

## Research Article

# Population-Based Analysis of Incidence Rates of Cancer and Noncancer Chronic Diseases in the US Elderly Using NLTCs/Medicare-Linked Database

I. Akushevich,<sup>1</sup> J. Kravchenko,<sup>2</sup> S. Ukraintseva,<sup>1</sup> K. Arbeev,<sup>1</sup> and A. Yashin<sup>1</sup>

<sup>1</sup>Center for Population Health and Aging, Duke University, P.O. Box 90408, Durham, NC 27708, USA

<sup>2</sup>Division of Surgical Sciences, Department of Surgery, Duke University Medical Center, Durham, NC 27705, USA

Correspondence should be addressed to I. Akushevich; igor.akushevich@duke.edu

Received 19 November 2012; Accepted 22 January 2013

Academic Editors: S. Vieira, Å. Wahlin, and R. Wirth

Copyright © 2013 I. Akushevich et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The age, disability, and comorbidity patterns of incidence rates of cancer and chronic noncancer diseases such as heart failure, diabetes mellitus, asthma, Parkinson's disease, Alzheimer's disease, skin melanoma, and cancers of breast, prostate, lung, and colon were studied for the US elderly population (aged 65+) using the National Long-Term Care Survey (NLTCs) data linked to Medicare records for 1991–2005. Opposite to breast cancer and asthma, incidence rates of heart failure and Alzheimer's diseases were increasing with age. Higher incidence rates of heart failure, diabetes, asthma, and Parkinson's and Alzheimer's diseases were observed among individuals with severe disabilities or/and comorbidities, while rates of breast and prostate cancers were higher among those with minor disabilities or fewer comorbidities. Our results were in agreement with those obtained from other epidemiological datasets, thus suggesting that Medicare administrative records can provide nationally representative incidence rates. Detailed sensitivity analysis that focused on the effects of alternative onset definitions, latent censoring, study design, and other procedural uncertainties showed the stability of reconstructed incidence rates. This Medicare-linked dataset can be used for studying highly debated effects of new medical technologies on aging-related diseases burden and future Medicare costs.

## 1. Introduction

The proportion of the elderly population in USA is constantly growing. That prioritizes the task of determining the national trends in health and vital status of older adults making it a major public health concern. To better address these health demands and to reduce economic burdens on society, it is important to understand the key factors driving the onset and progression of aging-related chronic diseases in humans. Unfortunately, the identification of disease age patterns with sufficient precision requires large population-based databases that are costly to collect. This is a reason why the studies on disease age patterns, along with investigations of factors affecting them, are not common in the US elderly population. Among aging-associated diseases, cancer incidences are better studied at a national level, predominantly due to the existence of the Surveillance Epidemiology and End

Results (SEER) Registry data [1]. Age patterns of incidence of cardiovascular, cerebrovascular, and neurodegenerative diseases in general population and in oldest adults were recently studied using the Cardiovascular Health Study (CHS) data [2, 3]. However, this database is not nationally representative. Other common elderly chronic diseases affecting well-being of this population group such as Alzheimer's disease and dementias [4] and diabetes mellitus [5] have been analyzed in details for general population, while their patterns among the US the elderly and the oldest old populations remain unclear. A number of studies of disease patterns in the elderly were conducted on local populations such as the Rochester, Minnesota, and Cache County studies on epidemiology of dementia and Alzheimer's disease [6, 7]. However, these datasets also cannot represent the spectrum of characteristics (e.g., age, race, male/female ratio, etc.) of the nationwide population. Furthermore, not only age patterns of disease

incidences are almost unknown, but also studies on the effects of their coexistence at the same individual are sparse. The older adults are known by their multimorbidity, thus making comorbidity issue one of the main factors limiting their well-being and survival. Patients with a single chronic disease or, more often, with several comorbid diseases are developing disability of varying severity. That became an important issue for policymakers as well as for health care agencies such as Medicare and Medicaid. Also, individuals with different disabilities could be at different (e.g., increased) risks of developing certain diseases; and these interrelations are poorly studied. Therefore, this study is motivated by the lack of comprehensive and representative analyses of age patterns of incidence of chronic diseases in the oldest adults at the national level, unclear interrelations between comorbid diseases, and sparse knowledge on how these diseases are affected by disability among the US older adults.

To calculate the incidence rates, data from the national registers or from the surveys representing the US population should be used. For surveys, special procedures are needed to generalize the obtained results to the national level. It could be done by using a weight function assigned to each individual in the sample such that the “weighted” sums (means) over individuals in the sample give the quantities at the national level. For such analysis the 1982–2005 National Long-Term Care Survey (NLTCS) can be used. It focuses on the US elderly (65 and older) population. Since NLTCS is linked with the Medicare records and the Vital Statistics files, it represents a unique opportunity for analysis of the data with continuous recording of health services provided, age of death, and detailed reassessment of health status which are performed by a survey every five years (except for the first two waves in 1982 and 1984). The participants of the NLTCS were drawn from the Medicare enrollment lists. In 1982–1999, in the NLTCS there were about 400,000 person-years of exposure over age 65, including over 100,000 person-years of exposure over age 85. The sample weights are provided by the US Census Bureau and are available for each year of followup. Thus, the NLTCS design provides an excellent opportunity to study the incidence patterns of aging-associated diseases in the US elderly population. Earlier analyses of the NLTCS data provided consistent results on functional disability, active life expectancy, and chronic disease prevalence in the US elderly; for example, for 1982–1999 period, a 15% decline in chronic disability (1.1% per annum) has been reported [8–11].

The Medicare service use files linked to the NLTCS (NLTCS-M) contain the information about disease diagnoses. This information can be used in a computational algorithm for identification of the dates of disease onsets. The primary goal of this study is to estimate age-adjusted and age-specific (as well as disability- and comorbidity-specific) disease incidence in the US elderly population applying such an algorithm for the NLTCS-M data. Specific attention is paid to detailed analysis of stochastic and systematic uncertainties in these estimates: for example, we compare the evaluated rates with those obtained using alternative approaches for evaluating the incidence rates using the Medicare files. Also, we compare our estimates with the results obtained from other studies. Thus, the disease incidence estimates for

advanced ages presented here should be very valuable for basic understanding of interrelations between chronic aging-related disease incidence and senescence and for practical implementations when analyzing the national health trends and forecasting future Medicare expenditures.

## 2. Data and Methods

*2.1. National Long-Term Care Survey (NLTCS), Medicare Files of Service Use, and Medical Cost.* The primary data to be analyzed are the six waves of the NLTCS [12] spanning the period from 1982 to 2004/5 together with the linked Medicare data. Two of the six waves, namely, cohorts of 1994 and 1999, are used in the analysis. These specific waves were chosen primarily because the high-quality Medicare follow-up data are available only beginning from 1991, and also because the complete 5-year followup after the NLTCS interview is accessible only for these two waves after 1991. The NLTCS uses a sample of individuals drawn from the national Medicare enrollment files. The NLTCS provides reported data on hundreds of variables including age, sex, and activities/instrumental activities of daily living allowing for disability measurements. The US Census Bureau was employed for collecting data over all waves—so, training methods and materials, survey administration and management procedures, field operations, computer processing, and editing procedures are consistent across the surveys. Together with the high (95%) response rates in all NLTCS waves, that minimizes bias in trend estimates. The 1982–2004 NLTCS files include information on 49,258 individuals. In total, 34,077 individuals were followed up between 1994 and 1999. The so-called screener weights released by the US Census Bureau with the NLTCS were used in this study to produce the national population estimates.

All individuals in the NLTCS are continuously tracked for Medicare Part A and Part B service use. Thus, for all persons we have continuous records of Medicare service use from 1991 (or from the time the person has passed the age of 65 after 1990) until his/her time of death. These records are available for each institutional (inpatient, outpatient, skilled nursing facility, hospice, or home health agency) and noninstitutional (carrier-physician-supplier and durable medical equipment providers) claim type.

The self-reported information about serious (i.e., activities of daily living; ADLs) and less serious (i.e., instrumental activities of daily living; IADLs) impairments was used to construct a disability index [10]. The disability index has six categories (nondisabled, IADL only, 1-2 ADLs, 3-4 ADLs, 5-6 ADLs, Institutionalized) and is measured at the date of interview, that is, at the beginning of the followup. Comorbidity measures are represented by the Charlson comorbidity index [13, 14] that was also evaluated at the date of interview using Medicare records during a year prior to the date of interview.

*2.2. Date of Onset Definitions.* The date of onset of chronic disease is not defined with the same precision as mortality and there is always certain arbitrariness in defining the date of onset. Computational approaches of different complexity have been used for reconstruction of onsets of cancer and

non-cancer diseases [15–17]. In this study we argue that the algorithm developed in our prior studies [18–21] can be applied for a spectrum of chronic diseases; therefore, this approach may be viewed as a unified definition of date of onset appropriate for population studies that use administrative data. Ten diseases were selected for analyses (the International Classification of Diseases (ICD-9) codes used for the considered diseases are shown in brackets): cancers of the lung (162.xx), colon (153.xx), breast (174.xx), and prostate (185.xx), skin melanoma (172.xx), heart failure (428.xx), Alzheimer’s disease (331.0), Parkinson’s disease (332.xx), diabetes mellitus (250.xx), and asthma (493.xx).

The computational algorithm is applied for each disease separately. First, the individual medical histories of an applicable disease were reconstructed from Medicare files combining all records with respective ICD-9 codes of that disease. Patients with the history of the disease before the date of interview in 1994 or in 1999 were excluded from the study for onset of this disease. Four variants of algorithms (Algorithms A, B, C, and D) of the identification of the disease onset from the disease-specific medical histories were considered. In Algorithm A, a date of a Medicare record (referred to as “*this record*” below in this subsection) is identified with the date of onset of applicable disease when both of the conditions mentioned below are met.

