Longitudinal analysis of blood markers reveals progressive loss of resilience and predicts ultimate limit of human lifespan

Timothy V. Pyrkov[a,*], Konstantin Avchaciov[a], Andrei E. Tarkhov[a,b],
Leonid I. Menshikov[a,c], Andrei V. Gudkov[d,e] and Peter O. Fedichev[a,f]
[a] *Gero LLC, 105064, Nizhny Susalny per. 5/4, Moscow, Russia*
[b] *Skolkovo Institute of Science and Technology, Bolshoy Boulevard 30, bld. 1, Moscow 121205, Russia*
[c] *National Research Center "Kurchatov Institute", 1, Akademika Kurchatova pl., Moscow 123182, Russia*
[d] *Roswell Park Comprehensive Cancer Center, Elm and Carlton streets, Buffalo, NY 14263, USA*
[e] *Genome Protection, Inc. 640 Ellicott St #444, Buffalo, NY 14203, USA and*
[f] *Moscow Institute of Physics and Technology, 141700, Institutskii per. 9, Dolgoprudny, Moscow Region, Russia*[*]

ABSTRACT

We analyzed aging trajectories of complete blood counts (CBC) and their association with the incidence of chronic diseases and death in cohorts of aging individuals registered in the UK Biobank and National Health and Nutrition Examination Survey (NHANES) studies. Application of a proportional hazards model to the CBC data allowed us to identify the log-transformed hazard ratio as a quantification of aging process, which we have named the dynamic morbidity index (DMI). DMI increased with age in the UK Biobank and NHANES cohorts, was associated with frailty, and predicted the prospective incidence of age-related diseases and death. To better understand the nature of DMI variations along individual aging trajectories, we acquired a sufficiently large longitudinal database of CBC measurements from a consumer diagnostics laboratory. We observed that population DMI distribution broadening can be explained by a progressive loss of physiological resilience measured by the DMI auto-correlation time. Extrapolation of this data suggested that DMI recovery time and variance would simultaneously diverge at a critical point of 120-150 years of age corresponding to a complete loss of resilience. We conclude that the criticality resulting in the end of life is an intrinsic biological property of an organism that is independent of stress factors and signifies a fundamental or absolute limit of human lifespan.

INTRODUCTION

Aging is manifested as a progressive functional decline leading to increasing prevalence [1, 2] and incidence of chronic age-related diseases (e.g., cancers, diabetes, cardiovascular diseases, etc. [3–5]) and disease-specific mortality [6]. Much of our current understanding of the relationship between aging and changes in physiological variables over an organism's lifespan originates from large cross-sectional studies, including the National Health and Nutrition Examination Survey (NHANES; 40592 subjects, age range $18-85$ y.o.) and UK Biobank (UKB; 471473 subjects, age range $39-73$ y.o.) studies. Analysis of such large datasets has identified a number of "biological clocks", including those reflecting age-related variations in blood markers [7], DNA methylation (DNAm) states [8, 9] or patterns of locomotor activity [10–12]. Typically, the physiological indices change from the levels observed in the young organism at approximately linear pace, much slower than would be expected from the exponential Gompertz mortality law [13, 14]. Most factors that serve as biological clocks are easily measurable and, at least in principle, may be modified pharmacologically. An improved understanding of the relationship between the linear physiological state dynamics and the exponentially increasing morbidity, frailty and mortality observed during aging is needed to facilitate the rational design, development, and clinical validation of anti-aging interventions.

High variability of physiological indices even within the reference ranges characteristic for healthy subjects poses a significant practical challenge for quantification of the aging process from medical test results. Deviations of a clinical biomarker from the norm may be diagnostically significant if they are either too strong or mild but present over a meaningfully long time. Typically, however, age-related changes in physiological variables are smaller than natural variability in healthy individuals, and therefore the aging drift of the organism state variables is only detectable in large cross-sectional studies or longitudinal studies of considerable length.

In this paper, we addressed these theoretical and practical issues by a systematic investigation of aging, organism state fluctuations, and gradual loss of resilience with the help of a sufficiently large study involving complete blood counts (CBC) measurements. We built a biomarker of aging in the form of the log-mortality risk predictor from CBC variables. First, we characterized its association with age, frailty, lifestyle, and its prog-

[*] To whom correspondence should be addressed. E-mail: tim.pyrkov@gero.com, peter.fedichev@gero.com

nostic performance for mortality and morbidity risks assessment. Then, we observed that the increase in the biomarker variability with age can be explained by the rise of the biomarker's auto-correlation time. The auto-correlation time is thus an independent biomarker of aging that is characteristic of resilience since it quantifies the organism state rigidity in response to stresses (such as the recovery time characterizing a transition from an illness to a healthy state). Our analysis shows that the auto-correlation time grows with age from about 2 weeks to over 4 weeks for cohorts aging from 40 to 70 years. Interestingly, by extrapolation, the recovery time would diverge at about $120 - 150$ years, thus indicating the absolute limit of human lifespan.

RESULTS

Quantification of aging based on complete blood counts (CBC) measurements

Complete blood count (CBC) measurements are included in standard blood tests and comprise the largest common subset of data available for both the NHANES and UKB study populations (see Table S1). In investigating whether CBC dynamics could be used to quantify aging, we found that co-clustering of age- and sex-adjusted CBC components revealed two dynamic subsystems associated with oxygenation and immune functions, represented by variations in red blood cell counts as well as total and mean corpuscular hemoglobin levels, and with platelet and white blood cell counts, respectively (Fig. 1A).

Since rates of morbidity and mortality increase exponentially with age, a log-linear risk model is a good starting point for quantification of the aging process [12, 15]. We used the death register of the NHANES study (3792 death events observed in the follow-up by year 2015) and trained the Cox proportional hazards model [16] using the CBC measurements and sex variables (but not age) from 23807 study participants aged 40 y.o. and older. We therefore assumed that the risk of death depended on the organism state at the time of CBC measurement and the follow-up time only. The model yielded a log-hazards ratio (log-HR) estimate (a linear combination of log-transformed CBC variables) for every participant (see Table S2 for description of the model). After adjustment for sex and age, this predictor was demonstrated to be equally well associated with mortality in the NHANES study (HR = 1.43) and in the independent UKB study (HR = 1.39; Table S3), which was used as a validation dataset.

For the NHANES cohort, the CBC log-HR of the Cox mortality model gradually increased with age (Pearson's correlation coefficient $r = 0.29$, $p < 10^{-100}$, Fig. 1B). This parameter changed at a comparable rate in the independent UKB cohort; however, the correlation coefficient was lower in this cohort due to the more limited age range of the subjects ($r = 0.13$, $p < 10^{-100}$, Fig. S1A).

