Health Heterogeneity in Older Adults: Exploration in the Canadian Longitudinal Study on Aging

Quoc Dinh Nguyen, MD, MA, MPH,[*†‡] *Erica M. Moodie, PhD,*[‡] *Marie-France Forget, MD, MSc,*[*] *Philippe Desmarais, MD, MHSc,*[*†] *Mark R. Keezer, MDCM, PhD,*[†§] *and Christina Wolfson, PhD*[‡¶‖]

See related editorial by Ferrucci et al. in this issue.

BACKGROUND: A widely held dictum in aging research is that heterogeneity in health increases with age, but the basis for this claim has not been fully investigated. We examined heterogeneity at different ages across health characteristics to describe variation and trends; we investigated the comparative importance of between-age versus within-age heterogeneity.

DESIGN: This was a cohort study.

SETTING: Community-dwelling older adults.

PARTICIPANTS: A total of 30,097 adults aged 45 to 86 years, from the Canadian Longitudinal Study on Aging, were included.

MEASUREMENTS: Thirty-four health characteristics in eight domains (physical measures, vital signs, physiological measures, physical performance, function/disability, chronic conditions, frailty, laboratory values) were assessed cross-sectionally. We used regression models to examine heterogeneity in health characteristics (using absolute deviation) and domains (using effective variance) in relation to age. Comparison between between-age and within-age heterogeneity was quantified by estimating the age threshold at which the former exceeds the latter.

RESULTS: Of the 34 health characteristics, 17 showed increased heterogeneity, 8 decreased, and 9 no association with age. The associations between heterogeneity and age increased generally but were nonlinear for most domains and nonmonotonic for some. We observed peak heterogeneity at approximately 70 years. Between-age heterogeneity, compared with within-age heterogeneity, was most important for forced expiratory volume in 1 second and grip strength but varied across characteristics.

CONCLUSION: Overall health heterogeneity increases with age but does not uniformly increase across all variables and domains. Heterogeneity in aging reinforces the need for geriatric assessment and personalized care, depending on which health characteristics are assessed, their measurement properties, and their referent group. Our findings suggest further research to develop improved single-dimension and multidimensional instruments, as well as specific vital and laboratory reference ranges for older adults. **J Am Geriatr Soc 69:678-687, 2021.**

Keywords: CLSA; measurement; variability; heterogeneity

INTRODUCTION

Aging is not uniform. Older adults have variable health states, embodied in different levels of functioning,[1] chronic conditions, and mortality rates.[2] A widely held dictum in aging research is that older adults show greater heterogeneity in health metrics (or "health heterogeneity") than their younger peers.[3,4] Heterogeneity can be defined as the quality or state of being diverse in character or content.[5] Although older adults differ from younger adults, heterogeneity in aging suggests that older adults would increasingly differ among themselves as they age, quantified as progressively greater deviation (spread) from the average value of their age group.

The ongoing search for a better understanding of heterogeneity[6] has provided the impetus for the development of age-related constructs such as multimorbidity, disability, and frailty.[1,7-11] Variability is a fundamental tenet in clinical practice, where it underpins reference ranges for vital signs and laboratory tests. However, the central premise that heterogeneity increases with chronological age seems

From the *Division of Geriatrics, Centre hospitalier de l'Université de Montréal, Montreal, Canada; †Innovation Hub, Centre de recherche du Centre hospitalier de l'Université de Montréal, Montreal, Canada; ‡Department of Epidemiology, Biostatistics, and Occupational Health, McGill University, Montreal, Canada; §Department of Neurosciences & Social and Preventative Medicine, Université de Montréal, Montreal, Canada; ¶Department of Medicine, McGill University, Montreal, Canada; and the ‖Neuroepidemiology Research Unit, Research Institute of the McGill University Health Centre, Montreal, Canada.

Address correspondence to Quoc Dinh Nguyen, MD, MA, MPH, 1000, Saint-Denis, Montreal, Quebec H2X0C1, Canada. E-mail: quoc.nguyen@mail.mcgill.ca. Twitter: @quocdngu

DOI: 10.1111/jgs.16919

0002-8614/21/$15.00

to have been assumed rather than substantiated.[3] Heterogeneity has often been alluded to and reported inconsistently in the literature in the last 50 years.[6,12-17] However, it has not been the focus of deliberate empirical research, particularly in contemporary populations.[4,18,19] In the absence of research, increasing heterogeneity as adults age may not pertain to all health characteristics, may not have a consistent trajectory, and may require consideration of the measurement scales used.[3,15]

In this study, using data from the Canadian Longitudinal Study on Aging (CLSA), our first objective was to explore whether, and how, heterogeneity increases in individuals from middle to advanced age. We examined heterogeneity in health characteristics and domains to describe variation and trajectories and explanations for changes in heterogeneity with aging. Our second objective was to compare the heterogeneity due to differences between age groups with the heterogeneity due to differences within age groups. The overarching objective of this study was to explore how features of heterogeneity in older adults, as a central tenet of aging, might inform geriatric practice and research.

METHODS

Cohort

We used cross-sectional data from the CLSA, which enrolled a nationally representative sample of over 51,000 participants aged 45 to 85 years, at baseline, into an in-person assessed cohort (Comprehensive) or a telephone-only assessed cohort.[20,21] Because data for physical assessments were required, we used the Comprehensive cohort, comprising 30,097 adults recruited and assessed from 2012 to 2015. Exclusion criteria for the CLSA were people living in the Canadian territories, on a First Nations reserve, or in institutions; being full-time members of the Armed Forces; having cognitive impairment (as assessed by interviewers); and being unable to respond in French or English.

Variables, Health Characteristics, and Domains

We examined the heterogeneity of 34 health characteristics in eight domains (physical measures, vital signs, physiological measures, physical performance, function and disability measures, chronic conditions, frailty, and laboratory values), summarized in Supplementary Methods S1 (and reported in Table 1). These characteristics were chosen based on their usage as heterogeneity-describing variables used in research or clinical practice. Physical measures, vital signs, physical performance, physiological measures, and laboratory values were assessed in person. Function, disability, and chronic conditions were self-reported; for details, see the CLSA protocol.[21] Based on a standard procedure previously described,[22] we derived a frailty index using 34 variables as reported in the Supplementary Methods S1.