- (i) *This record* is the earliest record with respective ICD code as a primary diagnosis in one of four Medicare sources (inpatient care, outpatient care, physician services, and skilled nursing facilities). This choice is in accordance with the general practice of reconstruction of the date at onset from Medicare data [15, 16].
- (ii) In addition to *this record*, there is another record with its respective ICD code as the primary diagnosis from one of the four Medicare sources listed in (i), which appeared with a date different from the date of *this record* and not later than  $\Delta = 0.3$  years after *this record*. Death occurred during this period (i.e., 0.3 years after *this record*) is also considered as the second record.

In Algorithm B, the confirmation by the occurrence of the second record is not required; that is, only the first condition is valid. In Algorithm C, all codes (not necessary being primary) are considered valid and the confirmation is also not required. In Algorithm D, death is not considered as the second record.

**2.3. Estimates of Age-Specific and Age-Adjusted Rates for the US Elderly Population.** Age patterns of incidence rates are assessed by stratifying the sample into relevant age categories (a year, or several years). Empirical age-specific risks ( $\lambda_a$ ) are calculated as a ratio of weighted numbers of cases to weighted person-years at risk:  $\lambda_a = n_a/P_a$ , where  $n_a = \sum_n w_n$ ,  $P_a = \sum_i w_i$ , and  $w_i$  is the individual weight;  $n$  runs over all disease onsets detected in the age group, and  $i$  runs over all individuals at risk in  $a$ th age group. The individual weights (the NLTCS weights have been calculated by the US Census and released with the NLTCS data) are necessary to have the estimates representative of the entire US elderly population,

that is, to take into account the effect of study design. The effects of study design also influence the calculation of standard error (SE) and confidence intervals of rate estimates. The used formula for SE is  $\sigma_E = \sqrt{\lambda_a(1 - \lambda_a)/P_{0a}}$ , where  $P_{0a}$  is the number of person-years estimated for unit weights. Thus, standard error is calculated based on the number of really measured individuals. A generalization of the formula for SE based on Wilson’s approach [22] is used when  $P_{0a}$  or  $\lambda_a$  is small. Comorbidity- and disability-specific patterns are calculated using the same equations for stratified population.

Age-adjusted rates (or directly standardized incidence rates) are the averages of age-specific rates. For population aged 66+, they are calculated as  $\lambda = \sum_{a=66}^{105+} \lambda_a P_a (\sum_{a'=66}^{105+} P_{a'})^{-1}$ . There are many ways to estimate SE for the age-adjusted rate [23–25]. In this study we use the simplest approach (in order to avoid dealing with uncertainties in SE estimation for low and even zero age-specific rates) based on the approximation suggested by Keyfitz [26], in which SE is estimated as  $SE = \lambda/\sqrt{n_0}$ , where  $n_0$  is the unweighted sum of cases.

### 3. Results

The numbers of individuals in the pooled cohort without prevalent cases for each disease are shown in Table 1. Age-adjusted disease incidence rates are presented in Table 2 for all cohorts (i.e., for cohort of years 1994, 1999, and pooled for both time periods) and genders (i.e., males, females, and total population). For males, the highest incidence rates were observed for heart failure and prostate cancer, and for females, for heart failure and diabetes. Because of using the weights, the results both for age-adjusted rates and standard errors are valid for the US elderly population.

Table 3 provides information about evaluated gender disparities and time trends in the age-adjusted incidence rates. The 5-year time trends (from 1994 to 1999) were evaluated as the difference between age-adjusted estimates of two compared rates (i.e., for 1994 and for 1999) divided by the square root of the sum of their standard errors squared. Thus, the approximate  $t$ -statistics are presented in Table 3 which can be compared to a normal distribution to estimate  $P$  values of respective hypotheses. The 5-year decline was significant for heart failure and prostate cancer, and the trends to decline were observed for male lung cancer ( $P = 0.07$ ) and female colon cancer ( $P = 0.06$ ). The 5-year increase was significant for diabetes and Alzheimer’s disease. Gender disparities were observed (i) for heart failure, skin melanoma, lung cancer, colon cancer, and Parkinson’s disease—incidence rates of all these diseases were higher in males—and (ii) for asthma, with higher incidence rates in females.

The age, disability, and comorbidity patterns were analyzed for all considered diseases. The results for cancer and non-cancer diseases are presented in Figures 1 and 2, respectively. These patterns can substantially differ for two genders and two NLTCS cohorts. As follows from Table 3, gender disparities were more pronounced for lung cancer, colon cancer, melanoma, Parkinson’s disease, and asthma; therefore, sex-specific age patterns were presented in Figures 1 and 2 for these diseases. Time trends were more pronounced

TABLE 1: The total numbers of individuals followed up for onsets of chronic diseases (i) in two waves of NLTCS-M, that is, waves formed in 1994 and 1999, (ii) without individuals who had an additional coverage by the Health Maintenance Organization (HMO) and prevalent cases, and (iii) with the registered onsets.

	Lung cancer	Colon cancer	Breast cancer*	Prostate cancer**	Melanoma	Heart failure	Alzheimer's disease	Parkinson's disease	Diabetes	Asthma
Total in two waves	34077	34077	20771	13306	34077	34077	34077	34077	34077	34077
Total without HMO	27607	27607	16985	10622	27607	27607	27607	27607	27607	27607
Total without HMO and prevalent cases	27480	27313	16494	9990	27509	25921	27431	27378	25399	27099
The number of onsets	752	525	546	605	189	3046	687	331	1231	415

\*Females only.

\*\*Males only.

TABLE 2: Age-adjusted incidence rates per 100,000 of geriatric diseases with standard errors (in brackets).

Sex	Year	Lung cancer	Colon cancer	Breast cancer	Prostate cancer	Melanoma	Heart failure	Alzheimer's disease	Parkinson's disease	Diabetes	Asthma
Males	1994	616	314		1146	163	1864	199	217	724	201
		(52)	(36)		(75)	(25)	(87)	(26)	(28)	(59)	(29)
Males	1999	491	311		924	117	1437	255	195	841	247
		(46)	(35)		(67)	(20)	(76)	(29)	(29)	(66)	(33)
Male	Combined	554	312		1033	139	1647	228	204	779	223
		(35)	(25)		(50)	(16)	(58)	(20)	(20)	(44)	(22)
Female	1994	293	262	555		53	1540	210	143	697	303
		(29)	(27)	(38)		(11)	(61)	(21)	(19)	(44)	(30)
Female	1999	331	196	508		71	1313	290	134	803	258
		(32)	(23)	(41)		(14)	(59)	(25)	(18)	(51)	(29)
Female	Combined	312	230	530		61	1426	250	139	744	282
		(21)	(18)	(28)		(9)	(42)	(16)	(13)	(33)	(21)
Total	1994	421	283			96	1667	205	173	707	263
		(27)	(22)			(12)	(50)	(16)	(16)	(35)	(21)
Total	1999	397	243			90	1366	276	159	818	253
		(27)	(20)			(12)	(47)	(19)	(16)	(40)	(22)
Total	Combined	410	264			92	1516	241	165	758	258
		(19)	(15)			(8)	(34)	(13)	(11)	(27)	(15)

TABLE 3: Estimates of  $t$ -statistics for 5-year time trend and male/female differences in the age-adjusted incidence rates of geriatric diseases. Absolute value of  $t$ -statistics larger than 2 approximately corresponds to significant differences with  $P < 0.05$ .

	Lung cancer	Colon cancer	Breast cancer	Prostate cancer	Melanoma	Heart failure	Alzheimer's disease	Parkinson's disease	Diabetes	Asthma
Time Trend										
Males	-1.8	-0.06		-2.21	-1.44	-3.7	1.44	-0.55	1.32	1.05
Female	0.88	-1.86	-0.84		1.01	-2.67	2.45	-0.34	1.57	-1.08
Total	-0.63	-1.35			-0.35	-4.39	2.86	-0.62	2.09	-0.33
Male/female differences										
1994	5.42	1.16			4.03	3.05	-0.33	2.19	0.37	-2.44
1999	2.86	2.75			1.88	1.29	-0.91	1.79	0.46	-0.25
Pooled	5.93	2.66			4.25	3.09	-0.86	2.72	0.64	-1.94

for prostate cancer, heart failure, Alzheimer's disease, and diabetes—so, we presented cohort-specific rates for the total population in Figures 1 and 2 for these diseases. Among all diseases, the incidence rates were increasing with age for heart failure and Alzheimer's disease; less prominent increase in incidence was observed for Parkinson's disease. Rates of breast cancer and asthma were decreasing with age. Rates of other diseases had tendency to be more stable for all ages.

For several diseases (e.g., heart failure, diabetes, asthma, and Parkinson's disease) the incidence rates were higher among individuals with severe disabilities, while for breast and prostate cancers the higher rate was registered among people with minor disabilities. The most dramatic increase of incidence with disability was for heart failure. Interestingly, that for many diseases institutionalized individuals have lower rates and for several (such as melanoma, lung cancer (males), colon cancer, and asthma) the lowest rates among all other disability groups including nondisabled individuals. However, for neurodegenerative diseases (i.e., Parkinson's (females) and Alzheimer's diseases) the rate for institutionalized individuals is the highest. Among individuals with high comorbidity indices (i.e., Charlson index) higher rates were observed for heart failure, melanoma, and Alzheimer's disease. The incidence rates of cancers of breast and prostate and of diabetes had decreasing trends with increasing comorbidity index.

