Supercentenarian and remarkable age records exhibit patterns indicative of clerical errors and pension fraud

Saul Justin Newman*[1,2]

[1]Leverhulme Center for Demographic Science, Nuffield Department of Population Health, University of Oxford
[2]University College, University of Oxford
ORCID 0000-0001-9841-1518
*Correspondence to: saul.j.newman@gmail.com

Abstract

The observation of individuals attaining remarkable ages, and their concentration into geographic sub-regions or 'blue zones', has generated considerable scientific interest. Proposed drivers of remarkable longevity include high vegetable intake, strong social connections, and genetic markers. Here, we reveal new predictors of remarkable longevity and 'supercentenarian' status. In the United States, supercentenarian status is predicted by the absence of vital registration. The state-specific introduction of birth certificates is associated with a 69-82% fall in the number of supercentenarian records. In Italy, England, and France, which have more uniform vital registration, remarkable longevity is instead predicted by poverty, low per capita incomes, shorter life expectancy, higher crime rates, worse health, higher deprivation, fewer 90+ year olds, and residence in remote, overseas, and colonial territories. In England and France, higher old-age poverty rates alone predict more than half of the regional variation in attaining a remarkable age. Only 18% of 'exhaustively' validated supercentenarians have a birth certificate, falling to zero percent in the USA, and supercentenarian birthdates are concentrated on days divisible by five: a pattern indicative of widespread fraud and error. Finally, the designated 'blue zones' of Sardinia, Okinawa, and Ikaria corresponded to regions with low incomes, low literacy, high crime rate and short life expectancy relative to their national average. As such, relative poverty and short lifespan constitute unexpected predictors of centenarian and supercentenarian status and support a primary role of fraud and error in generating remarkable human age records.

Introduction

The concentration of remarkable-aged individuals within geographic regions or 'blue zones'[1] has stimulated diverse efforts to understand factors driving survival patterns in these populations[2,3]. Both the overall population residing within these regions, and the individuals exceeding remarkable age cut-offs, have been subject to extensive analysis of lifestyle patterns[2,4–6], social connections[3,7], biomarkers[8,9], and genomic variants[10], under the assumption that these represent potential drivers behind the attainment of remarkable age.

However, alternative explanations for the distribution of remarkable age records appear to have been overlooked. Previous work has noted the potential of population illiteracy[11] or heterogeneity[12] to explain remarkable age patterns. Other investigations have revealed the potential role of errors[13–16], bad data[17], and operator biases[18] in generating old-age data and survival patterns. In turn, these findings prompted a response with potentially disruptive implications: that, under such models, the majority if not all remarkable age records may be errors[19].

There is a theoretical reason to expect that most or all old-age records are errors[13,15]. Consider, for example, a population of fifty-year-olds into which we introduce a vanishingly low rate of random, undetectable age-coding errors. These rare errors involve taking a forty-year-old, and making their documents state they are now fifty years old.

The 40-year-old 'young liar' errors will have over double the annual survival rate of the (error-free) 50-year-old population. They are, after all, biologically 10 years younger. As the 'young liar' errors are more likely to survive than the error-free data, age-coding errors constitute an exponentially larger fraction of the population over time. Eventually, this exponential growth overtakes the population and, even from a vanishingly low baseline error rates of (say) 0.001%[15], the entire population becomes 'young liar' errors at advanced ages[15]. Of course, the baseline error rate in most populations is not vanishingly low. The USA, which has more supercentenarian records than any other country, exhibits historical error rates of 17%-66%[20–23], even in the general population[20].

Here, we explore the possibility that extreme old-age data are dominated by age-coding errors by linking civil registration rates and indicators of poverty to per-capita estimates of remarkable age attainment, obtained from central population registries and validated supercentenarian databases, across the USA, France, Japan, England, and Italy.

These data reveal that remarkable age attainment is predicted by indicators of error and fraud, including illiteracy, poverty, high crime rates, short average lifespans, and the absence of birth certificates. As a result, these findings raise serious questions about the validity of an extensive body of research based on the remarkable reported ages of populations and individuals.

Such concerns were sufficiently serious that an extensive review was conducted, revealing evidence that brings into question key findings, widespread research practices, and the basic level of integrity in the field (Supplementary Materials).

Methods

The number and birthplace of all validated supercentenarians (individuals attaining 110 years of age) and semisupercentenarians (SSCs; individuals attaining 105 years of age) were downloaded from the Gerontology Research Group or GRG supercentenarian table[24] (updated 2017) and the International Database on Longevity or IDL [25]. These data were aggregated by subnational units for birth locations, which were provided for the IDL data, and obtained through biographical research for the GRG data. Populations were excluded due to incomplete subnational birthplace records (<25% complete) or countries with an insufficient number of provinces to fit spatial regressions (<15 total provinces), leaving population data on SSCs and supercentenarians in the USA, France, and England (Fig 1).

To quantify the distribution of remarkable-aged individuals in Italy, province-specific quinquennial life tables were downloaded from the Italian Istituto Nazionale di Statistica Elders.Stat database [26] to obtain age-specific cohort survivorship data (Fig 1c,f; S1 Code). Using contemporary data across Italian provinces, probabilities of survival (l_x) to ages 90-115, and life expectancy at age 100 were fit as dependent variables, and survival rates at age 55 and life expectancy at age 55 as independent variables, using a simple linear mixed model regression (S1 Code).

Japanese centenarian data were downloaded from the Japanese Ministry of Health, Labour, and Welfare[27] through the Statistics Japan portal[28] for all 47 prefectures (Fig S1), alongside data on prefectural income per capita (in 2011 yen), employment rates, age structure, survivorship, and a financial strength index, for 2010: the most complete recent year available for these data (S1 Code; Fig S1). These data were also linked to the most recent available prefecture-specific poverty rates[29].

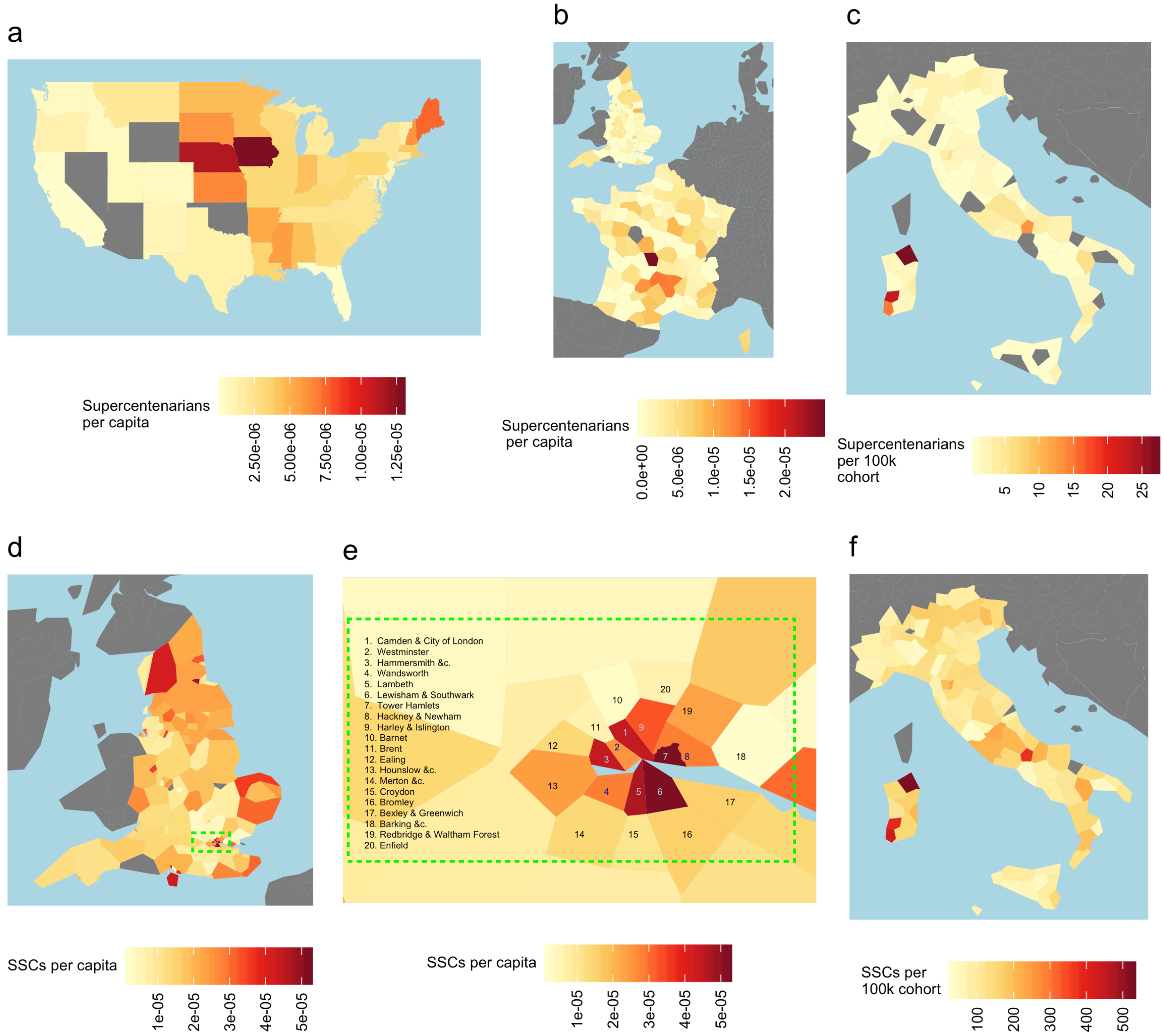

Figure 1. Regional distributions for the density of the oldest-old. A substantial majority of the oldest-old people are concentrated in a few countries exhibiting large regional variation in density of the oldest-old. Known supercentenarians are concentrated in the USA (a), with large numbers in France(b), England (b), and Italy (c; life table estimated rates). Marked variation exists in estimated SSC densities across England (d) – including (e) the 19-fold differences in SSC abundance across London (green box in d-e) between Tower Hamlets (region 7; most) and Barnet (region 10; least) – and across Italian provinces (f).

Supercentenarians recorded in the GRG database and born in the USA were matched to the 1900 census counts for state and territory populations[30], and linked to the National Center for Health Statistics estimates for the timing of complete birth and death certificate coverage in each US state and territory[31]. Both the number of supercentenarian births overall, and estimates of supercentenarians per capita, approximated by dividing supercentenarian number by state population size in the 1900 US census[30], were averaged across the USA and represented as discontinuity time series relative to the onset of complete-area birth registration (S1 Code).

To capture the geographic distribution of French supercentenarians, all 175 supercentenarians in the GRG database who were either born or deceased in France were linked to the smallest discoverable region of birth using biographical searches[24]. In addition, de-identified records in the IDL were already linked to birth locations encoded by the Nomenclature for Territorial Units level 3 codes (NUTS-3), which divide France into 101 regions [25]. These modern regions were linked manually to their corresponding Savoyard-era department to obtain historic region-specific estimates of life expectancy at birth[32] for the birth year and location of all supercentenarians in metropolitan France. For each supercentenarian, life expectancy at birth was then measured relative to the contemporary average life expectancy of metropolitan France (S1 Code).

The number of total supercentenarians and SSCs born into Eurostat NUTS-3 coded regions, either documented for French and English regions in the IDL or estimated for Italian regional cohorts by ISTAT, were linked to modern socioeconomic indicators available at this administrative level: total regional gross domestic product (GDP), GDP per capita, GDP per capita adjusted for purchase power scores (PPS), murder and employment rates per capita, and the number of 90+ year-olds, using the EUROSTAT regional database[33] (S1 Code).

Across England, additional data were obtained for the Index of Multiple Deprivation or IMD: a national metric used to indicate relative levels of deprivation, including income deprivation in people aged 60+, by the UK Office of National Statistics[34]. The IMD data are measured in 317 local authority districts, each of which is a subset of a single Eurostat NUTS-3 encoded region.

To capture the relative degree of deprivation within England, the IMD and its component scores were averaged within each of the 175 NUTS-3 regions (S1 Code).

Similar estimates of deprivation were obtained for French NUTS-3 regions, by downloading the regional poverty rates and poverty rates in the oldest available age group, ages 75 and over from the French National Institute of Statistics and Economic Studies INSEE [35].

To overcome the three orders of magnitude differences in population size across subnational geographic units, the number of centenarians, SSCs and supercentenarians were adjusted to per capita rates. However, the 'correct' adjustment for per capita rates of remarkable longevity is dependent on the *a priori* assumptions of their cause. For example, if the null hypothesis was that all supercentenarians are 'real', adjustment for birth cohort size 110+ years previously would be a more correct method for best predicting the population density of supercentenarians. However, if the null hypothesis is that supercentenarians are more frequently modern-era pension frauds or clerical mistakes, per capita correction for a birth cohort 110 years in the past is of uncertain value for predicting modern events. In this latter case, the occurrence of supercentenarians would be better and more accurately predicted by correcting for modern population sizes. This concern is mitigated by the non-migration of most SSCs and supercentenarians with age -- where data is available, two thirds died in their region of birth -- and the spatial stability of poverty patterns over time[36].

Historical rates were used whenever possible, favoring the 'real supercentenarians' assumption. Per capita rates of remarkable age attainment, calculated relative to the size of historical birth cohorts, were downloaded from the respective government statistical bureaus of Japan and Italy [26,27]. Due to the absence of birth certificates, USA supercentenarian data from the GRG were corrected to historical per capita rates based on population data in the 1900 US census[30]. However, individuals in France and England were located into geographic units that have only existed since 2003. As a result, there were no data on historical population sizes available for these geographic units. It was therefore necessary to estimate per capita rates using modern population sizes surveyed at the NUTS-3 geographic level within France and England.

To address this unavoidable difference in per capita rate calculations the number of the centenarians, SSCs, and supercentenarians were also corrected relative to the number of old-age residents in each modern geographic unit of Japan, England, and France (S1 Code). This adjustment was less susceptible to migration or large longitudinal shifts in population size, and better reflected the density of older people in modern geographic units after survival and migration processes. However, the insufficient granularity of birth cohorts within England and the considerable rearrangement of geographic units within France remains an important constraint on the upper accuracy of these models.

Collective socioeconomic indicators obtained for each country were used to develop linear mixed models across all regions at least one case, for centenarian in Japan, SSCs in Italy and England and supercentenarians in England and France (S1 Code), to predict the regional per capita and per 90+ year old density of the oldest available populations in each country. Linear mixed models were fit using either the population poverty rate (England, France, and Japan) or estimates of old-age poverty rates (percent in poverty over 75 in France, the IDOP index in England) as the single predictor variable, and the number of centenarians, SSCs and supercentenarians both per capita and per 90+ year old. These models were then extended by fitting, as interactive effects, basic socioeconomic indicators used as global indicators of health and deprivation available at a sufficient geographic level (S2 Code). Such models focused on capturing fundamental indicators, representing crime rates, health, and income, available at the NUTS-3 regional level in the EU and the prefectural level in Japan.

Where available, French supercentenarians were linked to regional estimates of life expectancy at birth, calculated quinquennially for each of the Savoyard-era departmental boundaries of France into which they were born[32]. These local rates were then corrected relative to the contemporary French national average life expectancy at birth to yield the relative life expectancy at birth, in years[32]. For example, Jeanne Calment was born in the Alpes-Maritime department in 1875, when average life expectancy at birth was just 33.4 years and the contemporary national French average life expectancy was 37.8 years: a relative regional life expectancy of −4.4 years. These rates were used to measure if regional life expectancy at birth was significantly higher or lower than the French national average, using a one sample t-test, for each supercentenarian.

To explore the potential for age manufacture amongst remarkable age records, birthdate data were aggregated within the GRG and IDL databases. Enrichment for specific birth days is usually indicative of nonrandom age selection due to fraud, error, and clerical uncertainty. This check, however, is limited in that it cannot detect diverse sources of error, such as identity fraud or failed death registrations, which retain a representative distribution of birth days.

As population representative birthdates were unavailable within the target populations, the distribution of births was tabulated by days of the month to remove the often poorly-categorized or undocumented effects of birth seasonality. This distribution was compared to both modern birthdate distributions from seventy million births in the US, which suffer distortion from the heaping of elective births and caesarean sections on certain dates, and to the distribution of birthdays under a uniform distribution of births.

To facilitate reproduction of these findings, all shareable data and code are available in a single structured file, with instructions and links for the non-shareable data, in S1 Data.

Results

Between the 1880 and 1900 census, a period covering 79% of US supercentenarian births, the US population increased by 150% and average life expectancy by 20%[30,37]. The introduction of complete vital registration in the USA coincided with this rapid increase in lifespan and population size and was expected to result in a large increase in the number of supercentenarian records per capita.

Instead, the introduction of state-wide birth certification coincides with a sharp reduction in the number of supercentenarians. In total, 82% of supercentenarian records from the USA predate state-wide birth certification. Forty-two states achieved complete birth certificate coverage during the survey period. When these states transition to state-wide birth registration, the number of supercentenarians falls by 80% per year (Fig 2a) and 69% per capita (Fig 2b).

The introduction of birth certificates in Italy largely predates the onset of supercentenarian records. Instead, the attainment of remarkable age in Italy is predicted by a short average lifespan. Within Italian cohorts, life expectancy at age 55 is positively correlated with life expectancy at all ages at all ages from 60 to 95. However, better early- and mid-life survival is predictive of worse mortality rates after age 95 ($r = 0.15$; $p = 0.1$; Fig S2a). Cohort survival to age 55 is increasingly negatively correlated with survival to ages 100 ($r = -0.17$; $p = 0.06$; Fig S2b), 105 ($r = -0.38$; $p = 0.00003$; Fig S2c) and 110 years ($r = -0.45$; $p < 0.000001$; Fig S2d), and with life expectancy at age 100 ($r = -0.4$; $p=0.00001$). Furthermore, individuals across Italy are more likely to attain supercentenarian ages if their province has higher unemployment rates (Fig S3a), fewer people above age 90 (Fig S3b), and a worse economy (Fig S3c).