Analysis

Heterogeneity by health characteristics and domains

We used boxplots by 10-year age bins and gender to examine the heterogeneity (spread) of each variable. The quantitative association between heterogeneity and age was evaluated in three sequential steps for each health characteristic. First, we derived a variable representing an individual's absolute deviation from his or her predicted age–gender–group mean value; this absolute deviation was regressed on age (1-year bins) and gender (Model 1). We assessed the statistical and clinical significance of the age coefficient to determine whether heterogeneity increases with age. Clinical significance was defined as a change in deviation over 40 years (age coefficient * 40) that was greater than the last clinically relevant unit of a measurement (e.g., 1 kg for weight, 0.01 m for height). Second, as some health characteristics may have a mean–variance relationship (i.e., increasing variance as the mean increases), we examined the relationship between the *mean* of a variable and age in linear regression models adjusted for gender (Model 2). If the variation in mean was in the same direction as the variation in deviation, we further adjusted Model 1 for this mean–variance relationship by adding the age- and gender-specific mean as a covariate (Model 3). If the age coefficient changed direction or was no longer clinically or statistically significant, we considered the change in heterogeneity to be explainable by a mean–variance relationship. Third, as the scaling of a measure may influence heterogeneity, we determined whether the health characteristic had a normative or clinical scaling through consensus (QDN, MFF, PD). Normative or clinical scaling was present when a measure had a floor or ceiling effect, or when the range of measurement and scaling of a characteristic was determined by clinical purpose. See Supplementary Figure S1 for a flowchart and descriptions of analytic steps and clinical significance.

To describe heterogeneity by health domains, we used effective variance, a summary measure of heterogeneity for multiple variables, which quantifies their average multivariate scatter of variables.[23] After stratifying by gender, we normalized all variables to ensure equal weighting when computing effective variance for each 1-year bin. Linear regression was used with effective variance regressed on age and gender. This was performed for each domain, all domains, and all domains excluding laboratory values. To detect nonlinear relationships with age, we tested for quadratic and cubic terms. We retained statistically significant terms to plot predicted effective variance by age and gender.

Heterogeneity between age groups and within groups

Finally, to determine whether heterogeneity was attributable to variation in age, we compared heterogeneity *between* age groups to that *within* the older age group (i.e., between individuals of that age group), using age 85 years as the reference group, with the older age group boundary being more relevant to gerontology and geriatric medicine. For each health characteristic, we predicted: (1) the individual differences as the average absolute deviation of individuals aged 85 years from their predicted gender–group mean value and (2) computed the linear coefficient of age predicting that characteristic. We then used this model to determine the age at which heterogeneity due to variation in chronological age exceeds that due to individual differences. The *closer* the age threshold is to 85 years, the more influential the between-age group variation is, compared with within-group variation. To avoid

Table 1. Characteristics of Community-Dwelling Older Adults in the Canadian Longitudinal Study of Aging Comprehensive Cohort by Age Group

	Overall	45–54 Years	55–64 Years	65–74 Years	75–86 Years
Sample size, n	30,097	7,595	9,856	7,362	5,284
Age, years	63.0 (10.2)	50.3 (2.7)	59.7 (2.8)	68.9 (2.8)	78.9 (2.9)
Gender, male, n	14,777 (49.1)	3,670 (48.3)	4,767 (48.4)	3,674 (49.9)	2,666 (50.5)
Race, White, n	28,771 (95.6)	7,098 (93.5)	9,463 (96.0)	7,097 (96.4)	5,113 (96.8)
Physical measures					
Weight, kg	79.7 (17.6)	80.9 (18.8)	81.3 (18.3)	79.6 (16.7)	75.4 (14.8)
Height, m	1.68 (0.10)	1.70 (0.09)	1.69 (0.10)	1.67 (0.10)	1.66 (0.10)
BMI, kg/m^2	28.1 (5.4)	27.8 (5.7)	28.4 (5.7)	28.3 (5.3)	27.3 (4.6)
Waist circumference, cm	94 (15)	92 (15)	95 (15)	96 (14)	95 (13)
Vital signs					
Pulse, beats per min	72 (12)	73 (11)	72 (12)	71 (12)	70 (12)
Systolic blood pressure, mmHg	122 (17)	116 (15)	121 (16)	126 (17)	128.6 (18)
Diastolic blood pressure, mmHg	74 (10)	76 (10)	76 (10)	73.6 (10)	71 (10)
Physiological measures					
FEV1, L	2.7 (0.8)	3.1 (0.7)	2.8 (0.7)	2.5 (0.7)	2.1 (0.6)
BMD, g/cm^2	1.01 (0.14)	1.03 (0.12)	1.00 (0.13)	1.00 (0.15)	0.99 (0.16)
Visual acuity[a]	0.9 (0.3)	1.0 (0.3)	0.9 (0.3)	0.8 (0.3)	0.7 (0.2)
Physical performance measures					
Gait speed, m/s	0.98 (0.20)	1.04 (0.18)	1.01 (0.20)	0.95 (0.19)	0.86 (0.19)
Grip strength, kg	35.2 (11.8)	39.4 (12.3)	36.2 (11.6)	33.4 (10.8)	29.0 (9.8)
Chair rise, s for one[b]	2.7 (0.8)	2.5 (0.7)	2.6 (0.7)	2.8 (0.8)	3.0 (0.9)
TUG, s	9.6 (2.6)	8.8 (1.7)	9.2 (2.4)	9.8 (2.3)	11.2 (3.4)
Function and disability					
ADL impairment, n[c]	837 (2.8)	95 (1.3)	210 (2.1)	198 (2.7)	334 (6.3)
IADL impairment, n[c]	1,566 (5.2)	170 (2.2)	345 (3.5)	371 (5.1)	680 (12.9)
OARS score[d]	27.8 (0.7)	27.9 (0.6)	27.9 (0.6)	27.8 (0.7)	27.6 (0.9)
Life Space Assessment score	85 (18)	91 (17)	87 (18)	83 (18)	77 (19)
PASE score	141 (74)	174 (83)	149 (73)	125 (60)	100 (53)
Chronic conditions, n					
Hypertension	11,101 (37.0)	1,495 (19.7)	3,364 (34.2)	3,369 (45.9)	2,873 (54.6)
Diabetes	2,957 (9.9)	370 (4.9)	963 (9.9)	923 (12.7)	701 (13.6)
Heart disease	4,232 (14.1)	317 (4.2)	1,003 (10.2)	1,383 (18.8)	1,529 (29.1)
Stroke or TIA	1,347 (4.5)	101 (1.3)	282 (2.9)	395 (5.4)	569 (10.8)
Arthritis	7,922 (26.4)	926 (12.2)	2,472 (25.1)	2,514 (34.2)	2010 (38.2)
Osteoporosis	2,689 (9.0)	154 (2.0)	738 (7.5)	931 (12.7)	866 (16.5)
Lung disease	5,094 (17.0)	1,265 (16.7)	1,687 (17.1)	1,268 (17.3)	874 (16.6)
Kidney disease	867 (2.9)	99 (1.3)	244 (2.5)	252 (3.4)	272 (5.2)
Cancer	4,637 (15.4)	427 (5.6)	1,270 (12.9)	1,454 (19.8)	1,486 (28.2)
Anxiety or depression	6,243 (20.8)	1736 (22.9)	2,356 (23.9)	1,464 (19.9)	687 (13.1)
Chronic condition count, mean	1.6 (1.4)	0.9 (1.0)	1.5 (1.3)	1.9 (1.4)	2.3 (1.5)
Frailty index	0.09 (0.08)	0.05 (0.06)	0.08 (0.07)	0.10 (0.08)	0.14 (0.09)
Laboratory values					
Hemoglobin, g/L	141 (13)	141 (13)	142 (13)	141 (13)	138 (14)
WBC count, 10^9/L	6.7 (2.2)	6.6 (1.9)	6.6 (1.8)	6.8 (2.3)	7.0 (3.1)
Platelet count, 10^9/L	222 (58)	227 (56)	225 (58)	220 (59)	211 (58)
Creatinine, μmol/L	82 (24)	79 (20)	80 (25)	83 (23)	88 (27)
GFR, mL/s/m^2	79 (15)	88 (13)	82 (13)	75 (14)	66 (14)
ALT, U/L	24 (14)	25 (14)	25 (15)	23 (13)	20 (12)
Albumin, g/L	40 (3)	40 (3)	40 (3)	40 (3)	39 (3)
Total cholesterol, mmol/L	5.1 (1.1)	5.2 (1.0)	5.3 (1.1)	5.0 (1.2)	4.8 (1.2)
HDL, mmol/L	1.5 (0.5)	1.5 (0.5)	1.5 (0.5)	1.5 (0.5)	1.5 (0.5)
LDL, mmol/L	2.8 (1.0)	3.0 (0.9)	3.0 (1.0)	2.8 (1.0)	2.6 (1.0)
TSH, mU/L	2.3 (2.2)	2.2 (1.8)	2.2 (2.6)	2.3 (1.7)	2.5 (2.4)
HbA1C, %	5.7 (0.8)	5.5 (0.7)	5.7 (0.8)	5.7 (0.7)	5.8 (0.7)
Ferritin, μg/L	158 (143)	142 (141)	162 (143)	168 (150)	161 (136)
Vitamin D, nmol/L	90 (38)	79 (34)	88 (37)	96 (39)	99 (38)
C-reactive protein, mg/L	2.6 (5.1)	2.2 (4.1)	2.5 (4.9)	2.7 (5.8)	2.9 (5.5)