As we discussed in Section 2.2 several definitions of date at onset can be done using large administrative databases. Table 4 presents the results of calculating age-adjusted rates using alternative approaches. From these results we can conclude that the calculated rates are relatively stable. Thus, columns V1 and V2 represent calculations without age standardization using standard population of 2000 (V1) and without using NLTCS sample weights (V2). In the alternative censoring scheme (V3), the last day of observation is the latest day among (i) Part B coverage, and (ii) Medicare record in Part A or Part B, (iii) response on interview in the next NLTCS wave, while in the basic calculation, the final date of observation is the earliest date among dates of disease onset or death and the last date of cohort observation. Only minor changes in incidence rates obtained within V1–V3 strategies were detected. The results of calculations V4 and V5 reflect the effect of removing individuals from the cohort with a different level of additional coverage by HMO (exactly, by different fractions of months covered by HMO denoted by  $\delta$ ). Since individuals covered by HMO are supposed to be healthier than the general elderly population, the obtained decline in the incidence rates under V4 and V5 strategies is expected. Other calculations represent less (V6–V11) or more conservative (V12) approaches to the definition of the date at onset. The higher variations are derived from V10 and V12 calculations types. In each of the approaches, one of the components of definition is replaced by an alternative one. Note that different choice of Medicare sources also influences the definition of prevalent cases, so using more Medicare sources does not always result in higher incidence rates. Overview of the results in Table 4 shows that the results on incidence rates of non-cancer diseases were more affected by the choice of calculation method than rates of cancers.

Observed variations in incidence rates display that different definition of incidence rate extracted from administrative data can be used and they result in significantly different incidence rates. However, male/female differences, estimates of time trends, and ratios of rates of various diseases remain stable across the different definitions of incidence rates.

*3.1. Comparison to Other Studies.* The age patterns (sex- and cohort-unspecific) calculated using two Medicare-based datasets (i.e., NLTCS-M as in this study and SEER-Medicare) were compared and an agreement was found for chronic diseases considered here [20]. Because of the same design for data collection and computation of disease incidences, the agreement between the rates obtained using these two datasets was expected. Further comparisons are required to compare our results with those obtained by other researchers in the studies of different designs and computational approaches. In such a comparison (described in detail below) we focus on each group of diseases, emphasizing (1) evaluated age-adjusted incidence rates, (2) the shape and level of the age pattern of incidence rates, and (3) sex differences. In general, the datasets which are used for disease analysis in elderly are predominantly disease-specific (i.e., on a single disease) rather than multidisease data. They are not specifically oriented toward the elderly population but include wide spectrum of age groups among which the people older than 65 years old (and especially 85+ years old) represent only a fraction of dataset.

*3.1.1. Cancer.* In USA, about 60% of cancers are diagnosed at age 65 and older [27]. The most detailed US data on cancer incidence come from the SEER registry [1]. Estimates of age-specific incidence rates obtained in our study on NLTCS-M data need to be compared with SEER estimates. Two types of disease cases were identified as cancers using the basic algorithm (i.e., Algorithm A): (i) confirmed in the second record and (ii) only one record existed, and death occurred just after the time of the first record. The second type of cases is not confirmed likely due to forthcoming death. The relative contributions of confirmed and nonconfirmed records differ for NLTCS-M and for SEER registry. The nonconfirmed cancer cases registered in the SEER registry are rare: about 95% of diagnosed cancers in the SEER registry are histologically confirmed and less than 2% of them are from the death certificate or at autopsy because of early death of a patient after the diagnosis [28–30]. Therefore, we use the scheme for onset identification which excluded the second type of events, that is, Algorithm D. This approach exactly corresponds to onset definition in model V12 used in sensitivity analysis (see Table 4). The age patterns of incidence rates calculated using Algorithms A and D are presented in Figure 3. As expected, the SEER rates are mostly between predictions of these two algorithms. All cancers, except melanoma, have demonstrated a good agreement of the age-specific incidence rates. Rates for melanoma do not exceed two standard errors from the SEER estimates. In summary, the results obtained in our study demonstrated a good agreement on cancer incidence disease patterns with SEER registry.

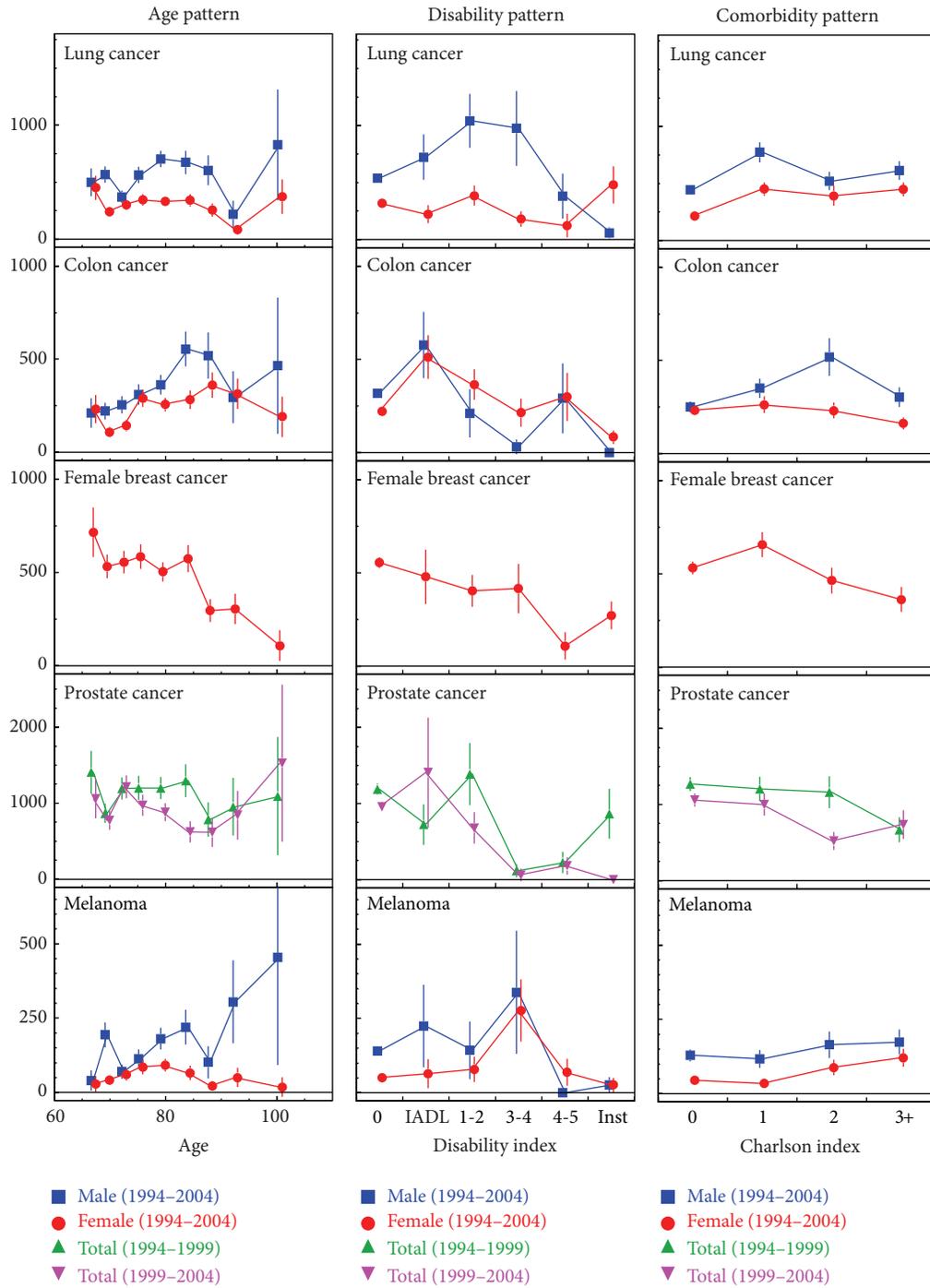


FIGURE 1: Age (left columns), disability (central columns), and comorbidity (right columns) patterns of the incidence rates (per 100,000) of cancers. Arguments are years (age), disability groups (nondisabled, IADL only, 1-2 ADLs, 3-4 ADLs, 5-6 ADLs, and institutionalized), and comorbidity group in the units of Charlson index (0, 1, 2, and 3 and more). The sex-specific rates are given for lung cancer, colon cancer, breast cancer, and melanoma. The time period-specific (i.e., cohort specific) rates are given for prostate cancer. The rates are calculated using the basic strategy for disease onset identification (Algorithm A).

