the NCDs was based on self-reported measures, and thus might be affected by measurement errors
55
56 or lack of accuracy. Nevertheless, some authors sustain self-reported diagnostics as a well-established
57
58 method for the measurement of multimorbidity in population-based studies [58]. Secondly,
59
60

1
2
3 participants with an incipient neurodegenerative disease may have been included in our analytical
4 sample. However, we excluded those participants who completed a proxy interview due to cognitive
5 problems, such as neurodegenerative diseases, which could affect the reliability of the data.
6
7 Nonetheless, participants with an incipient neurodegenerative disease may have been included in our
8 analytical sample because of the lack of strong diagnostic criteria for dementia in the included studies.
9
10 Thirdly, we could only focus on those diseases that were common across studies. Conditions such as
11 obesity, cancer, kidney disease, and neurological illness were not evaluated. This might have led to a
12 smaller number of latent classes, or to different patterns of multimorbidity. Fourthly, when performing
13 LCA, the three-class solution was forced. In order to determine whether the latent classes were
14 equivalent, invariance analysis should have been performed [59]. Nevertheless, this solution was
15 forced as we aimed to do comparisons among age subsamples and regions in terms of disease
16 prevalence as well as protective and risk factors. Finally, the use of multiple imputations could carry
17 some bias. Despite this, the use of multiple imputation procedures is widely advocated when missing
18 data occur in one or more covariates in a regression model and under a MAR assumption [60,61].
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

The results of this study suggest that NCDs cluster together in non-random associations across several regions worldwide. The three qualitatively distinct entities are also linked to several sociodemographic and economic characteristics, lifestyles, and health status variables. A deeper understanding of the interactions across regions and the studied variables is needed. Knowledge regarding broad patterns of conditions may contribute to the creation and implementation of guidelines that consider clusters of conditions instead of single diseases, since multimorbidity has become an unavoidable reality. Future efforts should focus on the underlying mechanisms of these clusters as well as their stability over time using longitudinal data. Moreover, cohort and age effects should be explored as might influence the likelihood of reporting some diagnosis and hence result in different multimorbidity patterns.

Acknowledgments

- The English Longitudinal Study of Ageing (ELSA):

ELSA is supported by the U.S. National Institute of Aging, the National Centre for Social Research, the University College London (UCL), and the Institute for Fiscal Studies. The authors gratefully acknowledge the UK Data Service and UCL who provided data for this paper.

- The Study on Global Ageing and Adult Health (SAGE):

The SAGE study is funded by the U.S. National Institute on Aging and has received financial support through Interagency Agreements (OGHA 04034785; YA1323-08-CN-0020; Y1-AG-1005-01) and Grants (R01-AG034479; IR21-AG034263-0182). The authors gratefully acknowledge the World Health Organization who provided data for this paper.

- The Survey of Health, Ageing and Retirement in Europe (SHARE):

The SHARE study is funded by the European Commission through FP5 (QLK6-CT-2001-00360), FP6 (SHARE-I3: RII-CT-2006-062193, COMPARE: CIT5-CT-2005-028857, SHARELIFE: CIT4-CT-2006-028812) and FP7 (SHARE-PREP: N°211909, SHARE-LEAP: N°227822, SHARE M4: N°261982). Additional funding from the German Ministry of Education and Research, the Max Planck Society for the Advancement of Science, the U.S. National Institute on Aging (U01_AG09740-13S2, P01_AG005842, P01_AG08291, P30_AG12815, R21_AG025169, Y1-AG-4553-01, IAG_BSR06-11, OGHA_04-064, HHSN271201300071C) and from various national funding sources is gratefully acknowledged (see www.share-project.org).

Funding: This work was supported by the five-year Ageing Trajectories of Health: Longitudinal Opportunities and Synergies (ATHLOS) project and the Centro de Investigación Biomédica en Red de Salud Mental (CIBERSAM). The ATHLOS project has received funding from the European Union's

1
2
3 Horizon 2020 research and innovation program under grant agreement No 635316. DF's work has been
4 supported by grant RTI2018-100927-J-I00 Retos Investigación from Ministerio de Ciencia e Innovación
5 (MCI), by Marsden grant E2987-3648 (Royal Society of New Zealand), and by grant 2017 SGR 622
6 (GRBIO) from the Departament d'Economia i Coneixement de la Generalitat de Catalunya (Spain). This
7 work, grant number RTI2018-100927-J-I00, is supported by the Ministerio de Ciencia e Innovación
8 (MCI, Spain), by the Agencia Estatal de Investigación (AEI, Spain) and by the European Regional
9 Development Fund FEDER (FEDER, UE). BO's work is supported by the PERIS program 2016-2020 "Ajuts
10 per a la Incorporació de Científics i Tecnòlegs" [grant number SLT006/17/00066], with the support of
11 the Health Department from the Generalitat de Catalunya.

12
13
14
15
16
17
18
19
20
21
22
23
24 **Competing interests:** The authors report no conflicts of interest.

25
26
27 **Patient consent for publication:** not required

28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
Ethics approval: The study protocol was approved by the Committee on the Ethics of Clinical Research, CEIC Fundació Sant Joan de Déu (Protocol No: PIC-22-15). All data were anonymized and EHR confidentially was respected in accordance with national and international law.

Data sharing statement: The original studies data are available on their respective websites: the Study on Global Ageing and Adult Health - SAGE (<https://www.who.int/healthinfo/sage/en/>), the English Longitudinal Study of Ageing – ELSA (<https://www.elsa-project.ac.uk/>), and the Survey of Health, Ageing and Retirement in Europe – SHARE (<http://www.share-project.org/home0.html>). R codes for harmonizing the included variables, as well as the STATA codes for the performed analysis are available on <https://github.com/athlosproject/athlos-project.github.io>

Additional files:

Additional file 1: **Table S1.** Proportion of missingness in the variables of interest in Mexico; **Table S2.** Overall proportion of missingness for indicators of latent classes and variables included in the imputation model; **Table S3.** Proportion of missingness for indicators of latent classes and variables included in the imputation model for region 1 (SAGE study - Africa: Ghana, South Africa); **Table S4.**

1
2
3 Proportion of missingness for indicators of latent classes and variables included in the imputation
4 model for region 2 (SAGE study - Eastern Asia: China); **Table S5**. Proportion of missingness for
5 indicators of latent classes and variables included in the imputation model for region 2 (SAGE study -
6 Southern Asia: India); **Table S6**. Proportion of missingness for indicators of latent classes and variables
7 included in the imputation model for region 2 (SAGE study - Eastern Europe: the Russian Federation);
8 **Table S7**. Proportion of missingness for indicators of latent classes and variables included in the
9 imputation model for region 2 (ELSA study- England); **Table S8**. Proportion of missingness for indicators
10 of latent classes and variables included in the imputation model for region 2 (SHARE study - Northern
11 Europe: Denmark, Sweden); **Table S9**. Proportion of missingness for indicators of latent classes and
12 variables included in the imputation model for region 2 (SHARE study - Southern Europe: Greece, Italy,
13 Spain); **Table S10**. Proportion of missingness for indicators of latent classes and variables included in
14 the imputation model for region 2 (SHARE study - Western Europe: Austria, Belgium, France, Germany,
15 Israel, Netherlands, Switzerland).

