ARIC Manuscript Proposal #4267

PC Reviewed: 6/13/23	**Status: _____**	**Priority: 2**
SC Reviewed: _____	**Status: _____**	**Priority: ____**

1.a. Full Title: The association between midlife and late life risk factor and chronic disease clustering and stroke incidence and stroke severity

b. Abbreviated Title (Length 26 characters):

2. Writing Group:
Writing group members:

First author: Marco Egle
Other authors: Anna Kucharska-Newton, Michelle Johansen, Renee Groechel, Rebecca Gottesman, Kamakshi Lakshminarayan, *others are welcome to join*

Last author: Silvia Koton

I, the first author, confirm that all the coauthors have given their approval for this manuscript proposal. _____ **[please confirm with your initials electronically or in writing]**

First author: Dr. Marco Egle
Address: Building 10, Room 4D37, 10 Center Drive, Bethesda, MD 20814

Phone: 220-234-2831
E-mail: eglemt@nih.gov

ARIC author to be contacted if there are questions about the manuscript and the first author does not respond or cannot be located (this must be an ARIC investigator).

Name: Dr. Rebecca Gottesman
Address: Building 10, Room 4D37, 10 Center Drive, Bethesda, MD 20814

Phone: 301-435-9321
E-mail: rebecca.gottesman@nih.gov

3. Timeline:

Manuscript prepared and submitted in ~5-6 months.

4. Rationale:

Co-occurrence of two or more chronic diseases is a major public health burden. This condition is frequently referred to as multimorbidity [1, 2]. However, the number and types of conditions significantly vary across studies and there is currently no consensus on how to best measure multimorbidity. [3] One problem in defining multimorbidity is that there are multiple risk factors underlying multimorbidity such as smoking, drinking and hypertension. It is unclear so far how to best consider them when studying multimorbidity. Assessing those risk factors in addition to the comorbid health conditions based on a simple count of conditions or based on weighted indices can pose significant limitations as they do not co-occur at random in a person but cluster according to specific patterns [4-6]. Distinct risk factor and chronic disease clusters related to vascular, metabolic, sensory, and psychiatric conditions have been shown to be associated with poorer health status, higher disability burden, and increased risk of dementia and mortality [7-10].

Studies in epidemiology have demonstrated that the accumulation of multiple unfavorable health conditions is associated with aging and that most individuals above age 65 suffer from multiple chronic conditions [2]. However, having multiple health conditions and risk factors is not simply a feature of aging: the Rochester Epidemiology Project (REP) showed that more people below age 65 (*N*= 17,838) than above 65 (*N*= 13,606) suffered from multiple risk factor and chronic disease conditions [11]. The difference in prevalence was not sex-specific but varied by race. More Black participants than White participants aged 50-59 already had multiple unfavorable health conditions in this US sample [11]. One significant predictor for early onset is low socioeconomic status [12, 13]. Evidence has shown that young and middle aged individuals in economically disadvantaged areas had similar rates of multiple chronic diseases and risk factors as people aged 10-15 years older in the most affluent areas [2].

Few studies have thus far investigated the impact of age-dependent onset of multiple chronic disease and risk factor conditions on clinical outcomes [14]. Recent evidence suggests that the onset

of multimorbidity in midlife is associated with significantly worse late life outcomes [15, 16]. In a large community-based study, participants with multimorbidity at age 55 had a significantly higher risk for late life dementia conversion than individuals with multimorbidity onset at age 70 [15].

It is unclear whether age-dependent onset of multiple chronic disease and risk factor conditions prior to the stroke event is also an important predictor for stroke incidence overall and incidence of more severe stroke events. Previous studies have shown that a significant proportion of stroke survivors have multiple chronic conditions and risk factors which are associated with a higher mortality rate and worse functional outcomes after stroke [17]. Commonly reported comorbid conditions in stroke patients include hypertension (35-89.8%), ischemic heart disease (14-38.1%), diabetes mellitus (9-35.6%), chronic pulmonary disease (8.1-30.1%), coronary heart disease (29.5%), clinical depression (5.9-21.9%), renal disease (13.9%), atrial fibrillation (11.0%), cancer (10.9- 18.3%), and congestive heart failure (8%). [17-21]. Understanding when these chronic disease and risk factor patterns first become apparent and whether these clusters can predict stroke incidence and stroke severity prior to a stroke event is unclear. Identifying joined contributions of chronic diseases and risk factors to stroke incidence and severity, beyond that imparted by individual health conditions, may provide a clinically applicable preventive approach to the reduction of the overall stroke burden in the population.