3.1.2. *Heart Failure.* Figure 4 presents the comparison of the results on heart failure incidence rates obtained in our study with the results from cohort studies (summarized in [31]) such as Atherosclerosis Risk in Communities (ARIC), Cardiovascular Health Study (CHS), and Framingham Health

Study (FHS). Figure 4 shows that our results for age-specific incidence rates of congestive heart failure (CHF) are in a good agreement for ARIC and FHS cohort studies estimates and significantly lower than those obtained from CHS. That could be due to the differences in disease incidence

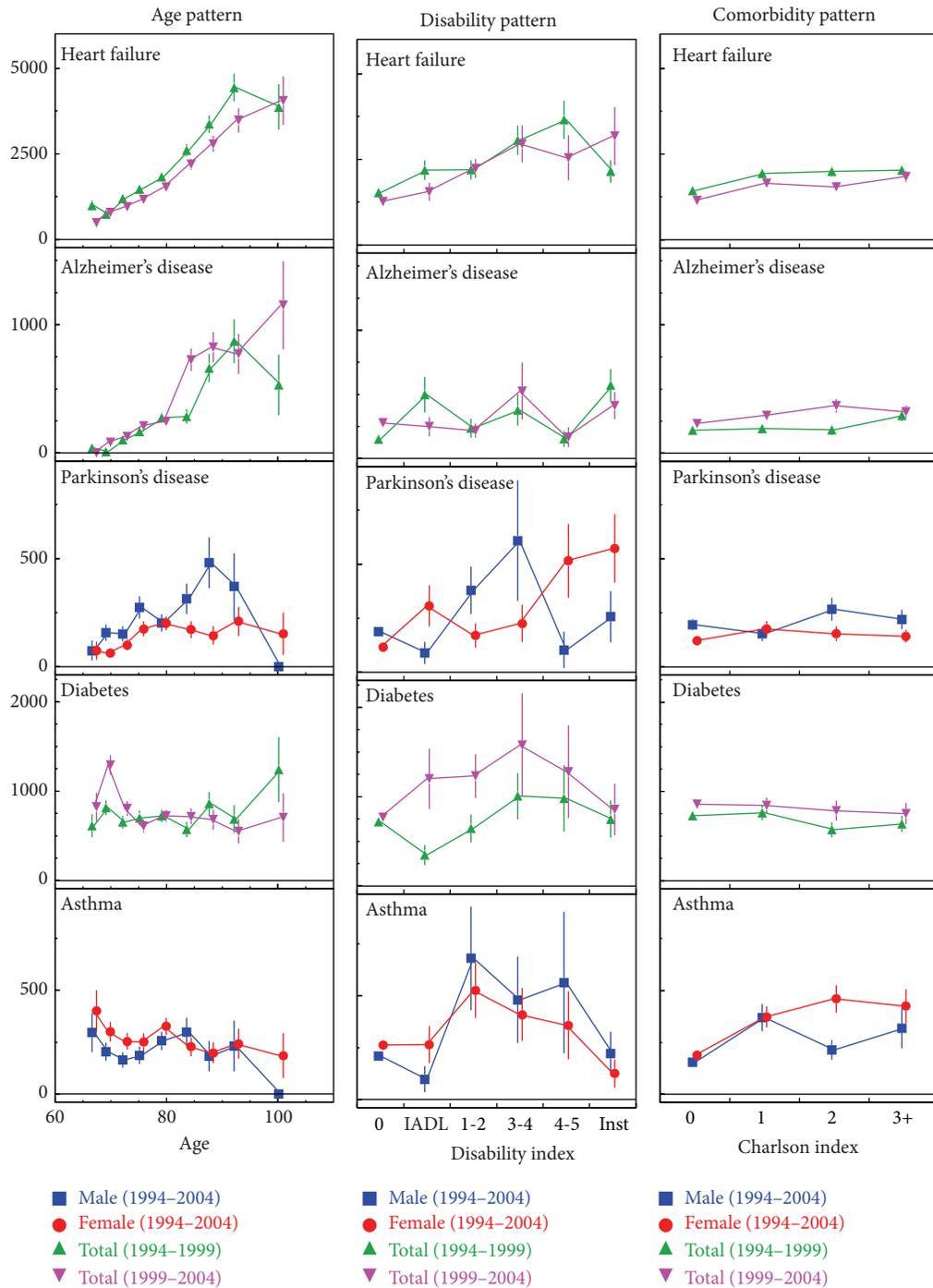


FIGURE 2: Age (left columns), disability (central columns), and comorbidity (right columns) patterns of the incidence rates (per 100,000) of non-cancer chronic diseases. Arguments are years (age), disability groups (nondisabled, IADL only, 1-2 ADLs, 3-4 ADLs, 5-6 ADLs, and institutionalized), and comorbidity group in the units of Charlson index (0, 1, 2, and 3 and more). The sex-specific rates are given for Parkinson's disease and asthma. The time period-specific (i.e., cohort specific) rates are given for heart failure, Alzheimer's disease, and diabetes. The rates are calculated using the basic strategy for disease onset identification (Algorithm A).

definitions in each study (i.e., what criteria were used for case selection/registration?), which could substantially influence the resulting incidence rates. For example, in the FHS cohort study the HF incident event is defined by combination of several major and minor criteria based on HF clinical symptoms,

in ARIC study HF incidence rates were selected based on the hospital discharge ICD-9 codes 428 or 518.4, and in CHS cohort study HF incidence events were defined as diagnosed by a physician plus the patients receiving specific medical treatment (such as diuretic plus either digitalis, vasodilator,

TABLE 4: Age-adjusted incidence rates per 100,000 under alternative approaches to the definition of age at onset. Standard calculation (V0) of age-adjusted incidence rates (per 100,000) was performed according to the aforementioned rules, that is, screener NLTCs population, using the NLTCs weights, the 4 basic Medicare sources, only the primary diagnosis, at least two records (or death) in  $\Delta = 0.3$  years, with cut on frequency of HMO coverage  $\delta = 0.05$ , age standardization using standard population of 1994. Other calculations are (V1) no age standardization; that is; age specific rates are averaged using population of respective year; (V2) No NLTCs weights; (V3) alternative censoring strategy; (V4)  $\delta = 0.5$ ; (V5)  $\delta = 1$ ; (V6) all Medicare sources; (V7)  $\Delta = 0.5$  years; (V8) no requirement for codes to be primary; (V9) no requirement of the second record (Algorithm B); (V10) no requirement from both V8 and V9 (Algorithm C); (V11) McBean's approach; and (V12) no death as a second event (Algorithm D).

	Sex	Year	V0	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	V12
Lung cancer	M	94	616	613	618	623	593	529	572	541	629	657	839	624	454
	M	99	491	490	488	495	484	432	511	462	496	612	711	532	373
	F	94	293	295	280	281	269	248	278	258	288	368	455	323	220
	F	99	331	330	332	333	322	289	342	302	324	393	454	330	264
Colon cancer	M	94	314	306	302	304	280	253	269	267	289	427	483	362	224
	M	99	311	311	321	326	322	296	308	300	356	440	526	377	270
	F	94	262	259	252	253	243	232	245	237	264	349	432	299	218
	F	99	196	195	194	195	192	177	200	187	200	282	361	220	170
Breast cancer	F	94	555	559	572	573	541	498	508	511	544	802	906	680	485
	F	99	508	509	512	514	496	432	454	442	453	580	674	491	424
Prostate cancer	M	94	1146	1148	1116	1125	1098	984	996	1033	1102	1649	2044	1419	944
	M	99	924	922	950	967	939	808	817	876	913	1305	1675	1131	792
Melanoma	M	94	163	159	161	162	149	131	131	137	142	225	277	147	110
	M	99	117	117	130	132	135	113	113	119	131	204	257	137	106
	F	94	53	51	49	49	45	42	44	46	57	139	166	73	36
	F	99	71	70	76	76	76	63	63	64	65	99	127	68	63
Heart fail.	M	94	1864	1795	1869	1881	1784	1612	1706	1767	2713	3055	5345	4484	1389
	M	99	1437	1438	1465	1486	1501	1312	1367	1445	2292	2659	4670	3554	1215
	F	94	1540	1484	1511	1515	1476	1346	1458	1511	2348	3009	5084	4128	1189
	F	99	1313	1313	1323	1329	1349	1185	1274	1301	2042	2500	4397	3166	1084
Parkinson disease	M	94	217	208	221	223	213	190	199	205	272	289	465	408	183
	M	99	195	195	186	189	179	156	159	171	313	282	531	416	150
	F	94	143	139	140	140	135	121	133	128	222	216	405	318	116
	F	99	134	134	127	127	130	117	130	130	216	209	413	283	107
Alzheimer disease	M	94	199	187	200	201	192	174	179	194	396	343	737	607	151
	M	99	255	257	245	247	242	219	246	262	462	540	992	794	183
	F	94	210	197	192	192	189	183	225	221	532	439	1048	876	157
	F	99	290	291	287	288	296	264	301	323	717	670	1457	1114	239
Diabetes	M	94	724	721	718	723	725	681	715	772	1303	1787	3403	2529	646
	M	99	841	841	801	816	884	784	878	936	1478	2130	4158	2654	739
	F	94	697	696	702	704	707	650	678	728	1179	1665	3197	2219	611
	F	99	803	804	797	801	799	713	818	827	1260	1835	3636	2195	673
Asthma	M	94	201	200	202	204	198	174	198	211	395	744	1484	804	157
	M	99	247	248	263	267	260	228	260	255	478	723	1591	818	206
	F	94	303	307	309	310	301	287	329	316	550	934	1692	1045	278
	F	99	258	257	228	229	239	213	274	232	525	842	1679	901	200

or angiotensin converting enzyme inhibitor) [31]. That can explain, at least in part, that incidence rates of HF in CHS cohort study were higher than those obtained from ARIC and FHS studies, as well as our results (see Figure 4).