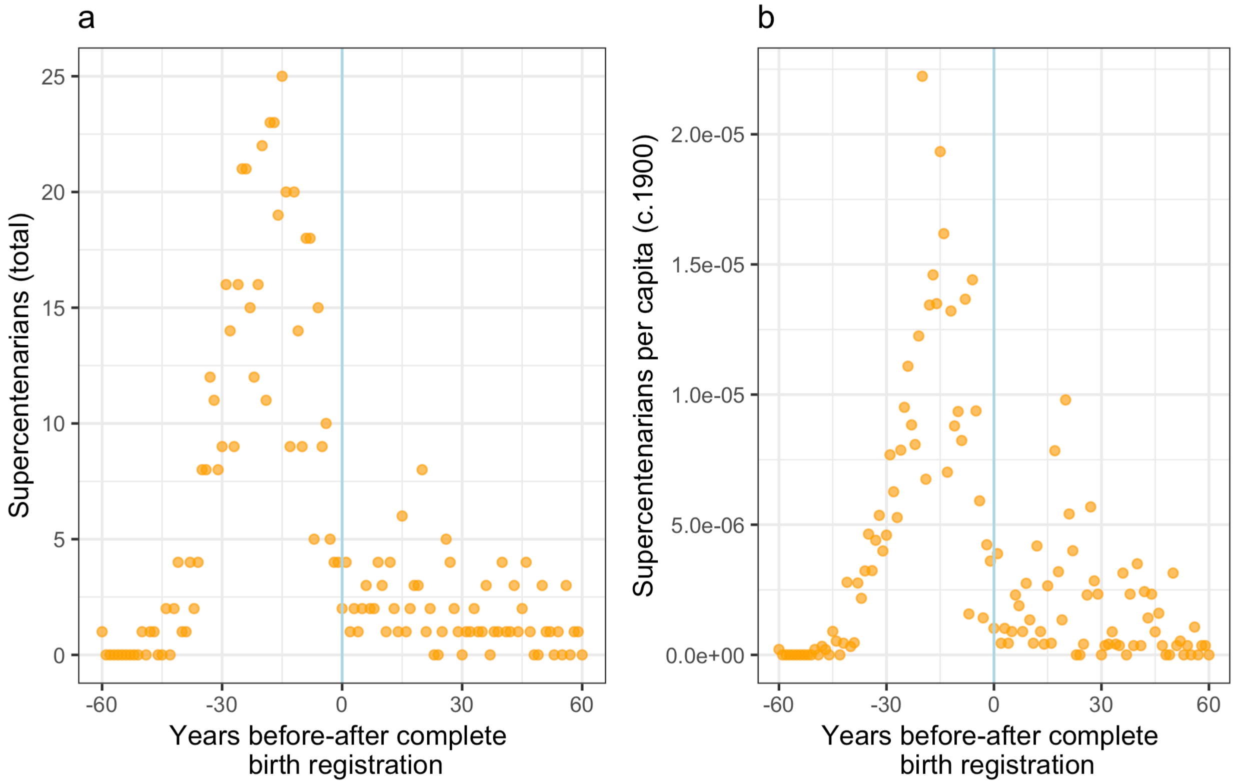

Figure 2. Number and per capita rate of attaining supercentenarian status across US states, relative to the introduction of complete-area birth registration. Despite the combined effects of rapid population growth and increasing life expectancy during this period the total number of US supercentenarians (a) falls dramatically after birth certificates achieve state-wide coverage (vertical blue line). This trend remains after adjusting for total population size c.1900 (b) within each state.

Findings from the Italian data support the hypothesis that centenarians, SSCs, and supercentenarians largely constitute a collection of errors and pension fraud[19]. For example, of Italy's 114 provinces Medio Campidano ranked second for supercentenarians, third for SSCs and fourth for centenarians per capita. However, the province also had the second-lowest employment rate of all Italian provinces and the single lowest GDP per capita (both unadjusted and PPS-adjusted; Table S1). Individuals born in Medio Campidano were the second-least likely to survive from birth to age 55, after the Blue Zone province of Ogliastra. By this metric, the province of Olbia-Tempio was the seventh-worst province for survival to age 55, and according to Eurostat had the eighth-fewest individuals alive over the age of 90, yet somehow also ranked as the best province for survival to ages 100,105 and 110 (Table S1).

French supercentenarians are over-represented in the overseas departments, former colonial holdings, and Corsica (which is included in metropolitan France; Table S2): regions that historically constitute some of the most neglected, least well-documented, and shortest-lived administrative regions of France. At the first reliable estimate of population size in 1950 these departments contained around 1.7% of French citizens. However, at least 11% (N=16) of the French supercentenarians in the GRG database originate there: a 6.5-fold over-representation. This number increases when integrating the deidentified IDL data, which only includes regions monitored by Eurostat (Réunion, Guadeloupe, Martinique and Corsica), to establish the minimum numbers of supercentenarians born in each region. Once these data are included, at a minimum the overseas and former colonial regions of France contain twenty-four (15.5%) of a total 155 supercentenarians: Guadeloupe and Martinique each contain eight, French Algeria four, and one each in French Guiana, Saint Barthélemy, Réunion, and New Caledonia.

When IDL and GRG data are combined, the top six (NUTS-2 level) regions of France by supercentenarians per capita are, respectively, Saint Barthélemy, Martinique, Guadeloupe, Corsica, New Caledonia, and French Guiana. These six regions contain eight times as many supercentenarians per capita (1.3 per 100,000), and more supercentenarians overall (N=21), than the region of Île-de-France (0.16 per 100,000; N=19). This is despite Île-de-France receiving more than double the per capita income, being the longest-lived region in mainland France, and containing seven times as many citizens including huge numbers of (non-supercentenarian) Guadeloupe- and Martinique-born internal migrants[38].

Within these regions, Martinique has the second-highest poverty rate both overall (29%) and for people aged 75 and over (31%) in all the 101 NUTS-3 coded provinces. Of the ranked regions only Réunion has higher poverty rates, with 39% below the poverty line. Amongst the unranked regions, French Guyana has a 44% unemployment rate, with 44% of households are listed as 'poor', and Guadeloupe a 24% unemployment rate[39]. The two departments of Corsica have the highest and second-murder rate, and the third and fifth highest poverty rates, of the 101 French departments.

When compared across the 101 modern NUTS-3 coded regions of France, Guadeloupe and Martinique are both equal second for total supercentenarians after Paris (Table S2), with at least eight supercentenarians each. However, these two provinces contained only 190,000 French citizens each c.1900: by comparison the Paris department contained 2.7 million people in 1901. The two small, remote departments of Guadeloupe and Martinique produced the same number of supercentenarians as the Nord department, despite the latter containing ten times as many citizens in 1901.

When measuring supercentenarians per capita using modern population sizes, due to incomplete data, Martinique and Guadeloupe ranked second and third for supercentenarians per capita out of the 101 NUTS-3 regions of France. Creuse ranked as having the highest percentage of oldest-old individuals in France. Creuse also ranks as having the fourth-highest poverty rate in older people, the 16th-worst poverty rate overall, and the fourth-lowest GDP per capita (both raw numbers and PPS adjusted; Table S2) of the 101 regions of France. Per capita rates in Creuse may have also been impacted by massive out-migration, with the Creuse population falling by 60% since 1901.

Region-specific estimates of life expectancy were complicated by large-scale historical changes in province number, boundaries, and even nations. From 1900 the number of French regions fell from 88 to eighteen, with just twelve mainland regions. Outside the overseas regions and Corsica, whose boundaries are largely unchanged, few of the GRG supercentenarians could be located within the Savoyard-era department boundaries required to gain an accurate estimate of life expectancy at birth. Most records could only be resolved to the 12 modern administrative regions, which contain up to thirteen Napoleonic/Savoyard era provinces each, resulting in a wide range of possible life expectancies and socioeconomic data for each individual. Combined with the initially small number of supercentenarians in France, this coarse resolution excluded a more detailed analysis of the GRG data.

There were notable disparities in the abundance of the oldest-old across England (Table S3). For example, the region with the most 90+ year-olds per capita[34], West Sussex, had only five SSCs in a population of 460,000. In contrast, in Tower Hamlets, the region with the least 90+ year olds in the country[34], records fifteen SSCs amongst 289,000 residents.

As a result, Tower Hamlets has the most SSCs per capita in England. Tower Hamlets also has the highest poverty rate[40], highest child poverty rate[40], highest income inequality[40], the shortest disability-free life expectancy by ten to 15 years[40], and the worst index of multiple deprivation[34] of the 32 London boroughs. According to the Income Deprivation Affecting Older People Index (or IDOP), which captures the financial stress and deprivation of older people relative to the national average, Tower Hamlets is the single most income-deprived population of older people across all 317 local authority districts[34]. Tower Hamlets has the smallest percentage of people aged 90 and over[34]: a notable discrepancy for a region that, according to the IDL figures, has the largest percentage of SSCs.

Of the 131 regions included in analysis, the top 20 regions for SSCs per capita contained the 1st and 3rd-8th most income-deprived regions for older people in England (Table S3). As such, high rankings for income deprivation in older people predicts higher rates of achieving ages 105 and over. Out of 175 regions in England, for example, Southwark and Lewisham ranked as the eighth most income-deprived district for older people[34], second for SSCs per capita, and first for SSCs overall (Table S3). Central Manchester produced 18 SSCs overall (equal sixth) and ranks 14th for SSCs per capita yet is the third most income-deprived district for older people, and has the highest crime index, third-worst population health index, fourth-worst index of multiple deprivation, and sixth smallest percentage of 90+ year old people of any region (Table S3). Such rankings are not recent, with inner-city Central Manchester the second most persistently deprived of all 317 local-authority districts[34].These specific cases reflect broader patterns across England, where regions with the highest number and per capita rate of SSCs also have lower incomes, worse health, more deprivation, and higher crime rates (Fig 3).

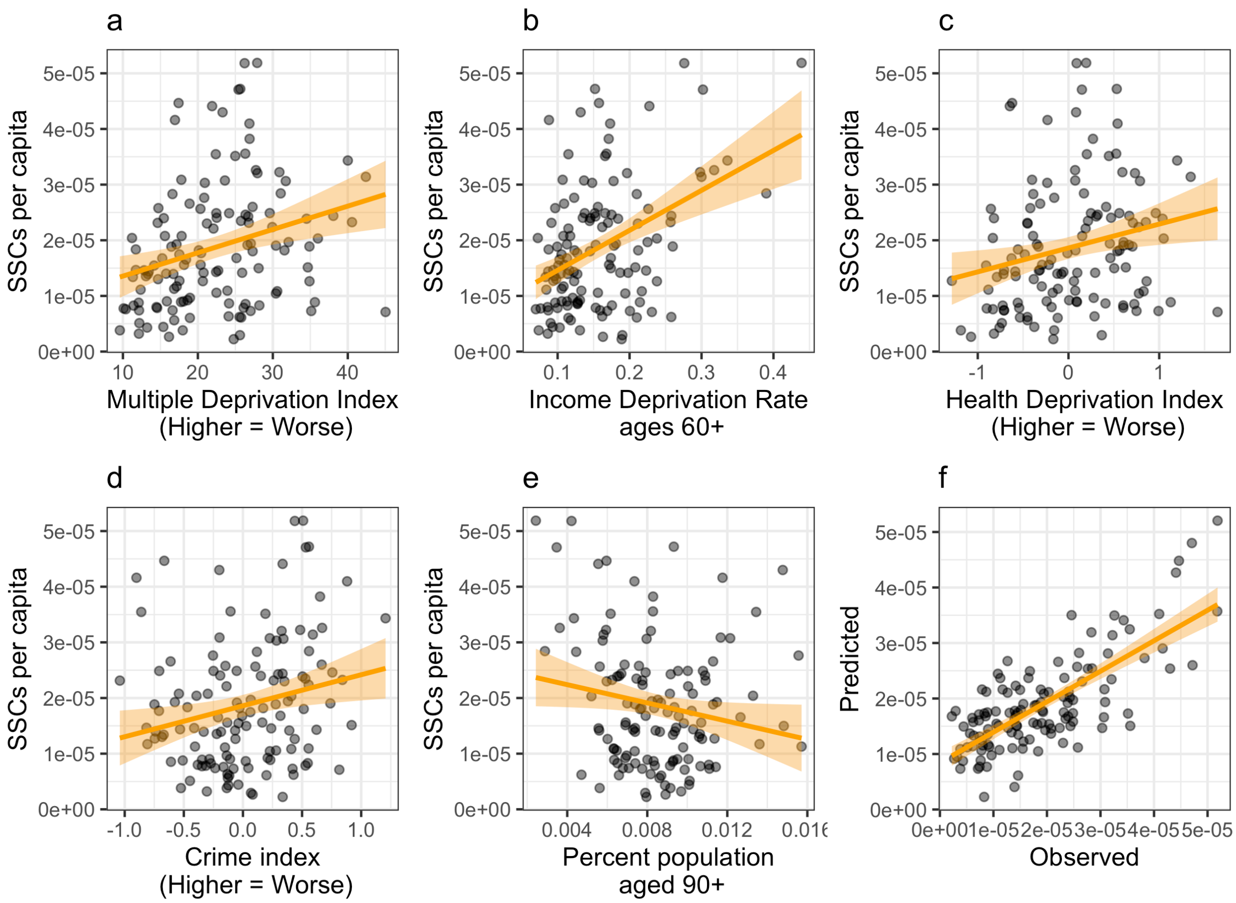

Figure 3. Social hardship and the distribution of the oldest-old in England. Across 128 regions containing at least one semisupercentenarian, the number of SSCs per capita is highest in regions where: (a) people are more deprived overall (r = 0.28; p = 0.001), (b) people over age 60 are more income-deprived (r = 0.42; p < 0.000001), (c) multidimensional health indices are worse (r = 0.23; p =0.01), and (d) crime rates are higher (r = 0.22; p= 0.01). In contrast, regions with a higher fraction of people surviving past 90 years of age (e), have significantly fewer SSCs per capita (r = -0.18, p = 0.04). Integrated into a simple linear mixed model (f), these five variables plus PPS-adjusted GDP and employment rates, collectively account for half of the regional differences in SSC abundance (adjusted R^2 = 0.39; p = 1e-06).

Higher poverty is a particularly good predictor of higher survival to extreme ages. For example, a single-variable linear model using only (general or old-age) poverty as a predictor captures a substantial amount of the variation in centenarian, SSC, and supercentenarian abundance. Higher poverty and deprivation are positively correlated with all age categories of the oldest-old in every country measured (Fig 3; Fig 4a-f).

The IDOP ranking is an accurate indicator of the distribution of income deprivation in older people across England. When aggregated into NUTS-3 regions containing at least one SSC, IDOP scores are positively correlated with SSC abundance per capita ($r = 0.42$; $p = 0.0000009$; Fig 4a). That is, higher income deprivation for residents over 60 predicts *higher* numbers of people surviving past age 105 (Fig 3b). This predictive accuracy improves markedly when predicting the number of SSCs per 90+ year old resident, rather than the number of SSCs per capita ($r = 0.70$; $p < 1x10^{-15}$; Fig 4d).

Like English data, the percentage of French supercentenarians in a population was predicted by higher rates of poverty in older people ($r = 0.42$; $p = 0.0004$; Fig 4b). Again, these associations were stronger when predicting the fraction of 90+ year-olds who lived over 110: across the 66 regions of France with sufficient data, the poverty rate over age 75 predicted half of the variation in supercentenarian abundance ($r = 0.51$, $p = 0.00001$; Fig 4e).

