Experimental Gerontology 110 (2018) 267–276

Contents lists available at ScienceDirect

Experimental Gerontology

journal homepage: www.elsevier.com/locate/expgero

Time trends in the prevalence of cancer and non-cancer diseases among older U.S. adults: Medicare-based analysis

Igor Akushevich[a,*], Julia Kravchenko[b], Arseniy P. Yashkin[a], Anatoliy I. Yashin[a]

[a] Biodemography of Aging Research Unit, Center for Population Health and Aging, Duke University, Durham, NC, United States of America
[b] Division of Surgical Sciences, Department of Surgery, Duke University School of Medicine, Duke University, Durham, NC, United States of America

ARTICLE INFO

Section Editor: Diana Van Heemst

Keywords:
Disease prevalence
Time trends
Medicare
Annual percent change
Epidemiology
Age-adjusted rate

ABSTRACT

Longer lifespan is accompanied by a larger number of chronic diseases among older adults. Because of a growing proportion of older adults in the U.S., this brings the problem of age-related morbidity to the forefront as a major contributor to rising medical expenditures. We evaluated 15-year time trends (from 1998 to 2013) in the prevalence of 48 acute and chronic non-cancer diseases and cancers in older U.S. adults aged 65+ and estimated the annual percentage changes of these prevalence trends using SEER-Medicare and HRS-Medicare data. We found that age-adjusted prevalence of cancers of kidney, pancreas, and melanoma, as well as diabetes, renal disease, limb fracture, depression, anemia, weight deficiency, dementia/Alzheimer's disease, drug/medications abuse and several other diseases/conditions increased over time. Conversely, prevalence of myocardial infarction, heart failure, cardiomyopathy, pneumonia/influenza, peptic ulcer, and gastrointestinal bleeding, among others, decreased over time. There are also diseases whose prevalence did not change substantially over time, e.g., a group of fast progressing cancers and rheumatoid arthritis. Analysis of trends of multiple diseases performed simultaneously within one study design with focus on the same time interval and the same population for all diseases allowed us to provide insight into the epidemiology of these conditions and identify the most alarming and/or unexpected trends and trade-offs. The obtained results can be used for health expenditures planning for growing sector of older adults in the U.S.

1. Introduction

During the recent decades a persistent increase in the length of life of successive birth-cohorts has been observed in the U.S. However, a longer life span is also associated with increased risk of contracting cancer as well as other chronic diseases associated with the aging process. This makes the ability to accurately estimate and model the changes in nationally representative age-specific prevalence proportions critically important for public health specialists and policy makers. To understand the patterns of the changes and use them in future projections and forecasts, information on disease incidence, mortality, and prevalence need to be analyzed. Extensive work has already been done on incidence and mortality patterns (Brauer et al., 2009; Cooper et al., 2011; Chen et al., 2011; Geiss et al., 2006; Yeh et al., 2010; Roger et al., 2004; Smigal et al., 2006; Levy et al., 2002), but studies focusing on diseases prevalence are relatively rare and often focus on diabetes, hypertension, obesity, and kidney disease (Egan et al., 2010; Cutler et al., 2008; Coresh et al., 2005; Menke et al., 2015). Furthermore, even when multiple disease are considered, these studies usually focus on mortality (e.g., the analysis of trends of leading causes of death (Jemal et al., 2005)) and/or study a group of diseases that are related to a single risk factor (e.g., smoking related mortality (Thun et al., 2013)).