3.1.3. *Diabetes*. Figure 5 shows that our results are in agreement with several studies performed on other cohorts, such as Canadian Study of Health and Aging [32], Zwolle Outpatient

Diabetes project Integrating Available Care (ZODIAC-1, The Netherlands, [33]), and UK Pooled Diabetes Study [34]. In general, the incidence of diabetes did not change significantly with age in males and females in both of the above-mentioned studies and also in our study (except for the first and the last points of the age pattern obtained in our study). In ZODIAC-1 study the diabetes type II incidence in 1998–2000 was slightly higher than that in our study while in the UK the Poole

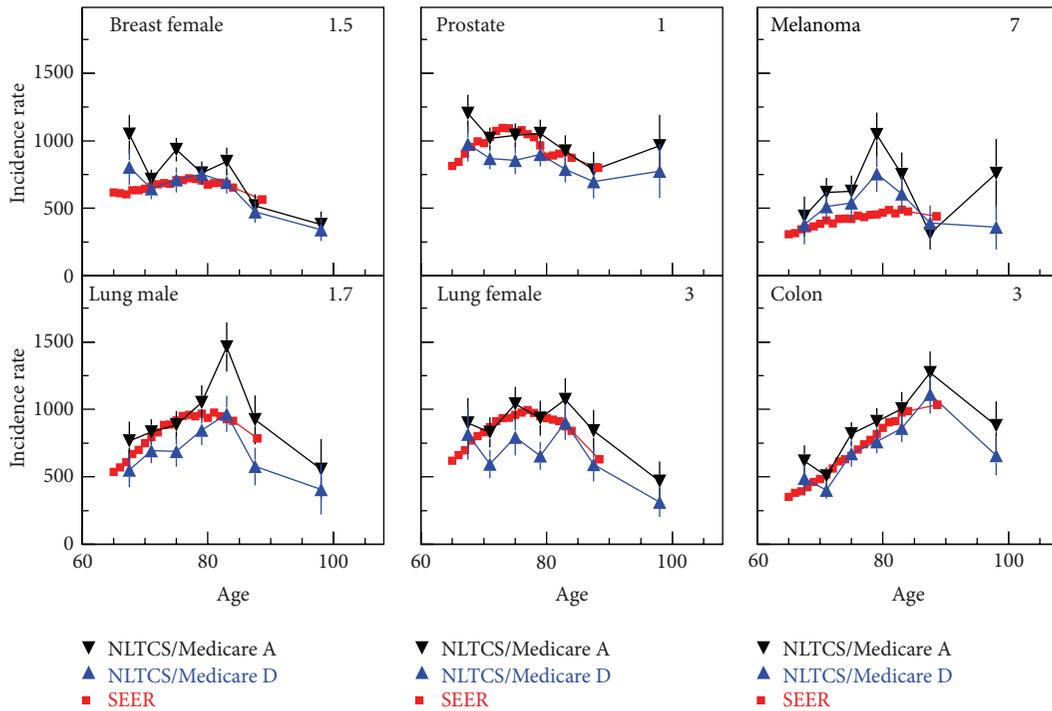


FIGURE 3: Age-specific cancer incidence rates: means and SE of NLTCS/Medicare (Algorithms A and D) and SEER data for 1994–2003. The rate for SEER above 85 is shown at mean age of cases above 85. The number in the right upper corner is the renormalization factor: true incidence rates are obtained by dividing plotted rates by this factor.

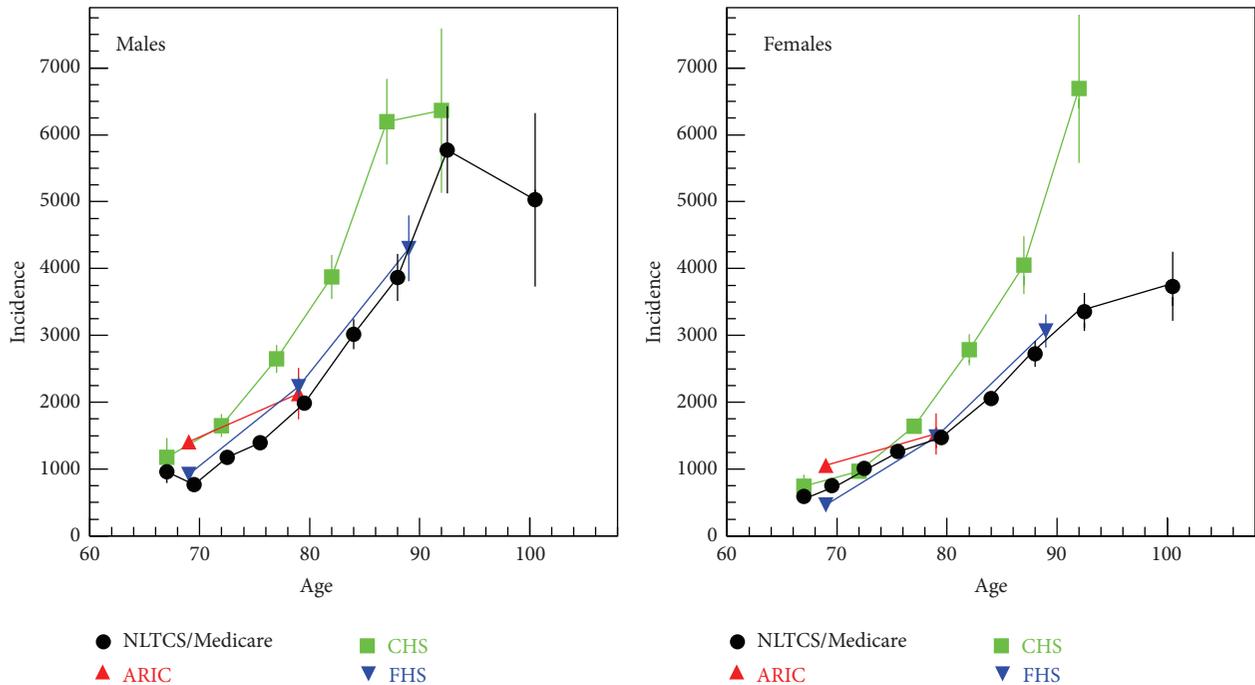


FIGURE 4: Age-specific incidence rates for heart failure obtained in this study (NLTCS/Medicare) in comparison with rates in other studies: the Atherosclerosis Risk in Communities (ARIC), the Cardiovascular Health Study (CHS), and the Framingham Health Study (FHS).

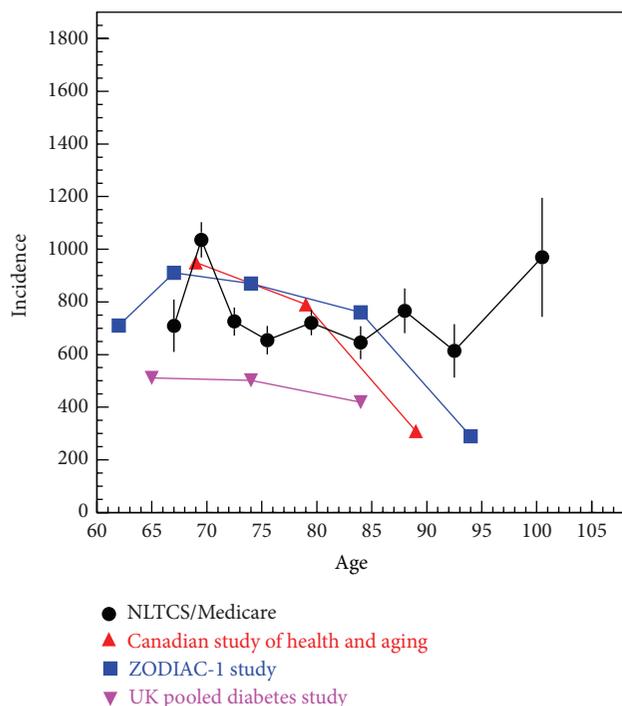


FIGURE 5: Age-specific incidence rates for diabetes mellitus obtained in this study (NLTCs/Medicare) in comparison with those studied for other countries.

Diabetes Study, those rates are slightly lower than in our study. Generally, the trend and absolute incidence rates in all considered studies correspond to our results. The age-specific predicted incidence rates of adult-onset diabetes were calculated in population-based retrospective study using community-based medical records in Rochester, Minnesota [35]: in 1985, the incidence rate per 100,000 person-years was about 600 for ages 70–74, and about 500 for ages 80–84. These results are in agreement with our estimates for these age groups.

Another study on diabetes was recently published by McBean et al. [5] who examined diabetes prevalence, incidence, and mortality from 1993 to 2001 among fee-for-service Medicare beneficiaries aged 67+ using a 5% random sample of enrollees. They estimated the rate at the level of 3,000 per 100,000. The reason of this disagreement could be in another scheme for diabetes onset identification [36] used for Medicare data. In this scheme, it is required that a second record with diabetes ICD must be observed only if the first one was registered as an ambulatory claim (i.e., a physician/supplier or hospital outpatient claim). This is a reason for the excess in the incidence rate in results of McBean et al. [5]. We recalculated incidence rates for diabetes using the approach used in [5, 36] and found almost 5-fold increase in the incidence rates comparing to the estimates given by Algorithm A. One of the alternative approaches for diseases onset identification model (model VII given in Table 4) was exactly the approach used by McBean et al. [5]. The results for age-adjusted rates found for this approach were about 2,700 for males and 2,300 for females, which as expected are close to those obtained by McBean et al. [5].

**3.1.4. Asthma.** Asthma may occur in the elderly more frequently than it is usually appreciated and be, therefore, underdiagnosed [37]. Few studies described asthma onset at ages 65+ and available studies are often limited by small numbers of patients and a tendency to group all patients older than 55 or 60 years old as a single category [38–41]. Age- and sex-specific incidence rates for definite and probable asthma per 100,000 person-years were analyzed in a population-based study in Rochester, Minnesota, for 1964–1983 [42]: incidence rates decreased with age the same way as in our study; however, the absolute incidence rates were lower than those obtained from our analysis—so, the rates were about 140 for males and 80 for females at ages 65–74, 110 for males and 70 for females at ages 75–84, and about 60 for males and 50 for females at ages 85+. Higher rates in our study could be due to the increased incidence of asthma over the decades, as well as due to the fact that data from the Rochester study were obtained from the medical records retrospectively, and it involved mostly Caucasians of a small midwestern city. Also the slight difference between male/female ratios in the two studies may also play a role. Also the results for incidence rate of asthma were reported by ARIC (summarized by the NIH/NHLBI in [31]) collected in 1987–2001. They were 225 for ages 65–74 and 398 (this estimate is considered unreliable) for ages 75–84, which is in a better agreement with our results. Female to male ratio in age-adjusted rates was similar in both studies: 1.49 (ARIC) and 1.27 (NLTCs-M). In a study conducted on Moscow population (Russia), which covered ages from the birth to age 85, asthma risk in adults declined steadily at ages 55+ in females and at ages 65+ in males and became very small among the oldest elderly [43, 44]. These declining trends are in agreement with our results.