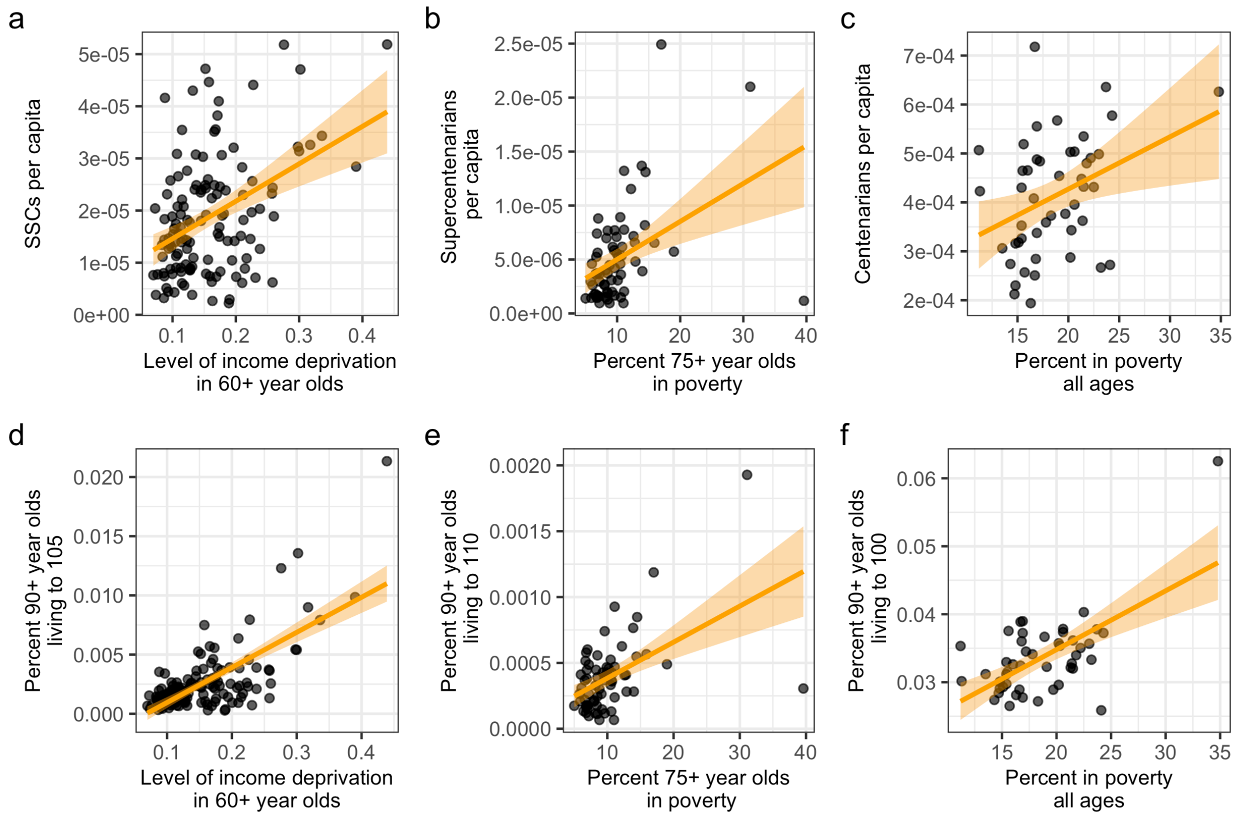

Figure 4. Old-age poverty and the density of the oldest-old. A higher percentage of people are centenarians, SSCs and supercentenarians in high-poverty and income-deprived regions of rich countries. Metrics of poverty and old-age poverty are positively correlated with the density of the oldest-old per capita across subnational data in England (a; r = 0.42; p < 0.000001; 127 regions), France (b; r = 0.42; p = 0.0004; 66 departments; r = 0.58 in the 63 mainland departments), and Japan (c; r = 0.36; p = 0.01; 47 prefectures). These relationships strengthen if density of the oldest-old is measured, not per capita, but by the fraction of 90+ year old people who are over age 105 (d, England; r = 0.71, p < 2e-16), age 110 (e, France; r = 0.51; p = 0.00001), or age 100 (f, Japan; r = 0.62, p = 0.000004).

Japan displayed similar anomalies in the regional distribution of extreme-age records. The first and second ranked Japanese prefectures for centenarians per capita, Shimane and Kochi, had the worst and second-worst regional economic rankings (Table S4), while extensive anomalies in third-ranked 'Blue Zone' of Okinawa are detailed in the supplementary materials (Supplementary Materials; Table S5). Japanese poverty rates in the general population were again positively correlated with attaining remarkable lifespans (r = 0.36; p = 0.01; Fig 4c), an interaction that strengthens when poverty rates are used to predict the fraction of 90+ year old people living past age 100 (r = 0.62; p = 4e-06; Fig 4f). The number of centenarians in Japan is also negatively correlated with income per capita (r = −0.44, p=0.001), the minimum wage (r = −0.64; p = 1e-09), and the Japanese financial strength index (r = −0.70; p = 3e-08) across all 47 Japanese prefectures. Prefectures that spend more money on old-age welfare per capita, a disincentive for welfare fraud, also produce fewer centenarians per capita (r = −0.49; p= 0.0004). These factors share latent drivers and are highly autocorrelated: prediction models for centenarians per capita based solely on poverty (R^2 = 0.37; p= 0.0001; S2 Code) approach the accuracy of linear mixed models containing all available socioeconomic variables (R^2 = 0.43; p = 3e-06) yet have a lower Akaike's information Criteria (S2 Code).

Linking French historical life expectancy at birth[32] to some 143 French supercentenarians with sufficient data revealed that the cohort life expectancy at birth of each supercentenarian was not significantly different to the contemporary national average (Fig S4). That is, mainland French supercentenarians were not, on average, born into regions with either significantly longer or shorter than expected life expectancy. Given it is only known where these individuals where born, and not where they lived most of their lives, it seems notable that the modern economic conditions and poverty rate of a birth location should be predictive of becoming the 'oldest-old' (Fig 4) while life expectancy at birth is not (Fig S4).

Enrichment for specific days of the month in birth dates is seen as indicative of nonrandom age selection due to fraud, error, and clerical uncertainty. This check, however, is limited in that it cannot detect diverse sources of error, such as identity fraud or failed death registrations, which retain a representative distribution of birth dates.

Days of the month (e.g. the 1st or 2nd of each month) are nearly uniformly distributed throughout the year. As a result, US births had a near-uniform distribution of births, with minimal deviation from random sampling. Even after the widespread uptake of induced births, which avoid weekends and public holidays, birthdays generally varying by less than two per cent across different days of the month (Figure S5a), such that the distribution of births across days of the month is nearly invariant in the modern USA. In contrast, supercentenarians in the GRG database are 1.4-fold more likely to be born on the first day of the month (Figure S5b) and 1.2-fold as likely to be born on a day that is divisible by five. The number of supercentenarians born on the first day of the month is 150% higher than the previous day. Given the near-complete absence of caesarean sections in these population, this pattern may be explained if a large percentage of people have non-randomly chosen or manufactured birthdates.

These patterns are broadly reflected in the IDL birthdates, with days divisible by five, excepting the 25th, being over-represented amongst supercentenarians and SSCs (Fig S5b). In contrast the first day of the month was not over-represented amongst the IDL data (Fig S5c): this was initially difficult to reconcile, given these databases overlap considerably and largely represent the same individuals.

However, this differential appears to be a result of dates of birth being missing from 48% (N=797) of the IDL supercentenarians: despite extensive records being available, every US supercentenarian in the IDL has their day and month of birth removed. In addition, all Japanese records are absent. In the GRG data, most of the signal for over-enrichment arises in Japanese and US birthdates (Code S1-S2), with Japanese supercentenarians 2.77 times more likely and US supercentenarians 1.57 times more likely to be born on the first day of the month. In the primarily European countries in the remaining GRG database, individuals instead have a 0.67 odds ratio of being born on the first day: oddly, the inverse of the US and Japanese pattern. As a result, the omission of certain dates and records from the IDL data, for reasons that are unclear, masks the age heaping apparent in the USA and Japan.

Discussion

Basic economic and social indicators in the modern economy, such as GDP per capita and poverty rates, predict the distribution of extreme age records. Despite constraints on model construction and accuracy, such as unavoidable differences in per capita adjustments, these basic models approached reasonable accuracy. However, the direction of these interactions is the opposite of rational expectations.

Diverse social and economic indicators that normally predict worse health outcomes, such as income deprivation, poverty, and high unemployment, are all positively associated with a higher probability of reaching an extreme age. These factors are linked to a lower probability of survival and worse health outcomes at every age below 90, for every population included in this study.

However, these 'anti-health' factors exhibit a consistent positive association with extreme longevity. In England, which contains the only national data with sufficiently granular regional health measures, even poor health itself is positively associated with attaining a remarkable lifespan (Fig 3c). Across England and Italy, a larger number of people over the age of 90 is a significantly predictor of *fewer* people over age 105 (Fig 3e; Fig S3g-i): a difficult pattern to explain, given the younger population contains the older, unless errors are a primary factor.

Observed rates of vital registration in extreme-age and supercentenarian data are even more difficult to explain. For example, 96% of all 105+ year olds died since 1990, overwhelmingly in populations with over 95% rates of death certification. During validation, the intensive search for vital registration documents and the removal of unsupported cases should have enriched this death certification rates *above* the 95% baseline. It is therefore unclear why only 8% of all SSCs and 1.4% of all US supercentenarians recorded in the IDL have a death certificate unless these data are mostly errors. A similar pattern holds for birth certificates: even an intensive search could not exceed 15% coverage[41].

Finally, claims to remarkable regional longevity are often contra-indicated by substantial, but curiously uncited, studies. For example, the Centres for Disease Control generated an independent estimate of average longevity across the USA: they found that Loma Linda, a Blue

Zone supposedly characterised by a 'remarkable' average lifespan 10 years above the national average, instead has an unremarkable average lifespan[29] (27th-75th percentile; Fig S6). European Blue Zones also enjoy independent estimates. According to the collective governments of the European Union[42], the Sardinian and Ikarian Blue Zones occupy regions that are 36-44th and 56-65th in the EU for old-age (85+) longevity.

Viewed in isolation these anomalies may, perhaps, be explained away by reference to unknown lifestyle factors. However, these findings should be considered in the context of other diverse and incongruous patterns observed in extreme old age studies. These concerns are so extensive that discussion of even the broad outlines requires a critical review (Supplementary Materials).

The problems of extreme-age research may, however, be focused into two critical points. The first is observational. Extremely high-frequency errors are not impossible, improbable, or hypothetical[19]: instead, they have been routinely discovered after long evading the notice of demographers, often in the most intensively studied populations in the world. Any statement that such error rates could not happen[19], for example due to the vigilance or rigour of demographic validation, are therefore belied by historical fact.

In the USA, for example, 27% ('white male') to 66% ('non-white female') of Americans had multiple official ages in 1960, with 8%-30% misreported by more than a decade[20]. At least 54% of US centenarians were revealed as errors in 1979[43], and over half of all decedent African-Americans had multiple official ages in 1985[22]. In 2003, Stone checked 550 US supercentenarians recorded from 1980-1999 and, after an exhaustive search, just 43 of 'supercentenarian' records had a birth certificate (14.5%), and only 217 (40%) had a consistently-reported age [41].

Substantial error rates were recently uncovered in every 'Blue Zone'. In 1997, thirty thousand Italian citizens were discovered to be claiming the pension whilst dead[44]. In 2008, 42% of Costa Rican 99+ year olds were revealed to have 'mis-stated' their age in the 2000 census[45] and, after limited error-correction, the Nicoya Blue Zone shrunk by around 90%[46] and old-age life expectancy plummeted from world-leading to 'near the bottom of the pack'[47]. In 2010, over 230,000 Japanese centenarians were discovered to be missing, imaginary, clerical errors, or dead[48,49] -- an error rate of 82% in data then considered among the best in the world[7,50].

Greece followed in 2012, when at least 72% of Greek centenarians reported in the census were discovered to be dead or, depending on your perspective, committing pension fraud[51]. Finally, at least 17% of centenarians in the USA were discovered to be non-centenarians in 2019, not through intensive validation or qualitative interviews, but by reading two plain-text files and finding the dates did not match[14].

Only two of these cases were discovered, and all seem generally ignored, by professional demographers[14,45]. Some cases were seemingly not even noticed *post-hoc*: when most Greek centenarians were found to not exist in the aftermath of the global financial crisis, their non-existence was published in newspapers[51], but was not mentioned or cited in the demographic literature.

If equivalent rates of fake data were discovered in any other field – for example, if 82% of people in the UK Biobank or 17% of galaxies detected by the Hubble telescope were revealed to be imaginary – a major scandal would ensue. In demography, however, such revelations seem to barely merit citation.

It seems worth asking, therefore, why the recurrent discovery of high rates of non-random age errors have been ignored by the scientific community, given the fundamental importance of accurate age data to fields like medicine, gerontology, the social sciences, and epidemiology.

A typical response to discovering most data is illusory has not been to stop trusting data built on identical methods and documents, but to instead call for better validation and a closer examination of documents[52]. This raises the second critical point. The examination of documents, the only measure of human age in virtually all demographic research, cannot detect routine errors.

Again, this is not an assertion but a reality. Consider, for example, the 1878 Taché investigation[53,54] that pioneered testing centenarian status using validation '*proved by authentic documents, examined with a rigorous scrutiny*': the unchanged state-of-the-art method. Of 421 claimed centenarians, 82 passed documentary validation, and only nine passed subsequent investigation, including Pierre Joubert: the world's oldest man and first validated supercentenarian[53,54]. The case endured over a century of scrutiny, and was continually included in world records, before Joubert was found to have died at age 65. It written on his death certificate the entire time[53,55].

Extreme-age demographers seem to have misunderstood the central lesson of such events: not that errors can be detected though increased scrutiny, but that acceptance of such cases is dependent on the often-arbitrary generation, survival, and detection of documents. Without all three, completely fake cases can remain 'validated' indefinitely, regardless of whether they are true.

Demographers have confounded consistency with accuracy by assuming that consistent and valid documents are accurate, or at least strongly indicative of accuracy. This assumption is deeply unsound, especially given the nonlinear accumulation of error rates with age[15] and the rate at which errors have previously passed validation. Showing that documents are consistent and valid does not mean that they are accurate, or that any connection to their supposed owner is real.

To understand this problem, consider a room containing 100 real Italian centenarians, each holding complete and validated documents of their age. One random centenarian is then exchanged for a younger sibling, who is handed their siblings' real and validated birth documents.

How could an independent observer discriminate this type I substitution from the 99 other real cases, using only documents as evidence? More problematically, how could that observer measure the frequency of such errors, and discriminate whether one, ten, or a hundred 'Italian Siblings' are in the room, given they each hold validated and consistent documents?

This simple class of error cannot be excluded based on documentary evidence: every document in the room is consistent, real, and validated. Any such type-I error therefore has the potential to indefinitely escape detection. Scrutinising documents, usually the only method open to demographers, provides no guarantee of accuracy.

Compounding this issue, any younger relative is also likely to have sufficient biographic knowledge to pass an interview (which almost never occur) with an accuracy indistinguishable from a plausibly-forgetful centenarian.

The 'Italian Sibling' though experiment illustrates just one reason that type I age-coding errors cannot be ruled out — or even necessarily measured — using documentary evidence alone. Instead, the experiment presents a substantial problem for remarkable-age databases, which can be embodied in a deliberately provocative, if seemingly absurd, hypothesis:

> Every 'supercentenarian' is an accidental or intentional identity thief, who owns real and validated 110+ year-old documents, and is passably good at deception.

This hypothesis cannot be invalidated by the further scrutiny of documents, or by models calibrated using document-informed ages[56,57]. Rather, invalidating this hypothesis requires a fundamental shift: it requires the measurement of biological ages from fundamental physical properties, such as amino acid chirality[58] or isotopic decay [59].

The long-hidden flaws of document-based age measurement may, however, be starting to unravel. New biometric methods of measuring ages[60,61] are revealing huge discrepancies with document-reported ages. For example, measurements of age based on epigenetic data uniformly predict the oldest-old are 8-20 years younger than their paper records suggest[62–65]. Researchers have dismissed these estimates as the result of slower aging in centenarians, of mysterious aetiology, even when centenarians were measured to have undifferentiated or 'non-significantly slower'[63] aging rates. It is hoped another explanation — the epigenetic ages are accurate, and the documents are not — may eventually be considered.

Until then, we can only hope that the development of new and better biometrics of age[59,60,66] may end our long reliance on documents, that we may finally stop ignoring the remarkable error rates observed in old-age demography.

Bibliography

1. Buettner D, Skemp S. Blue Zones: Lessons From the World's Longest Lived. Am. J. Lifestyle Med. 2016.

2. Poulain M, Pes GM, Grasland C, Carru C, Ferrucci L, Baggio G, et al. Identification of a geographic area characterized by extreme longevity in the Sardinia island: The AKEA study. Exp Gerontol. 2004;

3. Chrysohoou C, Panagiotakos DB, Siasos G, Zisimos K, Skoumas J, Pitsavos C, et al. Sociodemographic and lifestyle statistics of oldest old people (>80 years) living in ikaria island: The ikaria study. Cardiol Res Pract. 2011;

4. Rajpathak SN, Liu Y, Ben-David O, Reddy S, Atzmon G, Crandall J, et al. Lifestyle factors of people with exceptional longevity. J. Am. Geriatr. Soc. 2011.

5. Alfonso-Sánchez MA, Peña JA, Calderón R. Time trends and determinants of completed family size in a rural community from the Basque area of Spain (1800-1969). J Biosoc Sci. 2003;35:481–97.

6. Kwon IS, Kim C-H, Ko HS, Cho S Il, Choi YH, Park SC. Risk factors of cardiovascular disease in Korean exceptional longevity. J Korean Geriatr Soc. 2005;9:251–65.

7. Franceschi C, Bonafè M. Centenarians as a model for healthy aging. Biochem Soc Trans. 2004;

8. Heyn H, Li N, Ferreira HJ, Moran S, Pisano DG, Gomez A, et al. Distinct DNA methylomes of newborns and centenarians. Proc Natl Acad Sci. 2012;

9. Mondello C, Petropoulou C, Monti D, Gonos ES, Franceschi C, Nuzzo F. Telomere length in

fibroblasts and blood cells from healthy centenarians. Exp Cell Res. 1999;248:234–42.

10. Sebastiani P, Solovieff N, DeWan AT, Walsh KM, Puca A, Hartley SW, et al. Genetic signatures of exceptional longevity in humans. PLoS One. 2012;

11. Fries JF. Aging, natural death, and the compression of morbidity. Bull. World Health Organ. 2002. p. 245–50.

12. Vaupel JW, Manton KG, Stallard E. The Impact of Heterogeneity in Individual Frailty on the Dynamics of Mortality. Demography [Internet]. 1979;16:439. Available from: http://link.springer.com/10.2307/2061224

13. Preston SH, Elo IT, Stewart Q. Effects of age misreporting on mortality estimates at older ages. Popul Stud (NY). 1999;

14. Gavrilov LA, Gavrilova NS. Late-life mortality is underestimated because of data errors. PLoS Biol. 2019;17:e3000148.