To differentiate between the effects of chronic diseases and disease-free aging, we followed [17] and characterized the health status of each study participant based on their diagnosis with health conditions considered to be the most prevalent age-related conditions. The number of health conditions diagnosed for an individual was normalized to the total number of conditions included in the analysis to yield the "compound morbidity index" (CMI), a frailty index proxy with values ranging from zero to one. The list of health conditions common to the NHANES and UKB studies that were used for CMI determination is given in Table S3 and described in SI Appendix.

The association between frailty and estimated mortality risk is readily seen as the difference between the solid and dashed lines in Fig. 1B, which represent the mean log-HR values in the cohorts of healthy ("non-frail", CMI < 0.1) and "most frail" (CMI > 0.6) NHANES participants, respectively. Given the association between all-cause mortality risk and morbidity, we propose to refer to the log-HR of the mortality model as the "dynamic morbidity index" (DMI).

In groups stratified by increasing number of health condition diagnoses, the normalized distribution of DMI (after adjustment by mean DMI in age- and sex-matched cohorts of healthy subjects) exhibited a progressive shift to higher risk values and more variability (Fig. 1C for NHANES, Fig. S1B for UKB). For both NHANES and UKB, the largest shift was observed in the "most frail" (CMI > 0.6) population. The increasingly heavy tail at the high end of the DMI distribution in this group is characteristic of an admixture of a distinct group of individuals occupying the adjacent region in the configuration space corresponding to the increased DMI variable. Therefore, DMI displacement after adjustment for age and sex in cohorts of healthy subjects was expected to be determined by the fraction of "most frail" individuals in a cohort of any given age. This was confirmed to be true using the NHANES dataset (Fig. 2A; $r = 0.83$, $p = 8.8 \times 10^{-9}$). The fraction of surviving "most frail" subjects increased exponentially until the age corresponding to the end of healthspan was reached. The characteristic doubling rate constants for the "most frail" population fractions were 0.08 and 0.10 per year in the NHANES and the UKB cohorts, respectively, in comfortable agreement with the accepted Gompertz mortality doubling rate of 0.085 per year [18], see Fig. 2B.

We note that the prevalence of diseases in the NHANES cohort is consistently higher than that in the UKB population, in spite of the lack of comparable lifespan in the two countries. This may be a consequence of the enrollment bias in the UKB: according to comparison with life tables in [19], the UKB subjects appear to live longer than UK residents.

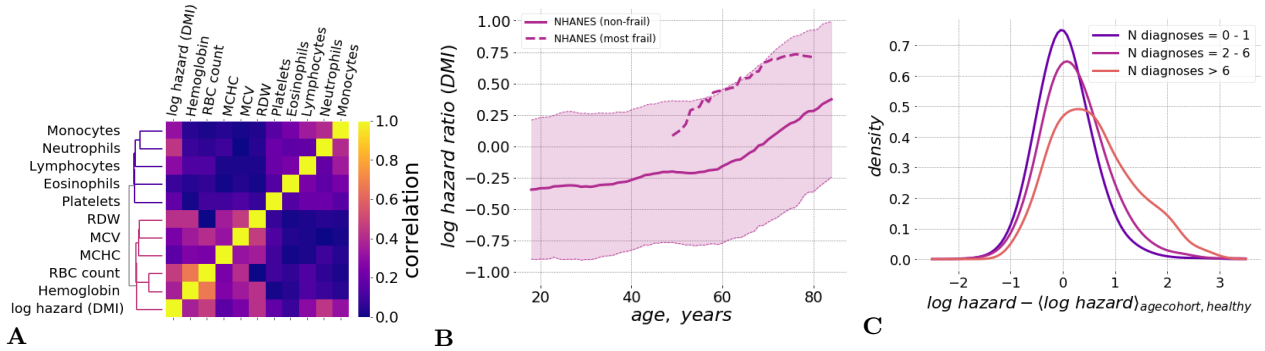

FIG. 1: **A.** Co-clustering of the age- and sex-adjusted CBC feature fluctuations in the NHANES dataset revealed two functionally related blood cell types, generally representing immune and oxygenation functions. We also included the dynamic morbidity index (DMI), the log-hazard ratio of a mortality risk model trained using the NHANES samples. As indicated by the vertical bar along the right-hand edge of the figure, the colors represent the absolute values of the Pearson's correlation coefficients between features. **B.** DMI mean values (solid line) and variance (shaded area) are plotted relative to age for the "non-frail" (combined morbidity index, CMI < 0.1) participants of NHANES study. The average DMI of the "most frail" (CMI > 0.6) individuals is shown with the dashed line. Data for other datasets investigated in this study are given in Supplementary Information (Fig. S1A). **C.** Distributions of sex- and age-adjusted DMI in cohorts of NHANES participants in different morbidity categories relative to the DMI mean in cohorts of "non-frail" (1 or no diagnoses, CMI < 0.1) individuals. Note that the distribution function in the "most frail" group (more than 6 diagnoses, CMI > 0.6) exhibited the largest shift and a profound deviation from the symmetric form.

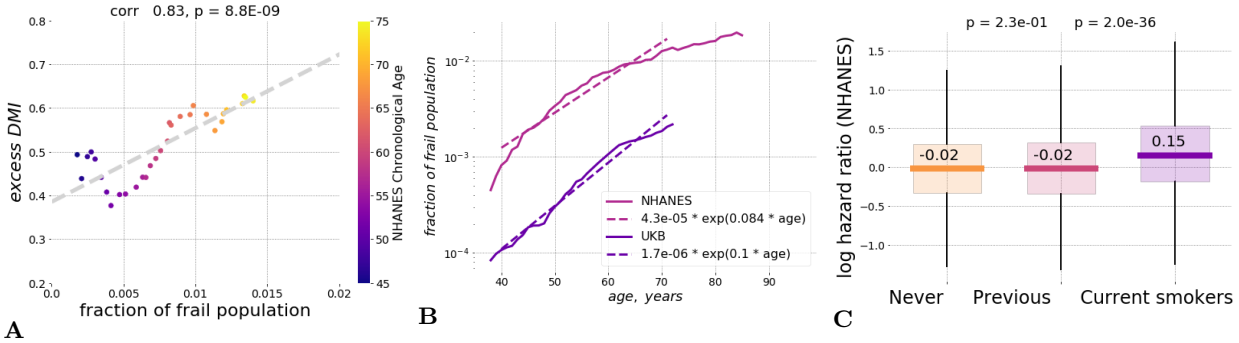

FIG. 2: **A.** Fraction of frail persons is strongly correlated with the excess DMI, that is the difference between the DMI of an individual and its average and the sex- and age-matched cohort in the "non-frail" population in NHANES. **B.** Exponential fit showed that until the age of 70 y.o. the fraction of the "most frail" individuals in the population grows approximately exponentially with age with the doubling rate constants of 0.08 and 0.10 per year in the NHANES and the UKB cohorts, respectively. **C.** Distribution of log hazards ratio in age- and sex-matched cohorts of NHANES participants who never smoked, smoked previously but quit prior to the time of study participation, or were current smokers at the time of the study.

