



Published in final edited form as:

Exp Gerontol. 2013 August ; 48(8): 824–830. doi:10.1016/j.exger.2013.05.056.

Recovery and Survival from Aging-Associated Diseases

Igor Akushevich, PhD¹, Julia Kravchenko, MD, PhD², Svetlana Ukraintseva, PhD^{1,2}, Konstantin Arbeev, PhD¹, and Anatoliy I. Yashin, PhD, Sc.D.^{1,2}

¹Center for Population Health and Aging, Duke University, Durham, NC 27708

²Duke Cancer Institute, Duke University, Durham, NC 27705

Abstract

Objectives—Considering disease incidence to be a main contributor to healthy lifespan of the US elderly population may lead to erroneous conclusions when recovery/long-term remission factors are underestimated. Using two Medicare-based population datasets, we investigated the properties of recovery from eleven age-related diseases.

Methods—Cohorts of patients who stopped visiting doctors during a five-year follow-up since disease onset were analyzed non-parametrically and using the Cox proportional hazard model resulted in estimated recovery and survival rates and evaluated the health state of recovered individuals by comparing their survival with non-recovered patients and the general population.

Results—Recovered individuals had lower death rates than non-recovered patients, therefore, patients who stopped visiting doctors are a healthier subcohort. However, they had higher death rates than in general population for all considered diseases, therefore the complete recovery does not occur.

Conclusion—Properties of recovery/long-term remission among the US population of older adults with chronic diseases were uncovered and evaluated. The results allow for a better quantifiable contribution of age-related diseases to healthy life expectancy and improving forecasts of health and mortality.

Keywords

Medicare; chronic disease onset; recovery/long-term remission; population-based analysis

Introduction

The question of particular importance for public health specialists is the relationship between increasing life span and the period of healthy life. Recent studies have demonstrated an improvement in the health status of the general population during the 20th century, including consistent evidence of improved disability status among the US elderly (Costa, 2002; Crimmins et al., 2009; Freedman et al., 2004; Manton et al., 2006; Schoeni et al., 2001; Waidmann and Liu, 2000), with a postponement of limitations in functioning despite an increased prevalence of certain chronic diseases. This improvement could be

© 2013 Elsevier Inc. All rights reserved.

Corresponding Author: Igor Akushevich, PhD, Center for Population Health and Aging, Duke University, 002 Trent Dr., Durham, NC 27708, Phone: (919) 668-2715, Fax: (919) 684-3861, igor.akushevich@duke.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

explained by earlier diagnosis, improved treatment, and amelioration of prevalent diseases (Christensen et al., 2009; Crimmins, 2004; Parker and Thorslund, 2007; Perenboom et al., 2004; Schoeni et al., 2008). However, while most of the statistics represent the changes in disability and long-term care, they do not account for the frequency of patients' rehabilitation, long-term remission of chronic diseases, and recovery (Campion, 1994). These events have been described as non-rare among the oldest adults: e.g., for asthma, a remission among elderly was observed in 30% of disease cases (Reed, 2006; Rhodes et al., 2002; Slavin et al., 2006; Taussig et al., 2003). Therefore, understanding age- and sex-specific patterns, as well as time trends in these outcomes among the US older adults, is of particular importance. While trends of geriatric disease prevalence suggested an increase in the chronic disease burden in the 1970s-1990s (Goff Jr et al., 2007), deeper insight into the patterns and trends of different phases of chronic diseases provide a background for optimizing the strategy of medical services, expenditures, insurance planning, and quality of life studies.

An increasing burden in certain disease prevalence and associated medical costs among the elderly US population does not necessarily mean an increase in the incidence rate. Instead, it may be due to the fact that the improvement in patients' survival contributes to disease prevalence greater than the decline in disease risk. In this context, the tremendous research potential of the data from the Medicare Files of Service Use (MFSU) for studying current and forecasting future health patterns of the US older adults remains unexplored. In this paper, we use two Medicare linked datasets to investigate the demographic and epidemiologic properties of the cohorts of survivors after certain chronic diseases were diagnosed. Part of such a cohort can "recover" or enter into long-term remission. In the majority of traditional forecasting models, such individuals are usually considered as still being ill (i.e., without a dynamic update of health status). This masks the effects of successful medical treatment on healthy life span and creates an illusion of "decompression of morbidity". Therefore, our study is focused on identification of patients who are "recovered" or entered into long-term remission, evaluation of the rates of recovery, and comparing characteristics of their survival with non-recovered patients and the general population. In this study we define a "recovery" not as a complete restoration of functioning of all systems and complete social rehabilitation but rather supposing that it is the absence of records on medical services related to the disease of interest which is coded by the ICD-9 codes. We focused on individuals with a diagnosed disease who are staying for long periods of time (say, one year or more) in such health conditions that they do not need to ask for medical services associated with the disease. For acute conditions (e.g., myocardial infarction or stroke) such health conditions are associated with recovery, and for chronic conditions (e.g., asthma or nephritis) they are associated with long-term remission. We use the term "recovery/long-term remission" or simply "recovery" to refer to this state.