16
17
18
19
20
21 Additional file 2: **Table S1**. Association between latent multimorbidity membership and outcomes in
22 individuals aged 50-64 and ≥ 65
23
24
25

26 **Author contributions:**

27
28 **IB:** Participated in the database management, drafted the paper, carried out the statistical analyses
29 and worked on the interpretation of data. She also gave final approval of the version to be published
30 and agreed to be accountable for all aspects of the work in ensuring that questions related to the
31 accuracy or integrity of any part of the work are appropriately investigated and resolved; **AS:**
32 Participated in the study design, database management, statistical support and critical revision of the
33 paper. He also gave final approval of the version to be published and agreed to be accountable for all
34 aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the
35 work are appropriately investigated and resolved; **LE:** Participated in the interpretation of data and
36 critical revision of the paper; **HN:** She participated in critical revision of the paper and gave final
37 approval of the version to be published and agreed to be accountable for all aspects of the work in
38 ensuring that questions related to the accuracy or integrity of any part of the work are appropriately
39 investigated and resolved; **MP:** Participated in the study design, and critical revision of the paper. He
40 also gave final approval of the version to be published and agreed to be accountable for all aspects of
41 the work in ensuring that questions related to the accuracy or integrity of any part of the work are
42 appropriately investigated and resolved; **DF:** Participated in the study design, database management,
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 statistical support and critical revision of the paper. He also gave final approval of the version to be
4 published and agreed to be accountable for all aspects of the work in ensuring that questions related
5 to the accuracy or integrity of any part of the work are appropriately investigated and resolved; **JMH:**
6 Participated in the study design, acquisition of data, interpretation of data and critical revision of the
7 paper. He also gave final approval of the version to be published and agreed to be accountable for all
8 aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the
9 work are appropriately investigated and resolved; **BO:** Participated in the acquisition of data, study
10 design, database management, and critical revision of the paper. She also gave final approval of the
11 version to be published and agreed to be accountable for all aspects of the work in ensuring that
12 questions related to the accuracy or integrity of any part of the work are appropriately investigated
13 and resolved.
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31

32 References

- 33
34
35 1 WHO. Ageing and Health. [https://www.who.int/news-room/fact-sheets/detail/ageing-and-](https://www.who.int/news-room/fact-sheets/detail/ageing-and-health)
36 health (11 December 2018, date last accessed)
37
38
39
40 2 Wang H, Naghavi M, Allen C, *et al*. Global, regional, and national life expectancy, all-cause
41 mortality, and cause-specific mortality for 249 causes of death, 1980–2015: a systematic
42 analysis for the Global Burden of Disease Study 2015. *Lancet* 2016;**388**:1459–544.
43
44
45
46 doi:10.1016/S0140-6736(16)31012-1
47
48
49
50 3 Multimorbidity: a priority for global health research. 2018.
51
52
53 4 Parekh AK, Barton MB. The challenge of multiple comorbidity for the us health care system.
54
55 *JAMA - J Am Med Assoc* 2010;**303**:1303–4. doi:10.1001/jama.2010.381
56
57
58 5 Mugisha J, Rosenbaum S, Ward PB, *et al*. Chronic physical conditions, multimorbidity and
59
60

- 1
2
3 physical activity across 46 low- and middle-income countries. *Int J Behav Nutr Phys Act*
4
5 2017;**14**. doi:10.1186/s12966-017-0463-5
6
7
8
9 6 Chatterji S, Byles J, Cutler D, *et al*. Health, functioning, and disability in older adults - Present
10 status and future implications. *Lancet* 2015;**385**:563–75. doi:10.1016/S0140-6736(14)61462-8
11
12
13
14 7 Arokiasamy P, Uttamacharya U, Jain K, *et al*. The impact of multimorbidity on adult physical
15 and mental health in low- and middle-income countries: What does the study on global
16 ageing and adult health (SAGE) reveal? *BMC Med* 2015;**13**:1–16. doi:10.1186/s12916-015-
17 0402-8
18
19
20
21
22
23
24 8 Bao J, Chua KC, Prina M, *et al*. Multimorbidity and care dependence in older adults: A
25 longitudinal analysis of findings from the 10/66 study. *BMC Public Health* 2019;**19**:1–10.
26 doi:10.1186/s12889-019-6961-4
27
28
29
30
31 9 Maximova T, Pati S, Biritwum RB, *et al*. The impact of multimorbidity on adult physical and
32 mental health in low- and middle-income countries: what does the study on global ageing and
33 adult health (SAGE) reveal? *BMC Med* 2015;**13**:1–16. doi:10.1186/s12916-015-0402-8
34
35
36
37
38
39 10 Afshar S, Roderick PJ, Kowal P, *et al*. *Applied demography and public health in the 21st*
40 *century*. Cham, Switzerland: : Springer International Publishing AG 2016.
41
42 doi:10.1249/mss.0000000000001652
43
44
45
46 11 Yen L, Valderas JM, McRae IS, *et al*. Multimorbidity and comorbidity of chronic diseases
47 among the senior Australians: prevalence and patterns. *PLoS One* 2014;**9**:e83783.
48 doi:10.1371/journal.pone.0083783
49
50
51
52
53
54 12 Baldo V, Corti MC, Boccuzzo G, *et al*. Multimorbidity patterns in high-need, high-cost elderly
55 patients. *PLoS One* 2018;**13**:e0208875. doi:10.1371/journal.pone.0208875
56
57
58
59 13 Larsen FB, Pedersen MH, Friis K, *et al*. A Latent class analysis of multimorbidity and the
60

- 1
2
3 relationship to socio-demographic factors and health-related quality of life. A national
4
5 population-based study of 162,283 Danish Adults. *PLoS One* 2017;**12**:1–17.
6
7 doi:10.1371/journal.pone.0169426
8
9
- 10
11 14 Alcalá-Nalvaiz JT, Sicras-Mainar A, Calderón-Larrañaga A, *et al.* Multimorbidity Patterns in
12
13 Primary Care: Interactions among Chronic Diseases Using Factor Analysis. *PLoS One*
14
15 2012;**7**:e32190. doi:10.1371/journal.pone.0032190
16
17
- 18
19 15 Guisado-Clavero M, Roso-Llorach A, López-Jimenez T, *et al.* Multimorbidity patterns in the
20
21 elderly: A prospective cohort study with cluster analysis. *BMC Geriatr* 2018;**18**:1–11.
22
23 doi:10.1186/s12877-018-0705-7
24
25
- 26
27 16 Sanchez-Niubo A, Egea-Cortés L, Olaya B *et al.* Cohort profile: the ageing trajectories of health
28
29 - longitudinal opportunities and synergies (ATHLOS) project. *Int J Epidemiol* 2019;**0**:1–11.
30
31 doi:org/10.1093/ije/dyz077
32
33
- 34
35 17 Kowal P, Chatterji S, Naidoo N, *et al.* Data resource profile: The World Health Organization
36
37 study on global AGEing and adult health (SAGE). *Int J Epidemiol* 2013;**41**:1639–49.
38
39 doi:10.1093/ije/dys210
40
41
- 42
43 18 Steptoe A, Breeze E, Banks J, *et al.* Cohort profile: The English Longitudinal Study of Ageing. *Int*
44
45 *J Epidemiol* 2013;**42**:1640–8. doi:10.1093/ije/dys168
46
47
- 48
49 19 Börsch-Supan A, Brandt M, Hunkler C, *et al.* Data resource profile: The survey of health,
50
51 ageing and retirement in europe (SHARE). *Int J Epidemiol* 2013;**42**:992–1001.
52
53 doi:10.1093/ije/dyt088
54
55
- 56
57 20 IDA TWBI. World Bank Open Data. <https://data.worldbank.org/> (5 March 2019, date last
58
59 accessed)
60
- 21 Fortier I, Raina P, Van den Heuvel ER, Griffith LE, Craig C, Saliba M, Doiron D, Stolk RP,