At the population level disease onset, progression, and patient survival interact to form disease prevalence. In turn, their trends are impacted by the time-trends of other assorted factors such as socio-economic characteristics, the demographic structure of the population, behavioral factors (i.e., smoking, physical activity, diet), as well as changes in prevention, diagnostic, and treatment strategies. For example, decreasing smoking prevalence in the U.S. is associated with a reduction in the rates of certain cancers (e.g., squamous cell carcinoma of lung, cancer of oral cavity, cancer of urinary bladder) as well as lower rates of chronic obstructive pulmonary disease and cardiovascular diseases (Islami et al., 2015; Burger et al., 2013; Freedman et al., 2011; Brown et al., 2012). Due to increasing rates of multi-morbidity, or the presence of multiple life threatening chronic diseases in a single patient, a single change in any of the above factors can have a synergistic effect on individual health. Furthermore, improvements in early diagnosis of cancer and the consequent increased efficacy of

* Corresponding author at: Biodemography of Aging Research Unit, Center for Population Health and Aging, Social Science Research Institute, Duke University, 2024 W. Main Str., Durham, NC 27708, United States of America.
E-mail address: igor.akushevich@duke.edu (I. Akushevich).

https://doi.org/10.1016/j.exger.2018.06.017
Received 3 October 2017; Received in revised form 30 April 2018; Accepted 17 June 2018
Available online 19 June 2018
0531-5565/

treatment increase patient survival of patients making it, for some cancers, compatible to survival for cardio- and cerebrovascular diseases (e.g., 5-year relative survival rate is 73–89% for breast cancer and 50–99% for prostate cancer vs. 50% for stroke and 62% for heart failure (Askoxylakis et al., 2010)). However, better cancer-specific survival does not imply better total survival. In fact, the mortality of cancer patients age 65+ from non-cancer causes substantially affects the total mortality trends as well as cancer prevalence. For example, during the past two decades the overall all-age cancer-specific mortality rate in the U.S. has declined by 30%–40% for common cancers such as cancers of the colorectum, female breast, male lung, and male prostate (Siegel et al., 2013). However, these improvements did not extend to patients age 65 and older (Owonikoko et al., 2007; Demicheli et al., 2007; Jatoi et al., 2007). One of the explanations for this observation could be that a substantial proportion of cancer patients die of other co-existing chronic diseases (Edwards et al., 2014; Baade et al., 2006). The close relationship between these two categories of chronic diseases suggests that the study of any individual disease in isolation from its co-morbidities is likely to paint a biased picture of the *status quo*.

The above concerns make the analysis and modeling of these complex processes with sufficient external validity a complicated task that requires large nationally representative medical data-sets that are not easily available. One way to overcome this difficulty is to use individual-level administrative health insurance claim records such as those provided by the Medicare social insurance system. Medicare is nationally representative of the U.S. population at ages 65+ and is the primary payer for institutional and professional health services for the U.S. elderly. Although not initially designed for research purposes, use of Medicare Administrative claims data has been validated for use in the analyses of incidence, survival, and prevalence of cancer, diabetes, cardio- and cerebrovascular diseases, diseases of renal and endocrine systems as well as other diseases (Warren et al., 2002; Hirsch et al., 2008; Foley et al., 2005; Weiss et al., 2011; McClellan et al., 2004; Brown et al., 2002).

In this study, we identified the time-trends in the prevalence and age patterns of several major cancers and common age-related non-cancer diseases among older U.S. adults. The list of cancers includes several site-specific cancers as well as groups of solid cancers characterized by slow and fast progression, cancers with metastasis, and non-solid malignancies. Non-cancer diseases are represented by a spectrum of cardio- and cerebrovascular, respiratory, mental, endocrine, digestive, renal, and blood system diseases, as well as infections and injuries. Rather than being preselected *a priori*, the disease list has been selected using a newly developed algorithm that takes into account the impact of each disease on mortality among older U.S. adults.

1.1. Data

Two complementary Medicare datasets are used: the Health and Retirement Study (HRS-Medicare) linked to administrative claims data from the Medicare program, and the Medicare-linked Surveillance, Epidemiology and End Results (SEER-Medicare) data. The HRS-Medicare is an ongoing, nationally representative study of individuals reaching age 50+ in the U.S. while SEER-Medicare is the largest and most detailed registry of cancer cases in the U.S. Mortality data from Medicare sources were supplemented with data on cause-specific mortality (not available in Medicare) drawn from the Multiple Cause of Death (MCD) data.