In this study, we aim to characterize chronic disease and risk factor clusters in midlife and late life in the context of stroke and to test whether there are clusters in midlife and late life that are associated with stroke incidence and severity of stroke. Our specific hypotheses were formulated based on our prior work in the Taiwan Stroke Registry study. The study, which is currently under review, showed that clusters related to chronic diseases, namely diabetes, peripheral artery disease and uremia, were associated with greater post-stroke disability burden as compared to other clusters with less representation of these conditions. The study also demonstrated that individuals who were part of a cluster related to smoking and drinking, were overall younger at the stroke event and had a faster post-stroke recovery rate. Based on these results, we would like to evaluate if similar associations are seen in a US population, but will also evaluate if novel clusters are relevant to stroke and its severity. Thus, we would like to test whether clusters related to chronic diseases, namely diabetes, peripheral artery disease and kidney disease in mid and late life are

associated with more severe strokes. This study also aims to assess whether clusters related to the risk factors smoking and drinking are associated with minor stroke, but not severe stroke. We furthermore want to determine whether cluster onset in midlife are overall more strongly associated with late life stroke than are late life clusters. Recognizing the importance of social and demographic factors on stroke incidence and severity, we will also explore whether the associations differ by age, sex, and race.

5. Main Hypothesis/Study Questions:

Primary hypotheses:

Stroke incidence:

- *Hypothesis 1*: There are chronic diseases and risk factor clusters in midlife and late life which are associated with stroke incidence. We hypothesize that those clusters containing diabetes, peripheral artery disease and/or kidney disease have the strongest associations with incident stroke, and we anticipate stronger associations when evaluated in midlife.

Stroke severity:

- *Hypothesis 2:* Clusters of chronic diseases associated with stroke severity are different in midlife and late life.

Stroke incidence and stroke severity in late life:

- *Hypothesis 3:* Among individuals alive at visit 5, particular chronic diseases and risk factor clusters identified in midlife (at visit 1) have stronger associations with post-visit 5 stroke incidence than clusters defined based on late life diseases and risk factors (at visit 5).
- *Hypothesis 4:* Chronic diseases and risk factor cluster onset in midlife (at visit 1) have stronger associations with severe strokes than clusters onset in late life (at visit 5).

Secondary hypothesis:

- Chronic diseases and risk factor clusters in midlife are not more strongly associated with stroke at age 65 years or younger than stroke at older age.
- The association of chronic diseases and risk factor clusters in mid and late life and stroke severity varies by sex. Chronic disease and risk factor clusters are more strongly associated with severe strokes in males than in females, although females experience more severe stroke events.
- Chronic disease and risk factor clusters in midlife are more strongly associated with stroke incidence and severe stroke events in individuals with a low educational level.
- Chronic disease and risk factor clusters in midlife are more strongly associated with stroke incidence and severe stroke events in Black than White individuals.

6. Design and analysis (study design, inclusion/exclusion, outcome, and other variables of interest with specific reference to the time of their collection, summary of data analysis, and any anticipated methodologic limitations or challenges if present).