- 1
2
3 Knoppers BM, Ferretti V, Granda P BP. Maelstrom Research guidelines for rigorous
4 retrospective data harmonization. 2017;;103–5. doi:10.1093/ije/dyw075
5
6
7
8
9 22 Ronald C. Kessler TBÜ. The World Mental Health (WMH) survey initiative version of the World
10 Health Organization (WHO) Composite Diagnostic Interview (CIDI). *Int J Methods Psychiatr Res*
11 2004;**13**:93–121. doi.org/10.1002/mpr.168
12
13
14
15
16 23 Radloff L S. The CES-D scale: A self report depression scale for research in the general
17 population. *Appl Psychol Meas* 1977;**1**:385–401. doi:10.1002/ardp.19552881111
18
19
20
21 24 Reischies F, Lobo A, Turrina C, *et al.* Development of the EURO–D scale – a European Union
22 initiative to compare symptoms of depression in 14 European centres. *Br J Psychiatry*
23 2008;**174**:330–8. doi:10.1192/bjp.174.4.330
24
25
26
27
28 25 Zivin K, Llewellyn DJ, Lang IA, *et al.* Depression among older adults in the United States and
29 England. *Am J Geriatr Psychiatry* 2010;**18**:1036–44. doi:10.1097/JGP.0b013e3181dba6d2
30
31
32
33
34 26 Prince MJ, Beekman ATF, Deeg DJH, *et al.* Depression symptoms in late life assessed during
35 the EURO-D scale. Effect of age, gender and marital status in 14 European centres. *Br J*
36 *Psychiatry* 1999;**174**:339–45. doi:10.1192/bjp.174.4.339
37
38
39
40
41 27 Morris JC, Heyman A, Mohs RC, Hughes JP, van Belle G, Fillenbaum G *et al.* The consortium to
42 establish a registry for Alzheimer’s disease (CERAD). Part I. Clinical and neuropsychological
43 assesment of Alzheimer’s disease. *Neurology* 1989;**39**:1159–65. doi:10.1212/wnl.39.9.1159
44
45
46
47
48
49 28 WHO. Country groupings.
50 https://www.who.int/quantifying_ehimpacts/global/ebdcountgroup/en/ (7 February 2019,
51 date last accessed)
52
53
54
55
56 29 United Nations Statistical Division (UNSD). Countries or areas/geographical regions.
57 <https://unstats.un.org/unsd/methodology/m49/> (7 February 2019, date last accessed)
58
59
60

- 1
2
3 30 Lanza ST, Dziak JJ, Huang L WA, Collins L. *LCA Stata plugin users ' guide (version 1.2.1)*. Penn
4 State: : University Park: The Methodology Center 2018.
5
6
7
8 31 Tein JY, Coxe S, Cham H. Statistical power to detect the correct number of classes in latent
9 profile analysis. *Struct Equ Model* 2013;**20**:640–57. doi:10.1080/10705511.2013.824781
10
11
12
13 32 Nylund-Gibson K, Choi AY. Ten frequently asked questions about latent class analysis. *Transl*
14 *Issues Psychol Sci* 2018;**4**:440–61. doi:10.1037/tps0000176
15
16
17
18 33 Masyn KE. Latent Class Analysis and Finite Mixture Modeling. Published Online First: 21 March
19 2013. doi:10.1093/OXFORDHB/9780199934898.013.0025
20
21
22
23 34 StataCorp. *Stata: Release 13. Statistical Software*. College St. 2013.
24
25
26
27 35 Rubin DB. *Multiple Imputation for nonresponse in surveys*. United States of America: : John
28 Wiley & Sons 1987. doi:10.2307/3172772
29
30
31
32 36 Garin N, Olaya B, Perales J, *et al*. Multimorbidity patterns in a national representative sample
33 of the Spanish adult population. *PLoS One* 2014;**9**:e84794. doi:10.1371/journal.pone.0084794
34
35
36
37 37 Olaya B, Moneta MV, Caballero FF, *et al*. Latent class analysis of multimorbidity patterns and
38 associated outcomes in Spanish older adults: A prospective cohort study. *BMC Geriatr*
39 2017;**17**. doi:10.1186/s12877-017-0586-1
40
41
42
43 38 Whitson HE, Johnson KS, Sloane R, *et al*. Identifying patterns of multimorbidity in older
44 americans: application of latent class analysis. *J Am Geriatr Soc* 2016;**64**:1668–73.
45 doi:10.1111/jgs.14201. Identifying
46
47
48
49 39 Prados-Torres A, Calderón-Larrañaga A, Hanco-Saavedra J, *et al*. Multimorbidity patterns: A
50 systematic review. *J Clin Epidemiol* 2014;**67**:254–66. doi:10.1016/j.jclinepi.2013.09.021
51
52
53
54 40 Alberti KGM, Zimmet P, Shaw J. Metabolic syndrome -a new world-wide definition. A
55
56
57
58
59
60