We evaluated the month-specific prevalence probabilities in the 1998–2013 period for both datasets. All individuals were longitudinally tracked for Medicare Part A and Part B service use. Records were available for institutional (inpatient, outpatient, skilled nursing facility, hospice, or home health agency) and non-institutional (physician and durable medical equipment) providers claim types. MCD dataset (1996–1998) was used as part of the algorithm used for disease selection.

2. Methods

The disease groups of interest for this study were identified based on the analysis of their impacts on mortality in older U.S. adults. Specifically, we evaluated i) disease rank as a leading cause of death in MCD data; ii) the strength of the association between the prevalence of a disease and mortality observed in SEER-Medicare and HRS-Medicare; and iii) the disease frequencies as secondary cause of death in patients who died from diseases selected in the first two steps.

Age-specific and age-adjusted disease prevalence were evaluated from Medicare administrative claims data using individual disease presence. An individual was considered to have a disease at any time t of his/her follow-up if there existed a primary or secondary record with a disease-specific ICD-9 code during the optimal disease-specific look-back period. This period was defined as the length of time before time t used to search for the presence of disease-specific code(s) in individual claim records. In this study, we use four base sources of Medicare data (physician, inpatient, outpatient and skilled nursing facility) and 12-month look-back period.

Estimates for the prevalence of acute conditions (e.g., myocardial infarction, limb fractures, gastrointestinal bleeding, and others) were obtained on a month-by-month basis. We estimated the prevalence of these conditions the same way we estimated the prevalence of chronic conditions. This approach provided us with opportunity to compare the results for both types of diseases as they were obtained from the same data using the same approach.

For both Medicare datasets, the empirical time- and age-specific prevalence, $p_a(t)$, were calculated as a fraction of individuals with disease presence at time t and age a. The age-adjusted rates (or directly standardized incidence rates) were calculated for the population age 66+ as $p(t) = \sum_{a=66}^{105+} p_a(t) P_a^{2000} \left(\sum_{a'=66}^{105+} P_{a'}^{2000}\right)^{-1}$, where P_a^{2000} are age-specific counts of US 2000 standard population. The standard error (SE) for the age-adjusted rate was estimated using the approach based on the approximation suggested by Keyfitz (1966): $SE = \lambda(t)/\sqrt{n_0}$, where n_0 is the sum of the cases.

Among many measures appropriate for the analysis of time trends we chose the average annual percentage change (APC) in prevalence proportions estimated using log-linear model in the form $\log(p(t)) = a + bt + \varepsilon_t$ or $p(t) = p_{2000} \exp(b(t - 2000))$, where t is calendar time, and ε_t is the error term of the regression. The estimate of the average APC is given by $100b$ and expressed in percent (Truelsen et al., 2003). The estimate of the prevalence at 2000 is given by p_{2000}. The model estimation is based on weighted least squares where the weights are the reciprocal of the variance estimated for each annual rate. This approach is preferable to the simple averaging of empirically estimated annual percentage changes because it: i) accounts for the standard errors of the prevalence proportion resulting in more precise estimates of APC, ii) provides estimates that are smoothed in respect to any random fluctuations that may arise as an artifact of the data gathering process in any individual year (e.g., for estimates of p_{2000}), and iii) results in the analytic formulae that represent a simple forecasting model of disease-specific prevalence proportions under the assumptions that the trend averaged over 1998–2003 will hold. Simultaneous visualizations of these estimates for all study conditions at the same time allow for a clear representation of integral picture of the epidemiological trends in prevalence.

Other approaches to identifying disease prevalence were tested in sensitivity analysis. These included: i) usage of all Medicare sources, ii) using only inpatient/outpatient records, iii) requiring a second verification claim during the look-back period, and iv) using only records with non-zero Medicare payments. No significant differences in the estimates of the APC were identified in sensitivity analysis.

3. Results

Diseases that were selected by our algorithm are presented in