15. Newman SJ. Errors as a primary cause of late-life mortality deceleration and plateaus. PLoS Biol. 2018;16:e2006776.

16. Gavrilova NS, Gavrilov LA. Mortality Trajectories at Extreme Old Ages: A Comparative Study of Different Data Sources on U.S. Old-Age Mortality. Living to 100 Monogr [Internet]. 2014;2014. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25664347%5Cnhttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4318539

17. Coale AJ, Kisker EE. Mortality crossovers: Reality or bad data? Popul Stud (NY). 1986;

18. Newman SJ. Plane inclinations: A critique of hypothesis and model choice in Barbi et al. PLoS Biol. 2018;16:e3000048.

19. Wachter KW. Hypothetical errors and plateaus: A response to Newman. PLoS Biol.

2018;16:e3000076.

20. Hambright TZ. Comparison of information on death certificates and matching 1960 census records: age, marital status, race, nativity and country of origin. Demography [Internet]. 1969 [cited 2024 Jan 2];6:413–23. Available from: https://pubmed.ncbi.nlm.nih.gov/21279795/

21. Preston SH, Elo IT. Black mortality at very old ages in official US life tables: A skeptical appraisal. Popul Dev Rev. 2006;

22. Preston SH, Elo I, Rosenwaike I, Hill M. African-American Mortality at Older Ages. Demography. 1996;33:193–209.

23. Gavrilov L. Mortality measurement at advanced ages: a study of the Social Security Administration Death Master File. North Am Actuar J [Internet]. 2011;15:432–47. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3269912/

24. Gerontology Research Group [Internet]. Table A Verif. Supercentenarians. 2019 [cited 2019 Jun 2]. Available from: http://www.grg.org/SC/SCListsTables.html

25. Vaupel J, Vallin J, Meslé F, Robine J-M, Jdanov DA, Grigoriev O. The International Database on Longevity [Internet]. 2020 [cited 2020 Feb 13]. Available from: https://www.supercentenarians.org/Home/Expand_IDL

26. Elders.Stat database: Istituto Nazionale di Statistica [Internet]. Ist. Naz. di Stat. 2019 [cited 2019 Feb 7]. Available from: http://dati-anziani.istat.it/

27. Director-General for Statistics and Information Policy, Ministry of Health L and W (organisation). Vital Statistics [Internet]. 2019 [cited 2019 Feb 7]. Available from: https://www.mhlw.go.jp/english/database/db-hw/index.html

28. Statistics Japan. e-Stat: Portal Site of Official Statistics of Japan [Internet]. 2020 [cited 2020 Jan 30]. Available from: https://www.e-stat.go.jp/en

29. Tomuro K. Trends Observed in Poverty Rates, Working Poor Rates, Child Poverty Rates and Take-Up Rates of Public Assistance Across 47 Prefectures in Japan [Internet]. Yamagata Univ. Fac. Humanit. Res. Annu. Rep. Yamagata; 2016. Available from: https://yamagata.repo.nii.ac.jp/?action=repository_uri&item_id=4067&file_id=17&file_no=1

30. Farnsworth Riche M, Benton B, Schnelder PJ, Norton AJ. Population of the States and Counties of the United States: 1790 to 1990 [Internet]. Washington D.C.; 1996. Available from: https://www.census.gov/population/www/censusdata/PopulationofStatesandCountiesoftheUnitedStates1790-1990.pdf

31. Hetzel AM. History and Organization of the Vital Statistics System. US Vital Stat Syst Major Act Dev 1950-95 [Internet]. Centers for Disease Control and Prevention; National Center for Health Statistics; 1997. p. 75. Available from: https://www.cdc.gov/nchs/data/misc/usvss.pdf

32. Bonneuil N. Table de mortalité France et par département et tables de migration nette par département, 1806-1906. Transform French Demogr Landsc 1806 - 1906. Oxford: Clarendon Press; 1997.

33. Lahti L, Huovari J, Kainu M, Biecek P. Retrieval and analysis of eurostat open data with the eurostat package. R J. 2017;

34. Ministry of Housing Communities & Local Government. English indices of deprivation 2019: research report [Internet]. 2019. Available from: https://www.gov.uk/government/publications/english-indices-of-deprivation-2019-research-report

35. Institut national de la statistique et des études économiques. Statistiques locales: Revenus-Pouvoir d’achat - comsommation Pauvreté. 2016.

36. Curtis KJ, Lee J, O’Connell HA, Zhu J. The Spatial Distribution of Poverty and the Long

Reach of the Industrial Makeup of Places: New Evidence on Spatial and Temporal Regimes. Rural Sociol [Internet]. 2019 [cited 2024 Jan 4];84:28. Available from: /pmc/articles/PMC6497161/

37. Gibson C, Jung K. Historical Census Statistics On Population Totals By Race, 1790 to 1990, and By Hispanic Origin, 1970 to 1990, For Large Cities And Other Urban Places In The United States. Washington D.C.; 2005.

38. Condon S, Ogden P. Emigration from the French Caribbean: the Origins of an Organized Migration. Int J Urban Reg Res [Internet]. 1991 [cited 2023 Dec 21];15:505–23. Available from: https://onlinelibrary.wiley.com/doi/full/10.1111/j.1468-2427.1991.tb00655.x

39. Pan American Health Organization. Country Report: French Guiana, Guadeloupe, and Martinique [Internet]. Heal. Am. 2017 [cited 2020 Feb 3]. Available from: https://www.paho.org/salud-en-las-americas-2017/?page_id=121

40. Aldridge H, Theo BB, Tinson A, MacInnes T. London's Poverty Profile 2015. Tackling Poverty Inequal. 2015;

41. Stone LF. Studies in the demography of supercentenarians in the United States [Internet]. University of Pennsylvania; [cited 2023 Dec 21]. Available from: https://www.proquest.com/docview/304711712?pq-origsite=gscholar&fromopenview=true&sourcetype=Dissertations & Theses

42. EUROSTAT. Life expectancy by age, sex and NUTS 2 region [Internet]. table demo_r_mlifexp. 2023 [cited 2023 Dec 3]. Available from: https://ec.europa.eu/eurostat/databrowser/product/page/demo_r_mlifexp

43. Rosenwaike I. A new evaluation of United States census data on the extreme aged. Demography. 1979;16:279–88.

44. Reuters. Italy's Dead Pensioners. New York Times. 1997;A00009.

45. Bixby L, Brenes G, Collado A. Tablas de vida para cálculo actuarial de rentas vitalicias y retiro programado Costa Rica circa 2000. Población y Salud en Mesoamérica [Internet]. 2004 [cited 2023 Dec 3];173–205. Available from: https://www.redalyc.org/pdf/446/44601204.pdf

46. Rosero-Bixby L. The vanishing advantage of longevity in Nicoya, Costa Rica: A cohort shift. Demogr Res. 2023;49.

47. Glei DA, Barbieri M, Santamaría-Ulloa C. Costa Rican Mortality 1950-2013:An Evaluation of Data Quality and TrendsCompared with Other Countries. Demogr Res [Internet]. 2019 [cited 2023 Dec 3];40:835. Available from: /pmc/articles/PMC6544445/

48. Fackler M. Japan, Checking on Its Oldest, Finds Many Gone. New York Times. 2010.

49. Japanese Ministry of Justice. About family register office work to affect location unknown elderly people [Internet]. 2010 [cited 2019 Jun 18]. Available from: http://www.moj.go.jp/MINJI/minji04_00008.html.

50. Willcox DC, Willcox BJ, He Q, Wang NC, Suzuki M. They really are that old: A validation study of centenarian prevalence in Okinawa. Journals Gerontol - Ser A Biol Sci Med Sci. 2008;

51. Koutantou A, Papachristou H. Greece pulls the plug on pensions for the dead. Reuters [Internet]. 2012; Available from: https://www.reuters.com/article/uk-greece-benefits/greece-pulls-the-plug-on-pensions-for-the-dead-idUKBRE83O14220120425

52. Poulain M. Exceptional longevity in Okinawa: A plea for in-depth validation. Demogr Res. 2011;

53. Desjardins B. Validation of Extreme Longevity Cases in the Past: The French-Canadian Experience. Valid Except Longev. 2003.

54. Government of Canada. Censuses of Canada 1608 to 1876. Ottowa; 1878.

55. Charbonneau H. Pierre Joubert a-t-il vécu 113 ans? Mémoires la Société généalogique Can. 1990;41:45–8.

56. Fleischer JG, Schulte R, Tsai HH, Tyagi S, Ibarra A, Shokhirev MN, et al. Predicting age from the transcriptome of human dermal fibroblasts. Genome Biol. 2018;

57. Malli RC, Aygun M, Ekenel HK. Apparent Age Estimation Using Ensemble of Deep Learning Models. IEEE Comput Soc Conf Comput Vis Pattern Recognit Work. 2016.

58. Ohtani S, Yamamoto T. Age estimation by amino acid racemization in human teeth. J Forensic Sci. 2010;

59. Nielsen J, Hedeholm RB, Heinemeier J, Bushnell PG, Christiansen JS, Olsen J, et al. Eye lens radiocarbon reveals centuries of longevity in the Greenland shark (Somniosus microcephalus). Science (80-) [Internet]. 2016;353:702–4. Available from: http://www.sciencemag.org/cgi/doi/10.1126/science.aaf1703

60. Horvath S. DNA methylation age of human tissues and cell types. Genome Biol [Internet]. 2013 [cited 2024 Jan 2];14:R115. Available from: /pmc/articles/PMC4015143/

61. Geng X, Zhou ZH, Smith-Miles K. Automatic age estimation based on facial aging patterns. IEEE Trans Pattern Anal Mach Intell. 2007;29:2234–40.

62. Daunay A, Hardy LM, Bouyacoub Y, Sahbatou M, Touvier M, Blanché H, et al. Centenarians consistently present a younger epigenetic age than their chronological age with four epigenetic clocks based on a small number of CpG sites. Aging (Albany NY) [Internet]. 2022 [cited 2024 Jan 3];14:7718. Available from: /pmc/articles/PMC9596211/

63. Komaki S, Nagata M, Arai E, Otomo R, Ono K, Abe Y, et al. Epigenetic profile of Japanese supercentenarians: a cross-sectional study. Lancet Heal Longev [Internet]. 2023 [cited 2024 Jan 3];4:e83–90. Available from: http://www.thelancet.com/article/S2666756823000028/fulltext

64. Horvath S, Pirazzini C, Bacalini MG, Gentilini D, Di Blasio AM, Delledonne M, et al. Decreased epigenetic age of PBMCs from Italian semi-supercentenarians and their offspring. Aging (Albany NY) [Internet]. 2015 [cited 2024 Jan 3];7:1159. Available from: /pmc/articles/PMC4712339/

65. Dec E, Clement J, Cheng K, Church GM, Fossel MB, Rehkopf DH, et al. Centenarian clocks: epigenetic clocks for validating claims of exceptional longevity. GeroScience [Internet]. 2023 [cited 2024 Jan 3];45:1817. Available from: /pmc/articles/PMC10400760/

66. Han H, Otto C, Jain AK. Age estimation from face images: Human vs. machine performance. 2013 Int Conf Biometrics [Internet]. 2013. p. 1–8. Available from: http://ieeexplore.ieee.org/document/6613022/

Statements and Declarations

Funding

The author declares that no funds, grants, or other support were received during the preparation of this manuscript.

Acknowledgements

The author would like to acknowledge Dr Elena Racheva for advice and encouragement, Zoe Campbell and Prof Heather Booth for providing much-needed feedback and editing advice, Chris Mulligan for providing inspiration to use US birth data, Sally Morell and Dr Jim Docherty for interesting leads, Dr Gideon Meyerowitz-Katz for the final nudge toward publication, and an extraordinary cross-section of the scientific community for providing commentary and support.

Author Contributions

SJN conceived, designed, analyzed, and wrote the study and supplementary critical review.

Competing Interests

The author has no relevant financial or non-financial interests to disclose.

Ethics approval

This study was performed in line with the principles of the Declaration of Helsinki constituting a reanalysis of published anonymized data from publicly accessible databases. No ethical approval was required.

Supplementary Figures and Tables

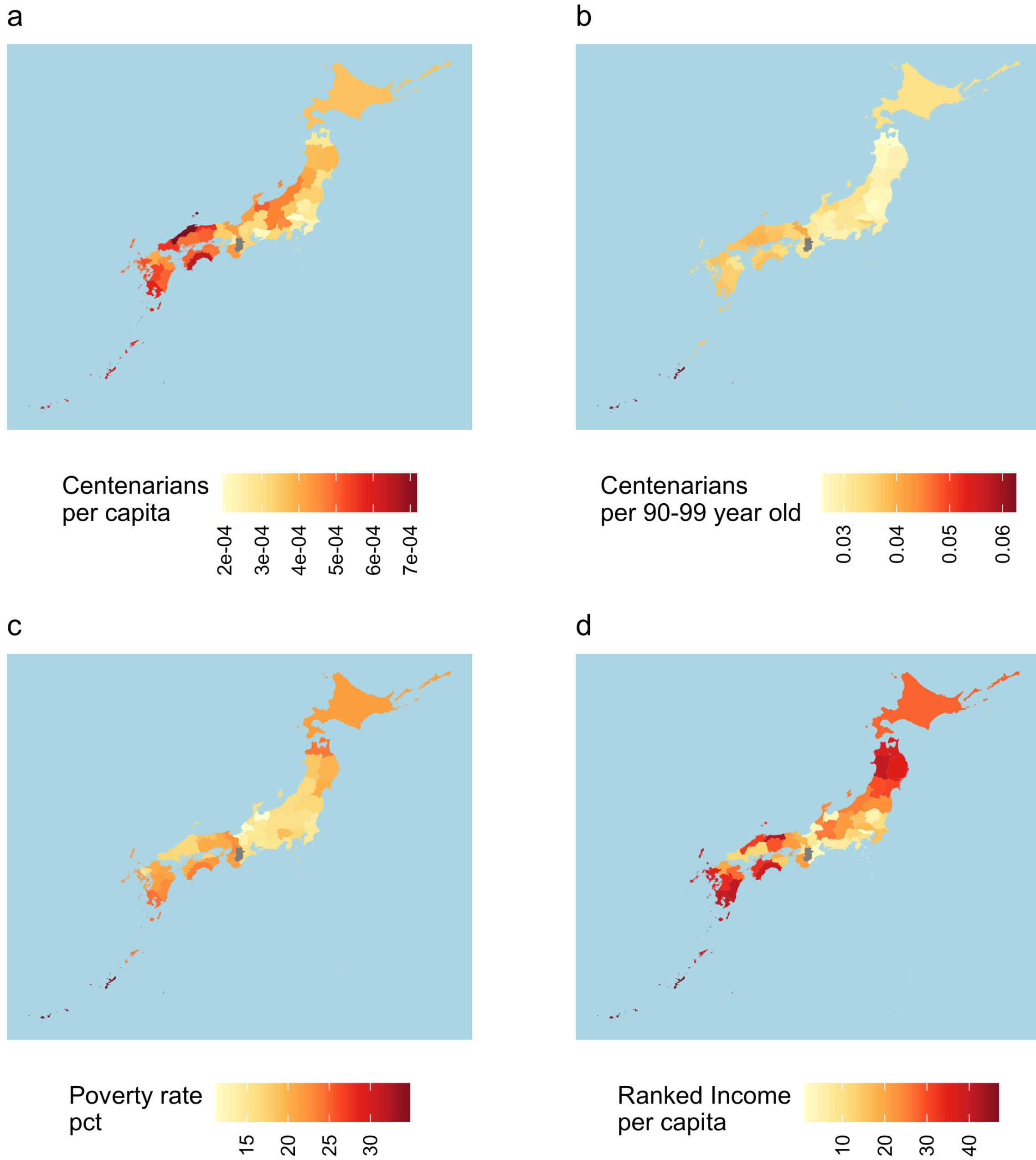

Figure S1. Poverty and the distribution of Japanese Centenarians. The 47 prefectures of Japan putatively contain over 48,000 centenarians, with generally higher concentration of centenarians per capita (a) and per 90-99 year-old (b) in prefectures with high poverty rates (c) and lower-ranked prefectural incomes (d).