Dynamic morbidity index and health risks

In the most healthy subjects, i.e. those with no diagnosed diseases at the time of assessment, the DMI predicted the future incidence of chronic age-related diseases observed during 10-year follow-up in the UKB study (Table S3) There was no relevant information available in NHANES. We tested this association using a series of Cox proportional hazard models trained to predict the age at the onset/diagnosis of specific diseases. We observed that the morbidity hazard ratios associated with the DMI relative to its mean in age- and sex-matched cohorts were statistically significant predictors for at least the most prevalent health conditions (those with more than 3000 occurrences in the UKB population). The effect size (HR $\approx 1.03 - 1.07$) was the same regardless of whether a disease was diagnosed first in a given individual or followed any number of other diseases. Therefore, we conclude that the DMI is a characteristic of overall health status that is universally associated with the risks of developing the most prevalent diseases and, therefore, with the end of healthspan as indicated by the onset of the first morbidity (HR ≈ 1.05 for the "First morbidity" entry in Table S3).

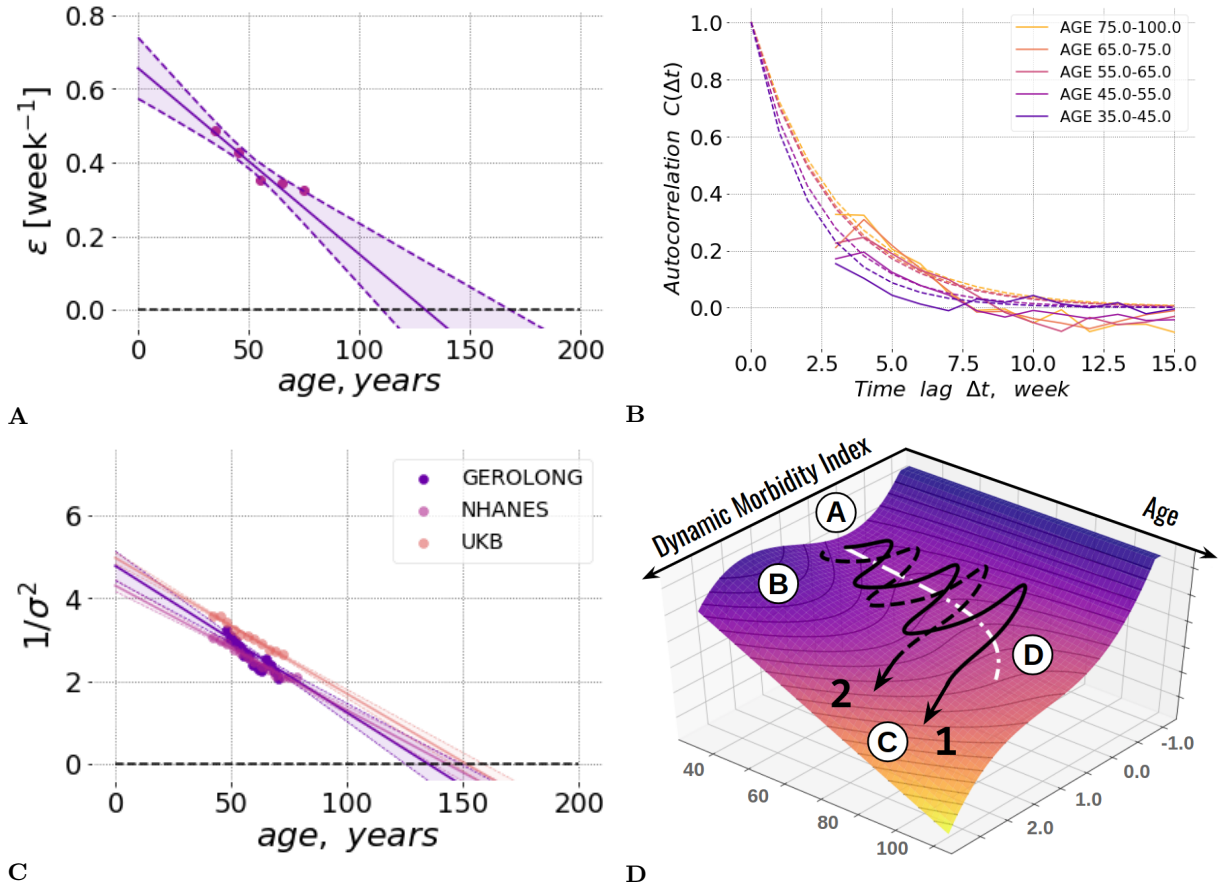

FIG. 3: **A.** The DMI relaxation rate (or the inverse characteristic recovery time) computed for sequential age-matched cohorts from the GEROLONG dataset decreased approximately linearly with age and could be extrapolated to zero at an age in the range of $\sim 110-170$ y.o. (at this point, there is complete loss of resilience and, hence, loss of stability of the organism state). The shaded area shows the 95% confidence interval. **B.** The auto-correlation function $C(\Delta t)$ of the DMI fluctuations during several weeks averaged in sequential 10-year age-cohorts of GEROLONG subjects showed gradual age-related remodelling. Experimental data and fit to autocorrelation function are shown with solid and dashed lines, respectively (see details in Supplementary Information). The DMI correlations are lost over time Δt between the measurements and, hence, the DMI deviations from its age norm reach the equilibrium distribution faster in younger individuals. **C.** The inverse variance of DMI decreased linearly in all three investigated datasets and its extrapolated value vanished (hence, the variance diverged) at an age in the range of $120-150$ y.o. We performed the linear fit for subjects 40 y.o. and older, excluding the "most frail" (CMI > 0.6) individuals. The shaded areas correspond to the 95% confidence intervals. **D.** Representative aging trajectories are superimposed over the potential energy landscape (vertical axis) representing regulatory constraints. The stability basin "A" is separated from the unstable region "C" by the potential energy barrier "B". Aging leads to a gradual decrease in the activation energy and barrier curvature and an exponential increase in the probability of barrier crossing. The stochastic activation into a dynamically unstable (frail) state is associated with acquisition of multiple morbidities and certain death of an organism. The white dotted line "D" represents the trajectory of the attraction basin minimum. Examples 1 (black solid line) and 2 (black dashed line) represent individual life-long stochastic DMI trajectories that differ with respect to the age of first chronic disease diagnosis.