**3.1.5. Neurodegenerative Diseases (NDD).** The most common disorders in the NDD group are Alzheimer’s disease (and other dementia) and Parkinson’s disease. They have been in the focus of several studies and meta-analyses which estimated their incidence rates and age patterns in elderly populations. We compare the results of our calculation to the meta-analysis [4] and to the analysis of CHS [3] (Figure 6). Two algorithms (A and C) were used to calculate our results. The results of both of those studies are in agreement with our results obtained by Algorithm C and much higher than our results obtained using the base algorithm. The results from the Bronx Aging Study have demonstrated that whereas dementia incidence continues to increase beyond age 85, the rate of the increase appears to slow relative to that of 65- to 85-year-olds, suggesting that dementia in the oldest old might be related not to the aging process itself but to the age-associated risk factors [45, 46]. Similar pattern for Alzheimer’s disease was observed in the study based on inpatient claims in the NLTCs-M data for 1984–2001: risk decline was registered at ages 90+ [47]. Epidemiological studies on Alzheimer’s disease incidence from Europe, North America, Asia, Africa, Australia, and South America were analyzed and average annual age-specific incidence rates per 1,000 person-years were estimated for ages 65–95: the increasing rates of disease are in agreement with our data, and the absolute rates were in general agreement with our Algorithm C data [45].

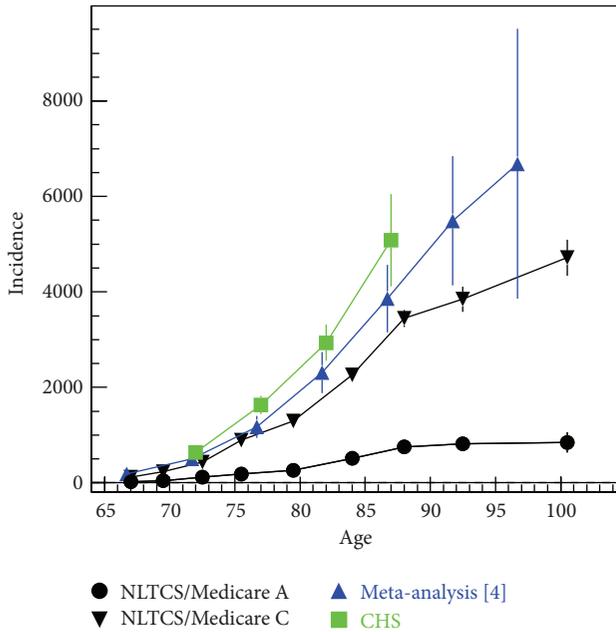


FIGURE 6: Age-specific incidence rates for Alzheimer's disease obtained in this study (NLTCS/Medicare) using the Algorithms A and C in comparison with other studies.

The difficulties with studying the incidence data on Parkinson's disease result from the difficulties in identifying a sufficiently large number of affected individuals in a well-defined or enumerated population. Low frequency of disease, difficulties in establishing diagnosis, and the absence of population-based disease registries contributed to the lack of its basic epidemiologic characteristics [48]. There is still few data on Parkinson's disease incidence, especially in the oldest elderly, and it remains controversial whether there is a progressive rise in late life or a decline in incidence [49, 50]. Parkinson's disease cases were analyzed for 1994-1995 for the members of the Kaiser Permanente Medical Care Program of Northern California: incidence rates per 100,000 were 38.8 at ages 60-69, 107.2 at ages 70-79, and 119.0 at ages 80-89, being more than twice higher in males than in females at ages 70+ [48]. Although our rates for ages above 80 are higher by a factor 1.5-2 and male to female ratio is 1.4, since statistical errors are large, we can conclude that generally these results are in agreement with those observed in our study.

## 4. Discussion

**4.1. NLTCS-M Information As a Source of Information about Referent Values of Incidence Rates for the US Elderly Population.** In this study the estimates on changes of incidence rates of cancer and non-cancer chronic diseases with (i) age among the older US adults (males and females), (ii) disability prevalence, and (iii) prevalence and severity of comorbidities were obtained using the NLTCS-Medicare-linked data. Disease incidences were analyzed for ten chronic conditions representing major groups of chronic conditions in the elderly; (i) circulatory: heart failure; (ii) cancer: breast, prostate, lung, and colon cancers and melanoma; (iii) neurodegenerative:

Parkinson's disease and Alzheimer's disease; (iv) diabetes; and (v) asthma. We have demonstrated that NLTCS-M dataset could be very useful for answering the spectrum of questions on the elderly health in the USA from both medical and economical perspectives. Also, this dataset allows for bringing additional information which cannot be obtained from the other datasets: for example, comorbidity and disability. These data are population-based, minimizing selection bias with respect to geographic region, urban versus rural location, racial health disparities with a whole spectrum of race- and ethnic-specific populations, and socioeconomic characteristics. Each of these factors is an important predictor of disease risk, progression, treatment availability, and response. This information is limited in databases from more restricted populations [15].

**4.2. Identified Properties of Incidence Rates.** Generally, the obtained results were in accordance with our expectations. The comparison of the age patterns with other studies, as well as their sex differences and time trends, demonstrated similarities of these patterns with those obtained in other population studies in USA and other countries. Patterns of the majority of diseases were well described by the base algorithm, the most important features of which included (i) the occurrence of primary diagnosis in one of four Medicare sources (inpatient care, outpatient care, physician services, and skilled nursing facilities) and (ii) the confirmation of the diagnoses by another record. The only exception was Alzheimer's disease: its patterns required certain corrections to the base algorithm to be adequately described; that is, only one record is sufficient and it needs not be primary. Note that for Alzheimer's disease the nonpostmortal diagnosis (i.e., largely subjective or included a subjective component in distinguishing between Alzheimer's disease and dementia) is still implemented; therefore, our result that an algorithm without confirmation better fits data could tell us that the incidence of Alzheimer's disease can be overestimated among the oldest elderly patients [51, 52]. Disability was measured using self-reported information, while comorbidity was estimated using Medicare records during the year prior the date of interview and the beginning of 5-year follow-up period. Because of using the special weight, the results are valid at the national level. The properties of these patterns were briefly discussed in Section 3.

Several types of age patterns were observed in our study. The first type was flat or plateau. The diseases manifesting this shape were prostate cancer, melanoma, and diabetes. Note that in the analyses of the shape of age patterns the first and the last point can be cut. The first point can still have a mixture of prevalence case and the last point typically has larger statistical uncertainty. The second type of the shape was monotonic increase with age. Diseases with this shape were heart failure and Alzheimer's disease. The third type had the shape with a maximum or inverted U-shape. Respective diseases were lung cancer, colon cancer, and Parkinson's disease. Age at maximal rate was in the region 80-90 years for all cases. The fourth shape appeared in the analysis had monotonic rates decline. Breast cancer and asthma possessed these shapes.

Occurrences of the shapes with a maximum and, especially, with monotonic decline contradict the hypothesis that risk of geriatric diseases correlates with accumulation of adverse health events (genetic mutations, deterioration of vascular system, immunosenescence, etc.). Three basic concepts could be considered to explain this phenomenon. The appearance of such effects can be attributed to the effect of selection [53], where frail individuals do not survive to advanced ages. Respective approaches are popular in cancer modeling and were successfully applied to the SEER data [54–56]. Another explanation is related to the possible under-registration of diagnoses at advanced ages; that, however, cannot be proved with available data [57, 58]. Other possible biologically motivated explanations are discussed in [59]: the authors suggested the mechanisms of possible contribution of individual aging to the shape of mortality curve by subdividing the individual age-associated changes into three components which had different influences on morbidity and mortality (i.e., basal, ontogenetic, and time dependent) and discussed how all three components of individual age-associated changes may interact in human organism and influence patterns of morbidity and mortality in population.

Also, note that the shape of age pattern can depend on how broad or, alternatively, how narrow the definition of the set of disease is. Akushevich et al. [60] considered broader definition of the diseases. For example, they considered total cancer disease including all malignant neoplasms (140–208) and neurodegenerative disorders (NDD) included psychoses (290–299), nonpsychotic mental disorders (300–316), and hereditary and degenerative diseases of the central nervous system (330–337). In such definition incidence rates for NDD monotonically increase with age. This is because NDD groups cover a wide range of specific diseases that are prevalent in the elderly population. Relatively few individuals at advanced ages do not have diagnoses corresponding to diseases from this group, significantly decreasing the number of unexposed persons. By focusing on specific diseases from these groups, we observe less pronounced increases—or even declines—of incidence rates at advanced ages that result in a peak in the age pattern.

*4.3. Addressing the Limitations of Medicare Data.* Medicare claim data have certain limitations that are with a matter of the determination of diagnoses. Sensitivity analysis is one possible way to deal with such uncertainties. In this paper we analyzed several sources of possible uncertainties such as different definitions of disease onset and different censoring schemes. The rates obtained for different schemes of onset identification can be significantly different, but this simply corresponds to different definitions of incidence rates, many of which are used in epidemiology (e.g., fatal and nonfatal incidence rates). The following specific sources of potential uncertainties were in the focus of this study: (i) approaches to identification of incident cases (all Medicare sources, keeping only primary diagnoses, different definitions of the onset), (ii) censoring strategies (dependence between recovery and death), (iii) disenrollment from Medicare and coverage by HMO, and (iv) study design effects (e.g., using NLTCs weights). Also, the impact of factors of observed

heterogeneity (e.g., age, disability, and comorbidity) on the incidence rates is also investigated.