Note: (%) or (SD) as appropriate.

Abbreviations: ALT, alanine aminotransferase; BMI, body mass index; FEV1, forced expiratory volume in 1 second; BMD, bone mineral density; ADL, activity of daily living; GFR, glomerular filtration rate; HDL, high-density lipoprotein; IADL, instrumental activity of daily living; LDL, low-density lipoprotein; OARS, Older Americans Resources and Services subscales for 7 activities of daily living (ADL) and 7 instrumental ADLs; PASE, physical activity scale for the elderly; TUG, timed up and go; TSH, thyroid-stimulating hormone; WBC, white blood cell.

[a]Fraction (e.g., 20/20 = 1).

[b]Average time required for one chair rise.

[c]Impairment on one or more ADL or iADL.

[d]Older Americans Resources and Services score of 14 ADL and iADLS (0–2).

Table 2. Variation in Heterogeneity and Mean by Health Characteristic in Relation to Chronological Age

Health characteristic	Variation in Heterogeneity by Age[a] (Deviation Slope, from Model 1)	Variation in Mean by Age[a] (Mean Slope, from Model 2)	Variation Explainable by Mean–Variation Relationship[b] (from Model 3)	Measured Characteristic has Normative or CLINICAL scaling
Physical measures				
Weight	–	–	No	No
Height	↔	–		No
BMI	–	–	No	No
Waist circumference	–	+	No	No
Vital signs				
Pulse	↔	–		No
Systolic blood pressure	+	+	No	No
Diastolic blood pressure	↔	–		No
Physiological measures				
FEV1	–	–	No	No
BMD	+	–	No	No
Visual acuity	–	–	Yes	No
Physical performance measures				
Gait speed	↔	–		No
Grip strength	–	–	Yes	No
Chair rise	+	+	Yes	Yes
TUG	+	+	Yes	Yes
Function and disability				
OARS	+	–	No	Yes
Life Space Assessment	+	–	No	Yes
PASE	–	–	Yes	Yes
Chronic condition count	+	+	Yes	Yes
Frailty index	+	+	Yes	Yes
Laboratory values				
Hemoglobin	+	–	No	No
WBC count	↔	↔		No
Platelet count	↔	–		No
Creatinine	+	+	No	No
GFR	+	–	No	No
ALT	–	–	No	No
Albumin	↔	–		No
Total cholesterol	+	–	No	No
HDL	↔	↔		No
LDL	+	–	No	No
TSH	+	+	Yes	No
HbA1C	+	+	Yes	No
Ferritin	+	+	Yes	No
Vitamin D	+	+	Yes	No
C-reactive protein	↔	+		No

Note: "+" indicates an increase, –" decrease, "↔" not clinically significant.

[a]Variation determined by linear regression of deviation (Model 1) or mean (Model 2) on age, adjusted for gender (gender-specific intercepts).

[b]Determined by examining the mean–deviation relationship and adjusting linear regressions for age and gender group mean; see Methods and Supplementary Figure S1 for details.

extrapolating beyond available data, we only report results where this age was between 45 and 85 years. Confidence intervals were computed via bootstrapping. Supplementary Figure S2 provides an explanation of the calculation.