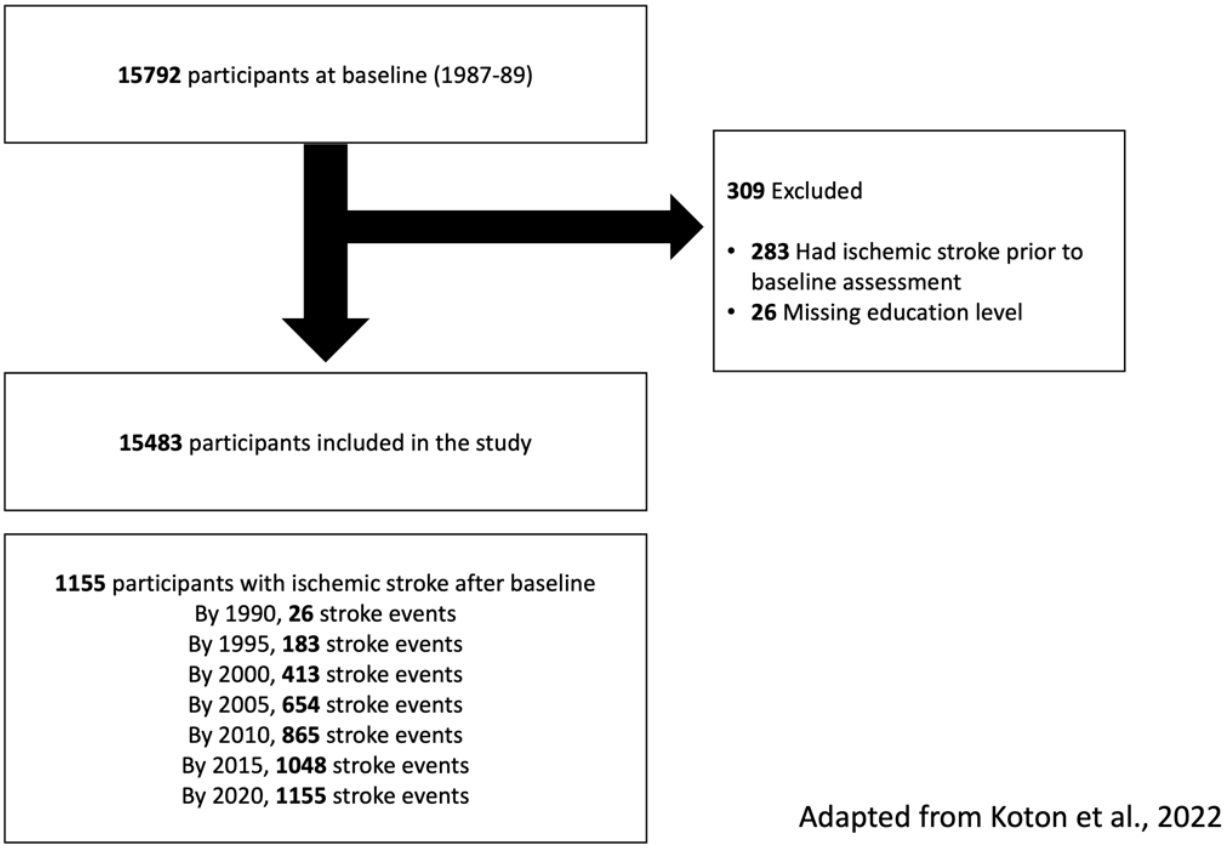
Design: longitudinal

Participant inclusion: Among the 15792 participants enrolled at ARIC baseline, 15483 participants will be included in this study [22]. Participants will be excluded if they had ischemic stroke prior to baseline (*N*= 283) or any missing key covariates. Figure 1 shows that there is a sufficient number of stroke events between 1990-2020 that can be used for the analysis.

Outcome: Three stroke outcome measures will be included in the analysis.

1. Stroke incidence- event: ischemic stroke, no ischemic stroke
2. Stroke severity incidence- event: no ischemic stroke, minor stroke (NIHSS $\leq$ 5), mild-moderate (NIHSS= 6-10), moderate-severe stroke (NIHSS $\geq$ 11) [23]

Figure 1. Flowchart showing the selection of ARIC participants included in the study



Demographic and risk factors

Demographic and risk factors at ARIC visit 1 and at visit 5 will be included in the study (**Table 1**). Demographic measures will include age, sex, race, and education level.

Table 1. Risk factor and chronic health conditions included in the cluster analysis.