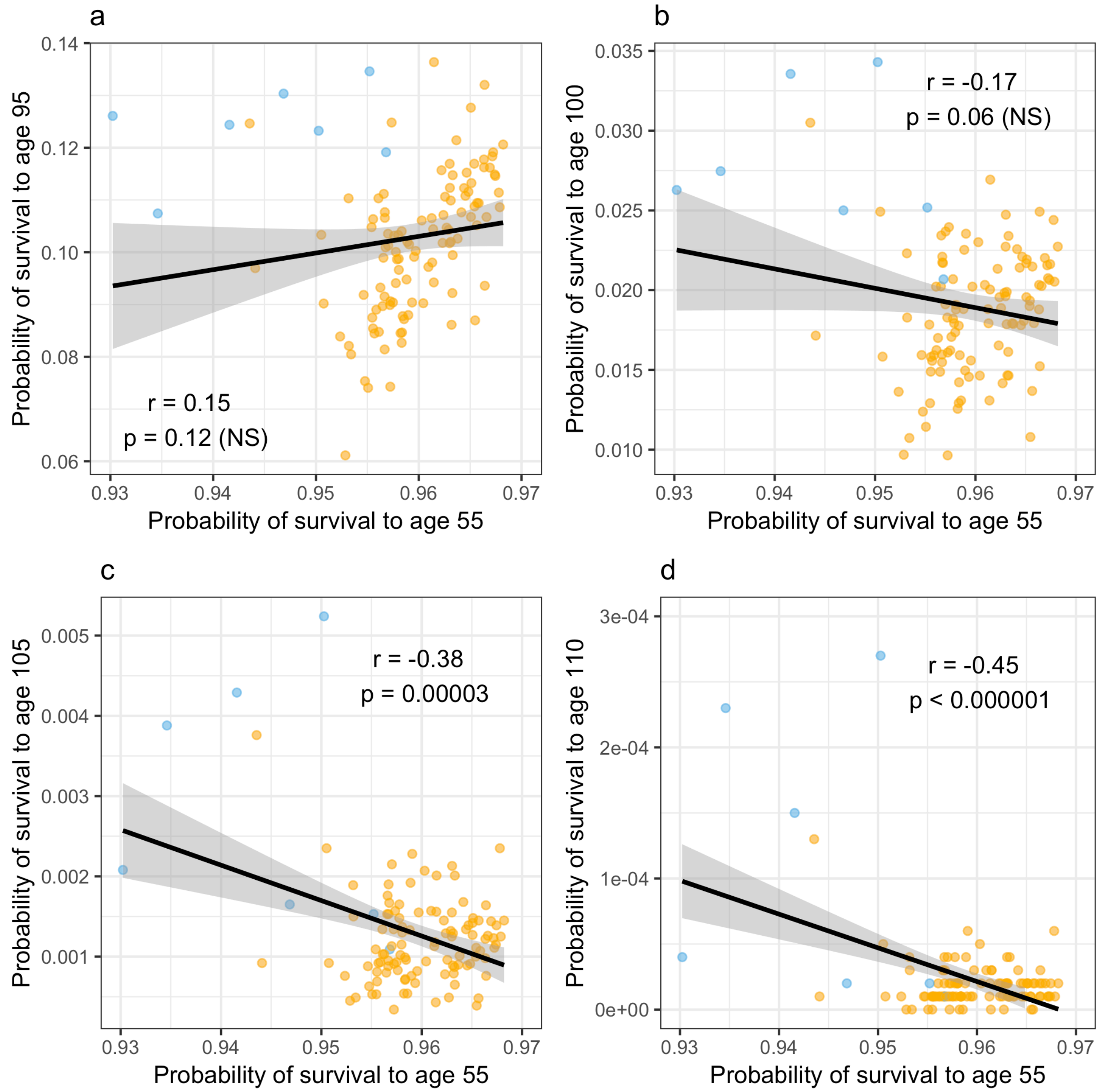

Figure S2. Relationship between mid-life and late-life survival across Italian provinces. Across Italian provinces (points), probabilities of survival in mid-life are positively correlated with the probability of survival at older ages until around age 95 (a; r = 0.15; p=0.1; N=116). However, this relationship inverts at advances ages: better mid-life and early-life probabilities of survival, and higher average longevity, are linked to significantly lower probabilities of survival at 100 years (b), 105 years (c), or 110 years (d) of age. Sardinian provinces shown in blue.

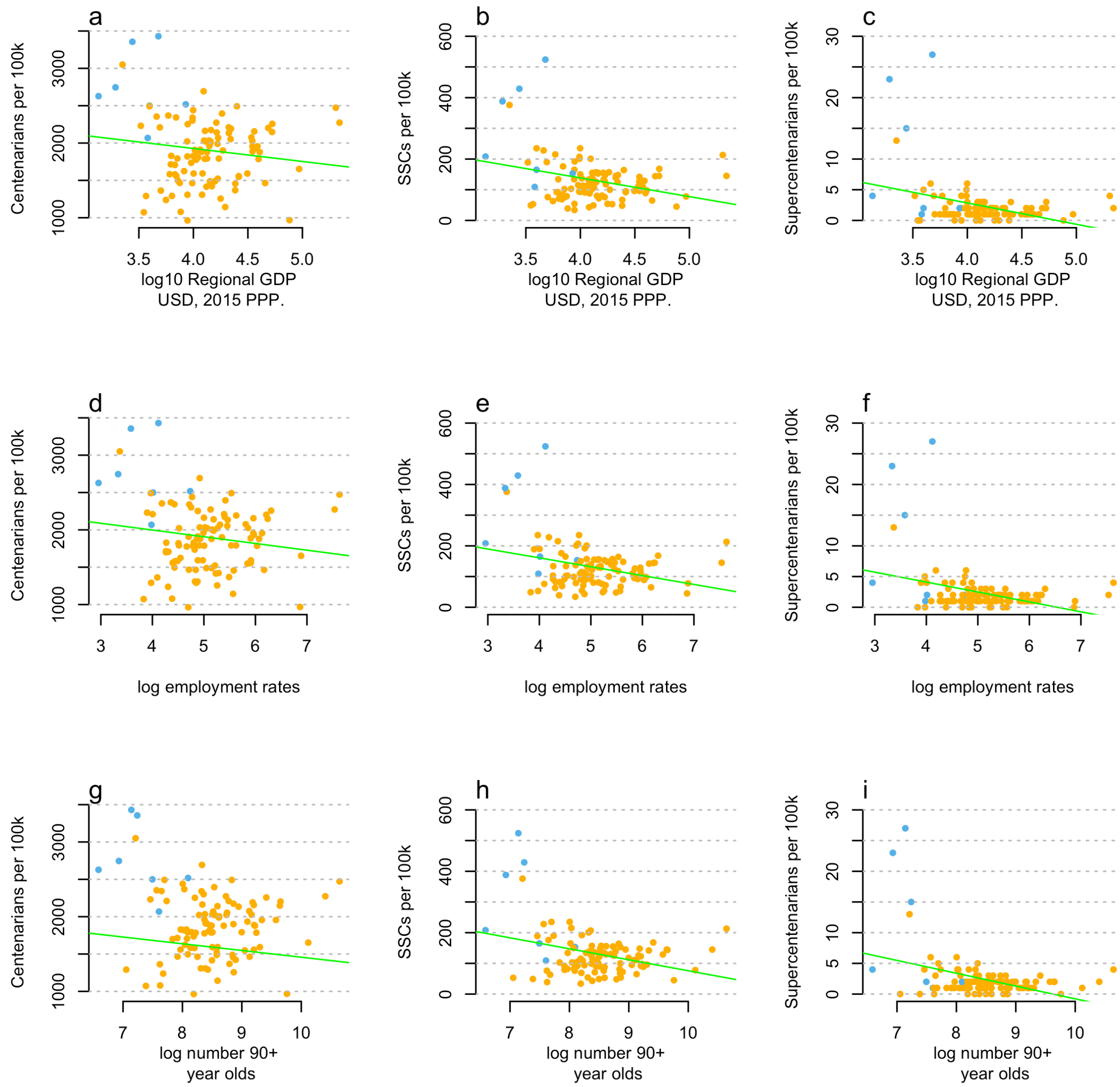

Figure S3. Italian provinces by GDP per capita and rates of the oldest-old. According to figures from the Italian national statistics office and regional GDP data from the OECD, purchase-power parity adjusted GDP is negatively correlated with the frequency of (a) centenarians, (b) SSCs and (c) supercentenarians per capita across Italy: a pattern repeated in employment rates (d-f). Furthermore, the total number of 90+ year olds, shown here in log scale, is also negatively associated with the per capita number of centenarians (g), SSCs (h) and supercentenarians (i) across Italy. Linear mixed model regressions shown in green, Sardinian provinces shown in blue.

Figure S4. Historical life expectancy in the home region of birth for French supercentenarians. Supercentenarians are born in regions with an 84-day shorter life expectancy at birth, a non-significant reduction relative to the national average (one sample t-test NS; p = 0.52; N=143). Comparisons for supercentenarians born in overseas provinces and imperial holdings are unavailable, data show metropolitan France only.

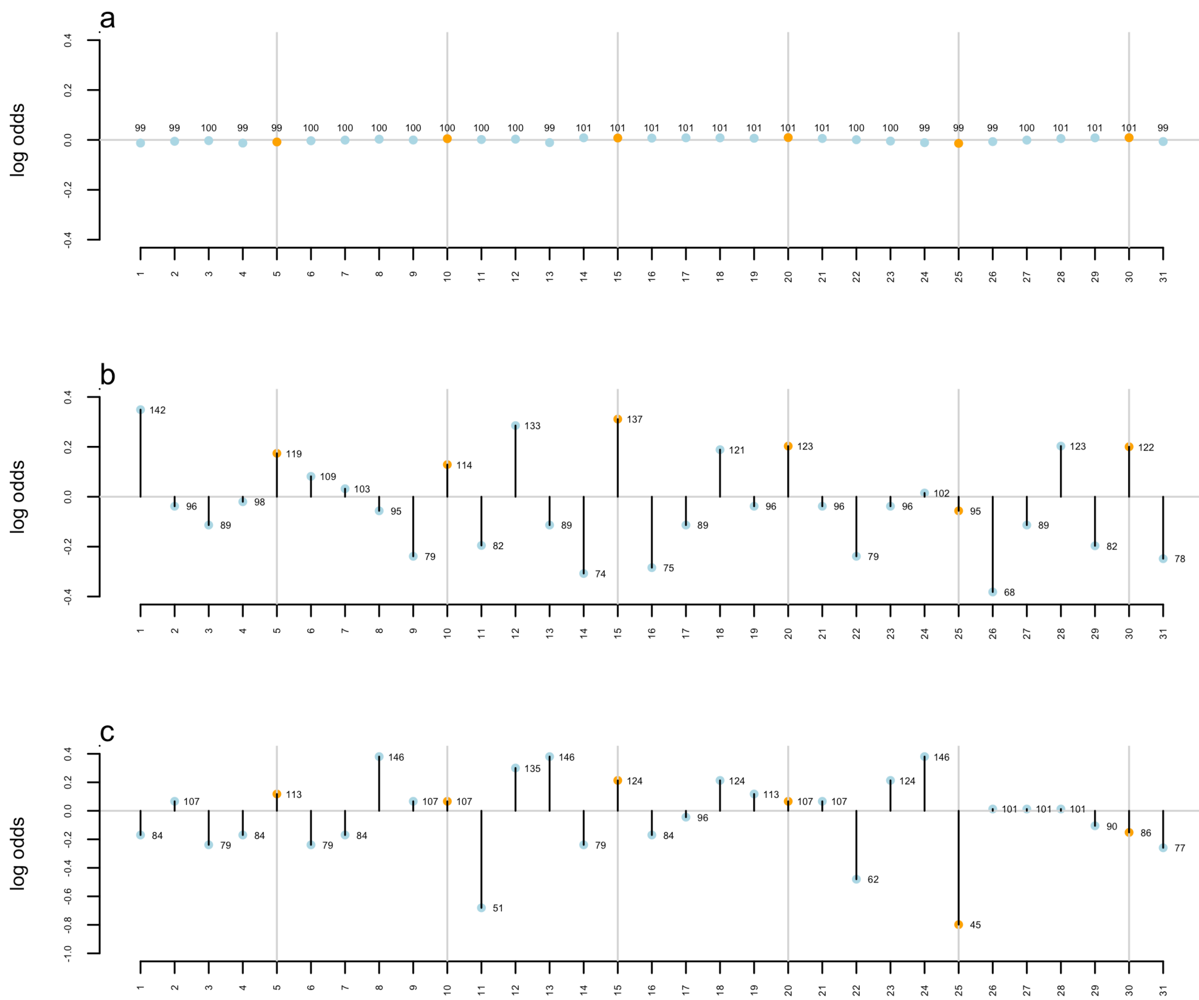

Figure S5. Age heaping of supercentenarian births. The distribution of modern birthdates (a), shown here by 70 million US birthdates observed from 1969-1988, display limited variation across days of the month. However, supercentenarian birth dates (N = 1739) are 1.42-fold more likely to be born on the first day of the month and 1.18-fold more likely to be born on days that are multiples of five (orange points) compared with randomly distributed births (b). Age heaping on the first day of the month or in multiples of five is not as clear in the IDL data (c), possibly because of the removal of US birth days and months, or differences in cultural patterns (the 25th is heavily under-represented) and data quality. Points are labeled by the percentage of births over- or under-represented, relative to random sampling.

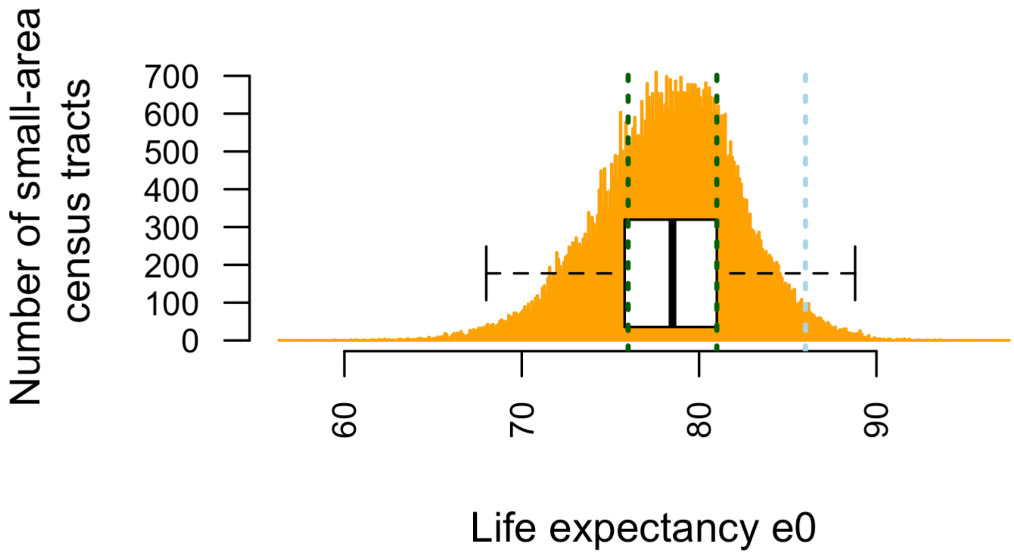

Figure S6. Distribution of life expectancy estimates in US small-area census tracts, including the Loma Linda 'blue zone'. When calculated independently by the Centers for Disease Control, the suburbs of Loma Linda range from the 27th to 75th percentiles of life expectancy at birth in the USA (green lines), relative to all other census tracts (orange). The absolute upper estimate, the female-only life expectancy calculated by 'blue zone' proponents (blue line), falls in the 98th percentile of life expectancy: high, but still behind 1401 longer-lived US census tracts and the nations of Japan, Singapore, Monaco, Spain, and South Korea.

Table S1. Top 20 Italian provinces and regions for supercentenarians per capita.

				National ranking, 1 = Worst, 116 = Best, ND = No Data			
Province	**Region**	**Super-centenarians per 100k (I_{110})**	**Rank I_{110}**	Survival to age 55	Unemployment	GDP per capita, PPS adjusted	90+ year-olds per capita (1=fewest)
Olbia-Tempio	Sardinia	27	1	7	14	41	8
Medio Campidano	Sardinia	23	2	2	2	1	33
Carbonia-Iglesias	Sardinia	15	3	3	4	13	43
Isernia	Molise	13	4	4	3	34	102
Matera	Basilicata	6	5	61	15	17	23
Prato	Tuscany	6	6	114	40	80	59
L'Aquila	Abruzzo	5	7	65	39	49	82
Rieti	Lazio	5	8	8	7	31	80
Belluno	Veneto	4	9	45	29	96	91
Campobasso	Molise	4	10	38	22	35	76
Catanzaro	Calabria	4	11	59	52	26	31
Gorizia	Friuli-Venezia Giulia	4	12	34	10	64	107
Ogliastra	Sardinia	4	13	1	1	23	66
Pistoia	Tuscany	4	14	87	37	59	84
Metropolitan City of Rome	Lazio	4	15	81	111	104	24
Vibo Valentia	Calabria	4	16	12	6	4	38
Brescia	Lombardy	3	17	90	106	94	25
Chieti	Abruzzo	3	18	64	64	53	71
Cremona	Lombardy	3	19	67	56	75	60
Fermo	Marche	3	20	35	ND	ND	ND

845 **Table S2. Top 20 French regions for the oldest-old per capita (metropolitan and ranked overseas territories only).**

				National Rankings 1 = Worst, 101 = Best, ND = No Data		
Region	**Super-centenarians (IDL)**	**Super-cent. per capita (rank)**	**Poverty rate ages 75+ (%)**	Poverty rate ages 75+ (rank)	Poverty rate	GDP per capita, PPS adjusted
Creuse	3	1	17.0	4	16	4
Martinique	8	2	31.1	2	2	46
Guadeloupe	8	3	ND	ND	ND	19
Cantal	2	4	13.9	10	60	21
Haute-Loire	3	5	11.1	26	77	10
Lozère	1	6	14.5	7	52	27
Lot	2	7	12.2	17	46	36
Indre	2	8	10.5	32	43	22
Yonne	3	9	7.0	75	44	39
Aude	3	10	14.4	8	5	11
Tarn	3	11	11.1	25	33	13
Paris	17	12	9.6	39	27	100
Jura	2	13	8.5	56	84	34
Maine-et-Loire	6	14	6.8	83	85	59
Aveyron	2	15	12.6	16	57	42
Hautes-Alpes	1	16	9.9	38	53	48
Orne	2	17	8.3	61	26	25
Ariège	1	18	12.8	15	14	6
Corse-du-Sud	1	19	15.9	5	19	75
Alpes-de-Haute-Provence	16	20	10.8	69	123	40

Table S3. Top 20 regions in England for the oldest-old per capita.