Data and Methods

Two datasets capable of providing estimates for the older adult population at the national level are used in this study: the Surveillance, Epidemiology, and End Results (SEER) Registry data linked to MFSU (SEER-M), and the National Long Term Care Survey (NLTCS-M) also linked to the MFSU. Using these extensive sources of information allows for identification of disease incidence and recovery events through elaboration and validation of specific computational algorithms of these events from administrative data.

The SEER-M data was the primary dataset analyzed in this study. The expanded SEER registry covers approximately 26% of the U.S. population. In total, the Medicare records for 2,154,598 individuals are available in SEER-M including individuals i) with diagnosed cancers: breast (n=353,285), colon (n=222,659), lung (n=342,961), prostate (n=448,410)

cancers and skin melanoma (n=101,123) and ii) from a random 5% sample of Medicare beneficiaries residing in the SEER areas who had none of the above mentioned cancers. For the majority of persons, we have continuous records of Medicare services use from 1991 (or from the time the person has passed the age of 65 after 1990) until his/her time of death. A small fraction of individuals (e.g., new patients diagnosed with cancer in 2003–2005) has Medicare records from 1998. Medicare records are available for each institutional (MedPAR, outpatient, hospice, or home health agency HHA) and non-institutional (Carrier-Physician-Supplier and durable medical equipment providers) claim type.

The NLTCS-M data contains two of the six NLTCS waves—namely, cohorts of 1994 and 1999. These two waves were chosen primarily because of the high-quality Medicare follow-up data available from 1991, and also because the complete 5-year follow-up after the NLTCS interview is accessible only for these two waves after 1991. In total, 34,077 individuals were followed-up between 1994 and 1999. So-called screener weights released with the NLTCS allowed for the production of the national population estimates (for recent discussion, see (Akushevich et al., 2012a)).

Disease selection and definition of the date of disease onset

To illustrate our approach, several diseases were selected for this study based on their prevalence in the US elderly population (i.e., the most prevalent) and those which were among the leading causes of death in the elderly: acute coronary heart disease (410.xx, 411.xx, 413.xx), stroke (431.xx, 433.x1, 434.x1, 436.xx), breast cancer (174.xx), prostate cancer (185.xx), skin melanoma (172.xx), lung cancer (162.xx), colon cancer (153.xx), ulcer (531.xx–534.xx), asthma (493.xx), nephritis/nephrosis (580.xx–585.xx, 587.xx–589.xx), and hip fracture (820.xx, 821.xx).

The ages at onsets of all diseases were reconstructed from the MFSU using the scheme described in details in (Akushevich et al., 2011a, b; Akushevich et al., 2012b, a). In the first step the individual medical histories of the applicable disease were reconstructed using all records with their respective ICD-9 codes. In the second step, a special computational procedure was applied for individuals with the history of the considered disease to separate incident and prevalent cases and to identify the age at disease onset. This procedure is based on two conditions applied to each medical history. The first condition allows for identification of the first occurrence of the disease code, and the second condition is required for confirmation of the disease's presence. The individual Medicare history contains all records with the respective disease ICD-9 code, however only records with the primary ICD-9 code and only from the so-called base Medicare sources (inpatient care, outpatient care, physician services, and skilled nursing facilities) are used for the disease onset identification. Several examples of individual trajectories illustrating our approach and basic definitions of used quantities are presented in Figure 1. Table 1 provides the number of individuals with disease onset identified for SEER-M and NLTCS-M datasets and several other characteristics of their distributions.

Only disease onsets detected in the first three years of NLTCS-M follow-up (i.e., in period from April 1, 1994 to April 1, 1997 and from April 1, 1999 to April 1, 2001) were collected. For detection of recovery the information from individual histories was used since the disease onset and till death or the end of follow-up that is the end of 2005.

