Evaluating Causes of Geographic Disparities in Life Expectancy Using Partitioning Approach

I. Akushevich[1], J. Kravchenko[2], A.Yashkin[1], A.I.Yashin[1]

[1]Biodemography of Aging Research Unit, Social Science Research Institute, Duke University, Durham, NC
[2]Department of Surgery, Duke University Medical Center, Durham, NC

Abstract

Persistent interregional disparities in health outcomes and life expectancy (LE) represent an important problem for the U.S. health system. We developed a computational approach involving four scenarios on how these disparities could be evaluated using health measures extracted from Medicare data: regions with lower LE exhibit i) higher disease incidence, ii) worse patient survival, iii) higher multimorbidity, and/or iv) worse health state of individuals aged 65 (time of entry into the Medicare system). Our four-step computational approach includes: i) ranking diseases according their effect on mortality, ii) evaluation of the contribution of the most lethal disease(s) to the disparity in LE, iii) iterative analyses of other disease contributions, and iv) explanation of how the disparities in all scenarios are generated in by clinic and non-clinic-related factors. Evaluation of factors contributing to interregional differences in LE will ultimately inform the design of the strategies to improve public health and healthcare systems.

Introduction

One important barrier to improving life expectancy (LE) in the U.S. are the persistent and growing[1] disparities in health outcomes. As a consequence, underperforming (lagging) regions decelerate the increase of the average LE in the U.S.[2], e.g., people in certain U.S. states live up to 3.5 years (males) and 4.6 years (females) less than in the states with better health outcomes. These geographic disparities are associated with increased burden of disease and increased health expenditures in the healthcare system and showcase an observed lag in health and longevity compared to other industrialized nations. Although research has identified many demographic, socioeconomic, behavioral, and access-to-care factors[3-5] that contribute to these disparities, they have been unable to fully account for the existing variations in LE[2,6,7]. In this paper we propose an approach allowing for the identification of the causes and mechanisms of existing geographic disparities in LE and clarification of the role different clinic- and non-clinic-related factors play in them. Specifically, we assume that regions lagging/leading in LE will show correspondingly higher/lower rates of disease incidence, prevalence, multimorbidity, and worse/better patient survival, and that the region-specific differences in these measures are explained by contributions of clinic-related (patterns of treatment, diagnostic procedures, comorbidity, etc.) and non-clinic-related (education, income, behavioral, and environmental) factors.

Our developments are based on the partitioning analyses recently developed by our group[8,9]. The standard partitioning approach is based on an explicit representation of prevalence and mortality at ages 65+ with no simplifying assumptions[8]. The resulting formulae for age-specific and age-adjusted prevalence and disease-specific mortality are expressed in terms of integrals over the age(s) at disease onset (see[8], eqs.(7-10)). Time trends are defined through derivatives that are calculated analytically by using semiparametric models for disease prevalence at 65, disease incidence, and relative survival (using B-splines or an appropriate parametric model if such a model is well established for a given measure) resulting in four terms that are interpreted as the contributions of following components to the time trend: i) prevalence at age 65; ii) the probability of relative survival of prevalent individuals at age 65; iii) the incidence rate; and iv) the probability of relative survival after disease diagnosis. In a recently published study[9], we applied the partitioning approach to the decomposition of trends of prevalence and mortality in diabetes mellitus using 5%-Medicare data (Figure 1).

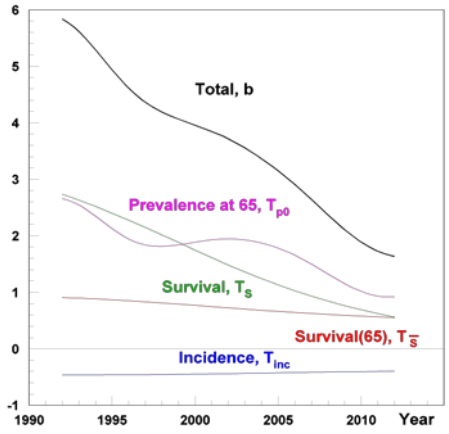

Figure 1 Time trends of diabetes prevalence represented by $b = P'(t)/P(t)$ where $P(t)$ is age-adjusted prevalence at time t (labeled as "Total, b") and its partitioning. $b = T_{p0} + T_{\bar{S}} + T_{inc} + T_S$.