Analyses were performed using R 3.6.1 (R Foundation). As our objectives were exploratory, we used available case analyses for variables with missing data, reported in Supplementary Table S1, without sampling weights.

RESULTS

Cohort Description and Mean Variation in Health Characteristics

Our study sample comprised 30,097 participants, of which 15,320 were women (50.1%). The mean age was 63.0 years (SD: 10.2; range 45–86 years (all enrolled at age 85 years or

younger)), and age groups were well represented. Table 1 describes all health characteristics examined. Impairment in activities of daily living (ADL, 2.8%) and instrumental ADL (iADL, 5.2%) was infrequent. Overall, the mean frailty index was 0.09. Hypertension (37.0%) and arthritis (26.4%) were the most prevalent reported chronic conditions. Most health characteristics across all domains showed mean changes with age, especially in participants aged 75 to 86 years. Older participants had lower mean values of physiological, physical performance, and functional measures. Chronic conditions accumulated with age (mean count of 0.9 in participants aged 45–54 years to 2.3 in those aged 75–86 years). The mean frailty index increased from 0.05 (0.06) in those aged 45 to 54 years to 0.14 (0.09) in those aged 75–86 years.

Heterogeneity by Health Characteristics and Domains

Boxplots for all variables by 10-year age group are reported in Supplementary Figure S3, and standard deviations of health characteristics by age group are reported in Table 1. When assessed qualitatively by quantile differences and SDs, the spread of many variables varied with age. Table 2 reports the linear relationship between heterogeneity and chronological age for each health characteristic. Of the 34 variables examined, 17 showed clinically significant increased heterogeneity, 8 showed decreased heterogeneity, and 9 showed no evidence of an association (Table 2, first column). By domains, physical measures (weight, body mass index, and waist circumference) showed decreasing heterogeneity with age. Conversely, number of chronic conditions and frailty index showed increasing heterogeneity. Within physiological and physical performance and functional measures, heterogeneity showed diverging associations with age. For example, heterogeneity in grip strength, physical activity scale for the elderly (PASE), and forced expiratory volume in 1 second (FEV1) decreased but increased in bone mineral density, timed up and go (TUG), and life space assessment (LSA). Of the 25 variables with clinically significant associations, 11 (8 with increasing heterogeneity, 3 decreasing) had potential mean–variation relationships that could explain the association between heterogeneity and age (Table 2, third column). Notably, heterogeneity in physical performance measures, chronic condition count, and frailty index was not associated with age after adjustment for the mean change. Measures of seven health characteristics had a normative or clinical scaling: chair rise and TUG times, functional measures (OARS [Older Americans Resources and Services subscales for activities of daily living [ADL] and instrumental ADLs], LSA, PASE), chronic conditions, and frailty index. Supplementary Tables S2–S4 detail the intermediate results used to reach these results.

Figure 1 shows the association between age and effective variance for each domain, all domains, and all domains excluding laboratory values. Overall, heterogeneity increases with age. When comparing between domains, vital signs had the largest increase in heterogeneity, followed by physical performance measures and laboratory values; heterogeneity in physical measures decreases, whereas heterogeneity in functional and in physiological measures seems stable. Figure 1A–F shows domains with significant associations between heterogeneity and age; these

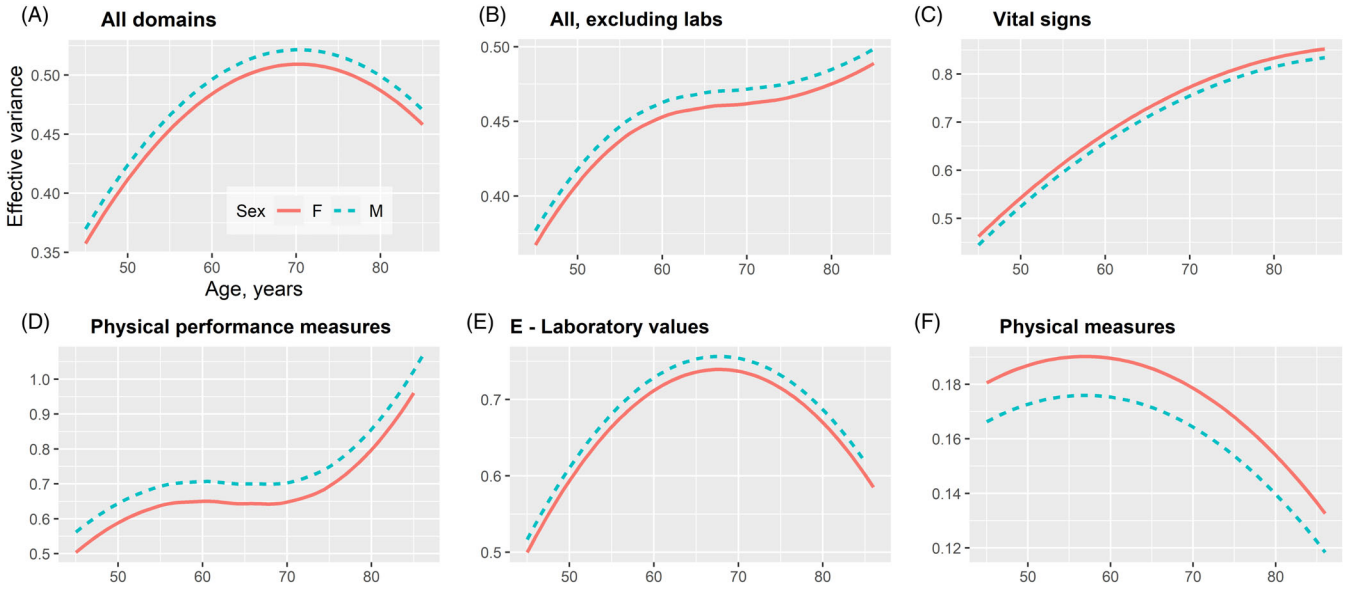

Figure 1. Nonlinear variation in effective variance by chronological age, overall and by health domains. Predicted effective variance curves overall and by domains, with gender-specific intercepts, are illustrated for significant (nonlinear) associations. The associations between chronological age and effective variance by *function and disability measures* and by *physiological measures* were nonsignificant. (A and E) Predicted effective variance curves reveal clearly nonmonotonic relationships where heterogeneity increases until approximately 70 years and then decreases for *all domains* and *laboratory values*. Although age is linearly associated with heterogeneity for both, assuming a linear relationship is misleading due to nonmonotonicity. (B, C, and D) Heterogeneity increases with age for *all*, *excluding laboratory values*, *vital signs*, and *physical performance measures*. Only the effective variance of *vital signs* increases approximately linearly. Effective variance for *physical performance measures* appears to stabilize between 55 and 70 years old. (F) Effective variance for *physical measures* appears to peak at around 57 years and then decrease.