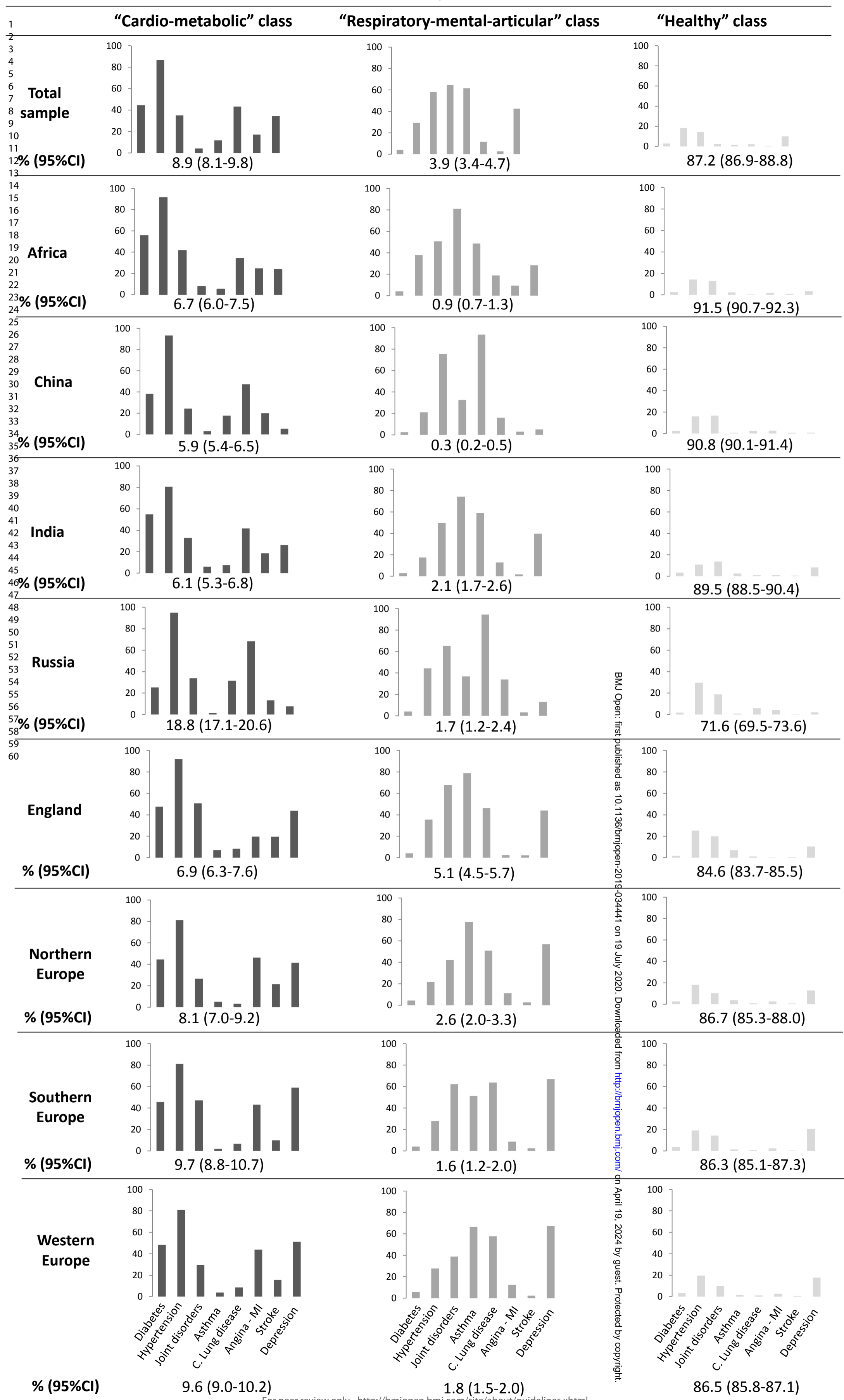
- 1
2
3 consensus statement from the International Diabetes Federation. *Abstr Appl Anal*
4
5 2006;**23**:469–80. doi:10.1155/2010/307409
6
7
8
9 41 Margaretten M, Julian L, Katz P, *et al*. Depression in patients with rheumatoid arthritis:
10 description, causes and mechanisms. *Int J Clin Rheumtol* 2011;**6**:617–23.
11
12 doi:10.2217/IJR.11.6.Depression
13
14
15
16 42 Druss BG, Walker ER. Mental disorders and medical comorbidity. 2011.
17
18 https://www.rwjf.org/content/dam/farm/reports/issue_briefs/2011/rwjf69438/subassets/rw
19
20 [jf69438_1](https://www.rwjf.org/content/dam/farm/reports/issue_briefs/2011/rwjf69438/subassets/rwjf69438_1)
21
22
23
24 43 Barbara Messner DB. Smoking and cardiovascular disease: Mechanisms of endothelial
25 dysfunction and early atherogenesis. 2014;**34**:509–15. doi:10.1161/ATVBAHA.113.300156
26
27
28
29 44 Meade T, Manolios N, Cumming SR, *et al*. Cognitive impairment in rheumatoid arthritis: A
30 systematic review. *Arthritis Care Res* 2018;**70**:39–52. doi:10.1002/acr.23243
31
32
33
34 45 Lam RW, Kennedy SH, McIntyre RS, *et al*. Cognitive dysfunction in major depressive disorder:
35 effects on psychosocial functioning and implications for treatment. *Can J Psychiatry*
36
37 2014;**59**:649–54. doi:10.1177/070674371405901206
38
39
40
41
42 46 Dodd JW. Lung disease as a determinant of cognitive decline and dementia. *Alzheimer's Res*
43
44 *Ther* 2015;**7**:1–8. doi:10.1186/s13195-015-0116-3
45
46
47
48 47 Rose Mayeda E, Whitmer RA, Yaffe K. Diabetes and cognition. *Clin Geriatr Med* 2015;**31**:101–
49
50 ix. doi:10.1016/j.cger.2014.08.021
51
52
53
54 48 Walker KA, Power MC, Gottesman RF. Defining the relationship between hypertension,
55
56 cognitive decline, and dementia: a review. 2018;**19**:24. doi:10.1007/s11906-017-0724-3
57
58
59
60 49 Deckers K, Schievink SHJ, Rodriguez MMF, *et al*. Coronary heart disease and risk for cognitive
impairment or dementia: Systematic review and meta-analysis. *PLoS One* 2017;**12**:1–15.

- 1
2
3 doi:10.1371/journal.pone.0184244
4
5
6 50 Leon DA, Shkolnikov VM, McKee M, *et al.* Alcohol increases circulatory disease mortality in
7
8 Russia: Acute and chronic effects or misattribution of cause? *Int J Epidemiol* 2010;**39**:1279–90.
9
10 doi:10.1093/ije/dyq102
11
12
13 51 Petrukhin IS, Lunina EY. Cardiovascular disease risk factors and mortality in Russia: Challenges
14
15 and barriers. *Public Health Rev* 2012;**33**:436–49. doi:10.1007/BF03391645
16
17
18 52 Y. He, M. Zhang, E. H. B. Lin *et al.* Mental disorders among persons with arthritis: results from
19
20 the World Mental Health Surveys. 2008;**38**:1639–50. doi:10.1017/S0033291707002474
21
22
23 53 Jamrozik E, Musk AW. Respiratory health issues in the Asia-Pacific region: An overview.
24
25 *Respirology* 2011;**16**:3–12. doi:10.1111/j.1440-1843.2010.01844.x
26
27
28 54 Sandeep Grover, Alakananda Dutt AA. An overview of Indian research in depression. *Indian J*
29
30 *Psychiatry* 2010;**52**:178–88. doi:10.4103/0019-5545.69231
31
32
33 55 Gu L, Xie J, Long J, *et al.* Epidemiology of major depressive disorder in Mainland China: A
34
35 Systematic review. *PLoS One* 2013;**8**:e65356. doi:10.1371/journal.pone.0065356
36
37
38 56 Hou Z, Jiang W, Yin Y, *et al.* The current situation on major depressive disorder in China:
39
40 Research on mechanisms and clinical practice. *Neurosci Bull* 2016;**32**:389–97.
41
42
43 doi:10.1007/s12264-016-0037-6
44
45
46 57 Castro-Costa E, Dewey M, Stewart R, *et al.* Ascertaining late-life depressive symptoms in
47
48 Europe: An evaluation of the survey version of the EURO-D scale in 10 nations. The SHARE
49
50 project. *Int J Methods Psychiatr Res* 2008;**17**:12–29. doi:10.1002/mpr.236
51
52
53 58 Martin LM, Leff M, Calonge N, *et al.* Validation of self-reported chronic conditions and health
54
55 services in a managed care population. *Am J Prev Med* 2000;**18**:215–8. doi:10.1016/S0749-
56
57 3797(99)00158-0
58
59
60