Variables	Variable definition at visit 1	Variable definition at visit 5
Cancer	*hom10f*: Cancer ever diagnosed.	*cancer51:* Having cancer at V5
Coronary heart disease	*prvchd05*: Prevalent CHD Defined as: • Myocardial infarction from adjudicated Visit 1 ECG data • History of Myocardial Infarction • having undergone heart or arterial surgery, coronary bypass, balloon angioplasty, angioplasty of coronary artery (ies)	*prvchd53:* Prevent CHD by end of Visit 5 derived from the baseline status of CHD (PRVCHD05) and the closed event years of ARIC Cohort Surveillance data through 2013, where the event occurred prior to 30AUG2013.
Current drinking	*curdrk02*: Current drinker	*curdrk51:* Current drinker
Current smoking	cursmk01: Current smoker	cursmk52: Current smoker
Diabetes	*diabts03*: Diabetes – Lower Cut point 126 mg/dL	*diabts54*: Diabetes – Lower Cut point 126 mg/dL
Heart failure	*prevhf01*: Defined as: • having taken any medication for heart failure, or • qualifies for the Gothenburg Criteria [24]	*prevhf52*: Prevalent Heart Failure (HFRC Recommended Definition as of 2018)
Hypercholesteremia	*tchsiu01:*	*tchsiu51*:

	Total cholesterol in SI Units ≥ 240 mg/dl [25]	Total Cholesterol in SI Units ≥ 240 mg/dl
Hypertriglyceridemia	*TRGSIU01*: Total Triglycerides in SI units > 2.0 mmol/L [26]	*TRGSIU51:* Total Triglycerides in SI units > 2.0 mmol/L
Hypertension	*hypert05:* Hypertension based on definition 5 Defined as: BP ≥140 mm Hg and diastolic BP ≥90 mm Hg or the use of antihypertensive medications.	*hypert55*: Hypertension based on definition 5 Defined as: BP ≥140 mm Hg and diastolic BP ≥90 mm Hg or the use of antihypertensive medications.
Obesity	*bmi01:* Body mass index KG/(M*M) ≥ 30	*bmi51*: Body mass index KG/(M*M) ≥ 30
Atrial fibrillation	Present if [27-29]: • ECG showing AF • Hospital discharge with an AF or atrial flutter	*prvaf52:* Prevalent Atrial Fibrillation/Flutter by end of V5
Depression	*HPBA01-21:* Maastricht questionnaire of vital exhaustion [30] *	*CESD51:* CES-Depression Scale
Peripheral artery disease	PAD02: Peripheral artery disease, def 2 ABI04 (Ankle Brachial Index, def 4) < 0.90	Peripheral artery disease, def 2 ABI04 (Ankle Brachial Index, def 4) < 0.90[31]
Renal dysfunction	Estimated GFR	Estimated GFR

** Maastricht questionnaire at visit 2*

Data analysis

The statistical analysis will be conducted using the R software programing language.

1. *Cluster analysis and cluster characterization*

An agglomerative hierarchical clustering analysis will be employed separately at ARIC visit 1 (midlife) and at visit 5 (late life) to examine the co-occurrences of chronic diseases and risk factors [32]. Gower's measure, i.e., Dice coefficient for factor variables, will be employed to measure the similarity of the patients' records. Clusters will be merged based on the minimum between-cluster distance employing the Ward's squared minimum variance clustering method. This clustering method tends to generate more compact clusters. To assess how close participant records in one cluster are to records in neighboring clusters, the average Silhouette width will be computed which ranges between -1 to +1. The optimal number of hierarchical clusters will be determined based on ocular inspection of the dendrogram which depicts the clustering tree and by determining the highest average silhouette width.

To characterize the clusters based on risk factor and chronic health condition profiles, the observed/expected ratios will be computed by dividing the prevalence of a given condition within a cluster by its prevalence in the overall population. A condition is associated with a cluster when the observed/expected ratio is ≥ 1.5, and when the intra-cluster prevalence of a condition is ≥ 0.25 [9]. Among the different generated clusters, one cluster will be chosen as a reference group which is characterized by the lowest chronic disease burden.