			National Rankings 1 = Worst, 131 = Best				
Region	**Total SSCs**	**SSCs per capita (rank)**	Income Deprivation Affecting Older People	90+ year-olds per capita (1=fewest)	Index of Multiple Deprivation	Crime rate	Health index
Tower Hamlets	15	1	1	1	26	23	52
Lewisham & Southwark	31	2	8	5	41	26	59
Darlington	5	3	62	85	46	14	32
Lambeth	15	4	5	3	49	18	56
Camden & City of London	11	5	55	18	90	119	110
Kensington & Chelsea and Hammersmith & Fulham	15	6	16	10	67	37	112
Isle of Wight	6	7	72	128	58	92	61
East Cumbria	11	8	120	121	94	130	82
Southampton	10	9	42	42	34	3	30
Portsmouth	8	10	44	65	33	10	48
Derby	9	11	48	64	40	75	33
North and West Norfolk	9	12	90	126	64	129	51
Telford and Wrekin	6	13	49	21	51	53	38
Manchester	18	14	3	6	4	1	3
Haringey & Islington	16	15	4	4	28	8	65
Leicester	11	16	7	17	17	25	27
Tyneside	27	17	29	62	25	39	16
Liverpool	15	18	6	16	2	12	2
Suffolk	23	19	101	120	86	91	97
East Kent	16	20	69	123	57	49	64

Table S4. Top 20 Japanese prefectures for centenarians per capita in 2015.

Prefecture	**Centenarians (total)**	**Rank Centenarians per capita**	**Poverty rate (%)**	**National ranking, 1 = Worst, 47 = Best**		
				Poverty rate (rank)	Financial strength	Income per capita 2011
Shimane	515	1	16.7	30	1	13
Kochi	486	2	23.7	4	2	6
Okinawa	872	3	34.8	1	5	1
Kagoshima	985	4	24.3	2	6	5
Tottori	334	5	18.9	21	3	2
Yamaguchi	806	6	16.9	27	23	36
Kumamoto	972	7	21.5	10	16	11
Saga	441	8	15.6	35	13	10
Toyama	554	9	11.2	47	25	44
Okayama	980	10	20.6	15	29	17
Ehime	720	11	20.2	17	20	12
Miyazaki	566	12	23.0	6	9	3
Nagasaki	700	13	22.2	8	7	7
Hiroshima	1395	14	16.9	26	35	35
Kagawa	482	15	17.2	24	27	28
Tokushima	377	16	21.8	9	8	33
Niigata	1105	17	16.0	33	19	21
Nagano	1000	18	15.5	36	26	25
Yamanashi	392	19	19.1	20	18	34
Oita	536	20	21.3	13	15	19

855

Table S5. Claims and evidence from the Okinawan Blue Zone.

Claim	Observed Evidence	Rank in 47 prefectures
“Move Naturally” [1]	Body Mass Index[2]	1st
*“The world’s longest-lived people don’t pump iron, run marathons or join gyms. Instead, they live in environments that **constantly** nudge them into **moving without thinking** about it. They **grow gardens** and don’t have mechanical conveniences for house and yard work.”* [1]	Percent population that engages in gardening[3]	45th
	Rest and Relaxation Time[3]	46th
“Purpose” [1]	Suicide rate[4]	9th
“The Okinawans call it “Ikigai” and the Nicoyans call it “plan de vida;” for both it translates to “why I wake up in the morning.” Knowing your sense of purpose is worth up to seven years of extra life expectancy.” [1]	Suicide rate (2019) in over-65s[5]	4th
“Plant slant” [1]	Consumption of … root vegetables[6,7]	47th (last)
*“**Beans**, including fava, black, soy and lentils, are the cornerstone of most centenarian diets. **Meat—mostly pork**—is eaten on average only **five times per month**. Serving sizes are **3-4 oz.**, about the size of a deck of cards.”*[1]	leafy vegetables[6,7]	47th (last)
	pickled vegetables[6,7]	47th (last)
	Taro[6,7]	47th (last)
Note: 3-4oz of meat a serve, five times a month, is **5.1-6.8kg of meat a year**- without fish, shellfish, or other meats, actual consumption is **40kg a year**.	Sweet potato[6,7]	47th (last)
	String beans[6,7]	44th
	Soybean paste[6,7]	44th
	Bean curd[6,7]	46th
19.1kg/year	Pork[6,7]	20th
6.2kg/year	Beef[6,7]	23rd
12.0kg/year	Chicken[6,7]	46th
2.6kg/year	‘Other raw meat’[6,7]	3rd
“Wine at 5” [1]	All (pure) alcohol[8]	4th (8.8 L/year)
*“People in all blue zones (except Adventists) drink alcohol **moderately*** and regularly. **Moderate drinkers outlive non-drinkers.** The trick is to drink 1-2 glasses per day* (preferably Sardinian Cannonau wine), with friends and/or with food. And no, you can’t save up all week and have 14 drinks on Saturday.”* [1]	Wine[8]	8th (3.6 L/year)
	Beer[8]	2nd (61.0L/year)
“Belong” [1]	Percent Atheist / No religion[9]	1st (93.4% atheist)
*“All but five of the 263 centenarians we interviewed **belonged to some faith-based community**. Denomination doesn’t seem to matter. Research shows that attending faith-based services four times per month will add 4-14 years of life expectancy.”* [1]	Religious attendances (top 3 religions)[10]	45th (third last)

*Substantially exceeds CDC & NHS guidelines for ‘heavy’ drinking.

Table S5 cont.

Claim	Observed Evidence	Rank in 47 prefectures
"Loved Ones First" [1] *"Successful centenarians in the blue zones put their families first. This means **keeping aging parents** and grandparents **nearby or in the home** (It lowers disease and mortality rates of children in the home too.). They **commit to a life partner** (which can add up to 3 years of life expectancy) and **invest in their children** with time and love (They'll be more likely to care for you when the time comes)."* [1]	Percent population married[11]	46th
	Percentage of households with members 65 and over[4,11]	46th
	Single parent households[11]	5th
	Percent Divorced** – male 50-59[11]	1st
	Percent Divorced** – female 50-59[11]	14th
	Percent never married – women 45-49[11]	22nd
	Percent never married – men 45-49[11]	14th
"Right Tribe" [1] *"The world's longest lived people chose – or were born into – social circles that supported healthy behaviors, Okinawans created "moais"–groups of five friends that committed to each other for life. Research from the Framingham Studies shows that **smoking, obesity, happiness**, and even **loneliness** are contagious. So the social networks of long-lived people have favorably shaped their health behaviors."* [1]	While the idea is essentially untestable..	
	Body Mass Index[2]	1st
	Suicide rate[4]	9th
	Suicide rate (2019) in over-65s[5]	4th
	Smoking - male[7,11]	43rd (37.4%)
	Smoking - female[7,11]	23rd (9.7%)

** The overall divorce rate in Okinawa has remained well above the national average in every measured[4] year since 1911.

Table S5 Bibliography

1. Buettner, D. Power 9 Reverse Engineering Longevity. https://www.bluezones.com/2016/11/power-9/# (2023).
2. Ikeda, N., Nakaya, T., Bennett, J., Ezzati, M. & Nishi, N. Trends and Disparities in Adult Body Mass Index Across the 47 Prefectures of Japan, 1975–2018: A Bayesian Spatiotemporal Analysis of National Household Surveys. *Front. Public Heal.* **10**, 830578 (2022).
3. Statistics Bureau of Japan. Survey on Time Use and Leisure Activities. https://www.stat.go.jp/english/data/shakai/2016/h28kekka.html (2016).
4. Vital statistics Division of the Ministry of Health Labor and Welfare. Trends in Vital Statistics by prefecture in Japan, 1899-1998. https://www.mhlw.go.jp/english/database/db-hw/vs_8/index.html.
5. Ministry of Internal Affairs and Communications Statistics Bureau. *System of Social and Demographic Statistics (SSDS)*. https://www.e-stat.go.jp/en/regional-statistics/ssdsview (2019).
6. Statistics Bureau of Japan. *Family Income and Expenditure Survey*. https://www.stat.go.jp/english/data/kakei/index.html (2016).
7. Ministry of Health Labor and Welfare. *National Health and Nutrition Survey Report.* https://www.nibiohn.go.jp/eiken/kenkounippon21/en/eiyouchousa/ (2016).
8. National Tax Agency. *Statistics on imposition of liquor tax 2016.* https://www.nta.go.jp/taxes/sake/tokei/mokuji.htm (2016).
9. Public Opinion Research Division of the NHK Broadcasting Culture Research Institute. *The Survey on Japanese Value Orientations, 1973-2008*. (2023) doi:10.34500/SSJDA.0814.
10. Ministry of Education Culture Sports Science and Technology. *Annual Statistics of Religion*. https://www.mext.go.jp/b_menu/toukei/chousa07/shuukyou/1262852.htm (2016).
11. Ministry of Internal Affairs and Communications Statistics Bureau. System of Social and Demographic Statistics (SDSS). https://www.e-stat.go.jp/en/regional-statistics/ssdsview (2023).

Supplementary Materials: Critical review

Issues raised in the analysis of extreme-age records suggest a remarkable situation: either that the anomalous pattern of extreme-age records can somehow be explained, or that extreme-age demography suffers extraordinary and systematic problems. Given the highly disruptive outcomes suggested by this analysis, the latter scenario requires some examination.

Indicators of systemic cultural problems in extreme-age demography

While individual experiences do not necessarily inform an entire research field, the author has independently uncovered multiple, recent, high-profile, and uncorrected issues that suggest deep-seated problems in the research culture of extreme-age demography. It is disturbing and noteworthy that that one individual can uncover broad-ranging problems, largely unaided, over such a short period.

First, a prominent Nature[1] paper on extreme ages rounded most of their old-age data down to zero[2] by accidentally using the wrong column in a life table (l_x not q_x) and treating zero as equal to one in a log transform. The authors then concluding that the rate of change in this data, which they had accidentally rounded off to 0 then turned into 1, was lower than the rate of change in unaffected data at younger ages[1]. Using the correct column of the same life table generates the opposite result[2] and contradicts the entire study. Despite these issues being revealed through open code and data, the paper remains unretracted and its many criticisms[2–7] ignored.

Barbi *et al.* then promoted an opposing view in Science[8] after picking the most empirically biased and worst-fit model possible from 861 alternatives[9,10]. Barbi *et al.* precipitated this remarkable coincidence by selecting ‘middle age’ as the ages 65-85 in their regression[9]. Fitting exactly the same regression to any other age range eliminates the claimed plateau, contradicts every research finding in the paper, and provides a far better fit of their own data[9]. The author highlighted these problems through a peer-reviewed, reproducible, and open analysis[9,10] and confronted the lead authors who explained their choices in a way that indicated research fraud. Again, however, the paper remains unretracted and uncorrected.

Further published explanations of late-life mortality patterns were then uncovered that are similarly problematic – such as work by Alvarez *et al.* that simply hides old-age data, but not the error bars, above the y-axis (see their figure 2) before eyeballing the remainder and claiming the data '*levels off'*[11]. Other research merely stretches credulity, such as the idea promoted by leading demographers that inbreeding is somehow good for survival[12,13], but some striking and unchallenged explanations push outright scientific racism.

Recently, for example, Vallin[14] claimed that extreme age records on Guadeloupe and Martinique are a positive outcome of the African slave trade's "*selection of the strongest individuals*" through "*the tremendous health selection effect of slavery*", and that such hypothetical effects are further enriched "*due to low immigration and high fertility among black people compared to white*"[14]. Vallin then proposes, without providing sources, that gaps in supercentenarian status between these two islands and Réunion exist because mortality rates from slavery was markedly lower in the Indian Ocean: a claim directly contradicted by contemporary and historical mortality estimates[15,16].

Even ignoring the neo-eugenics, such explanations make no historical or biological sense. Over 5.3 million people, 8% of the mainland French population, are descendants of the transatlantic slave trade[17,18]. This population is over ten times larger than Guadeloupe and Martinique combined, and includes hundreds of thousands of Guadeloupe- and Martinique-born internal migrants[19]. However, despite enjoying better healthcare, higher wages, lower poverty, and more complete vital registration records of mainland France, this larger slave-descendant population has apparently not produced a single supercentenarian. Given 16% of birth certificates in supercentenarian families are missing (Vallin Table 6) it seems the claim that slavery enriches for 'robust' individuals is just racist window-dressing of poor-quality data[14].

This view is unfortunately not an anomaly, but a continuation of openly propagated racism in extreme-age demography. A general argument that late-life mortality crossovers[20,21] develop when weaker individuals are 'selected out' at a young age[22], an argument that seems unlikely given high early-life mortality predicts higher late-life mortality[20], has long been co-opted for racist ideas. In the words of Markides and Machalek "*While errors in the data are recognized as a possible explanatory factor, a consensus seems to have developed that the* [old-age mortality] *crossover is indeed real and results from higher early mortality that removes less hardy blacks*"[23]: a casually racist idea that has enjoyed decades of support, often because the primary alternative is acknowledging residual undetected errors in demographic data.

The director of the GRG Robert Young typifies this thinking, with a desire to resurrect scientific racism through his "biological superiority hypothesis" developed at Georgia State University[24]. Young proposes that supercentenarians are more likely to be African-American because the GRG director can (apparently) see that African-Americans have '*thicker skin than Caucasian-Americans*' which '*protects their internal organs*' because "'*Black don't crack*'"[24]. His other explanations include that "*greater selection pressures, and thus faster evolution*" caused supercentenarian status in African-Americans because of the '*pruning*' benefits of slavery (as with Vallin) and the forced breeding of slaves[24] because "*often the most-healthy African American males were selected to be the 'stud' that would then be mated with the healthy females; less-healthy slaves would be discouraged from breeding*"[24].

Young extends his racist theory to explain why the USA has supercentenarians and Africa does not. He first states that racial 'pruning' must be moderate to be effective, as "*clearly, the Jewish population in Auschwitz had a very short life expectancy (akin to cutting a tree down, not pruning it)*'[24]. He then uses this argument as an explanation for the lack of African-resident supercentenarians[24] by equating living standards in modern Africa to being Jewish in Auschwitz.

It is hard to adequately capture the full scale of reprehensible ideas proposed by this key Supercentenarian researcher, who has been dramatically promoted since publication of these ideas, or by other individuals who use scientific racism as an explanation for longevity patterns. It is equally hard to explain why these ideas have been propagated for over 15 years, loudly and in public, without any comment or push-back from the demographic community.

Young is responsible for compiling one of the two major supercentenarian databases as director of the GRG, contributes to the IDL, and assembles many of the Guinness World Records for longevity using his qualitative assessment of supercentenarian data. That his qualitative assessment includes staring at African-American supercentenarians to see if they have 'thicker skin'[24] should ring alarm bells. In contrast, Young has been promoted to the top of the field *after* publishing such views.

Collectively, these issues suggest fundamental problems with the research culture of old-age demography. It is not that racist ideas and largely fake data are present in the literature, although these are extremely serious problems, but that such issues are met by a resounding absence of criticism or action. Excepting Dong *et al.*[1] none of the above dubious explanations, apparent misconduct, dubious models, or outright racism have been met with any perceptible response from the broader demographic community. That needs to change, rapidly, if we seek to understand the actual patterns of old-age human survival.

Indicators of fraud and error in population-wide data and 'Blue Zones'

Indications of a problematic research culture and a high error rate in extreme-age demography can be evidenced at all scales: broad population patterns, population studies, individual case evaluations, and even the basic logical evaluation of data and evidence.

Emblematic of these systematic issues are the 'Blue Zones'. Blue Zones (BZs) are five regions – in Okinawa, Costa Rica, California, Sardinia and Greek Ikaria – with claimed high concentrations of extreme age records[25–27]. The primary claim of BZ research is simple: that these regions have remarkable rates of achieving extreme old-age survival, or remarkable average survival in the case of Loma Linda and Costa Rica. In particular, it is routinely claimed these regions have unexpectedly high rates of reaching centenarian status[27].

The supposed enrichment for extreme-age survival in BZs is then subject to a host of secondary claims, each aimed at explaining the primary pattern of extraordinary longevity. Old-age survival in the BZs is supposed to result from diverse causes such as 'moderate' drinking at twice the NHS heavy-drinking guidelines[28], plant-based diets[25–27] and inbreeding[12,13]. Most prominent amongst these secondary claims is the idea that nine specific behaviours, the 'power nine'[29], underpin the achievement of extreme longevity (see Table S5).

The BZ longevity claims have been made, very publicly, for decades and have seen widespread support in the general demographic community. However, evaluation of the central BZ claims to extreme longevity, and secondary claims of their supposed cause, reveal the general neglect of error processes as a potential generative factor in remarkable age records, and more broadly reveal a lack of critical appraisal of extreme-age claims.

Consider the first BZ, the one most amenable to measurement: extreme longevity in Okinawa. In Okinawa the central claim of remarkable survival to extreme ages was established in 2004, before 82% of Japanese records were debunked. Okinawa then had the highest number of centenarians per 90-99-year-old of any Japanese prefecture and was world-famous for remarkable longevity.

However, according to the statistics bureau of Japan, Okinawa also has the highest murder rate per capita, the worst over-65 dependency ratio, the second-lowest median income, the highest percentage (60%) of older people on welfare[30], and the highest unemployment rate of all 47 Japanese prefectures[31]. Despite prior claims of dietary benefits based on high vegetable[29], oily fish, and sweet potato consumption[26,27], Okinawa has the lowest per capita intake of sweet potato, fruits, vegetables, seafood, taro, shellfish, root vegetables, pickled vegetables, and oily fish such as sardines and yellowtail[32] of any prefecture.