- 1
2
3 59 Finch H. A comparison of statistics for assessing model invariance in latent class analysis. *Open*
4
5 *J Stat* 2015;**5**:191–210. doi:10.4236/ojs.2015.53022
6
7
8 60 White IR, Carlin JB. Bias and efficiency of multiple imputation compared with complete-case
9 analysis for missing covariate values. *Stat Med* 2010;**29**:2920–31. doi:10.1002/sim.3944
10
11
12
13 61 Sterne JAC, White IR, Carlin JB, *et al*. Multiple imputation for missing data in epidemiological
14 and clinical research: Potential and pitfalls. *BMJ* 2009;**339**:157–60. doi:10.1136/bmj.b2393
15
16
17
18
19
20
21

22 **Figure 1** Prevalence of diseases in the three latent classes in the total sample and by regions (subsample 50-64
23 years)

24
25
26
27 **Figure 2** Prevalence of diseases in the three latent classes in the total sample and by regions (subsample +65
28 years)
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60



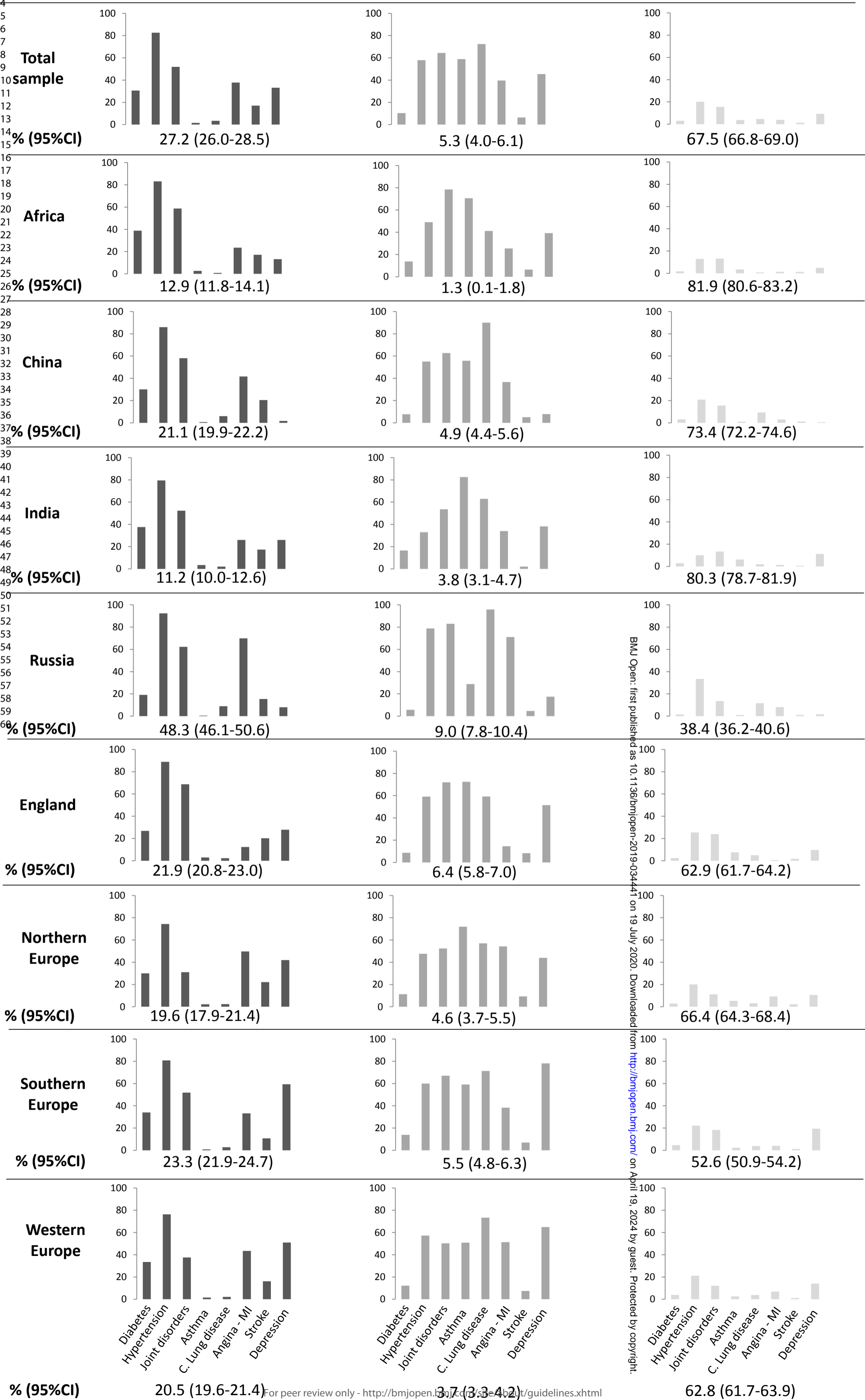
BMJ Open: first published as 10.1136/bmjopen-2019-034441 on 19 July 2020. Downloaded from <http://bmjopen.bmj.com/> on April 19, 2024 by guest. Protected by copyright.

Multimorbidity latent classes

“Cardio-metabolic” class

“Respiratory-mental-articular” class

“Healthy” class



BMJ Open: first published as 10.1136/bmjopen-2019-034441 on 19 July 2020. Downloaded from <http://bmjopen.bmj.com/> on April 19, 2024 by guest. Protected by copyright.

Table S1. Proportion of missingness in the variables of interest in Mexico

(a) Indicators of multimorbidity classes	
	n (%) missingness
Diabetes	2814 (51.65)
Hypertension	2815 (51.67)
Joint disorders	2814 (51.65)
Asthma	2814 (51.65)
Chronic lung disease	2814 (51.65)
Myocardial infarction – Angina	2814 (51.65)
Stroke	2814 (51.65)
Depression	2815 (51.67)
(b) Variables used in the regression model	
Sex	2706 (49.67)
Age	2707 (49.69)
Country	0 (0.00)
Study	0 (0.00)
Marital status	2811 (51.60)
Education	3264 (59.91)
Wealth	1922 (35.28)
Loneliness	2819 (51.74)
Ever smoked	2814 (51.65)
Vigorous exercise in last 2 weeks	2814 (51.65)
Self-rated health	2811 (51.6)
ADL – Using the toilet	2814 (51.65)
ADL – Bathing or showering	2815 (51.67)
ADL – Getting dressed	0 (0.00)
ADL - Eating	2814 (51.65)
ADL – Getting in or out of bed	2813 (51.63)
ADL – Moving around the house	2817 (51.71)
Memory: Immediate recall	2854 (52.39)
Memory: Delayed recall	2854 (52.39)
Verbal fluency	2852 (52.35)