In "non-frail" individuals with life-shortening lifestyles/behaviors, such as smoking, the DMI was also elevated, indicating a higher level of risks of future diseases and death (Fig. 2C). Notably, however, this effect appeared to be reversible: while the age- and sex-adjusted DMI means were higher in current smokers compared to non-smokers, they were indistinguishable between groups of individuals who never smoked and who quit smoking (c.f. [12, 20]).

Physiological state fluctuations and loss of resilience

To understand the nature of forces shaping the dynamics of the aging process at higher age/time-resolution, we acquired a large set of longitudinal CBC measurements from a clinical diagnostics laboratory. This dataset, which we referred to as GEROLONG, included 629 male and 1800 female subjects aged $35 - 90$ with complete CBC analyses that were sampled $4 - 20$ times within a period of up to 42 months.

As seen with the NHANES and UKB cohorts, DMI also increased with age in the longitudinal GEROLONG cohort. The average DMI value and its population variance at any given age were, however, considerably larger than those in the reference "non-frail" groups from the NHANES and UKB studies (see Fig. S1A). This difference likely reflects an enrollment bias: many of the GEROLONG blood samples were obtained from patients visiting clinic centers, presumably due to health issues. This could explain why the GEROLONG population appeared generally more frail in terms of DMI than the reference cohorts of the same age from other studies (Fig. S1A, compare the relative positions of the solid blue line and the two dashed lines representing the GEROLONG cohort and the frail cohorts of the NHANES and UKB studies, respectively). There was no medical condition information available for the GEROLONG subjects. Hence, we used the mean DMI of the "most frail" NHANES and UKB participants (which coincided, approximately, with the mean DMI of all GEROLONG subjects) as the cutoff value to select "non-frail" GEROLONG individuals.

Within the "non-frail" GEROLONG population, serial CBC measurements from individuals over periods of time of up to three years revealed large stochastic fluctuations of the DMI around its mean values, which differed between individual study participants. The averaged DMI auto-correlation function is a basic property of a stochastic process (see e.g., [21]) and decayed exponentially as a function of the time delay between measurements within approximately a month (see Fig. 3B). As described in the SI Appendix, for a stationary process the inverse auto-correlation time is an indicator of the relaxation (recovery) rate, characterizing the time scale involved in the equilibration of a system's state in response to external perturbations. We therefore propose using this quantity as a measure of an organism's "resilience", the capacity of an individual organism to resist and recover from the effects of physiological or pathological stresses [22, 23]).

We fitted the auto-correlation functions to an exponential function of the time delay and observed that recovery rates obtained from fitting to data in the subsequent age-cohorts decreased approximately linearly with age (Fig 3A). Extrapolation to older ages suggested that the equilibration rate vanishes and, hence, the recovery time becomes formally infinite, at an age of approximately $120 - 150$ y.o. (95% CI, $110 - 170$ y.o.).

The variance of DMI increased with age in every dataset evaluated in this study. Since the dynamic range of random fluctuations would be inversely proportional to the recovery force (or the recovery rate, see [21]), we plotted the inverse variance of the DMI computed in sex- and age-matched cohorts of healthy persons (Fig. 3C). Again, extrapolation suggested that, if the tendency holds at older ages, the population variability would increase indefinitely at an age of approximately $120 - 150$ y.o.

DISCUSSION

In this study, we investigated aging trajectories of human CBC values and their association with risks of chronic age-related diseases, mortality and life-shortening lifestyles. We produced a proportional hazards mortality model using NHANES dataset and defined its log-hazard ratio prediction as the dynamic morbidity index (DMI). This quantitative parameter displayed all of the expected properties of a biomarker of aging in several large independent datasets: DMI increased with age, was predictive of the prospective incidence of age-related diseases and death, and was associated with typical life-shortening lifestyles, such as smoking, and frailty. These findings support the idea that predictors from log-linear mortality or morbidity risk models can be effectively used to quantify the progress of aging and effects of lifestyles and diseases [12, 15, 24].

The simultaneous divergence of the organism state recovery times (critical slowing down) and the range of DMI variations (critical fluctuations) is characteristic of proximity of a critical point [21] at some advanced age over 100 y.o. Under these circumstances, the organism state dynamics are stochastic and dominated by the variation of the single dynamic variable associated with the criticality, the DMI (Fig. 3D). Schematically, far from the critical point (at younger ages), the organism state perturbations can be thought of as confined to the vicinity of a possible stable equilibrium state in a potential energy basin (A). Initially, the dynamic stability is provided by a sufficiently high potential energy barrier (B) separating this stability basin from the inevitably present dynamically unstable regions (C) in the space of physiological parameters. While in stability basin, an organism follows the trajectory (D) of the equilibrium state, which is gradually displaced with aging even for the successfully aging individuals.

The DMI auto-correlation times (one-two months, see Fig. 3B) are much shorter than lifespan. The dramatic separation of time scales makes it very unlikely that the linear decline of the recovery force measured by the recovery rate in Fig. 3A can be explained by the dynamics of the organism state captured by the DMI variation alone. Therefore, we conclude that the progressive remodeling of the attraction basin geometry reflects adjustment of the DMI fluctuations to the slow independent process that is aging itself. In this view, the aging drift of the DMI

mean in cohorts of healthy individuals (as in Fig. 1B) is the adaptive organism-level response to ever increasing stress produced by the aging process.

The longitudinal analysis in this work demonstrated that the DMI undergoes a random walk under the influence of stochastic stress factors and slowly increases, on average, as the organism ages. The dynamic range of the DMI fluctuations is inversely proportional to the recovery rate of the DMI fluctuations, and hence the two parameters are dependent quantities. Therefore, the mean DMI, on the one hand, and the dependent variables, the DMI variability and the auto-correlation time, on the other, comprise the minimum pair of independent biomarkers of human aging.