2. *Association between the clusters in mid and late life and stroke incidence in the population*

In the primary analysis, the association between the clusters in mid and late life and stroke incidence will be tested by two separate Cox proportional hazard regression models. These models will be adjusted for age, sex, race, and educational level. The assumptions underlying the regression models such as the proportional hazards assumption and the linearity between log hazard and the covariates will be tested. For the midlife analysis, all incident stroke case events

between 1990-2020 will be included. For the analysis in late life, only incident stroke at visit 5 and onwards will be included. By adding an interaction term between cluster and sex, race or education levels, it will also be assessed whether the association between the clusters in mid or in late life and stroke incidence differs by the demographic levels.

Given that mortality is a strong competing risk factor in the analysis, a sensitivity analysis will be conducted which tests the associations between the clusters and stroke incidence with a competing risk model (first ischemic stroke event vs. no stroke event) using the function "crr" of the "cmprsk" R package and of the "tidycmprsk" package extension. Competing risk analysis is a type of survival analysis which estimates the marginal probability of an outcome event in the presence of one or several competing events based on the cumulative incidence function (CIF) [33]. In this analysis, the competing event is lifetime mortality. The competing risk models will be adjusted by age, sex, race, and educational level. Similarly, as in the Cox regression analysis, an interaction term between the clusters and the demographic measures will be added to assess whether the association between the clusters in mid or in late life and stroke incidence differs by the demographic levels.

3. *Association between the clusters in mid and late life and stroke severity*

In the primary analysis, the association between the clusters in mid (*hypothesis 2)* and late life and stroke severity levels vs. no stroke will be tested with multiple separate Cox proportional hazard regression models. These models will be adjusted by age, sex, race, and educational level. The assumptions underlying the regression models such as the proportional hazards assumption and the linearity between log hazard and the covariates will be tested. By adding an interaction term between the clusters and sex, race, or education, it will also be assessed whether the association between the clusters in mid or in late life and stroke severity differs by the demographic factor levels. In case the number of severe stroke events is too low, a cox regression model with Firth bias correction will be employed which is a statistical approach to mitigate the bias resulting from only rare events in the cohort data.

The association between the clusters in mid and late life on stroke severity vs. no stroke will also be tested by separate competing risk models using the "cmprsk", "survival" and "tidycmprsk" R packages. All 3 stroke severity outcomes (NIHSS ≤ 5; NIHSS= 6-10; NIHSS ≥ 11) will in turn be coded as failure types of interest in 3 separate Fine-Gray regression analyses. Mortality will be treated as a competing risk event. The association will be accounted for by age, sex, race, and education level. By adding an interaction term, it will also be tested whether the association between the clusters in mid or in late life and stroke severity differs by sex, race, or educational levels.

In a separate analysis focusing on the association of clusters with stroke severity, only participants experiencing a stroke event will be included. The proportion of cases belonging to each cluster group stratified by stroke severity will be computed and graphically displayed using the "ggplot2" package. Using a permutational ANOVA model, differences in age by cluster group and stroke severity will be tested (R package "lmPerm"). Employing a multinomial logistic regression model, the association between the clusters in mid-life and late life and stroke severity levels will be assessed while adjusting for the demographic covariates. Minor stroke (NIHSS ≤ 5) will be coded as the baseline category. An interaction term, i.e., cluster x sex; cluster x race, will be included in the regression model to test whether the association between clusters and stroke severity varies by demographic covariates education and race. Employing a binomial logistic regression analysis, it will also be determined whether different clusters in midlife are more strongly associated with midlife (< age 65) than late life stroke ($\geq$ age 65).

4. *Assessing the association of clusters in mid vs late life on stroke incidence and stroke severity after ARIC visit 5*

Only stroke incidences at visit 5 and onwards will be included. The data set will be reshaped in long format and a new variable will be created that gives combined information on cluster type and ARIC visit (e.g., Cluster1_visit1, Cluster2_visit1… etc.). The information of cluster type and visit will be combined assuming that an individual will not always belong to the same cluster type at both visits. In addition, the overall number of clusters at visit 1 and visit 5 may not be the same.