domains had at least second-order polynomial associations: quadratic for all domains, physical measures, laboratory values, and vital signs; cubic for all, excluding laboratory values and physical performance. The relationship was nonmonotonic (i.e., varied in direction) for all domains, laboratory values (inverted-U shape), and physical measures. Overall peak heterogeneity is reached around 70 years of age but increases continuously when excluding laboratory values (Figures 1A,B).

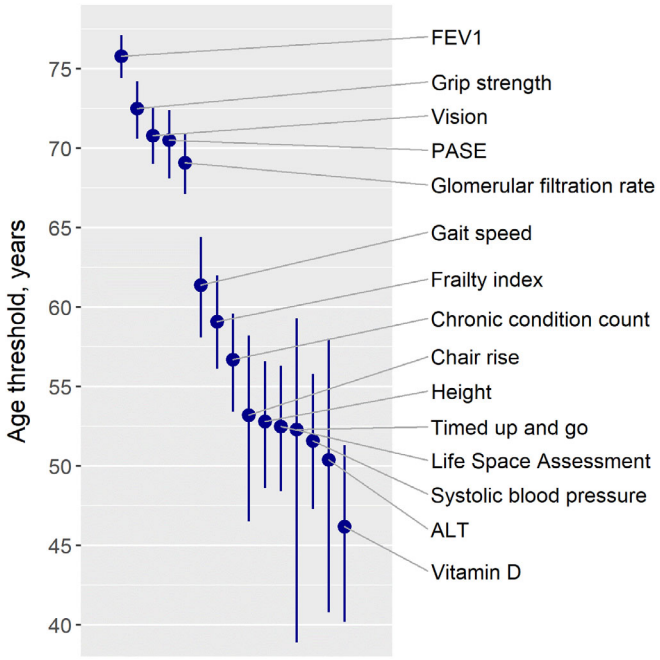

Figure 2. Age thresholds at which mean deviation between an age group and the 85-year-old age group exceeds mean deviation within the group of 85-year-old individuals. Abbreviations: FEV1, forced expiratory volume in 1 second; PASE, physical activity scale for the elderly. Vertical lines indicate bootstrapped 95% confidence intervals. The age thresholds give an indication of the relationship between variation *within* the 85-year-old age group compared with variation *between* age groups for each health characteristic. Once this age threshold is crossed, individuals below that age can be considered more different from those 85 years old than 85-year-olds between themselves. For example, for *FEV1*, starting at 76 years, the deviation between the average 85-year-old individual and the average 76-year-old individual ("between-group" deviation) exceeds the mean deviation between 85-year-old individuals themselves ("within-group" deviation). The analogous threshold is crossed at 46 years for *vitamin D*. Fifteen health characteristics cross the threshold between 45 and 85 years. The remaining health characteristics do not: *albumin*, *BMI*, *bone mineral density*, *cholesterol*, *CRP*, *creatinine*, *diastolic blood pressure*, *ferritin*, *HbA1C*, *HDL*, *hemoglobin*, *LDL*, *OARS28*, *platelet*, *pulse*, *TSH*, *waist circumference*, *weight*, *WBC*. For these health characteristics, individuals at 85 years continue to show more deviation within their age group than between their age group and the age group of individuals aged 45 years. For most variables, the variation between older adults is greater than between age groups: this requires clinicians to customize management based on specific individual values of prognostic health characteristics rather than relying on chronological age. See Methods and Supplementary Figure S2 for more details. [Color figure can be viewed at wileyonlinelibrary.com]

Comparing between-Age Group and Within-Older Age Group Heterogeneity

Figure 2 shows the age thresholds where the between-age group deviation exceeds the within-group deviation of individuals aged 85 years. Fifteen variables crossed this threshold between 45 and 85 years, with the thresholds for FEV1 (75.8 years), grip strength (72.5), vision (70.8), and PASE (70.5) closest to 85 years. Once this age threshold is crossed, individuals below that age can be considered more different from those 85 years old than between 85-year-olds themselves (within-age group). Physical performance measures (grip strength, gait speed, chair rise, and TUG), chronic condition count, frailty index, and LSA crossed the threshold before 45 years, but not ADL-iADLs (measured by OARS).

DISCUSSION

Features of Heterogeneity in Aging

Overall, our results confirm the widely held dictum that health heterogeneity increases with chronological age.[3,4,19] Older adults are, *in general*, more heterogeneous among themselves than younger adults. However, our analyses reveal that this statement requires many caveats. In line with, and extending previous work,[13,15,16] half of the 34 variables examined showed increased variability, but 8 showed decreased variability, and for 9, variability did not appear to change with age. Except for physical measures, heterogeneity tended to increase for all domains, but associations were mostly nonlinear, and nonmonotonic for overall domains, laboratory values, and physical measures. Our findings suggest multiple heterogeneity trajectories,[3] including an inverted-U trajectory for laboratory values. Of the 17 variables with increasing heterogeneity, 8 could be attributable to mean–variation relationships and 5 to normative or clinical scaling of measures. What is measured and how it is measured influences heterogeneity. Supplementary Table S5 presents six key features that clarify the description and understanding of heterogeneity: group, spread, measure, specificity, monotonicity, and mean–variation. Heterogeneity in aging is itself heterogeneous and multifaceted: in what follows, we wish to highlight how features of heterogeneity in older adults are relevant to clinical practice and research, as summarized in Table 3.