Table S2. Overall proportion of missingness for indicators of latent classes and variables included in the imputation model

(a) Indicators of multimorbidity classes	
	n (%) missingness
Diabetes	62 (0.08)
Hypertension	154 (0.21)
Joint disorders	31 (0.04)
Asthma	82 (0.11)
Chronic lung disease	51 (0.07)
Myocardial infarction – Angina	63 (0.08)
Stroke	41 (0.05)
Depression	1233 (1.70)
(b) Variables used in the regression model	
Sex	0 (0.00)
Age	0 (0.00)
Country	0 (0.00)
Study	0 (0.00)
Marital status	106 (0.14)
Education	11475 (15.91)
Wealth	844 (1.17)
Loneliness	10985 (15.23)
Ever smoked	201 (0.27)
Vigorous exercise in last 2 weeks	258 (0.35)
Self-rated health	218 (0.30)
ADL – Using the toilet	355 (0.49)
ADL – Bathing or showering	340 (0.47)
ADL – Getting dressed	304 (0.42)
ADL - Eating	400 (0.55)
ADL – Getting in or out of bed	310 (0.42)
ADL – Moving around the house	345 (0.47)
Memory: Immediate recall	1442 (1.99)
Memory: Delayed recall	1448 (2.00)
Verbal fluency	1641 (2.27)

Table S3. Proportion of missingness for indicators of latent classes and variables included in the imputation model for region 1 (SAGE study - Africa: Ghana, South Africa)

(a) Indicators of multimorbidity classes	
	n (%) missingness
Diabetes	1 (0.01)
Hypertension	1 (0.01)
Joint disorders	3 (0.03)
Asthma	1 (0.01)
Chronic lung disease	1 (0.01)
Myocardial infarction – Angina	3 (0.03)
Stroke	1 (0.01)
Depression	17 (0.21)
(b) Variables used in the regression model	
Sex	0 (0.00)
Age	0 (0.00)
Country	0 (0.00)
Study	0 (0.00)
Marital status	89 (1.11)
Education	3705 (46.6)
Wealth	23 (0.28)
Loneliness	48 (0.60)
Ever smoked	7 (0.08)
Vigorous exercise in last 2 weeks	12 (0.15)
Self-rated health	15 (0.18)
ADL – Using the toilet	40 (0.50)
ADL – Bathing or showering	32 (0.40)
ADL – Getting dressed	27 (0.33)
ADL - Eating	31 (0.38)
ADL – Getting in or out of bed	32 (0.40)
ADL – Moving around the house	35 (0.44)
Memory: Immediate recall	36 (0.45)
Memory: Delayed recall	31 (0.38)
Verbal fluency	9 (0.11)

Table S4. Proportion of missingness for indicators of latent classes and variables included in the imputation model for region 2 (SAGE study - Eastern Asia: China)

(a) Indicators of multimorbidity classes	
	n (%) missingness
Diabetes	35 (0.27)
Hypertension	127 (0.98)
Joint disorders	2 (0.01)
Asthma	53 (0.41)
Chronic lung disease	23 (0.17)
Myocardial infarction – Angina	34 (0.26)
Stroke	13 (0.10)
Depression	146 (1.13)
(b) Variables used in the regression model	
Sex	0 (0.00)
Age	0 (0.00)
Country	0 (0.00)
Study	0 (0.00)
Marital status	8 (0.06)
Education	3105 (24.18)
Wealth	18 (0.14)
Loneliness	74 (0.57)
Ever smoked	2 (0.01)
Vigorous exercise in last 2 weeks	45 (0.35)
Self-rated health	16 (0.12)
ADL – Using the toilet	51 (0.39)
ADL – Bathing or showering	52 (0.40)
ADL – Getting dressed	41 (0.31)
ADL - Eating	46 (0.35)
ADL – Getting in or out of bed	36 (0.28)
ADL – Moving around the house	48 (0.37)
Memory: Immediate recall	317 (2.46)
Memory: Delayed recall	399 (3.10)
Verbal fluency	335 (2.60)

Table S5. Proportion of missingness for indicators of latent classes and variables included in the imputation model for region 2 (SAGE study - Southern Asia: India)

(a) Indicators of multimorbidity classes	
	n (%) missingness
Diabetes	0 (0.00)
Hypertension	1 (0.01)
Joint disorders	1 (0.01)
Asthma	0 (0.00)
Chronic lung disease	0 (0.00)
Myocardial infarction – Angina	0 (0.00)
Stroke	1 (0.01)
Depression	1 (0.01)
(b) Variables used in the regression model	
Sex	0 (0.00)
Age	0 (0.00)
Country	0 (0.00)
Study	0 (0.00)
Marital status	0 (0.00)
Education	0 (0.00)
Wealth	38 (0.57)
Loneliness	10 (0.15)
Ever smoked	1 (0.01)
Vigorous exercise in last 2 weeks	0 (0.00)
Self-rated health	0 (0.00)
ADL – Using the toilet	19 (0.28)
ADL – Bathing or showering	12 (0.18)
ADL – Getting dressed	11 (0.16)
ADL - Eating	73 (1.11)
ADL – Getting in or out of bed	7 (0.10)
ADL – Moving around the house	24 (0.36)
Memory: Immediate recall	88 (1.34)
Memory: Delayed recall	88 (1.34)
Verbal fluency	71 (1.08)

Table S6. Proportion of missingness for indicators of latent classes and variables included in the imputation model for region 2 (SAGE study - Eastern Europe: the Russian Federation)

(a) Indicators of multimorbidity classes	
	n (%) missingness
Diabetes	5 (0.12)
Hypertension	4 (0.10)
Joint disorders	3 (0.07)
Asthma	6 (0.15)
Chronic lung disease	5 (0.12)
Myocardial infarction – Angina	5 (0.12)
Stroke	5 (0.12)
Depression	14 (0.36)
(b) Variables used in the regression model	
Sex	0 (0.00)
Age	0 (0.00)
Country	0 (0.00)
Study	0 (0.00)
Marital status	6 (0.15)
Education	41 (1.05)
Wealth	5 (0.12)
Loneliness	79 (2.03)
Ever smoked	1 (0.02)
Vigorous exercise in last 2 weeks	0 (0.00)
Self-rated health	5 (0.12)
ADL – Using the toilet	47 (1.20)
ADL – Bathing or showering	46 (1.18)
ADL – Getting dressed	27 (0.69)
ADL - Eating	52 (1.33)
ADL – Getting in or out of bed	37 (0.95)
ADL – Moving around the house	40 (1.02)
Memory: Immediate recall	153 (3.93)
Memory: Delayed recall	81 (2.08)
Verbal fluency	208 (5.35)

Table S7. Proportion of missingness for indicators of latent classes and variables included in the imputation model for region 2 (ELSA study- England)