Furthermore, different risk factor characteristics may be attributed at visit 1 vs. visit 5. To test the association between clusters in mid and late life on stroke incidence in late life, a mixed effect Cox regression model will be employed (R packages “coxme”; “survival”). To account for the correlated observational data structure by participant, a random intercept for participant ID will be included. In addition to cluster type by visit, the covariates age, sex, race, and education will also be added as fixed effects. The underlying assumptions of the mixed effect model including the normality of the random effects will be tested. Using the deviance statistic of the “anova()” R function (R package “car”), the fit of the mixed effect model will be compared to a baseline model which includes all demographic covariates but not the cluster by visit factor.

If enough cases for each stroke severity category exist at visit 5 and onwards, the mixed effect Cox regression analysis will be repeated for all events: minor (NIHSS $\leq$ 5) vs. no stroke, mild-moderate vs. no stroke (NIHSS= 6-10), and moderate-severe (NIHSS $\geq$ 11) vs. no stroke.

Anticipated methodologic limitations or challenges?

- Information regarding stroke severity was not prospectively assessed but is based on data in the hospitalization records from admission. However, the information was collected by trained physicians, experienced in evaluation of stroke patients. Severity was categorized based on all the data available in the hospitalization charts and following a validated algorithm for chart-based retrospective evaluation of stroke severity [34, 35]
- The way we define disease burden in this proposal is not conventional. Instead of quantifying the overall health burden (2 or more diseases) in a patient, we are interested in a data-driven approach which identifies distinct groups of individuals that are characterized by specific disease clusters. These clusters will hopefully provide valuable information on stroke risk and stroke severity.
- Hierarchical clustering is an unsupervised machine learning approach. The results may therefore be unpredictable or difficult to understand. However, this cluster approach has been carried out multiple times before, when measuring the combined health burden in a population.

- The risk factor hypertension will likely be an important predictor for stroke incidence in the analysis. It is likely that hypertension will play a role in each of the identified clusters.
- We recognize that the measures referring to the chronic disease and risk factor conditions in this study are not always the same at visit 1 and visit 5.

7.a. Will the data be used for non-ARIC analysis or by a for-profit organization in this manuscript? ____ Yes X No

b. If Yes, is the author aware that the current derived consent file ICTDER05 must be used to exclude persons with a value RES_OTH and/or RES_DNA = "ARIC only" and/or "Not for Profit" ? ____ Yes ____ No
(The file ICTDER has been distributed to ARIC PIs, and contains the responses to consent updates related to stored sample use for research.)

8.a. Will the DNA data be used in this manuscript? ___ _ Yes __x_ No

8.b. If yes, is the author aware that either DNA data distributed by the Coordinating Center must be used, or the current derived consent file ICTDER05 must be used to exclude those with value RES_DNA = "No use/storage DNA"? ___ Yes ____ No

9. The lead author of this manuscript proposal has reviewed the list of existing ARIC Study manuscript proposals and has found no overlap between this proposal and previously approved manuscript proposals either published or still in active status.
ARIC Investigators have access to the publications lists under the Study Members Area of the web site at: http://www.cscc.unc.edu/aricproposals/dtSearch.html

__x___ Yes _____ No

10. What are the most related manuscript proposals in ARIC (authors are encouraged to contact lead authors of these proposals for comments on the new proposal or collaboration)?

1. MP #3490. Koton et al., 2019. Thirty-year trends in stroke severity on admission in the Atherosclerosis Risk in Communities (ARIC) 1987-2017
2. MP #3672. Koton et al. 2020. Stroke incidence and severity at risk factors for dementia and MCI in in the Atherosclerosis Risk in Communities (ARIC) cohort study
3. MP #3801. Johansen et al. 2021. The association between ischemic stroke subtype and stroke severity: The ARIC study
4. MP #4180. Mediano et al. Physical activity at midlife and severity of ischemic stroke in the ARIC participants

5. MP #4022 Kucharska-Newton et al. Multimorbidity- Natural history, risk, and resilience factors; disparities by sex, race, and socioeconomic context

11.a. Is this manuscript proposal associated with any ARIC ancillary studies or use any ancillary study data? _ Yes _X__ No

11.b. If yes, is the proposal

A. primarily the result of an ancillary study (list number* 2017.22**)**

___ B. primarily based on ARIC data with ancillary data playing a minor role (usually control variables; list number(s)* _____ _____ _____)

*ancillary studies are listed by number https://sites.cscc.unc.edu/aric/approved-ancillary-studies

12a. Manuscript preparation is expected to be completed in one to three years. If a manuscript is not submitted for ARIC review at the end of the 3-years from the date of the approval, the manuscript proposal will expire.