The DMI recovery rate characterizes fluctuations of DMI on time scales of a few months or more, decreases with age and thus indicates the progressive loss of physiological resilience. Such age-related remodeling of recovery rates has been previously observed in studies of various physiological and functional parameters in humans and other mammals. For example, in humans, a gradual increase in recovery time required after macular surgery was reported in sequential 10-year age cohorts [25] and age was shown to be a significant factor for twelve months recovery and the duration of hospitalization after hip fracture surgery [26, 27], coronary artery bypass [28], acute lateral ankle ligament sprain [29]. A mouse model suggested that the rate of healing of skin wounds also can be a predictor of longevity [30].

In a reasonably smooth potential energy landscape forming the basin of attraction, the activation energy required for crossing the protective barrier (B) decreases along with the curvature at the same pace, that is, linearly with age. Whenever the protective barrier is crossed, dynamic stability is lost (see example trajectories 1 and 2 in Fig. 3D, which differ by the age of crossing) and deviations in the physiological parameters develop beyond control, leading to multiple morbidities, increasing frailty, and, eventually, death.

On a population level, activation into such a frail state is driven by stochastic forces and occurs approximately at the age corresponding to the end of healthspan, understood as "disease-free survival". Since the probability of barrier crossing is an exponential function of the required activation energy (i.e., the barrier height) [21], the weak coupling between DMI fluctuations and aging is then the dynamic origin of the Gompertz mortality law. Since the remaining lifespan of an individual in the frail state is short, the proportion of frail subjects at any given age is proportional to the barrier crossing rate, which is an exponential function of age (see Fig. 2B).

The end of healthspan can therefore be viewed as a form of a nucleation transition [21], corresponding in our case to the spontaneous formation of states corresponding to chronic diseases out of the metastable phase corresponding to healthy organisms. The DMI is then the order parameter associated with the organism-level stress responses at younger ages and plays the role of the "reaction coordinate" of the transition to the frail state later in life. All chronic diseases contributing to frailty and death in our model originate from the dynamic instability associated with single protective barrier crossings. This is, of course, a simplification and yet the assumption could naturally explain why mortality and the incidence of major age-related diseases increase exponentially with age at approximately the same rate [3].

The DMI is an organism level variable and, as such, is not a property of any specific functional subsystem or organism compartment. Indeed, the associations of individual CBC components with the DMI extend over the functional clustering shown in Fig. 1A. In the vicinity of a critical point, fluctuations associated with the critical mode become amplified and hence the DMI should be identifiable with the signal components explaining most of the variance in virtually every biological signal (see, e.g., [7, 31] for reviews including performance assessment of variance-based biological age models (e.g. based on Principal Component Analysis). For example, in our recent study [12], we observed that the first principal component score in the configuration space spanning the physical activity acceleration/deceleration statistics was strongly associated with mortality, incidence of first morbidity, and health status.

According to the presented model, external stresses (such as smoking) or diseases produce perturbations that modify the shape of the effective potential leading to the shift of the equilibrium DMI position. For example, the mean DMI values in cohorts of individuals who never smoked or who quit smoking are indistinguishable from each other, yet significantly different from (lower than) the mean DMI in the cohort of smokers (Fig. 2C). Thus, the effect of the external stress factor is reflected by a change in the DMI and is reversed as soon as the factor is removed. These findings agree with earlier observations suggesting that the effects of smoking on remaining lifespan and on the risks of developing diseases are mostly reversible once smoking is ceased well before the onset of chronic diseases [12, 20]. The decline in the lung cancer risk after smoking ablation [32] is slower than the recovery rate reported here. This may be the evidence suggesting that long-time stresses may cause hard-to repair damage to the specific tissues and thus produce lasting effects on the resilience.

The reversible character of the DMI deviations in healthy subjects suggests a complex relationship between the DMI and survival. On the one hand, the elevation of physiological variables associated with the DMI indicates reversible activation of the most generic protective stress responses at younger ages, when the organism state is dynamically stable. Moderately elevated DMI levels are therefore a marker of generic stress that can measured by molecular markers (e.g., C-reactive protein) and affects general physical and mental health status [24]. On the other hand, the excessive DMI observed in older individuals can be thought of as an aberrant activation of stress-responses beyond the dynamic stability range. This is

a characteristic of frailty, multiple chronic diseases and death.

We propose that therapies targeting frailty-associated phenotypes (e.g., inflammation) would, therefore, produce distinctly different effects in disease-free versus frail populations. In healthy subjects, who reside in the region of the stability basin (B) (see Fig. 3D), a treatment-induced reduction of DMI would quickly saturate over the characteristic auto-correlation time and lead to a moderate decrease in long-term risk of morbidity and death without a change in resilience. Technically, this would translate into an increase in healthspan, although the reduction of health risks would be transient and disappear after cessation of the treatment. In frail individuals, however, the intervention could reduce frailty, thus increasing lifespan beyond healthspan.

The emergence of chronic diseases and frailty out of increasingly unstable fluctuations of the organism state provides the necessary dynamic argument to support the derivation of the Gompertz mortality law in the Strehler-Mildvan theory of aging [33]. In [34, 35], the authors suggested that the exponential growth of disease burden observed in the National Population Health Survey of Canadians over 20 y.o. could be explained by an age-related decrease in organism recovery in the face of a constant rate of exposure to environmental stresses. Our study provides evidence suggesting that vanishing resilience cannot be avoided even in the most successfully aging individuals and, therefore, could explain the very high mortality seen in cohorts of super-centennials characterized by the so-called compression of morbidity (late onset of age-related diseases [36]). Formally, such a state of "zero-resilience" at the critical point corresponds to the absolute zero on the vitality scale in the Strehler-Mildvan theory of aging, thus representing a natural limit on human lifespan.

The proximity to the critical point indicates that the apparent human lifespan limit is not likely to be improved by therapies aimed against specific chronic diseases or frailty syndrome. Thus, no strong life extension is possible by preventing or curing diseases without interception of the aging process, the root cause of the underlying loss of resilience. We do not foresee any laws of nature prohibiting such an intervention. Therefore, further development of the aging model presented in this work has may eventually lead to experimental demonstration of a dramatic life-extending therapy.

MATERIALS AND METHODS

Full details for the materials and methods used in this study, including information of the CBC parameters, Cox proportional hazards model, health conditions, and analysis of physiological state fluctuations are provided in the Supplementary Information Appendix.

[1] Arnold Mitnitski and Kenneth Rockwood. The rate of aging: the rate of deficit accumulation does not change over the adult life span. *Biogerontology*, 17(1):199–204, 2016.

[2] Ruby Yu, Wan-Chi Wu, Jason Leung, Susan C Hu, and Jean Woo. Frailty and its contributory factors in older adults: a comparison of two asian regions (hong kong and taiwan). *International journal of environmental research and public health*, 14(10):1096, 2017.