(a) Indicators of multimorbidity classes	
	n (%) missingness
Diabetes	2 (0.01)
Hypertension	2 (0.01)
Joint disorders	3 (0.02)
Asthma	3 (0.02)
Chronic lung disease	3 (0.02)
Myocardial infarction – Angina	2 (0.01)
Stroke	2 (0.01)
Depression	372 (3.23)
(b) Variables used in the regression model	
Sex	0 (0.00)
Age	0 (0.00)
Country	0 (0.00)
Study	0 (0.00)
Marital status	0 (0.00)
Education	1023 (8.88)
Wealth	330 (2.86)
Loneliness	356 (3.09)
Ever smoked	173 (1.50)
Vigorous exercise in last 2 weeks	176 (1.52)
Self-rated health	175 (1.51)
ADL – Using the toilet	173 (1.50)
ADL – Bathing or showering	173 (1.50)
ADL – Getting dressed	173 (1.50)
ADL - Eating	173 (1.50)
ADL – Getting in or out of bed	173 (1.50)
ADL – Moving around the house	173 (1.50)
Memory: Immediate recall	360 (3.12)
Memory: Delayed recall	373 (3.23)
Verbal fluency	362 (3.14)

Table S8. Proportion of missingness for indicators of latent classes and variables included in the imputation model for region 2 (SHARE study - Northern Europe: Denmark, Sweden)

(a) Indicators of multimorbidity classes	
	n (%) missingness
Diabetes	2 (0.04)
Hypertension	2 (0.04)
Joint disorders	2 (0.04)
Asthma	2 (0.04)
Chronic lung disease	2 (0.04)
Myocardial infarction – Angina	2 (0.04)
Stroke	2 (0.04)
Depression	78 (1.70)
(b) Variables used in the regression model	
Sex	0 (0.00)
Age	0 (0.00)
Country	0 (0.00)
Study	0 (0.00)
Marital status	0 (0.00)
Education	52 (1.13)
Wealth	6 (0.13)
Loneliness	1384 (30.26)
Ever smoked	4 (0.08)
Vigorous exercise in last 2 weeks	6 (0.13)
Self-rated health	2 (0.04)
ADL – Using the toilet	3 (0.06)
ADL – Bathing or showering	3 (0.06)
ADL – Getting dressed	3 (0.06)
ADL - Eating	3 (0.06)
ADL – Getting in or out of bed	3 (0.06)
ADL – Moving around the house	3 (0.06)
Memory: Immediate recall	68 (1.48)
Memory: Delayed recall	66 (1.44)
Verbal fluency	83 (1.81)

Table S9. Proportion of missingness for indicators of latent classes and variables included in the imputation model for region 2 (SHARE study - Southern Europe: Greece, Italy, Spain)

(a) Indicators of multimorbidity classes	
	n (%) missingness
Diabetes	12 (0.16)
Hypertension	12 (0.16)
Joint disorders	12 (0.16)
Asthma	12 (0.16)
Chronic lung disease	12 (0.16)
Myocardial infarction – Angina	12 (0.16)
Stroke	12 (0.16)
Depression	194 (2.59)
(b) Variables used in the regression model	
Sex	0 (0.00)
Age	0 (0.00)
Country	0 (0.00)
Study	0 (0.00)
Marital status	1 (0.01)
Education	20 (0.26)
Wealth	95 (1.27)
Loneliness	2715 (36.37)
Ever smoked	3 (0.04)
Vigorous exercise in last 2 weeks	4 (0.05)
Self-rated health	0 (0.00)
ADL – Using the toilet	3 (0.04)
ADL – Bathing or showering	3 (0.04)
ADL – Getting dressed	3 (0.04)
ADL - Eating	3 (0.04)
ADL – Getting in or out of bed	3 (0.04)
ADL – Moving around the house	3 (0.04)
Memory: Immediate recall	125 (1.67)
Memory: Delayed recall	125 (1.67)
Verbal fluency	176 (2.35)

Table S10. Proportion of missingness for indicators of latent classes and variables included in the imputation model for region 2 (SHARE study - Western Europe: Austria, Belgium, France, Germany, Israel, Netherlands, Switzerland)

(a) Indicators of multimorbidity classes	
	n (%) missingness
Diabetes	5 (0.02)
Hypertension	5 (0.02)
Joint disorders	5 (0.02)
Asthma	5 (0.02)
Chronic lung disease	5 (0.02)
Myocardial infarction – Angina	5 (0.02)
Stroke	5 (0.02)
Depression	411 (2.36)
(b) Variables used in the regression model	
Sex	0 (0.00)
Age	0 (0.00)
Country	0 (0.00)
Study	0 (0.00)
Marital status	2 (0.01)
Education	165 (0.95)
Wealth	329 (1.89)
Loneliness	6319 (36.42)
Ever smoked	10 (0.05)
Vigorous exercise in last 2 weeks	15 (0.08)
Self-rated health	5 (0.02)
ADL – Using the toilet	19 (0.10)
ADL – Bathing or showering	19 (0.10)
ADL – Getting dressed	19 (0.10)
ADL - Eating	19 (0.10)
ADL – Getting in or out of bed	19 (0.10)
ADL – Moving around the house	19 (0.10)
Memory: Immediate recall	295 (1.7)
Memory: Delayed recall	285 (1.64)
Verbal fluency	397 (2.28)

Table S1. Association between latent multimorbidity membership and outcomes in individuals aged 50-64 and ≥65