12b. The NIH instituted a Public Access Policy in April, 2008 which ensures that the public has access to the published results of NIH funded research. It is **your responsibility to upload manuscripts to PubMed Central** whenever the journal does not and be in compliance with this policy. Four files about the public access policy from http://publicaccess.nih.gov/ are posted in http://www.cscc.unc.edu/aric/index.php, under Publications, Policies & Forms. http://publicaccess.nih.gov/submit_process_journals.htm shows you which journals automatically upload articles to PubMed central.

Bibliography

1. Johnston MC, Crilly M, Black C, Prescott GJ, Mercer SW. Defining and measuring multimorbidity: a systematic review of systematic reviews. Eur J Public Health 2019;29:182-189.
2. Barnett K, Mercer SW, Norbury M, Watt G, Wyke S, Guthrie B. Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study. The Lancet 2012;380:37-43.
3. Ho IS, Azcoaga-Lorenzo A, Akbari A, et al. Examining variation in the measurement of multimorbidity in research: a systematic review of 566 studies. Lancet Public Health 2021;6:e587-e597.
4. Marengoni A, Roso-Llorach A, Vetrano DL, et al. Patterns of Multimorbidity in a Population-Based Cohort of Older People: Sociodemographic, Lifestyle, Clinical, and Functional Differences. J Gerontol A Biol Sci Med Sci 2020;75:798-805.
5. Prados-Torres A, Calderon-Larranaga A, Hancco-Saavedra J, Poblador-Plou B, van den Akker M. Multimorbidity patterns: a systematic review. J Clin Epidemiol 2014;67:254-266.
6. Salive ME. Multimorbidity in older adults. Epidemiol Rev 2013;35:75-83.
7. Zheng DD, Loewenstein DA, Christ SL, et al. Multimorbidity patterns and their relationship to mortality in the US older adult population. PLoS One 2021;16:e0245053.
8. Rivera-Almaraz A, Manrique-Espinoza B, Avila-Funes JA, et al. Disability, quality of life and all-cause mortality in older Mexican adults: association with multimorbidity and frailty. BMC Geriatr 2018;18:236.
9. Vetrano DL, Roso-Llorach A, Fernandez S, et al. Twelve-year clinical trajectories of multimorbidity in a population of older adults. Nat Commun 2020;11:3223.
10. Grande G, Marengoni A, Vetrano DL, et al. Multimorbidity burden and dementia risk in older adults: The role of inflammation and genetics. Alzheimers Dement 2021;17:768-776.
11. Rocca WA, Boyd CM, Grossardt BR, et al. Prevalence of multimorbidity in a geographically defined American population: patterns by age, sex, and race/ethnicity. Mayo Clin Proc 2014;89:1336-1349.
12. Skou ST, Mair FS, Fortin M, et al. Multimorbidity. Nat Rev Dis Primers 2022;8:48.
13. Low LL, Kwan YH, Ko MSM, et al. Epidemiologic Characteristics of Multimorbidity and Sociodemographic Factors Associated With Multimorbidity in a Rapidly Aging Asian Country. JAMA Netw Open 2019;2:e1915245.
14. Cezard G, McHale CT, Sullivan F, Bowles JKF, Keenan K. Studying trajectories of multimorbidity: a systematic scoping review of longitudinal approaches and evidence. BMJ Open 2021;11:e048485.
15. Ben Hassen C, Fayosse A, Landre B, et al. Association between age at onset of multimorbidity and incidence of dementia: 30 year follow-up in Whitehall II prospective cohort study. BMJ 2022;376:e068005.
16. Nicholson K, Griffith LE, Sohel N, Raina P. Examining early and late onset of multimorbidity in the Canadian Longitudinal Study on Aging. J Am Geriatr Soc 2021;69:1579-1591.
17. Gallacher KI, Jani BD, Hanlon P, Nicholl BI, Mair FS. Multimorbidity in Stroke. Stroke 2019;50:1919-1926.
18. Schmidt M, Jacobsen JB, Johnsen SP, Bøtker HE, Sørensen HT. Eighteen-year trends in stroke mortality and the prognostic influence of comorbidity. 2014;82:340-350.
19. Gruneir A, Griffith L, Fisher K, et al. Increasing comorbidity and health services utilization in older adults with prior stroke. Neurology 2016;87:2091-2098.
20. Gallacher K, Batty G, McLean G, et al. Stroke, multimorbidity and polypharmacy in a nationally representative sample of 1,424,378 patients in Scotland: implications for treatment burden. BMC Med 2014;12.
21. Johansen HL, Wielgosz A, Nguyen K, Fry RN. Incidence, comorbidity, case fatality and readmission of hospitalized stroke patients in Canada. Can J Cardiol 2006;22:65-71.
22. Koton S, Pike JR, Johansen M, et al. Association of Ischemic Stroke Incidence, Severity, and Recurrence With Dementia in the Atherosclerosis Risk in Communities Cohort Study. JAMA Neurol 2022;79:271-280.
23. Koton S, Patole S, Carlson JM, et al. Methods for stroke severity assessment by chart review in the Atherosclerosis Risk in Communities study. Sci Rep 2022;12:12338.
24. Rosamond WD, Chang PP, Baggett C, et al. Classification of heart failure in the atherosclerosis risk in communities (ARIC) study: a comparison of diagnostic criteria. Circ Heart Fail 2012;5:152-159.
25. Szklo M, Chambless LE, Folsom AR, et al. Trends in plasma cholesterol levels in the atherosclerosis risk in communities (ARIC) study. Prev Med 2000;30:252-259.
26. Dron JS, Hegele RA. Genetics of Hypertriglyceridemia. Front Endocrinol (Lausanne) 2020;11:455.
27. Ghelani KP, Chen LY, Norby FL, Soliman EZ, Koton S, Alonso A. Thirty-Year Trends in the Incidence of Atrial Fibrillation: The ARIC Study. J Am Heart Assoc 2022;11:e023583.