Clinical Implications

Greater heterogeneity with age for most health characteristics and domains justifies greater attention when managing older adults.[24] The greater probability of finding clinically relevant differences in older adults compared with their younger peers strongly supports the careful and potentially time-consuming comprehensive geriatric assessment, particularly in oncological or perioperative settings where these differences are highly predictive of outcomes.[25,26]

However, this increased heterogeneity was not found for all variables and was especially important for physical performance measures, chronic condition count, frailty

Table 3. Summary of Major Findings on Heterogeneity in Aging and Clinical and Research Implications

Findings	Clinical Implications	Research Implications and Themes
Overall heterogeneity in health increases with chronological age.	- Heterogeneity underlies the need and relevance of age appropriate care and management.	—
Heterogeneity does not increase uniformly and may be attributable measurement properties (e,g., scaling of measures, mean–variation relationship).	—	- Variability in heterogeneity in aging precludes uniform statement about heterogeneity in older adults. - Investigate the impact of measurement properties and selective survival on variation in heterogeneity in aging.
Clinically age-relevant variables are more heterogeneous with age.	- Heterogeneity reinforces the importance of clinically age-relevant variables in the CGA (chronic conditions, function and disability, physical performance measures, frailty, etc.).	- Develop heterogeneity capturing measures that use optimal scaling for older adults.
The scaling of a measure determines, to a great extent, the amount of heterogeneity detected.	- Clinicians should select measures that use clinically relevant scaling for its intended purpose, for example, PASE vs OARS. - Heterogeneity of healthcare costs with age is driven by clinically relevant measures.	
Heterogeneity in aging can be decomposed into differences between age groups and differences between older individuals (within age group).	- Geriatric care is based on managing older adults differently from younger adults, as well as differently between older adults themselves. - Care for older adults must account for age, also going beyond age as a surrogate mean marker for relevant prognostic factors.	- Refine and develop multidimensional constructs to stratify subgroups older adults by using variables that are most heterogeneous among them. - Measures of biological age may benefit from using variables that are less heterogeneous with aging.
Deviation from the mean (and heterogeneity) will vary by the specific group of reference selected. This is especially true when there are important differences between the mean values by age group.	- The reference group selected is essential to interpret disease states and conditions that are based on a statistical distribution, for example, osteoporosis (T-score vs Z-score), vital signs, or anemia. - Underdiagnosis or overdiagnosis may occur if a younger referent group is used without relevant clinical justification.	- Identify participants for research as those outlying within their chronological age group. - Investigate alternative and clinically useful ranges for vital signs and laboratory results for older adults.
Laboratory values attain peak heterogeneity in the late 60s.	—	- Develop and integrate laboratory biomarkers that are better suited to assess and differentiate older adults.

index, and—to a lesser extent—functional measures. These variables have in common an age-related focus and an underlying normative or clinical scaling. Assessment using age-related and clinically relevant measures will uncover greater heterogeneity in older adults. In practice, this reinforces the chief importance of physical performance, multimorbidity, frailty, and functional measures as core dimensions of the comprehensive geriatric assessment, beyond other health characteristics generally considered in the medical setting.

Our findings indicate that the scaling of measures influences the amount of heterogeneity captured within a dimension: PASE and OARS both measure function and disability, yet we show decreasing heterogeneity of PASE and increasing heterogeneity of OARS with age. PASE assesses function from extremely active to no activity, whereas OARS measures ADLs, which are only impaired with clinically significant functional decrease. As a reduction in PASE at the higher range of functional capacity has a lower impact on quality of life than a reduction in OARS, the latter should be favored when evaluating older adults. Clinicians caring for older adults should select and incorporate measures that are optimally scaled for this population to better characterize heterogeneity and improve decision-making. Most age-related health characteristics are clinically scaled, which may explain increasing heterogeneity with age and drive heterogeneity in healthcare costs.[19]

We show that increased heterogeneity in older adults can be decomposed into that between older and younger adults and between one older adult and others of the same age. Geriatric expertise and teaching are premised on this dual difference between ages and between aged individuals. These two differences contribute distinct knowledge to geriatric care: the first informs how care should be different by age groups (between the younger adults and older adults as a *group*), and the second highlights the critical importance of personalizing care beyond chronological age (as an *individual* older adult in their age group). The relative importance of these two differences depends on the specific variable considered. FEV1, visual acuity, or grip strength are variables

where between-age variation dominates and thus where chronological age (being "old") captures most of the variation. However, for most variables, the variation between older adults themselves is greater. Even in geriatric practice, chronological age can only be used as a gross surrogate marker for the mean[16]: clinicians should customize management of older adults beyond age by considering an individual's specific levels of prognostic factors as they will often be discordant with average age category levels. This may be contrasted with pediatric medicine, where the greater part of differences is between children and adults rather than between children themselves.

Determination of heterogeneity in aging is contingent on the referent group used. Typically, norms for vital signs and laboratory values are similar in adults regardless of age: the same referent group and center location are used, from which the spread of each individual is calculated. We demonstrated that both mean value and spread from the mean change with age for systolic blood pressure and the majority of laboratory values. This provides compelling arguments for the implementation of age-specific ranges.[27-29] Moreover, individual-specific ranges to determine "normal" values could also be more widely considered, as has been recommended for temperature.[30] Using an all-age or age-specific referent group may be specific to the context and depends on the expected benefit of interventions. For example, using the T-score (all-age referent group) or the Z-score (age-specific referent group) for bone mineral density (BMD) will identify individuals who may benefit differentially from treatment. For the many distribution-determined conditions, such as osteoporosis or anemia, clinicians should carefully question the appropriateness of the referent group and the clinical relevance of the absolute threshold of what is considered abnormal.

Research Implications

Our results showing variability in increases of heterogeneity in aging preclude uniform statements about heterogeneity in older adults. They show that its quantification is intertwined with the measurement properties of the instruments used (scaling and mean–variation relationship), which future research should seek to disentangle from true variation in aging. Selective survival, whereby mortality occurs in a nonrandom segment of a population,[31] may influence variation in heterogeneity. As extreme values of health characteristics are more strongly associated with mortality, the attrition of individuals may result in the underestimation heterogeneity. Homeostenosis of aging, with decreased resistance and redundancy to stressors, may translate into greater variability[32,33] but only to a threshold above which death ensues. Future research could leverage the longitudinal design of the ongoing CLSA or other cohorts to examine variability at the cohort and individual levels, and its association with mortality.