[3] Aleksandr Zenin, Yakov Tsepilov, Sodbo Sharapov, Evgeny Getmantsev, Leonid Menshikov, Peter Fedichev, and Yurii Aulchenko. Identification of 12 genetic loci associated with human healthspan. *bioRxiv*, page 300889, 2018.

[4] Dmitriy I Podolskiy, Alexei V Lobanov, Gregory V Kryukov, and Vadim N Gladyshev. Analysis of cancer genomes reveals basic features of human aging and its role in cancer development. *Nature communications*, 7:12157, 2016.

[5] Teresa Niccoli and Linda Partridge. Ageing as a risk factor for disease. *Current Biology*, 22(17):R741–R752, 2012.

[6] Nir Barzilai and Gad Rennert. The rationale for delaying aging and the prevention of age-related diseases. *Rambam Maimonides medical journal*, 3(4), 2012.

[7] M. E. Levine. Modeling the rate of senescence: can estimated biological age predict mortality more accurately than chronological age? *J. Gerontol. A Biol. Sci. Med. Sci.*, 68(6):667–674, Jun 2013.

[8] G. Hannum, J. Guinney, L. Zhao, L. Zhang, G. Hughes, S. Sadda, B. Klotzle, M. Bibikova, J. B. Fan, Y. Gao, R. Deconde, M. Chen, I. Rajapakse, S. Friend, T. Ideker, and K. Zhang. Genome-wide methylation profiles reveal quantitative views of human aging rates. *Mol. Cell*, 49(2):359–367, Jan 2013.

[9] S. Horvath. DNA methylation age of human tissues and cell types. *Genome Biol.*, 14(10):R115, 2013.

[10] Philippe Terrier and Fabienne Reynard. Effect of age on the variability and stability of gait: a cross-sectional treadmill study in healthy individuals between 20 and 69 years of age. *Gait & posture*, 41(1):170–174, 2015.

[11] Tim Althoff, Jennifer L Hicks, Abby C King, Scott L Delp, Jure Leskovec, et al. Large-scale physical activity data reveal worldwide activity inequality. *Nature*, 547(7663):336, 2017.

[12] Timothy V. Pyrkov, Evgeny Getmantsev, Boris Zhurov, Konstantin Avchaciov, Mikhail Pyatnitskiy, Leonid Menshikov, Kristina Khodova, Andrei V. Gudkov, and Peter O. Fedichev. Quantitative characterization of biological age and frailty based on locomotor activity records. *Aging*, 10(10):2973–2990, 2018.

[13] Benjamin Gompertz. A sketch of an analysis and notation applicable to the value of life contingencies. *Philosophical Transactions of the Royal Society*, 110:214–294, 1820.

[14] William Matthew Makeham. On the law of mortality and construction of annuity tables. *The Assurance Magazine and Journal of the Institute of Actuaries*, 8(06):301–310,

1860.

[15] Zuyun Liu, Pei-Lun Kuo, Steve Horvath, Eileen Crimmins, Luigi Ferrucci, and Morgan Levine. Phenotypic age: a novel signature of mortality and morbidity risk. *bioRxiv*, page 363291, 2018.

[16] David R Cox. Regression models and life-tables. In *Breakthroughs in statistics*, pages 527–541. Springer, 1992.

[17] K Rockwood, JM Blodgett, O Theou, MH Sun, HA Feridooni, A Mitnitski, RA Rose, J Godin, E Gregson, and SE Howlett. A frailty index based on deficit accumulation quantifies mortality risk in humans and in mice. *Scientific Reports*, 7, 2017.

[18] Benjamin Gompertz. On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies. *Philosophical transactions of the Royal Society of London*, 115:513–583, 1825.

[19] Andrea Ganna and Erik Ingelsson. 5 year mortality predictors in 498 103 uk biobank participants: a prospective population-based study. *The Lancet*, 386(9993):533–540, 2015.

[20] Donald H Taylor Jr, Vic Hasselblad, S Jane Henley, Michael J Thun, and Frank A Sloan. Benefits of smoking cessation for longevity. *American journal of public health*, 92(6):990–996, 2002.

[21] LD Landau and EM Lifshitz. Physical kinetics, vol. 10. *Course of Theoretical Physics*, 1981.

[22] Gregory Hicks and Ram R Miller. Physiological resilience. In *Resilience in Aging*, pages 89–103. Springer, 2011.

[23] N Jennifer Klinedinst and Alisha Hackney. Physiological resilience and the impact on health. In *Resilience in Aging*, pages 105–131. Springer, 2018.

[24] Timothy V Pyrkov and Peter O Fedichev. Biological age is a universal marker of aging, stress, and frailty. *bioRxiv*, page 578245, 2019.

[25] Yonguk Kim, Eung Suk Kim, Seung-Young Yu, and Hyung Woo Kwak. Age-related clinical outcome after macular hole surgery. *Retina*, 37(1):80–87, 2017.

[26] Jana M Mossey, Elizabeth Mutran, Kathryn Knott, and Rebecca Craik. Determinants of recovery 12 months after hip fracture: the importance of psychosocial factors. *American journal of public health*, 79(3):279–286, 1989.

[27] Kenneth J Koval, Mary Louise Skovron, Gina B Aharonoff, and Joseph D Zuckerman. Predictors of functional recovery after hip fracture in the elderly. *Clinical orthopaedics and related research*, 1(348):22–28, 1998.

[28] NT Artinian, C Duggan, and P Miller. Age differences in patient recovery patterns following coronary artery bypass surgery. *American Journal of Critical Care*, 2(6):453–461, 1993.

[29] Jacqueline Yewande Thompson, Christopher Byrne, Mark A Williams, David J Keene, Micheal Maia Schlussel, and Sarah E Lamb. Prognostic factors for recovery following acute lateral ankle ligament sprain: a systematic review. *BMC musculoskeletal disorders*, 18(1):421, 2017.

[30] Hagai Yanai, Arie Budovsky, Robi Tacutu, and Vadim E Fraifeld. Is rate of skin wound healing associated with aging or longevity phenotype? *Biogerontology*, 12(6):591–597, 2011.

[31] Linpei Jia, Weiguang Zhang, Rufu Jia, Hongliang Zhang, and Xiangmei Chen. Construction formula of biological age using the principal component analysis. *BioMed research international*, 2016, 2016.

[32] Hilary A Tindle, Meredith Stevenson Duncan, Robert A Greevy, Ramachandran S Vasan, Suman Kundu, Pierre P Massion, and Matthew S Freiberg. Lifetime smoking history and risk of lung cancer: Results from the framingham heart study. *JNCI: Journal of the National Cancer Institute*, 110(11):1201–1207, 2018.