28. Alonso A, Agarwal SK, Soliman EZ, et al. Incidence of atrial fibrillation in whites and African-Americans: the Atherosclerosis Risk in Communities (ARIC) study. Am Heart J 2009;158:111-117.
29. Soliman EZ, Prineas RJ, Case LD, Zhang ZM, Goff DC, Jr. Ethnic distribution of ECG predictors of atrial fibrillation and its impact on understanding the ethnic distribution of ischemic stroke in the Atherosclerosis Risk in Communities (ARIC) study. Stroke 2009;40:1204-1211.
30. Bogle BM, Sotoodehnia N, Kucharska-Newton AM, Rosamond WD. Vital exhaustion and sudden cardiac death in the Atherosclerosis Risk in Communities Study. Heart 2018;104:423-429.
31. Wu A, Coresh J, Selvin E, et al. Lower Extremity Peripheral Artery Disease and Quality of Life Among Older Individuals in the Community. J Am Heart Assoc 2017;6.
32. Maechler M, Rousseuw P, Struyf A, Hubert M, Hornik K. Cluster: Cluster Analysis Basics and Extensions. R packages 2019.
33. Scrucca L, Santucci A, Aversa F. Regression modeling of competing risk using R: an in depth guide for clinicians. Bone Marrow Transplant 2010;45:1388-1395.
34. Lindsell CJ, Alwell K, Moomaw CJ, et al. Validity of a retrospective National Institutes of Health Stroke Scale scoring methodology in patients with severe stroke. J Stroke Cerebrovasc Dis 2005;14:281-283.
35. Williams LS, Yilmaz EY, Lopez-Yunez AM. Retrospective assessment of initial stroke severity with the NIH Stroke Scale. Stroke 2000;31:858-862.