Mortality rates in Okinawa 'cross over' after age 50, such that older individuals and cohorts have age-specific mortality rates far below the national average[33]: a pattern indicative of unreliable data and misreported ages[20,34]. Okinawa also has the second-highest per capita intake of beer and the highest per-capita intake of KFC, the highest child poverty rate at 36% (15% higher than the next-highest prefecture)[35], the most 'flophouses' and shotgun weddings[36], and according to USDA estimates Okinawan residents each consume an average 14 cans of SPAM per year[37]. Okinawa has the second-lowest minimum wage (by one yen)[31,36], the lowest household savings[31,36], the highest percentage of over-65s on income assistance[31,36], the highest poverty rate[35], and the worst average body mass index (BMI) of all 47 prefectures[36,38].

These rankings do not represent a recent sudden shift away from 'traditional lifestyles', and have not changed substantially for extended periods before, during, or after the BZ surveys[35]. For example, the body mass index of the over-75 residents of each prefecture has been continuously assessed since 1975, with the first measurement comparing residents born in 1900 or before to every other 75+ year old in Japan[38]. Every single year, over-75 Okinawans have retained the worst BMI [38]. Similar patterns hold across other nutritional indicators: in no case, at any point, have the massive population-representative surveys of Okinawa supported the claimed vegetarian dietary patterns pushed by BZ proponents[27,29,39].

These results become even more questionable given the capacity to directly assess each of the 'power nine' claims for BZ longevity[29] in Okinawa (Table S5). All nine claimed drivers of extreme longevity are assessable through data measured by the government of Japan. The 'power nine' claims are directly contradicted in every single case, usually through population-representative surveys of hundreds of thousands of people, with levels of inaccuracy that border on farce (Table S5).

If nothing else, these surveys reveal both astounding attention to detail by the Japanese government, and the scale of data that has been overlooked by BZ researchers. The older residents of Okinawa are not filled with purpose or "Ikigai" at remarkable rates: over-65 Okinawans have the 4th-highest suicide rate in Japan [40]. Older Okinawans do not **"**grow gardens"[29]: they self-report the lowest rate of gardening in the country, beating only the apartment-dominated Tokyo and Osaka megacities. Okinawans do not eat "Meat … only five times per month"[29] in 3-4oz. servings, which would total 5.1-6.8kg a year: they consume well over 40kg of meat a year[32,41] without including seafood. Nor do Okinawans overwhelmingly "belong to some faith-based community"[29]: they are 93.4% atheist[42], the most irreligious population in Japan, ranking third-last in the country for religious attendance[43].

What is most astounding about such claims is not that they have been used to generate profit by selling the fountain-of-youth repackaged for a modern wellness audience. Rather, it is that the demographic community, after hundreds of papers on the subject, does not seem to have read or cited the world-leading, comprehensive, population-representative surveys[35,38,42–45] that refute the BZ claims.

In Okinawa, the anomalous pattern of centenarians is even compounded by other error-generative processes. Japanese birth and marriage records are not generated by a central bureaucracy, but instead are hand-recorded by family members as ‘Koseki’ documents, which are then filed in local town halls and government offices. This combination of citizen self-reporting and government filing allows the propagation of errors without requiring fraud. On top of this self-report data, the large-scale bombing and invasion of Okinawa involved the destruction of entire cities and towns, obliterating around 90% of Koseki records[33] with almost universal losses outside of Miyako and the Yaeyama archipelago[46]. Post-war Okinawans subsequently requested replacement documents, described from memory[46], through a US-led military government that largely spoke no Japanese and which used a different calendar. As observed by Poulain[33] the number of replacement documents issued, a proxy measure of American bombing and shelling intensity, predicts 79% of the variation in centenarian status across Okinawa.

Despite the enrichment of regular error-generating processes through American firebombing, Okinawa is not an atypical case amongst the BZs. It is, rather, just the best-surveyed BZ region. Every proposed BZ displays patterns that suggest a dominant role of error, fraud, and (to phrase it generously) researcher degrees of freedom in explaining the distribution of extreme-age records.

As detailed in the manuscript, Sardinia, Ikaria, and Okinawa represent deprived regions of a rich high-welfare state. These BZ regions rank amongst the least educated, poorest, highest-crime and least healthy regions of their respective countries (Table S1-S2, S4). Therefore, it may be hypothesized that the relatively low incomes and high poverty rates of these populations generate age-reporting errors and pension fraud, and therefore remarkable age records.

The two remaining blue zones, Loma Linda and the Nicoya Peninsula, are considered exceptional due to their high average longevity rather than the presence of the oldest-old[25,26].

As such, these claims are not relevant to assessments of supercentenarian status, yet their analysis reinforces the evidence for broader problems in extreme longevity research.

Loma Linda is a Californian suburb containing just 23,000 people, designated as a BZ because of a supposed average lifespan of 86 years for females and 83 years for males. Even if taken at face value, this average lifespan is matched or exceeded by the 125 million citizens of Japan, the seven million citizens of Hong Kong, and the seven and a half million citizens of Singapore[47]. However, when assessed independently by the Centers for Disease Control (CDC) the five small-area survey tracts covering Loma Linda instead have an average life expectancy of 76 to 81 years[48]: the 27th to 75th percentiles of US life expectancy (Fig S6). This means, at best, the independent CDC estimates rank Loma Linda as the 16,101st most long-lived neighborhood in the USA (Fig. S6; S1 code). As such, it is again unclear why the lifespan of this community has been considered remarkable.

Like most BZs, Loma Linda is not a standard census tract or statistical area, but a custom-selected region within a larger geographic area: the largest US county of San Bernadino, which has an average lifespan of 78 years[48,49]. The Sardinian BZ was similarly delineated: by drawing circles on a map[12,25] and cut across the two standard, independently surveyed regions in Italy that have the lowest and sixth-lowest probability of survival to age 55[50]. The Nicoya peninsula in Costa Rica, where independent estimates are currently not available, is also a non-standard region cut from several independent census units of moderate life expectancy. Even worse, the goalposts have frequently moved on the BZ locations. For example, the Costa Rican BZ became two regions, and now one of the two Costa Rican BZs has now disappeared, and the other has shrunk by ~90% in population, while a third another has apparently appeared 300km away near Nicaragua[51].

Given the lower life expectancy of their encompassing regions and the uncertain basis of their ascertainment, it remains somewhat questionable that these custom-selected and frequently redefined regions should be considered valid outliers for human longevity.

Indicators of error and fraud in health studies

Like anomalous population-scale patterns, indicators of poverty and fraud and contra-indications of health are regularly ignored or downplayed in targeted studies of extreme age. For example, smoking rates of *e.g.* 17-50% and illiteracy rates of 50-80% are often observed in samples of the oldest-old[52,53]. Surveying the BZ of Ikaria, Chrysohoou *et al.* observed that the Ikarian oldest-old (ages 80+) have a below-median income in over 95-98% of cases, moderate to high alcohol consumption (5.1-8.0 L/ year), a 10% illiteracy rate, an average 7.4 years of education, and a 99% rate of smoking in men[54].

Seemingly more moderate rates of smoking occur in the Tokyo study of exceptional longevity, where "*few*" centenarians smoked, some 15.4%, of the mostly (78%) female centenarian sample were current smokers[55]. However, according to the national statistics bureau of Japan, only 3.9% of Japanese women and 19.3% of men over the age of 80 are smokers[56]. Tokyo centenarians therefore smoke at around twice the rate that could be expected in a younger, 80+ year old cohort with an identical sex ratio. Likewise, 80% of the 'exceptional' health-status centenarian population were daily drinkers, followed by 49% of the 'normal' and less than 40% of the 'frail or fragile' centenarians, resulting in "*a* [significant] *positive relationship between drinking habits and functional status*"[55]. In contrast with these figures, Japanese government surveys estimate only 2.8% of women and 23% of men aged 80+ drink every day[56]. Daily drinking peaks at 36.7% in men aged 60-69, the heaviest-drinking cohort in Japan[56]. As such, Tokyo centenarians drink at higher rates than any other age group, and smoke at rates equal to a 45-year younger population[56].

In the USA only 8.4% of general population over the age of 65 currently smoke[57], and in Europe 4.1% population over the age of 75 currently smoke[58] and 31% formerly smoked: figures that continue to fall with age due to two-fold higher mortality rates in smokers[59,60]. However, in the US and Europe individuals over the age of 100 often smoke and drink at higher rates than any other age group: in one US centenarian study 60% of people over age of 95 were former smokers[61], compared to just 25% of individuals over the age of 65 in the broader USA.

This latter study is notable as it compares lifestyle factors in the oldest-old to earlier surveys of the same cohort. This comparison revealed that centenarians have similar or worse body mass index, rates of physical activity, smoking and alcohol consumption than the younger baseline population from which they were drawn[61]: a striking equivalence given the baseline population was 35 years younger. For these two samples of the same cohort to have a similar BMI, smoking, and drinking rates would require that these factors were somehow *not* risk factors for mortality across the intervening 35 years. Alternately, it is possible that smoking, obesity, and drinking cause mortality but that the decline in these factors with age is offset by rising vital statistics error rates with age[10].

Suitable alternatives explanations may be that these centenarians had worse health, lower exercise and much higher rates of drinking and smoking than the general population, or that they constitute a population of younger individuals with lifestyle patterns and smoking rates that reflect higher rates of fraud, illiteracy, or chronic government neglect.

While centenarian studies likely contain cleaner samples than supercentenarian populations, they often reveal broad assumptions on sample integrity. For example, the benchmark New England Centenarian study masks a widespread lack of birth certificates. Only 30% of the total sample had an official birth document discovered[62]. Of the 80 'centenarians' who were reported to be alive in 1995, roughly 8% were an entire century younger than reported, thirteen (16%) had reported a false birthday, and eleven refused to enroll but were investigated anyway[63]. Even in this post-screening sample, just 45 of the 67 investigated centenarians had a birth certificate. This rate of possessing documents is, inexplicably, substantially lower than the background population, which achieved complete coverage for birth certificates in all New England states[64] by 1897. For the remaining individuals who lacked birth certificates, age validation was instead carried out using documents including "*military certificates, an old passport, school report card, family bible, and baptismal or other church certificate*"[63].

Missing birth certificates are also common in other samples of the oldest-old. For example, in the IDL database none of the 797 supercentenarians from the USA is listed as having a birth certificate[65] while 41% of US cases had the possibility of a birth certificate explicitly ruled out. Even an intensive search of 297 US supercentenarians by Stone[66] only managed, as an absolute best, 43 birth certificates (of which three were inaccurate, generating 13% coverage). This absence of birth registration data extends to the larger sample of supercentenarians and SSCs across the IDL data: only 19% of all supercentenarians, and 20% of the 'exhaustively' validated supercentenarians, have either an original or copy of a birth certificate[65] (Supplementary Code 1). Across all individuals in the IDL database, only 6.6% have an original birth certificate listed, and 74% of cases have no listed birth documents of any kind. These low rates of birth certification have not addressable in intensive follow-on searches: the most intensive search of vital registration data, conducted in US supercentenarians, discovered 43 birth certificates amongst 550 starting candidates[66].

Death certificates are also largely missing. The first follow-up survey of the population-representative NHANES I cohort, representative of the US population, found 3.8% of decedent men and 5.7% of decedent women did not have death certificates and remained alive on paper while actually dead[67]. However, for supercentenarians listed as dying in the USA under the IDL database, only seven are listed as having death certificates: some 98.7% do not have listed death certificates. In total only 15% of supercentenarians and 8% of SSCs in the total IDL database are listed as having death certificates[65].

Indicators of error and fraud in individual cases

The potential for bias during case validation is of marked importance, especially in the absence of such fundamental evidence as birth certification. However, individual case studies often highlight the potential for personal judgement during age validation. These problems extend beyond cases detailed in the manuscript, such as the century-long acceptance, and then rejection of the Taché investigation.

For example, the world's oldest man, Jiroemon Kimura, is now widely considered to be a valid supercentenarian case. Evidence for case validation used primary source material collected and vetted by biological relatives of Kimura, who stood to gain from his validation[68]. As such, Jiroemon Kimura has at least three wedding dates to the same wife, was conscripted to the same military three times in four years despite the mandatory conscription period being three years long[69], has three dates of graduation from the same school, and has at least three birthdays[68]. In addition, for the first 20 years of his life all of Kimura's birthdates and school records are actually those of 'Kinjiro Miyake', a name whose connection to Kimura is not attested to by an official document, but by an unofficial hand-written note from a Korean mail and telephone company[68]. Under interview, Kimura explained one of his extra birthdays in a way that was "*not feasible*"[68], and Gondo *et al.* concluded the birth date had been deliberately forged[68].

Gondo *et al.* then concluded the case validation by assuming any conflicting official records were mistakes and, amongst the diverse birth, wedding, conscription, and graduation dates, selected those they felt were accurate. The paper describing the world's oldest man then lists a new birthday for Kimura as 1987, and his brother's birthday as 1985: two undetected and uncorrected errors that rather undermine the claimed capacity of demographers to spot erroneous dates[68]. The study concludes that "no critical discordances were discovered" [68] and the case is still considered valid[70,71].

The validity of the Kimura case has been widely accepted under the assumption that age discrepancies can be discarded through the qualitative judgement of demographers. This reliance on qualitative judgement during case validation reflects general practice. For example, concerns surrounding the validation of ages are often met with a response that biographical inconsistencies, detected during interview by a demographer, will result in cases being removed from the record.

However, this sentiment can be difficult to reconcile with observed practice. For example, former smoker and occasional drinker Adele Dunlap, who "*ate anything she wanted*" and "*never went out jogging or anything*" was validated by the GRG and IDL as the oldest woman in the USA. This was despite Dunlap consistently maintaining under interview that she was a decade younger: when "asked how it felt to be 113, Dunlap… looked her questioner in the eye and answered: *'I'm 104'*"[72]. Despite consistently maintaining her age was incorrect, for years, Dunlap remains validated as a supercentenarian on the basis of documents contradicted by her own testimony.

Further indications of the role of opinion, especially when ignoring contra-indications of health, are highlighted by other prominent cases. When interviewing the then-oldest man and woman in the world – Christian Mortensen and Jeanne Calment – demographers from the Max Planck Institute for Demographic Research issued the contradictory statement that Jeanne Calment smoked both one and two cigarettes a day for an entire century[73], followed by the justification that Calment "*possibly did not inhale at all*"[73]. It was then observed that, from age 20 to age 117, Christian Mortensen smoked "*mainly a pipe and later on cigars, but almost never cigarettes… he had also chewed tobacco…but never inhaled*"[73]. Why two people might choose to smoke for an accumulated 190 years, but never inhale, was not questioned.

Similar opinions aimed at explaining the questionable health habits of the oldest old are widely considered satisfactory. A notable fraction of supercentenarians smoke and drinking, yet these anomalous health patterns are routinely downplayed. Of the five oldest men ever recorded, Kimura and Mortensen are detailed above, Emiliano Del Toro (3rd) smoked for 76 years, Mathew Beard (4th) was busted for drink-driving at age 90, and Walter Breuning (5th) smoked cigars until he was 108. The oldest man in the UK stated his secret to health as "*cigarettes, whiskey and wild, wild women*"[74], while the former oldest man in the USA smoked 12-18 cigars and drank alcohol every day, which routinely started with "*a little bourbon in* [his] *coffee*"[75]. At least three of the ten oldest women drank every day, two smoked every day, and four are of unknown smoking status, while Jeanne Calment smoked daily, drank daily, and ate around a kilogram of chocolate a week.

These instances of poor lifestyle choices are not rare but constitute a substantial fraction of all supercentenarian cases. As summarized by Coles, supercentenarians lifestyle are characterized by "*heavy smoking, heavy drinking, or both, failure to exercise on a regular basis, and no conscious effort to eat nutritiously*"[76]. Instead of prompting skepticism, under the relatively safe assumption that smoking, drinking, poverty, lack of exercise, poor nutrition, and illiteracy should not enrich for remarkable longevity records, these contra-indications of survival are routinely ignored or downplayed. For example, the study by Chrysohoou *et al.* concluded that "*physical activity, dietary habits, smoking cessation, and midday naps*" predict extreme longevity in Ikaria: a conclusion that questionably re-shapes past smoking status as a positive indicator of survival[54].

Finally, a troubling aspect of case validation is not restricted to the questionable, and largely irreproducible qualitative choices made by extreme-age demographers. Nor is it the repeat capacity to overlook extraordinarily high rates of later-invalidated cases, or the overt racism of individuals responsible for case validation[24]. Instead, the rather simpler challenge is that cases are almost never subject to even cursory investigation, let alone intensive interview. For example, the IDL now lists only 365 validations for 9,878 French SSCs, of which none have an original birth certificate [65]. It is unclear how many of these 365 individuals were interviewed, even under the irreproducible and qualitative-driven framework of a field that does not encourage pre-registered findings or the publishing of non-results.

Summary

Individually these issues raised here may be, in some cases, answerable. Yet in aggregate such patterns reveal a devastating reality. In extreme-age demography, not only are critical questions not answered: they are not asked.