[33] Bernard L Strehler and Albert S Mildvan. General theory of mortality and aging. *Science See Saiensu*, 132, 1960.

[34] Arnold Mitnitski, Xiaowei Song, and Kenneth Rockwood. Assessing biological aging: the origin of deficit accumulation. *Biogerontology*, 14(6):709–717, 2013.

[35] Arnold Mitnitski and Kenneth Rockwood. Aging as a process of deficit accumulation: its utility and origin. In *Aging and Health-A Systems Biology Perspective*, volume 40, pages 85–98. Karger Publishers, 2015.

[36] Stacy L Andersen, Paola Sebastiani, Daniel A Dworkis, Lori Feldman, and Thomas T Perls. Health span approximates life span among many supercentenarians: compression of morbidity at the approximate limit of life span. *Journals of Gerontology Series A: Biomedical Sciences and Medical Sciences*, 67(4):395–405, 2012.

SUPPLEMENTARY INFORMATION

Complete Blood Count Datasets

NHANES CBC data were retrieved from the category "Complete Blood Count with 5-part Differential - Whole Blood" for NHANES surveys 1999 − 2010. Corresponding UKB CBC data fields with related database codes are listed in Table S1. Samples with missing (or filled with zero) data for any of the used CBC components were discarded. Differential white blood cell percentages were converted to cell counts by multiplication by $0.01 \times$ WBC. Neutrophils (NEU) data field was calculated as $0.01 \times \text{WBC} \times (100\% - \text{MONO} - \text{LYM} - \text{EOS})$ (thus actually included neutrophils and basophils). After these calculations, all CBC parameters were log-transformed.

Hazards model

The Cox proportional hazards model was trained using NHANES 2015 Public-Use Linked Mortality data. After filtering out samples with missing CBC data, this yielded a NHANES dataset slice of 40592 participants aged $18 - 85$ y.o. Cox model was trained based on data of participants aged $40 - 85$ y.o. (11731 male and 12076 female) with 3792 recorded death events during follow-up until the year 2015 ($1999 - 2014$ surveys). CBC components and the biological sex label were used as covariates. The model was well-predictive of all-cause mortality and yielded a concordance index value of CI = 0.76 in NHANES and CI = 0.72 in UKB (samples collected $2007 - 2011$, 216250 male and 255223 female participants aged $39 - 75$ y.o., 13162 recorded death events during follow-up until the year 2016). The Cox proportional hazards model was used as implemented in lifelines package (version 0.14.6) in python. The parameters of model fitting are given in Table S2. The model was then applied to calculate the hazards ratio for all samples in the GEROLONG, UKB and NHANES cohorts (including individuals younger than 40 y.o.).

The most prevalent chronic diseases and health status

We quantified the health status of individuals using the sum of major age-related medical conditions that they were diagnosed with, which we termed the compound morbidity index, CMI. The CMI is similar in spirit to the frailty index suggested for NHANES [17]. We were not able to use the frailty index because it was based on Questionnaire and Examination data that were not consistent between all NHANES surveys. Also, we did not have enough corresponding data for the UKB dataset. For CMI determination, we followed [36] and selected the top 11 morbidities strongly associated with age after the age of 40. The list of health conditions included cancer (any kind), cardiovascular conditions (angina pectoris, coronary heart disease, heart attack, heart failure, stroke, or hypertension), diabetes, arthritis and emphysema. Notably, we did not include dementia in the list of diseases since it occurs late in life and hence is severely underrepresented in the UKB cohort due to its limited age range. We categorized participants who had more than 6 of those conditions as the "most frail" (CMI > 0.6), and those with CMI < 0.1 as the "non-frail". NHANES data for diagnosis with a health condition and age at diagnosis is available in the questionnaire category "Medical Conditions" (MCQ). Data on diabetes and hypertension was retrieved additionally from questionnaire categories "Diabetes" (DIQ) and "Blood Pressure & Cholesterol" (BPQ), respectively.

UK Biobank does not provide aggregated data on these medical conditions. Rather, it provides self-reported questionnaire data (UKB, Category 100074) and diagnoses made during hospital in-patient stay according to ICD10 codes (UKB, Category 2002). We aggregated self-reported and ICD10 (block level) data to match that of NHANES for transferability of the results between populations and datasets. We used the following ICD10 codes to cover the health conditions in UK Biobank: hypertension (I10-I15), arthritis (M00-M25), cancer (C00-C99), diabetes (E10-E14), coronary heart disease (I20-I25), myocardial infarction (I21, I22), angina pectoris (I20), stroke (I60-I64), emphysema (J43, J44), and congestive heart failure (I50).

Analysis of physiological state fluctuations

The data presented in this manuscript suggests that DMI fluctuations are governed by external factors as well as intrinsic forces that can be restorative or disruptive. The fluctuations in DMI are fast compared to the rate of aging changes and, therefore, the dynamics of the organism state can be described by the stochastic Langevine equation

$$\dot{x}(t) = -\varepsilon x(t) + f(t), \tag{1}$$

here $\varepsilon > 0$, and $x(t) = \text{DMI}(t) - \text{DMI}_0$ is the deviation of the DMI from its equilibrium value DMI_0, determined by the individual organism, environment and life history properties. The recovery rate ε is proportional to inverse equilibration time and characterizes deterministic forces responsible for the organism state maintenance. The inevitable stochastic factors, $f(t)$, stand for the effects of endogenous and external stress, are assumed zero-mean and uncorrelated, $\langle f(t)f(t')\rangle = B\delta(t - t')$ with B being the power of the stochastic stress ($\langle ... \rangle$ stands for the averaging over realization of the random process).

The stationary solution of Eq. (1) is a random function with zero mean and variance $\sigma^2 = \langle \delta x^2 \rangle = B/\varepsilon$. That is why we choose to plot the recovery rate from the

auto-correlation function and the inverse DMI fluctuations variance $1/\sigma^2$ and in Figs. 3A and 3C, respectively.

To obtain an estimate for the recovery rate ε, we used the normalized auto-correlation function of the time lag Δt between the measurements,

$$C(\Delta t) = \left\langle \Big(x(t+\Delta t) - x(t) \Big)^2 \right\rangle_t \sigma^{-2} = \Big(1 - e^{-\varepsilon \Delta t} \Big).$$

Here $\langle ... \rangle_t$ stands for the everaging along the individual trajectories.

TABLE S1: CBC data used in the study.