Geographic disparities in life expectancy

The development is based on the assumption that there are four scenarios (and their combinations) of how the disparities could be expressed in the health measures/records extracted from Medicare data: i) incidence of certain disease(s) associated with high mortality is higher in lagging U.S. regions; ii) survival after disease

onset is worse in lagging regions; iii) multimorbidity is higher in lagging regions; and iv) the pre-Medicare population is more morbid in lagging regions, resulting in a higher proportion of unhealthy individuals at the age of Medicare eligibility (65 years old). Each scenario is based on an epidemiological process: disease incidence, disease-specific survival, multimorbidty, or/and morbidity at age 65. Differences in these four respective measures result in the observed disparities in LE. We demonstrate in our approach that the relationship between disease-specific mortality (and, therefore, total mortality and LE described through the sum of contributing diseases) and these four measures is exact and, therefore, the disparity in LE must be represented through one or more (combination) of these four components. Our computational approach includes four steps: i) ranking diseases according their effect on mortality in selected geographic regions, ii) evaluation of the contribution to the geographic disparities in LE from the most severe/lethal disease, iii) iterative analyses of other disease contributions, and iv) explanation of how the disparities in all scenarios are generated in terms of clinic- and non-clinic-related factors. Implementation and applying this approach to Medicare data allow the researchers to: i) identify the high-impact diseases that have the greatest effects on generating disparities, ii) test whether each scenario can (and in what extent) explain the existing geographic disparities in mortality; iii) identify the scenario that is dominant; and iv) evaluate the extent to which the real-life situations can be represented by a mixture of these scenarios. Ultimately, identification of lagging (with LE lower than the U.S. average) and leading (with LE above the U.S. average) states, and evaluation of contributing factors behind them will inform the design of the strategies to improve public health and healthcare systems.

Computational approach

In the first step we rank all causes of death according their effects on mortality in U.S. Specifically, we identify indicators of chronic diseases using an algorithm previously developed[10] and applied by our group[11-17]. Then we ranked each condition according to the size of the disease-specific death hazard ratio evaluated in a univariable or multivariable Cox model. The disease list includes 48 age-related diseases identified in[16] as well as external causes, injury and poisoning (ICD-9 codes 800-999 and E800-E999) and symptoms and signs (780-799).

In the second step, we evaluated the geographic disparities attributable to the disease ranking first in the list (i.e., the most lethal disease). We considered cohorts of individuals from the leading and lagging states starting from age $x_0 = 65$. Prevalence of the disease at age x is $P(x) = P(x_0)S(x,x_0) + \int_{x_0}^{x} I(\tau)S(x,\tau)d\tau$, where $I(\tau)$ is cohort incidence or density of incident cases, $S(x,x_0)$ and $S(x,\tau)$ are the survival function of individuals prevalent at x_0 and incident at τ. The expression for mortality is obtained by substitution of the survival function to respective density: $M(x) = P(x_0)d(x,x_0) + \int_{x_0}^{x} I(\tau)d(x,\tau)d\tau$. In the original papers[8,9], we demonstrated how to write the exact equations for prevalence and mortality in terms of usual epidemiologic functions. A disparity in mortality due to the disease is defined as the difference between geographic region-specific mortality/prevalence (by indexes 1 and 2) and represented in terms of four terms

$$\Delta M_{dis} \equiv M_1(x) - M_2(x) = \left(P_1(x_0) - P_2(x_0)\right)\bar{d}(x,x_0) + \bar{P}(x_0)\left(d_1(x,x_0) - d_2(x,x_0)\right)$$
$$+ \int_{x_0}^{x}\left(I_1(\tau) - I_2(\tau)\right)\bar{d}(x,\tau)d\tau + \int_{x_0}^{x}\bar{I}(\tau)\left(d_1(x,\tau) - d_2(x,\tau)\right)d\tau$$

The "bar" on the top of quantity denotes averaging: e.g., $\bar{P}(x_0) = \left(P_1(x_0) + P_2(x_0)\right)/2$. Four terms in the decomposition are interpreted as relative contributions to the geographic disparities in mortality from: i) initial morbidity, ii) survival of patient with the disease at 65, iii) incidence after 65, and iv) survival of individuals who became sick after age 65. Thus, we decompose mortality (as well as prevalence, age-adjusted mortality/prevalence, and, with the help of life tables, LE at 65) caused by the most severe/lethal disease (according to the rank identified above) into the four terms. Summing over cohorts provides time dependence of this decomposition as in Figure 1.

In the third step we iteratively analyze the contributions of the remaining diseases from the list. Individuals identified as having died from a previously considered disease are excluded from the following iterations of the analysis. This step results in a representation of the geographic disparities in age-adjusted all-cause mortality

and life expectancies at age 65 through the sum of contributions of all diseases in the list, thus evaluating relative contributions of scenarios #1, 2, & 4, e.g., disparities in age-adjusted mortality is $\sum_{dis} \Delta M_{dis}^{age.adj.}$. The analysis of relative contributions in the sum will allow us to identify the diseases with the strongest contributions to the disparities. We can observe the situation when there will be not a subgroup of diseases with stronger contributions to disparities but the disparities are due to moderate contributions of majority of diseases. In this case we conclude that multimorbidity (Scenario #3) is responsible for the geographic disparities.The initial disease list could be modified by combining diseases of similar etiology if statistical power is not sufficient to evaluate their effects separately. For diseases selected as contributors to the disparities and, therefore, for which further analyses is planned, we repeat analyses for the total population i.e. without the exclusion of individuals identified as having died due to other conditions in the iterative process of step 3.