In addition to research on heterogeneity itself, our findings suggest lines of enquiry that use heterogeneity to enhance clinical management. Heterogeneity can refine the selection of variables used to develop constructs to stratify subgroups of older adults specifically. Age-related constructs, most importantly frailty, seek to capture heterogeneity among older adults as a broad group,[4] rather than distinguishing individuals among smaller subgroups of older age. We show that the heterogeneity by age decreases for many variables (e.g., FEV1, visual acuity, grip strength). To distinguish the more robust 85-year-olds from others of the same age, using or developing novel scales for variables that show increasing heterogeneity by age should be considered. Conversely, measures of biological age that seek to capture the latent aging process might benefit from including variables that have *decreasing* heterogeneity by age.

Heterogeneity may inform the selection of participant subgroups for research. An epistemological and clinical assumption is that large deviation from the mean may hold potential for discovery and intervention. Modifiable health states or trajectories are more likely to be identified in individuals with outlying characteristics *from their age group* rather than outlying from all adults, especially because adults of considerable age may all be outliers from the general population.

Overall heterogeneity appears to have an inverted U-shape with maximum variability at approximately 70 years. This inverted shape is strongly driven by laboratory values, leading to the possibility that standard laboratory measures optimally distinguish younger older adults and that other better-suited biomarkers should be developed for older age groups.

Limitations

First, due to the large number of variables, our exploratory results may be prone to multiple testing issues. Nonetheless, most reported associations had strong statistical significance and were also clinically significant. Second, participants were community-dwelling older adults without cognitive impairment. Our analyses may underestimate heterogeneity if institutionalized and/or cognitively impaired older adults have more extreme variable values. However, the age range of participants from 45 to 85 years allowed exploration of heterogeneity in younger age groups, where the proportion of excluded participants was low. Third, although we attempted to choose variables representative of clinical practice, our selection of health characteristics may have influenced our findings. We focused on health states, but heterogeneity has also been described on psychological and social levels.[15,16] Fourth, because our analyses were cross-sectional, we cannot disentangle period or cohort effects from the true aging process per se.[3] Our findings should not be considered from the perspective of mechanistic or biologic aging but from a perspective of descriptive aging, which holds a predictive and clinically relevant meaning as discussed above. Along the same lines, we did not account for clinical management, which may decrease the "natural" variability of some variables (e.g., HbA1C, TSH, LDL). From a descriptive standpoint, treatment can be understood as a valid modifier of observed variation in heterogeneity with *chronological* age: medical conditions are increasingly prevalent with age but are also treated.

CONCLUSION

Overall health heterogeneity increases with age but does not uniformly increase across all variables and domains. Heterogeneity in aging reinforces the need for geriatric assessment and care, depending on which health characteristics are assessed, their measurement properties, and their referent group. Like the older adults it seeks to describe,

heterogeneity in aging is itself heterogeneous, suggesting that further research is necessary to develop improved single-dimension and multidimensional instruments, as well as specific vital and laboratory reference ranges for older adults.

ACKNOWLEDGMENTS

The authors thank Drs Mark Clarfield and Han Ting Wang for their suggestions to improve the manuscript. This research was made possible using the data/biospecimens collected by the Canadian Longitudinal Study on Aging (CLSA). Funding for the CLSA is provided by the Government of Canada through the Canadian Institutes of Health Research (CIHR) under Grant Reference: LSA 94473 and the Canada Foundation for Innovation. This research has been conducted using the CLSA dataset Comprehensive V4.0, under Application Number 190209. The CLSA is led by Drs Parminder Raina, Christina Wolfson, and Susan Kirkland. The opinions expressed in this manuscript are the authors' own and do not reflect the views of the CLSA.

Conflict of Interest: None.

Author Contributions: QDN, EMM, and CW designed the study. QDN analyzed data and drafted the manuscript. All authors contributed significantly to the content, critically reviewed, and approved the final manuscript for publication.

Sponsor's Role: The sponsor had no role in the design and conduct of this study; analysis and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