The list is extensive. Why have population and lifestyle surveys of Japan, which cover over 94% of households, not been used to evaluate population and lifestyle patterns in Japan after twenty years? Why are a Nature paper that rounded its data off to zero, and a Science paper using the most empirically biased model possible, unretracted? Why do demographers trust a single metric of age – documents – when documents have routinely generated observed error rates above 50%, and exciting alternatives are being developed? How might documentary validation or interview methods ever *reproducibly* solve the 'Italian sibling' problem? Why are leading demographers allowed to state that African-American people live longer because they 'have evolved 10% more genes'[24], or are more 'robust' due to 'faster breeding'[14], whilst being promoted to senior roles in major institutes, without a solitary word of criticism?

Instead of being asked, such questions are ignored, basic critical analysis is derided, and fundamental problems written in reproducible code have no effect. Such responses may be effective in removing criticism and have resulted in the author leaving the field. But the future of demography is saddening indeed if it cannot bring itself to answer these questions alone.

I may leave demography with a final question. Half of all births in India, for example, have no birth certificate: these people are excluded *de facto* from age-stratified medical and epidemiological models, because we don't know their age. How much better could our collective medical, demographic, and epidemiological models perform, if we replaced document-measured ages with accurate biometric measurements of age across every population that, like India, suffers from systematically unreliable or absent documents[77,78]?

If the 'Italian sibling' problem were taken seriously, and solved, we may find out. Given the reaction and conduct of the demographic community so far, however, hope for such an answer seems remote.

Bibliography

1. Dong X, Milholland B, Vijg J. Evidence for a limit to human lifespan. Nature. 2016;538:257–9.

2. Newman SJ, Easteal S. The dynamic upper limit of human lifespan. F1000Research [Internet]. 2017;6. Available from: https://f1000research.com/articles/6-569/v2

3. Rozing MP, Kirkwood TBL, Westendorp RGJ. Is there evidence for a limit to human lifespan? Nat 2017 5467660 [Internet]. 2017 [cited 2023 Dec 21];546:E11–2. Available from: https://www.nature.com/articles/nature22788

4. De Beer J, Bardoutsos A, Janssen F. Maximum human lifespan may increase to 125 years. Nat 2017 5467660 [Internet]. 2017 [cited 2023 Dec 21];546:E16–7. Available from: https://www.nature.com/articles/nature22792

5. Lenart A, Vaupel JW. Questionable evidence for a limit to human lifespan. Nat 2017 5467660 [Internet]. 2017 [cited 2023 Dec 21];546:E13–4. Available from: https://www.nature.com/articles/nature22790

6. Hughes BG, Hekimi S. Many possible maximum lifespan trajectories. Nat 2017 5467660 [Internet]. 2017 [cited 2023 Dec 21];546:E8–9. Available from: https://www.nature.com/articles/nature22786

7. Brown NJL, Albers CJ, Ritchie SJ. Contesting the evidence for limited human lifespan. Nat 2017 5467660 [Internet]. 2017 [cited 2023 Dec 21];546:E6–7. Available from: https://www.nature.com/articles/nature22784

8. Barbi E, Lagona F, Marsili M, Vaupel JW, Wachter KW. The plateau of human mortality: Demography of longevity pioneers. Science (80-). 2018;360:1459–61.

9. Newman SJ. Plane inclinations: A critique of hypothesis and model choice in Barbi et al. PLoS Biol. 2018;16:e3000048.

10. Newman SJ. Errors as a primary cause of late-life mortality deceleration and plateaus. PLoS Biol. 2018;16:e2006776.

11. Alvarez J-A, Villavicencio F, Strozza C, Camarda CG. Regularities in human mortality after age 105. PLoS One [Internet]. 2021 [cited 2021 Aug 31];16:e0253940. Available from: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0253940

12. Poulain M, Pes GM, Grasland C, Carru C, Ferrucci L, Baggio G, et al. Identification of a geographic area characterized by extreme longevity in the Sardinia island: The AKEA study. Exp Gerontol. 2004;

13. Caselli G, Pozzi L, Vaupel JW, Deiana L, Pes G, Carru C, et al. Family clustering in Sardinian longevity: A genealogical approach. Exp Gerontol. 2006;41:727–36.

14. Vallin J. Why are supercentenarians so frequently found in French Overseas Departments? The cases of Guadeloupe and Martinique. Genus [Internet]. 2020 [cited 2023 Dec 19];76:1–17. Available from: https://genus.springeropen.com/articles/10.1186/s41118-020-00086-3

15. Fillot J-M. La Traite des esclaves vers les Mascareignes au XVIIIe siècle [Internet]. Paris: Office de la Recherche Scientifique et Technique Outre-Mer; 1974. Available from: https://horizon.documentation.ird.fr/exl-doc/pleins_textes/pleins_textes_6/Mem_cm/06931.pdf

16. Deschamps H. Histoire de Madagascar. 4th ed. Paris: Berger-Levrault; 1972.

17. Karima L. North African Diaspora in France and Colonial Legacies. In: Hewitt N, Geary D, editors. Diasporas Movements Cult. Nottingham: Critical, Cultural and Communication Press; 2007.

18. Varus R. The Role of France and the French in North Africa. Annu Am Acad Polit Soc Sci. 1956;306:17–25.

19. Condon S, Ogden P. Emigration from the French Caribbean: the Origins of an Organized Migration. Int J Urban Reg Res [Internet]. 1991 [cited 2023 Dec 21];15:505–23. Available from: https://onlinelibrary.wiley.com/doi/full/10.1111/j.1468-2427.1991.tb00655.x

20. Coale AJ, Kisker EE. Mortality crossovers: Reality or bad data? Popul Stud (NY). 1986;

21. Pearl R. The Biology of Death. Osterhout JLATHMAWJ V, editor. J.B. Lippincott Co. Washington Square Press, Philadelphia; 1922.

22. Vaupel JW, Manton KG, Stallard E. The Impact of Heterogeneity in Individual Frailty on the Dynamics of Mortality. Demography [Internet]. 1979;16:439. Available from: http://link.springer.com/10.2307/2061224

23. Markides KS, Machalek R. Selective survival, aging and society. Arch Gerontol Geriatr. 1984;3:207–22.

24. Young RD. African American Longevity Advantage: Myth or Reality? A Racial Comparison of Supercentenarian Data. Gerontol Theses [Internet]. 2008 [cited 2024 Jan 1]; Available from: https://scholarworks.gsu.edu/gerontology_theses/10

25. Poulain M, Herm A, Pes G. The blue zones: Areas of exceptional longevity around the world. Vienna Yearb Popul Res. 2013;

26. Buettner D. The Blue Zones: lessons for living longer from those who've lived the longest. Washington, D.C.; 2010.

27. Jeter C. Living to 100: Secrets of the Blue Zones [Internet]. Netflix; 2023. Available from: https://www.netflix.com/gb/title/81214929

28. National Institute on Alcohol Abuse and Alcoholism (NIAAA). Drinking Levels Defined [Internet]. 2023 [cited 2023 Dec 3]. Available from: https://www.niaaa.nih.gov/alcohol-health/overview-alcohol-consumption/moderate-binge-drinking

29. Buettner D. Power 9 Reverse Engineering Longevity [Internet]. 2023 [cited 2023 Dec 3]. Available from: https://www.bluezones.com/2016/11/power-9/#

30. Japanese Institute for Health and Welfare. Number of aged persons assisted by livelihood protection

per 1,000 persons 65 years old and over. [Internet]. [cited 2024 Jan 2]. Available from: https://www.e-stat.go.jp/en/stat-search/file-download?statInfId=000040019850&fileKind=0

31. Director-General for Statistics and Information Policy, Ministry of Health L and W (organisation). Vital Statistics [Internet]. 2019 [cited 2019 Feb 7]. Available from: https://www.mhlw.go.jp/english/database/db-hw/index.html

32. Ministry of Health Labor and Welfare. National Health and Nutrition Survey Report. [Internet]. 2016. Available from: https://www.nibiohn.go.jp/eiken/kenkounippon21/en/eiyouchousa/

33. Poulain M. Exceptional longevity in Okinawa: A plea for in-depth validation. Demogr Res. 2011;

34. Preston SH, Elo IT, Stewart Q. Effects of age misreporting on mortality estimates at older ages. Popul Stud (NY). 1999;

35. Tomuro K. Trends Observed in Poverty Rates, Working Poor Rates, Child Poverty Rates and Take-Up Rates of Public Assistance Across 47 Prefectures in Japan [Internet]. Yamagata Univ. Fac. Humanit. Res. Annu. Rep. Yamagata; 2016. Available from: https://yamagata.repo.nii.ac.jp/?action=repository_uri&item_id=4067&file_id=17&file_no=1

36. Statistics Japan. e-Stat: Portal Site of Official Statistics of Japan [Internet]. 2020 [cited 2020 Jan 30]. Available from: https://www.e-stat.go.jp/en

37. Motomura C. Global Agricultural Information Network Regional Report - Okinawa. 2014.

38. Ikeda N, Nakaya T, Bennett J, Ezzati M, Nishi N. Trends and Disparities in Adult Body Mass Index Across the 47 Prefectures of Japan, 1975–2018: A Bayesian Spatiotemporal Analysis of National Household Surveys. Front Public Heal. 2022;10:830578.

39. Buettner D, Skemp S. Blue Zones: Lessons From the World's Longest Lived. Am. J. Lifestyle Med. 2016.

40. Ministry of Internal Affairs and Communications Statistics Bureau. System of Social and Demographic Statistics (SSDS) [Internet]. 2019. Available from: https://www.e-stat.go.jp/en/regional-statistics/ssdsview

41. Statistics Bureau of Japan. Family Income and Expenditure Survey [Internet]. Tokyo; 2016. Available from: https://www.stat.go.jp/english/data/kakei/index.html

42. Public Opinion Research Division of the NHK Broadcasting Culture Research Institute. The Survey on Japanese Value Orientations, 1973-2008. 2023.

43. Ministry of Education Culture Sports Science and Technology. Annual Statistics of Religion [Internet]. 2016. Available from: https://www.mext.go.jp/b_menu/toukei/chousa07/shuukyou/1262852.htm

44. Statistics Bureau of Japan. Survey on Time Use and Leisure Activities [Internet]. 2016. Available from: https://www.stat.go.jp/english/data/shakai/2016/h28kekka.html

45. Vital statistics Division of the Ministry of Health Labor and Welfare. Trends in Vital Statistics by prefecture in Japan, 1899-1998 [Internet]. [cited 2023 Dec 3]. Available from: https://www.mhlw.go.jp/english/database/db-hw/vs_8/index.html

46. Go M. Special feature: Thinking about modern society from Okinawa. Geppō shihōshoshi [Internet]. 2004;June. Available from: https://web.archive.org/web/20070825091736/http://www.shiho-shoshi.or.jp/web/publish/geppou/200406/2004_06_021.html

47. CIA. CIA World Factbook. World Factb 2013. 2017;

48. Arias E, Escobedo LA, Kennedy J, Fu C, Cisewski JA. U.S. small-area life expectancy estimates project : methodology and results summary [Internet]. (U.S.) NC for HS, editor. Hyattsville, MD; 2018. Available from: https://stacks.cdc.gov/view/cdc/58853

49. Chetty R, Stepner M, Abraham S, Lin S, Scuderi B, Turner N, et al. The Association Between Income and Life Expectancy in the United States, 2001–2014. Jama [Internet]. 2016;315:1750–66. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4866586/

50. Elders.Stat database: Istituto Nazionale di Statistica [Internet]. Ist. Naz. di Stat. 2019 [cited 2019 Feb 7]. Available from: http://dati-anziani.istat.it/

51. Rosero-Bixby L. The vanishing advantage of longevity in Nicoya, Costa Rica: A cohort shift. Demogr Res. 2023;49.

52. Afonso RM, Ribeiro O, Vaz Patto M, Loureiro M, Loureiro MJ, Castelo-Branco M, et al. Reaching 100 in the Countryside: Health Profile and Living Circumstances of Portuguese Centenarians from the Beira Interior Region. Curr Gerontol Geriatr Res. 2018;2018.

53. Kwon IS, Kim C-H, Ko HS, Cho S Il, Choi YH, Park SC. Risk factors of cardiovascular disease in Korean exceptional longevity. J Korean Geriatr Soc. 2005;9:251–65.

54. Chrysohoou C, Panagiotakos DB, Siasos G, Zisimos K, Skoumas J, Pitsavos C, et al. Sociodemographic and lifestyle statistics of oldest old people (>80 years) living in ikaria island: The ikaria study. Cardiol Res Pract. 2011;

55. Gondo Y, Hirose N, Arai Y, Inagaki H, Masui Y, Yamamura K, et al. Functional status of centenarians in Tokyo, Japan: Developing better phenotypes of exceptional longevity. Journals Gerontol - Ser A Biol Sci Med Sci. 2006;

56. Nakamura T, Tano J, Arai N, Jimbou K. Summary Report of Comprehensive Survey of Living Conditions 2016 [Internet]. 2017. Available from: https://www.mhlw.go.jp/english/database/db-hss/dl/report_gaikyo_2016.pdf

57. Centres for Disease Control and Prevention. Current Cigarette Smoking Among Adults in the United States | CDC [Internet]. Smok. Tob. Use Fact Sheet. [cited 2024 Jan 12]. Available from: https://www.cdc.gov/tobacco/data_statistics/fact_sheets/adult_data/cig_smoking/index.htm#anchor_1549

902047693

58. EUROSTAT. Daily smokers of cigarettes by sex, age and educational attainment level: table hlth_ehis_sk3e [Internet]. 2014 [cited 2024 Jan 12]. Available from: https://ec.europa.eu/eurostat/databrowser/view/hlth_ehis_sk3e__custom_9320277/default/table?lang=en&page=time:2019

59. Gellert C, Schöttker B, Brenner H. Smoking and All-Cause Mortality in Older People. Arch Intern Med. 2012;

60. Taylor DH, Hasselblad V, Henley SJ, Thun MJ, Sloan FA. Benefits of smoking cessation for longevity. Am J Public Health. 2002;

61. Rajpathak SN, Liu Y, Ben-David O, Reddy S, Atzmon G, Crandall J, et al. Lifestyle factors of people with exceptional longevity. J. Am. Geriatr. Soc. 2011.

62. Sebastiani P, Perls TT. The genetics of extreme longevity: Lessons from the new england centenarian study. Front Genet. 2012;

63. Perls TT, K B, Freemen M, Alpert L, Silver MH. Age Validation in the New England Centenarian Study. Valid Except Longev. Odense; 1999.

64. Farnsworth Riche M, Benton B, Schnelder PJ, Norton AJ. Population of the States and Counties of the United States: 1790 to 1990 [Internet]. Washington D.C.; 1996. Available from: https://www.census.gov/population/www/censusdata/PopulationofStatesandCountiesoftheUnitedStates1790-1990.pdf

65. Vaupel J, Vallin J, Meslé F, Robine J-M, Jdanov DA, Grigoriev O. The International Database on Longevity [Internet]. 2020 [cited 2020 Feb 13]. Available from: https://www.supercentenarians.org/Home/Expand_IDL

66. Stone LF. Studies in the demography of supercentenarians in the United States [Internet]. University of Pennsylvania; [cited 2023 Dec 21]. Available from: https://www.proquest.com/docview/304711712?pq-origsite=gscholar&fromopenview=true&sourcetype=Dissertations & Theses

67. Madans JH, Kleinman JC, Cox CS, Barbano HE, Feldman JJ, Cohen B, et al. 10 Years after NHANES I: Report of initial followup, 1982-1984. Public Health Rep. 1986;

68. Gondo Y, Hirose N, Yasumoto S, Arai Y, Saito Y. Age verification of the longest lived man in the world. Exp Gerontol. 2017;

69. Takata Y, Ogawa G. Conscription system in Japan. New York, NY: Oxford University Press; 1921.

70. Gerontology Research Group [Internet]. Table A Verif. Supercentenarians. 2019 [cited 2019 Jun 2]. Available from: http://www.grg.org/SC/SCListsTables.html

71. Robine J-M, Gampe J, Vaupel J. IDL, the International Database on Longevity. Gerontology. 2005;1–

30.

72. Levin J. Oldest American is 113-year-old Jersey girl [Internet]. USA Today. 2016 [cited 2020 Feb 11]. Available from: https://www.usatoday.com/story/news/nation-now/2016/07/21/oldest-american-year-old-jersey-girl/87373424/

73. Maier H, Gampe J, Jeune B RJ and VJ. Supercentenarians. Demogr Res Monogr. 2010;7.

74. Smith D. RAF flypast for Europe's oldest man [Internet]. Guard. 2008 [cited 2020 Feb 11]. Available from: https://www.theguardian.com/uk/2008/jun/01/military.firstworldwar

75. Christie J. The oldest living US veteran, age 108, credits drinking whiskey and smoking a dozen cigars every day for his long life [Internet]. Dly. Mail. 2014 [cited 2020 Feb 11]. Available from: https://www.dailymail.co.uk/news/article-2830888/The-oldest-living-veteran-smokes-drinks-whiskey-day-drives-does-yard-work-s-108-YEARS-OLD.html

76. Coles LS. Aging: The Reality: Demography of Human Supercentenarians. Journals Gerontol Ser A Biol Sci Med Sci. 2004;59:B579–86.

77. Cappa C, Gregson K, Wardlaw T, Bissell S. Birth registration: A child's passport to protection. Lancet Glob. Heal. 2014.

78. Hambright TZ. Comparison of information on death certificates and matching 1960 census records: age, marital status, race, nativity and country of origin. Demography [Internet]. 1969 [cited 2024 Jan 2];6:413–23. Available from: https://pubmed.ncbi.nlm.nih.gov/21279795/