*4.4. Incidence Rates and Forecasting Models.* Modern models of forecasting the health state and the associated medical costs include three essential components or submodels: (i) the model of medical cost projections conditional on health state, (ii) health state projections, and (iii) description of initial health state of a cohort to be projected [61–64]. In this paper we investigated the risks of the chronic diseases and their age patterns as well as associations of the risks with potential covariates like comorbidity and disability indices. A detailed investigation of the model of medical cost projections conditional on health state was presented by us in [19]. Given the developed models of health state and the medical cost projections (as well as a quantified description of the initial state of a cohort), the projection for a population can be constructed by simulating individual trajectories. The total expenditures are calculated as a sum of medical costs obtained for each individual. If the health state needs to be described by a number of chronic diseases, then the detailed stratification over respective categorized variables or use of multivariate regression models allows for better description of health states; however, it can result in an abundance of model parameters to be estimated. One way to overcome these difficulties is to use an approach where the model components would be some demographically based aggregated characteristics allowing to mimic the effects of specific states. In this case the information about comorbidity- and disability-specific patterns of incidence rates has to be incorporated into the model. The model developed using this concept (i.e., the use of comorbidity index rather than the set of multiple correlated categorical variables representing the health state) allows for essential reduction in the degrees of freedom of the problem. Note that time trends and covariates are naturally implemented into such an approach.

## 5. Conclusion

The age-, disability-, and comorbidity-specific incidence rates of ten highly prevalent aging-related chronic diseases were analyzed using the NLTCs-M data. The most appropriate approach for identification of the disease onset required forthcoming occurrence of repeated claims containing chosen ICD codes as a prime diagnosis in basic Medicare sources. Comparing the age patterns obtained using this computational approach with those available in the literature showed a good agreement for the majority of diseases. Thus, the national incidence rates can be adequately evaluated from the Medicare service use files. Usefulness of the Medicare data for evaluation of the national incidence rates is very important because of limited data sources for evaluation of incidence patterns at advanced ages in the national population. These timely results can inform current scientific and policy debates about the effects of biomedical research and therapeutic innovations on disease incidence at increasingly advanced ages.

## Acknowledgments

The research reported in this paper was supported by the National Institute on Aging Grants R01AG027019, R01AG032319, and R01AG028259. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute on Aging or the National Institutes of Health.

## References

- [1] S. F. Altekruse, C. L. Kosary, M. Krapcho et al., "SEER Cancer Statistics Review, 1975–2007," 2009, [http://seer.cancer.gov/csr/1975\\_2007/](http://seer.cancer.gov/csr/1975_2007/).
- [2] A. M. Arnold, B. M. Psaty, L. H. Kuller et al., "Incidence of cardiovascular disease in older Americans: the cardiovascular health study," *Journal of the American Geriatrics Society*, vol. 53, no. 2, pp. 211–218, 2005.
- [3] A. L. Fitzpatrick, L. H. Kuller, D. G. Ives et al., "Incidence and prevalence of dementia in the cardiovascular health study," *Journal of the American Geriatrics Society*, vol. 52, no. 2, pp. 195–204, 2004.
- [4] S. Gao, H. C. Hendrie, K. S. Hall, and S. Hui, "The relationships between age, sex, and the incidence of dementia and Alzheimer disease: a meta-analysis," *Archives of General Psychiatry*, vol. 55, no. 9, pp. 809–815, 1998.
- [5] A. M. McBean, S. Li, D. T. Gilbertson, and A. J. Collins, "Differences in diabetes prevalence, incidence, and mortality among the elderly of four racial/ethnic groups: whites, blacks, Hispanics, and Asians," *Diabetes Care*, vol. 27, no. 10, pp. 2317–2324, 2004.
- [6] E. Kokmen and C. Mary Beard, "Epidemiology of dementia in Rochester, Minnesota," *Mayo Clinic Proceedings*, vol. 71, no. 3, pp. 275–282, 1996.
- [7] R. A. Miech, J. C. S. Breitner, P. P. Zandi, A. S. Khachaturian, J. C. Anthony, and L. Mayer, "Incidence of AD may decline in the early 90s for men, later for women: the Cache County study," *Neurology*, vol. 58, no. 2, pp. 209–218, 2002.
- [8] K. G. Manton, L. Corder, and E. Stallard, "Chronic disability trends in elderly United States populations: 1982–1994," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 94, no. 6, pp. 2593–2598, 1997.
- [9] K. G. Manton, E. Stallard, and L. S. Corder, "The dynamics of dimensions of age-related disability 1982 to 1994 in the U.S. elderly population," *Journals of Gerontology. Series A*, vol. 53, no. 1, pp. B59–B70, 1998.
- [10] K. G. Manton and X. Gu, "Changes in the prevalence of chronic disability in the United States black and nonblack population above age 65 from 1982 to 1999," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 98, no. 11, pp. 6354–6359, 2001.
- [11] K. G. Manton, X. Gu, and V. L. Lamb, "Change in chronic disability from 1982 to 2004/2005 as measured by long-term changes in function and health in the U.S. elderly population," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 103, no. 48, pp. 18374–18379, 2006.
- [12] NLTC, "National Long Term Study, public use dataset," Produced and distributed by the Duke University Center for Demographic Studies with funding from the National Institute on Aging under Grant No. U01-AG007198.
- [13] M. E. Charlson, P. Pompei, K. A. Ales, and C. R. MacKenzie, "A new method of classifying prognostic comorbidity in longitudinal studies: development and validation," *Journal of Chronic Diseases*, vol. 40, no. 5, pp. 373–383, 1987.
- [14] H. Quan, V. Sundararajan, P. Halfon et al., "Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data," *Medical Care*, vol. 43, no. 11, pp. 1130–1139, 2005.
- [15] A. B. Nattinger, P. W. Laud, R. Bajorunaite, R. A. Sparapani, and J. L. Freeman, "An algorithm for the use of medicare claims data to identify women with incident breast cancer," *Health Services Research*, vol. 39, no. 6 I, pp. 1733–1749, 2004.
- [16] A. B. Nattinger, P. W. Laud, R. Bajorunaite, R. A. Sparapani, and J. L. Freeman, "Clarification note to an algorithm for the use of medicare claims data to identify women with incident breast cancer," *Health Services Research*, vol. 41, no. 1, p. 302, 2006.
- [17] F. A. Sloan, D. S. Brown, E. S. Carlisle, J. Ostermann, and P. P. Lee, "Estimates of incidence rates with longitudinal claims data," *Archives of Ophthalmology*, vol. 121, no. 10, pp. 1462–1468, 2003.
- [18] A. Yashin, I. Akushevich, S. Ukraintseva, L. Akushevich, K. Arbee, and A. Kulminski, "Trends in survival and recovery from stroke: evidence from the national long-term care survey/medicare data," *Stroke*, vol. 41, no. 3, pp. 563–565, 2010.
- [19] I. Akushevich, J. Kravchenko, L. Akushevich, S. Ukraintseva, K. Arbee, and A. I. Yashin, "Medical cost trajectories and onsets of cancer and noncancer diseases in US elderly population," *Computational and Mathematical Methods in Medicine*, vol. 2011, Article ID 857892, 14 pages, 2011.
- [20] I. Akushevich, J. Kravchenko, S. Ukraintseva, K. Arbee, and A. I. Yashin, "Age patterns of incidence of geriatric disease in the U.S. elderly population: medicare-based analysis," *Journal of the American Geriatrics Society*, vol. 60, no. 2, pp. 323–327, 2012.
- [21] I. Akushevich, J. Kravchenko, L. Akushevich, S. Ukraintseva, K. Arbee, and A. Yashin, "Cancer risk and behavioral factors, comorbidities, and functional status in the US elderly population," *ISRN Oncology*, vol. 2011, Article ID 415790, 9 pages, 2011.
- [22] L. D. Brown, T. T. Cai, and A. DasGupta, "Interval Estimation for a Binomial Proportion," *Statistical Science*, vol. 16, no. 2, pp. 101–133, 2001.
- [23] A. J. Dobson, K. Kuulasmaa, E. Eberle, and J. Scherer, "Confidence intervals for weighted sums of Poisson parameters," *Statistics in Medicine*, vol. 10, no. 3, pp. 457–462, 1991.
- [24] M. P. Fay and E. J. Feuer, "Confidence intervals for directly standardized rates: a method based on the gamma distribution," *Statistics in Medicine*, vol. 16, no. 7, pp. 791–801, 1997.
- [25] N. E. Breslow and N. E. Day, "Statistical methods in cancer research. Volume II—the design and analysis of cohort studies," *IARC Scientific Publications*, no. 82, pp. 1–406, 1987.
- [26] N. Keyfitz, "3. Sampling variance of standardized mortality rates," *Human Biology*, vol. 38, no. 3, pp. 309–317, 1966.
- [27] M. Hewitt and J. Simone, Eds., *Enhancing Data Systems to Improve the Quality of Cancer Care*, The National Academies Press, 2000.
- [28] A. Bleyer, M. O'Leary, R. Barr, and L. A. G. Ries, Eds., *Cancer Epidemiology in Older Adolescents and Young Adults 15 to 29 Years of Age, Including SEER Incidence and Survival: 1975–2000*, NIH Pub. No. 06-5767, National Cancer Institute, Bethesda, Md, USA, 2006.
- [29] C. Johnson and M. Adamo, *SEER Program Coding and Staging Manual 2007*, National Cancer Institute, National Institutes of Health, US Department of Health and Human Services, Bethesda, Md, USA.