Outcomes ^a	Aged 50-64		Aged ≥65	
	“Cardio-metabolic” class	“Respiratory-mental-articular” class	“Cardio-metabolic” class	“Respiratory-mental-articular” class
Sex				
Woman	1.0	1.0	1.0	1.0
Male	0.9 (0.8-1.0)	0.6 (0.5-0.7)	0.7 (0.7-0.8)	0.9 (0.8-1.0)
Age (years)	1.1 (1.1-1.1)	1.1 (1.0-1.1)	1.0 (1.0-1.0)	1.0 (1.0-1.0)
Marital status				
Single	1.0	1.0	1.0	1.0
Married	0.9 (0.8-1.1)	0.7 (0.5-1.0)	1.2 (1.00-1.4)	1.2 (0.9-1.6)
Divorced	1.1 (0.9-1.4)	1.7 (1.2-2.4)	1.1 (1.0-1.4)	1.8 (1.2-2.6)
Widowed	1.4 (1.1-1.7)	1.1 (0.8-1.6)	1.5 (1.3-1.8)	1.6 (1.2-2.3)
Education level				
Primary or less	1.0	1.0	1.0	1.0
Secondary	0.9 (0.9-1.0)	0.7 (0.6-0.8)	1.0 (0.9-1.0)	0.7 (0.6-0.8)
Tertiary	0.8 (0.7-0.9)	0.6 (0.5-0.7)	0.9 (0.8-1.0)	0.6 (0.5-0.7)
Wealth				
1 st (worse)	1.0	1.0	1.0	1.0
2 nd	1.0 (0.9-1.1)	0.8 (0.6-1.0)	1.1 (1.0-1.2)	1.1 (0.9-1.2)
3 rd	0.9 (0.9-1.1)	0.7 (0.5-0.8)	1.1 (1.0-1.2)	1.0 (0.9-1.2)
4 th	0.9 (0.8-1.0)	0.6 (0.5-0.8)	1.1 (1.0-1.2)	0.8 (0.7-1.0)
5 th (best)	0.8 (0.7-0.9)	0.5 (0.4-0.6)	1.2 (1.0-1.2)	0.7 (0.6-0.9)
Region				
Africa	1.0	1.0	1.0	1.0
China	0.9 (0.8-1.0)	0.3 (0.2-0.5)	1.8 (1.6-2.1)	4.1 (3.0-5.7)
India	0.9 (0.8-1.1)	2.2 (1.5-3.2)	0.9 (0.8-1.0)	2.9 (2.0-4.2)
Russia	3.6 (3.0-4.2)	2.2 (1.4-3.5)	8.0 (7.0-9.2)	14.5 (10.3-20.3)
England	1.1 (0.9-1.3)	5.6 (4.1-7.7)	2.2 (2.0-2.5)	6.2 (4.5-8.5)
Northern Europe	1.3 (1.1-1.5)	2.8 (1.9-4.1)	1.9 (1.6-2.2)	4.2 (2.9-6.1)
Southern Europe	1.5 (1.3-1.8)	1.7 (1.2-2.5)	2.8 (2.5-3.2)	6.5 (4.7-9.0)
Western Europe	1.5 (1.3-1.7)	1.9 (1.3-2.6)	2.1 (1.9-2.3)	3.6 (2.6-5.0)
Loneliness (yes/no)	1.9 (1.8-2.1)	2.5 (2.1-3.0)	1.4 (1.4-1.6)	1.9 (1.7-2.2)
Ever smoked (yes/no)	1.0 (0.9-1.1)	1.7 (1.5-2.0)	0.8 (0.8-0.8)	1.4 (1.3-1.6)
Physical activity (yes/no)	0.5 (0.5-0.5)	0.6 (0.5-0.7)	0.5 (0.5-0.6)	0.5 (0.4-0.6)
Limitations in ADL (yes/no)	2.5 (2.3-2.7)	3.2 (2.7-3.7)	1.8 (1.7-1.9)	2.8 (2.5-3.1)
Self-rated health				
Good	1.0	1.0	1.0	1.0
Moderate	3.7 (3.3-3.9)	2.5 (2.1-2.9)	2.5 (2.3-2.7)	3.9 (3.3-4.6)
Poor	8.5 (7.7-9.4)	7.4 (6.1-8.8)	4.3 (4.6)	10.9 (9.3-12.8)
Memory: Immediate recall	0.9 (0.9-0.9)	0.9 (0.9-1.0)	0.9 (0.9-0.9)	0.9 (0.9-0.9)
Memory: Delayed recall	0.9 (0.9-0.9)	0.9 (0.8-0.9)	0.9 (0.9-0.9)	0.9 (0.9-0.9)
Verbal fluency	1.0 (1.0-1.0)	1.0 (1.0-1.0)	1.0 (0.9-1.00)	1.0 (0.9-1.0)

Note: The reference group for the multimorbidity group variable was the “Healthy” class
Relative Risk Ratios (95% confidence interval) from multinomial logistic regression models
Models were run in 100 imputed datasets and results combined using Rubin’s rules
ADL Activities of Daily Living
^aUnadjusted analyses.

STROBE CHECKLIST: “MULTIMORBIDITY PATTERNS IN LOW-MIDDLE AND HIGH INCOME REGIONS: A MULTI-REGION LATENT CLASS ANALYSIS USING ATHLOS HARMONIZED COHORTS”

1. Title and abstract

- (a) Indicate the study’s design with a commonly used term in the title or the abstract **(page 1 and 3)**
- (b) Provide in the abstract an informative and balanced summary of what was done and what was found **(page 3)**

Introduction

2. Background/rationale: Explain the scientific background and rationale for the investigation being reported **(page 5-6)**
3. Objectives: State specific objectives, including any prespecified hypotheses **(page 6, lines 10-13)**

Methods

4. Study design: Present key elements of study design early in the paper **(page 6, lines 16-19)**
5. Setting: Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection **(page 6, lines 20-24, page 7, lines 1-3)**.
6. Participants: Give the eligibility criteria, and the sources and methods of selection of Participants **(page 7, lines 4-9)**.
7. Variables: Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable. **(page 7, lines 13- page 8)**.
8. Data sources/measurement: For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group **(page 7, lines 13- page 8)**.
9. Bias: Describe any efforts to address potential sources of bias **(page 9, lines 19-23)**.
10. Study size: Explain how the study size was arrived at **(page 7, lines 4-9)**.
11. Quantitative variables: Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why. **(page 7, lines 13- page 8)**.
12. Statistical methods **(page 9-10)**:
 - (a) Describe all statistical methods, including those used to control for confounding
 - (b) Describe any methods used to examine subgroups and interactions
 - (c) Explain how missing data were addressed
 - (d) If applicable, describe analytical methods taking account of sampling strategy
 - (e) Describe any sensitivity analyses

Results

13. Participants **(page 7, lines 4-9)**:

- 1
- 2
- 3 (a) Report numbers of individuals at each stage of study—eg numbers potentially eligible,
- 4 examined for eligibility, confirmed eligible, included in the study, completing follow-up,
- 5 and analysed
- 6 (b) Give reasons for non-participation at each stage
- 7 (c) Consider use of a flow diagram
- 8
- 9

10 14. Descriptive data (**page 10, lines 8-18; page 11, lines 1-7**):

- 11 (a) Give characteristics of study participants (eg demographic, clinical, social) and
- 12 information on exposures and potential confounders
- 13 (b) Indicate number of participants with missing data for each variable of interest
- 14 (**Supplementary file 1**)
- 15
- 16

17 15. Outcome data: Report numbers of outcome events or summary measures (**pages 12-15**).

18 16. Main results (**page 15, lines 14-21; to page 17**):

- 19 (a) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their
- 20 precision (eg, 95% confidence interval). Make clear which confounders were adjusted
- 21 for and why they were included
- 22 (b) Report category boundaries when continuous variables were categorized
- 23 (c) If relevant, consider translating estimates of relative risk into absolute risk for a
- 24 meaningful time period
- 25
- 26
- 27
- 28

29 17. Other analyses: Report other analyses done—eg analyses of subgroups and interactions, and

30 sensitivity analyses. **Not applicable.**

31 **Discussion**

32 18. Key results: Summarise key results with reference to study objectives (**page 17, lines 5-8**).

33 19. Limitations: Discuss limitations of the study, taking into account sources of potential bias or

34 imprecision. Discuss both direction and magnitude of any potential bias (**page 21, lines 19-24,**

35 **page 22, 1-14**).

36 20. Interpretation: Give a cautious overall interpretation of results considering objectives,

37 limitations, multiplicity of analyses, results from similar studies, and other relevant evidence

38 (**pages 18-21**).

39 21. Generalisability: Discuss the generalisability (external validity) of the study results (**page 19,**

40 **13-15**).

41 **Other information**

42 22. Funding: Give the source of funding and the role of the funders for the present study and, if

43 applicable, for the original study on which the present article is based (**page 23, lines 20-22,**

44 **page 24, lines 1-6**).

45

46

47

48

49

50

51

52

53

54

55

56

57

58

59

60