REFERENCES

1. Freedman VA, Spillman BC. Disability and care needs among older Americans. Milbank Q. 2014;92(3):509-541. https://doi.org/10.1111/1468-0009.12076.
2. Vaupel JW, Manton KG, Stallard E. The impact of heterogeneity in individual frailty on the dynamics of mortality. Demography. 1979;16(3):439-454. https://doi.org/10.2307/2061224.
3. Dannefer D, Sell RR. Age structure, the life course and "aged heterogeneity": prospects for research and theory. Compr Gerontol B. 1988;2(1):1-10.
4. Mitnitski A, Howlett SE, Rockwood K. Heterogeneity of human aging and its assessment. J Gerontol A Biol Sci Med Sci. 2016;72(7):877-884. https://doi.org/10.1093/gerona/glw089.
5. Oxford English Dictionary. *Heterogeneity*. https://www.lexico.com/definition/heterogeneity.
6. Carnes BA, Olshansky SJ. Heterogeneity and its biodemographic implications for longevity and mortality. Exp Gerontol. 2001;36(3):419-430. https://doi.org/10.1016/S0531-5565(00)00254-0.
7. Mitnitski AB, Mogilner AJ, Rockwood K. Accumulation of deficits as a proxy measure of aging. Sci World. 2001;1:323-336. https://doi.org/10.1100/tsw.2001.58.
8. Fried LP, Tangen CM, Walston J, et al. Frailty in older adults: evidence for a phenotype. J Gerontol A Biol Sci Med Sci. 2001;56(3):808-813. https://doi.org/10.1093/gerona/56.3.M146.
9. Fabbri E, Zoli M, Gonzalez-Freire M, Salive ME, Studenski SA, Ferrucci L. Aging and multimorbidity: new tasks, priorities, and frontiers for integrated gerontological and clinical research. J Am Med Dir Assoc. 2015;16(8):640-647. https://doi.org/10.1016/j.jamda.2015.03.013.
10. Nguyen QD, Wu C, Odden MC, Kim DH. Multimorbidity patterns, frailty, and survival in community-dwelling older adults. J Gerontol A Biol Sci Med Sci. 2019;74(8):1265-1270. https://doi.org/10.1093/gerona/gly205/5088142.
11. Hoogendijk EO, Afilalo J, Ensrud KE, Kowal P, Onder G, Fried LP. Frailty: implications for clinical practice and public health. Lancet. 2019;394(10206):1365-1375. https://doi.org/10.1016/S0140-6736(19)31786-6.
12. Botwinick J, Thompson LW. A research note on individual differences in reaction time in relation to age. J Genet Psychol. 1968;112:73-75.
13. Maddox GL, Douglass EB. Aging and individual differences: a longitudinal analysis of social, psychological, and physiological indicators. J Gerontol. 1974;29(5):555-563. https://doi.org/10.1093/geronj/29.5.555.
14. Bornstein R, Smircina MT. The status of the empirical support for the hypothesis of increased variability in aging populations. Gerontologist. 1982;22(3):258-260. https://doi.org/10.1093/geront/22.3.258.
15. Nelson AE, Dannefer D. Aged heterogeneity: fact or fiction? The fate of diversity in gerontological research. Gerontologist. 1992;32(1):17-23. https://doi.org/10.1093/geront/32.1.17.
16. Stone ME, Lin J, Dannefer D, Kelley-Moore JA. The continued eclipse of heterogeneity in gerontological research. J Gerontol B Psychol Sci Soc Sci. 2017;72(1):162-167. https://doi.org/10.1093/geronb/gbv068.
17. Santoni G, Angleman S, Welmer AK, Mangialasche F, Marengoni A, Fratiglioni L. Age-related variation in health status after age 60. PLoS One. 2015;10(3):1-10. https://doi.org/10.1371/journal.pone.0120077.
18. Mungas D, Beckett L, Harvey D, et al. Heterogeneity of cognitive trajectories in diverse older persons. Psychol Aging. 2010;25(3):606-619. https://doi.org/10.1037/a0019502.
19. Lowsky DJ, Olshansky SJ, Bhattacharya J, Goldman DP. Heterogeneity in healthy aging. J Gerontol A Biol Sci Med Sci. 2014;69(6):640-649. https://doi.org/10.1093/gerona/glt162.
20. Raina P, Wolfson C, Kirkland S, et al. Cohort profile: the Canadian longitudinal study on aging (CLSA). Int J Epidemiol. 2019;48(6):1752-1753j. https://doi.org/10.1093/ije/dyz173.
21. Raina PS, Wolfson C, Kirkland SA. *Canadian Longitudinal Study on Aging – Protocol*. https://clsa-elcv.ca/doc/511. Accessed November 26, 2019.
22. Searle SD, Mitnitski A, Gahbauer EA, Gill TM, Rockwood K. A standard procedure for creating a frailty index. BMC Geriatr. 2008;8(1):24. https://doi.org/10.1186/1471-2318-8-24.
23. Peña D, Rodríguez J. Descriptive measures of multivariate scatter and linear dependence. J Multivar Anal. 2003;85(2):361-374. https://doi.org/10.1016/S0047-259X(02)00061-1.
24. Parker SG, McCue P, Phelps K, et al. What is comprehensive geriatric assessment (CGA)? An umbrella review. Age Ageing. 2018;47(1):149-155. https://doi.org/10.1093/ageing/afx166.
25. Chow WB, Rosenthal RA, Merkow RP, Ko CY, Esnaola NF. Optimal preoperative assessment of the geriatric surgical patient: a best practices guideline from the American college of surgeons national surgical quality improvement program and the American geriatrics society. J Am Coll Surg. 2012;215(4):453-466. https://doi.org/10.1016/j.jamcollsurg.2012.06.017.
26. Wildiers H, Heeren P, Puts M, et al. International society of geriatric oncology consensus on geriatric assessment in older patients with cancer. J Clin Oncol. 2014;32(24):2595-2603. https://doi.org/10.1200/JCO.2013.54.8347.
27. Chester JG, Rudolph JL. Vital signs in older patients: age-related changes. J Am Med Dir Assoc. 2011;12(5):337-343. https://doi.org/10.1016/j.jamda.2010.04.009.
28. Vadiveloo T, Donnan PT, Murphy MJ, Leese GP. Age- and gender-specific TSH reference intervals in people with no obvious thyroid disease in Tayside, Scotland: the thyroid epidemiology, audit, and research study (TEARS). J Clin Endocrinol Metab. 2013;98(3):1147-1153. https://doi.org/10.1210/jc.2012-3191.
29. Santoni G, Calderón-Larrañaga A, Vetrano DL, Welmer A-K, Orsini N, Fratiglioni L. Geriatric health charts for individual assessment and prediction of care needs: a population-based prospective study. J Gerontol A Biol Sci Med Sci. 2020;75(1):131-138. https://doi.org/10.1093/gerona/gly272.
30. High KP, Bradley SF, Gravenstein S, et al. Clinical practice guideline for the evaluation of fever and infection in older adult residents of long-term care facilities: 2008 update by the Infectious Diseases Society of America. Clin Infect Dis. 2009;48(2):149-171. https://doi.org/10.1086/595683.
31. Markides KS, Machalek R. Selective survival, aging and society. Arch Gerontol Geriatr. 1984;3(3):207-222. https://doi.org/10.1016/0167-4943(84)90022-0.
32. Ferrucci L, Windham BG, Fried LP. Frailty in older persons. Genus. 2005;61(1):39-53.
33. Lipsitz LA, Goldberger AL. Loss of "complexity" and aging: potential applications of fractals and chaos theory to senescence. JAMA. 1992;267(13):1806-1809. https://doi.org/10.1001/jama.1992.03480130122036.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article.

Supplementary Table S1: Missing data and proportion for health characteristics.

Supplementary Table S2: Coefficients for Change in Deviation by Age and Statistical and Clinical Significance.

Supplementary Table S3 Coefficients for Change in Mean by Age and Statistical and Clinical Significance.

Supplementary Table S4: Explainability of Change in Deviation by Mean–Variability Relation.

Supplementary Table S5: Features and Ontology of Heterogeneity and Examples Related to Aging.

Supplementary Figure S1: Flowchart for Determining the Association Between Heterogeneity and Age and Explainability by Mean–Variation Relation and by Measurement Scaling.

Supplementary Figure S2: Illustrative Scatterplots Comparing Within-Group Deviation at 85 Years vs Between-Age Group Deviation, for Grip Strength and HbA1C.

Supplementary Figure S3: Boxplots of Health Characteristics by Age and Sex.

Supplementary Methods S1: Health characteristics examined for heterogeneity in aging.