In the fourth step, we considered diseases that contribute to the geographic disparities separately and demonstrate how clinic- and non-clinic-related factors from Medicare and other data generate the disparities through the four scenarios and are ultimately responsible for the observed regional/geographic disparities. For example, Medicare records are used to evaluate the effects of treatment patterns, adherence to prescribed treatment, diagnostic/screening procedures to explain the disparities from disease-specific survival. Individual non-clinic-related measures such as income, education, and behavioral factors (all are available in HRS-Medicare data) and area-based measures of SES, access to medical care, and environmental factors are used to further clarify geographic disparities both in disease incidence and patient survival. Medicare records also provide additional information on access to and quality of care. All these measures can have the direct effects of these factors and can modulate the effects of clinic-related factors studied on the geographic disparities and specific scenarios.

Numeric illustration on how the four-step approach works for selected leading and lagging regions will be presented at PAA 2019.

References.

1. Ezzati M, Friedman AB, Kulkarni SC, Murray CJ. The reversal of fortunes: trends in county mortality and cross-county mortality disparities in the United States. *PLoS medicine.* 2008;5(4):e66.
2. Council NR, Population Co. *International differences in mortality at older ages: Dimensions and sources.* National Academies Press; 2011.
3. Murray CJ, Kulkarni SC, Michaud C, et al. Eight Americas: investigating mortality disparities across races, counties, and race-counties in the United States. *PLoS Medicine.* 2006;3(9):e260.
4. Banks J, Marmot M, Oldfield Z, Smith JP. Disease and disadvantage in the United States and in England. *Jama.* 2006;295(17):2037-2045.
5. Marengoni A, Angleman S, Melis R, et al. Aging with multimorbidity: a systematic review of the literature. *Ageing research reviews.* 2011;10(4):430-439.
6. Woolf SH, Aron LY. The US health disadvantage relative to other high-income countries: findings from a National Research Council/Institute of Medicine report. *Jama.* 2013;309(8):771-772.
7. Council NR, Population Co. *Explaining divergent levels of longevity in high-income countries.* National Academies Press; 2011.
8. Akushevich I, Yashkin A, Kravchenko J, et al. Theory of partitioning of disease prevalence and mortality in observational data. *Theoretical Population Biology.* 2017;114:117-127.
9. Akushevich I, Yashkin AP, Kravchenko J, et al. Identifying the Causes of the Changes in the Prevalence Patterns of Diabetes in Older US Adults: A New Trend Partitioning Approach. *Journal of Diabetes and its Complications.* 2018;32(4):362-367.
10. Akushevich I, Kravchenko J, Ukraintseva S, Arbeev K, Yashin AI. Age Patterns of Incidence of Geriatric Disease in the US Elderly Population: Medicare-Based Analysis. *Journal of the American Geriatrics Society.* 2012;60(2):323-327.
11. Akushevich I, Kravchenko J, Ukraintseva S, Arbeev K, Yashin AI. Recovery and survival from aging-associated diseases. *Exp Gerontol.* 2013;48(8):824-830.
12. Akushevich I, Kravchenko J, Ukraintseva S, Arbeev K, Yashin AI. Circulatory Diseases in the US Elderly in the Linked National Long-Term Care Survey-Medicare Database Population-Based Analysis of Incidence, Comorbidity, and Disability. *Research on Aging.* 2013;35(4):437-458.

13. Akushevich I, Kravchenko J, Ukraintseva S, Arbeev K, Yashin AI. Time trends of incidence of age-associated diseases in the US elderly population: Medicare-based analysis. *Age and ageing.* 2013;42(4):494-500.
14. Akushevich I, Kravchenko J, Arbeev KG, Ukraintseva SV, Land KC, Yashin AI. Health Effects and Medicare Trajectories: Population-Based Analysis of Morbidity and Mortality Patterns. *Biodemography of Aging*: Springer; 2016:47-93.
15. Worni M, Akushevich I, Greenup R, et al. Trends in treatment patterns and outcomes for ductal carcinoma in situ. *Journal of the National Cancer Institute.* 2015;107(12):djv263.
16. Akushevich I, Kravchenko J, Yashkin AP, Yashin AI. Time trends in the prevalence of cancer and non-cancer diseases among older US adults: Medicare-based analysis. *Experimental Gerontology.* 2018.
17. Akushevich I, Yashkin AP, Kravchenko J, Ukraintseva S, Stallard E, Yashin AI. Time Trends in the Prevalence of Neurocognitive Disorders and Cognitive Impairment in the United States: The Effects of Disease Severity and Improved Ascertainment. *Journal of Alzheimer's Disease.* 2018(Preprint):1-12.