- [30] L. A. G. Ries, J. L. Young, G. E. Keel, M. P. Eisner, Y. D. Lin, and M.-J. Horner, Eds., *SEER Survival Monograph: Cancer Survival Among Adults: U.S. SEER Program, 1988–2001, Patient and Tumor Characteristics*, NIH Pub. No. 07-6215, National Cancer Institute, SEER Program, Bethesda, Md, USA.
- [31] NIH/NHLBI, “Incidence and Prevalence: 2006 Chart Book on Cardiovascular and Lung Diseases,” 2006, [http://www.nhlbi.nih.gov/resources/docs/06a\\_ip\\_chtbk.pdf](http://www.nhlbi.nih.gov/resources/docs/06a_ip_chtbk.pdf).
- [32] K. Rockwood, E. Awalt, C. MacKnight, and I. McDowell, “Incidence and outcomes of diabetes mellitus in elderly people: report from the Canadian Study of Health and Aging,” *Canadian Medical Association Journal*, vol. 162, no. 6, pp. 769–772, 2000.
- [33] L. J. Ubink-Veltmaat, H. J. G. Bilo, K. H. Groenier, S. T. Houweling, R. O. Rischen, and B. Meyboom-De Jong, “Prevalence, incidence and mortality of type 2 diabetes mellitus revisited: a prospective population-based study in The Netherlands (ZODIAC-1),” *European Journal of Epidemiology*, vol. 18, no. 8, pp. 793–800, 2003.
- [34] W. Gatling, R. N. Guzder, J. C. Turnbull, S. Budd, and M. A. Mullee, “The Poole Diabetes Study: how many cases of type 2 diabetes are diagnosed each year during normal health care in a defined community?” *Diabetes Research and Clinical Practice*, vol. 53, no. 2, pp. 107–112, 2001.
- [35] C. L. Leibson, P. C. O’Brien, E. Atkinson, P. J. Palumbo, and L. J. Melton, “Relative contributions of incidence and survival to increasing prevalence of adult-onset diabetes mellitus: a population-based study,” *American Journal of Epidemiology*, vol. 146, no. 1, pp. 12–22, 1997.
- [36] P. L. Hebert, L. S. Geiss, E. F. Tierney, M. M. Engelgau, B. P. Yawn, and A. M. McBean, “Identifying persons with diabetes using medicare claims data,” *American Journal of Medical Quality*, vol. 14, no. 6, pp. 270–277, 1999.
- [37] D. K. Banerjee, G. S. Lee, S. K. Malik, and S. Daly, “Underdiagnosis of asthma in the elderly,” *British Journal of Diseases of the Chest*, vol. 81, no. 1, pp. 23–29, 1987.
- [38] H. Y. Lee and T. B. Stretton, “Asthma in the elderly,” *British Medical Journal*, vol. 4, no. 832, pp. 93–95, 1972.
- [39] M. L. Burr, T. J. Charles, K. Roy, and A. Seaton, “Asthma in the elderly: an epidemiological survey,” *British Medical Journal*, vol. 1, no. 6170, pp. 1041–1044, 1979.
- [40] S. S. Braman, J. T. Kaemmerlen, and S. M. Davis, “Asthma in the elderly: a comparison between patients with recently acquired and long-standing disease,” *American Review of Respiratory Disease*, vol. 143, no. 2, pp. 336–340, 1991.
- [41] R. M. Ford, “Aetiology of asthma: a review of 11,551 cases (1958 to 1968),” *Medical Journal of Australia*, vol. 1, no. 12, pp. 628–631, 1969.
- [42] R. Ballard-Barbash et al., “Obesity and body composition,” in *Cancer Epidemiology and Prevention*, D. Schottenfeld and J. F. Fraumeni, Eds., Oxford University Press, New York, NY, USA, 2006.
- [43] S. V. Ukraintseva and A. S. Sergeev, “Population risk of bronchial asthma in Moscow,” *Genetika*, vol. 31, no. 2, pp. 264–267, 1995.
- [44] S. V. Ukraintseva and A. S. Sergeev, “Analysis of genetic heterogeneity of bronchial asthma as related to the age of onset,” *Genetika*, vol. 36, no. 2, pp. 266–270, 2000.
- [45] R. de la Fuente-Fernández, “Impact of neuroprotection on incidence of Alzheimer’s disease,” *PLoS ONE*, vol. 1, no. 1, article e52, 2006.
- [46] C. B. Hall, J. Verghese, M. Sliwinski et al., “Dementia incidence may increase more slowly after age 90: results from the Bronx Aging Study,” *Neurology*, vol. 65, no. 6, pp. 882–886, 2005.
- [47] S. Ukraintseva, F. Sloan, K. Arbeeve, and A. Yashin, “Increasing rates of dementia at time of declining mortality from stroke,” *Stroke*, vol. 37, no. 5, pp. 1155–1159, 2006.
- [48] S. K. Van Den Eeden, C. M. Tanner, A. L. Bernstein et al., “Incidence of Parkinson’s disease: variation by age, gender, and race/ethnicity,” *American Journal of Epidemiology*, vol. 157, no. 11, pp. 1015–1022, 2003.
- [49] R. Mayeux, K. Marder, L. J. Cote et al., “The frequency of idiopathic Parkinson’s disease by age, ethnic group, and sex in Northern Manhattan, 1988–1993,” *American Journal of Epidemiology*, vol. 142, no. 8, pp. 820–827, 1995.
- [50] D. M. Morens, J. W. Davis, A. Grandinetti, G. W. Ross, J. S. Popper, and L. R. White, “Epidemiologic observations on Parkinson’s disease: incidence and mortality in a prospective study of middle-aged men,” *Neurology*, vol. 46, no. 4, pp. 1044–1050, 1996.
- [51] M. Fotuhi, “The challenge and public health implications of Alzheimer overdiagnosis in the oldest old,” *Archives of Neurology*, vol. 67, no. 7, p. 899, 2010.
- [52] R. S. Wilson, D. R. Weir, S. E. Leurgans et al., “Sources of variability in estimates of the prevalence of Alzheimer’s disease in the United States,” *Alzheimer’s and Dementia*, vol. 7, no. 1, pp. 74–79, 2011.
- [53] J. W. Vaupel, J. R. Carey, K. Christensen et al., “Biodemographic trajectories of longevity,” *Science*, vol. 280, no. 5365, pp. 855–860, 1998.
- [54] J. Trussell and T. Richards, “Correcting for unmeasured heterogeneity in hazard models using the Heckman-Singer procedure,” in *Sociological Methodology*, vol. 15, pp. 242–276, Jossey-Bass, San Francisco, Calif, USA, 1985.
- [55] K. Manton, I. Akushevich, and J. Kravchenko, “Cancer mortality and morbidity patterns in the U.S. population: an interdisciplinary approach,” in *Statistics for Biology and Health*, M. Gail, K. Krickeberg, J. Samet, A. Tsiatis, and W. Wong, Eds., Springer, New York, NY, USA, 1 edition, 2009.
- [56] A. I. Yashin, I. Akushevich, K. Arbeeve, L. Akushevich, A. Kulminski, and S. Ukraintseva, “Studying health histories of cancer: a new model connecting cancer incidence and survival,” *Mathematical Biosciences*, vol. 218, no. 2, pp. 88–97, 2009.
- [57] P. L. Enright, R. L. McClelland, A. B. Newman, D. J. Gottlieb, and M. D. Lebowitz, “Underdiagnosis and undertreatment of asthma in the elderly,” *Chest*, vol. 116, no. 3, pp. 603–613, 1999.
- [58] P. R. Solomon and C. A. Murphy, “Should we screen for Alzheimer’s disease? A review of the evidence for and against screening for Alzheimer’s disease in primary care practice,” *Geriatrics*, vol. 60, no. 11, pp. 26–31, 2005.
- [59] S. V. Ukraintseva and A. I. Yashin, “How individual age-associated changes may influence human morbidity and mortality patterns,” *Mechanisms of Ageing and Development*, vol. 122, no. 13, pp. 1447–1460, 2001.
- [60] I. Akushevich, A. Kulminski, L. Akushevich, and K. G. Manton, “Age Patterns of Disease Incidences in the U.S. Elderly: Population-Based Analysis,” TRENDS Working Paper Series 2006; No. 06-06, 2006, <http://www.psc.isr.umich.edu/pubs/pdf/tr06-6.pdf>.
- [61] D. P. Goldman, P. G. Shekelle, J. Bhattacharya et al., “Health status and medical treatment of the future elderly,” Final Report TR-169-CMS, RAND, Santa Monica, Calif, USA, 2004.

- [62] D. P. Goldman, B. Shang, J. Bhattacharya et al., “Consequences of health trends and medical innovation for the future elderly,” *Health Affairs*, vol. 24, supplement 2, pp. W5R5–W5R17, 2005.
- [63] D. P. Goldman, D. M. Cutler, B. Shang, and G. F. Joyce, “The value of elderly disease prevention,” in *Forum for Health Economics & Policy*, 2006.
- [64] D. P. Goldman, Y. Zheng, F. Girosi et al., “The benefits of risk factor prevention in Americans aged 51 years and older,” *American Journal of Public Health*, vol. 99, no. 11, pp. 2096–2101, 2009.



# Hindawi

Submit your manuscripts at  
<http://www.hindawi.com>