CBC component	Abbreviation	NHANES	UKB
Hemoglobin	HB	LBXHGB	Haemoglobin concentration (30020)
Red blood cell count	RBC	LBXRBCSI	Red blood cell (erythrocyte) count (20010)
Mean corpuscular volume	MCV	LBXMCVSI	Mean corpuscular volume (30040)
Mean corpuscular hemoglobin concentration	MCHC	LBXMC	Mean corpuscular haemoglobin concentration (30060)
Red blood cell distribution width	RDW	LBXRDW	Red blood cell (erythrocyte) distribution width (30070)
Platelets	PLT	LBXPLTSI	Platelet count (30080)
Monocytes, %	MONO	LBXMOPCT	Monocyte percentage (30190)
Lymphocytes, %	LYM	LBXLYPCT	Lymphocyte percentage (30180)
Eosinophils, %	EOS	LBXEOPCT	Eosinophill percentage (30210)
White blood cell count	WBC	LBXWBCSI	White blood cell (leukocyte) count (30000)

TABLE S2: NHANES Cox proportional mortality hazards model parameters.

Covariate	HR (95% CI)	p-value (multivariate)	
Haemoglobin	0.81 (0.59 - 1.11)	$p = 0.51$	
Red blood cell count	0.99 (0.72 - 1.36)	$p = 0.98$	
Mean corpuscular volume	1.50 (1.21 - 1.86)	$p = 0.058$	
Mean corpuscular haemoglobin concentration	0.88 (0.80 - 0.96)	$p = 0.14$	
Red blood cell distribution width	1.37 (1.35 - 1.38)	$p = 8.4E - 122$	***
Platelets	0.86 (0.84 - 0.87)	$p = 3.6E - 26$	***
White blood cell count	1.27 (1.25 - 1.29)	$p = 8.3E - 39$	***
Monocytes, %	1.20 (1.18 - 1.23)	$p = 1E - 22$	***
Lymphocytes, %	0.79 (0.77 - 0.80)	$p = 2.5E - 45$	***
Eosinophils, %	1.07 (1.05 - 1.09)	$p = 6.1E - 05$	***
Sex	1.20 (1.18 - 1.22)	$p = 6.2E - 28$	***

TABLE S3: Significance of prediction of acquiring a health condition based on estimated log hazards ratio (adjusted for age and gender). Only UKB subjects with none of the listed health conditions at the time of survey were considered; the total number of subjects evaluated for each condition was 263956. The numbers in parentheses in the far right column indicate the occurrence of the disease being the first diagnosis in an individual.

Condition	HR (95% CI)	p-value	n_{events} ($n_{is\ first\ morbidity}$)
Death	1.35 (1.33 - 1.37)	4.9E-110	4745 (927)
First morbidity	1.05 (1.05 - 1.06)	3E-40	68126 (68126)
Hypertension	1.04 (1.04 - 1.05)	1.2E-13	31143 (25681)
Arthritis	1.07 (1.07 - 1.08)	1.7E-32	28745 (24451)
cancers	1.03 (1.02 - 1.04)	2.7E-05	18838 (15860)
Coronary heart disease	1.05 (1.04 - 1.06)	2.9E-05	7422 (5500)
Diabetes	1.03 (1.02 - 1.05)	0.0084	6605 (5265)
Angina pectoris	1.02 (1.00 - 1.03)	0.35	3747 (2164)
Emphysema	1.48 (1.45 - 1.51)	4.5E-108	2382 (1508)
Heart attack	1.05 (1.03 - 1.08)	0.012	2186 (1605)
Stroke	1.10 (1.08 - 1.13)	4.1E-05	1686 (1168)
Heart failure	1.32 (1.29 - 1.36)	6.2E-26	1209 (583)
Bronchitis	1.18 (1.12 - 1.25)	0.0034	280 (177)

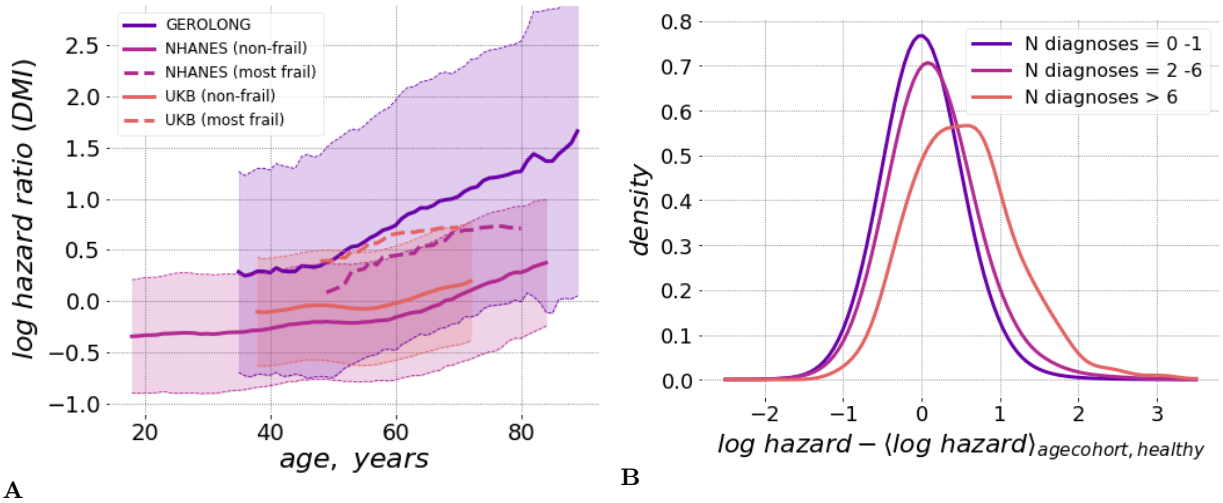

FIG. S1: **A.** DMI mean values (lines) and variance (shaded areas) are plotted relative to age for the NHANES (same as in Fig. 1B), UKB and GEROLONG datasets (color-matched with respect to each study). For NHANES and UKB the solid line and shaded regions mark the population average ad the range spanned by one standard deviation from it for the "non-frail" (CMI < 0.1) participats. The population mean for the "most frail" (CMI > 0.6) individuals is shown with dashed lines. **B.** Distributions of sex- and age-adjusted DMI in cohorts of UKB participants in different morbidity categories relative to the DMI mean in cohorts of "non-frail" (one or no diagnoses, CMI < 0.1) individuals. Note that the distribution function in the "most frail" group (more than 6 diagnoses, CMI > 0.6) exhibited the largest shift and a profound deviation from the symmetric form, similarly as it was seen in NHANES.