Definition of recovery and long-term remission

An individual was considered to be recovered (or started to have a sustained remission) at a given date if he/she did not have a Medicare record containing the respective ICD code in his/her medical history during a time period τ_d (e.g., one, two, or three years) after this date

(Yashin et al., 2010). All Medicare records from the individual medical histories (i.e., records appeared in any Medicare sources and not necessary was a primary) are used for the definition of recovery. The time periods τ_d are referred to as recovery times. An individual was censored at a date if that date plus the recovery time exceeded the date of the end of follow-up (i.e., in this case an individual does not have a room for recovery).

Note, that since the identification of the date at onset requires confirmation by a record in another day, recovery at the day of diagnosis is not possible even if the length of service is day 0. Another property of the recovery rate is that a recovery event within the three-year strategy (i.e., for $\tau_d=3$ years) implies a recovery event within a one- or two-year strategy, but not otherwise.

Evaluation of the effects of recovery/remission

The Cox proportional hazard model with the covariates represented the recovery event and age at diagnosis as a cofactor was designed to evaluate the effect of recovery or long-term remission on survival. Individual follow-up starts from the date of diagnosis and ends by the date of death from any causes or censoring date (the end of 2005). Recovery is described by the binary indicator, $x=1$ if $t > t_{rec}$, and $x=0$ otherwise, where t_{rec} is a period between the diagnosis and recovery. Importantly, that part of the follow-up period that was used for identification of the recovery event (i.e., time spot τ_d started at the time of recovery) has to be excluded from parameter estimates because the survival probability in this period is exactly one. Technically, the hazard ratios are estimated by considering two time periods as two observations with different time periods (and different times of the beginning) of follow-up: between diagnosis and recovery for the first and between recovery time plus τ_d and death/censoring for the second. The option of varying beginning of the follow-up is available in major packages for statistical analysis (e.g., in Proc PHREG of SAS 9.2, that was used in our analysis). This consideration is exactly equivalent to the more intuitive (but not feasible using standard software) approach when i) follow-up of all individuals starts from the date of diagnosis, ii) one observation represents one (and only one) observation, iii) the recovery state is represented by the time-dependent indicator (as in Figure 1), and iv) individuals with the current follow-up time appeared in time spot τ_d are excluded from the set in individuals contributed to the denominator of the partial likelihood.

In addition, the Cox proportional hazard model is used to compare the survival of recovered individuals with survival in the general population. In this model age is used as a follow-up variable, therefore the estimated hazard ratio represents the death hazard ratios for recovered individuals and the general population of the same age. Similarly to the above model the option of varying beginning of the follow-up is used, and the time period τ_d used for recovery identification is excluded, i.e., the initial age for recovered individuals is the age of recovery plus τ_d .

Results

The results of analyses of survival and recovery/remission for males and females are shown in Figure 2 for five-year follow-up since disease diagnosis. Both survival and recovery are shown in each plot. The first is presented by the Kaplan-Meier estimate of the survival curve and the second is presented by a similar estimate of the probability of not-yet-recovery. Three scenarios of the “recovery time” are shown (i.e., when the patient has not seen medical specialists for one, two, and three years since the last disease-specific record in MFSU). Because of limited statistical power for non-parametric analyses using NLTCs-M, only results for SEER-M are presented in the figure and the results for recovery rates from analysis of both SEER-M and NLTCs-M are shown in Supplementary Figure 1. The

obtained results on the recovery/remission rates calculated using NLTCs-M and SEER-M data were similar for all studied diseases.

For the studied diseases, several patterns of recovery/remission and survival could be described: i) substantial decrease of prevalence of “not-yet-recovered” individuals during 1st and 2nd years and slow decrease of survival over a 5-year period (e.g., for ACHD and asthma), ii) continuous decline of the “not-yet-recovered” patients during the follow-up period with a pronounced decrease of survival of a concave shape (e.g., for stroke, nephritis/nephrosis, and colon cancer), iii) dramatic decrease of the “not-yet-recovered” persons during 1st year and slow decrease of the survival over a 5-year period (e.g., for ulcer, hip fracture, and skin melanoma), and iv) continuous decrease of the “not-yet-recovered” which stayed high even after 5 years and continuous decrease of survival which stayed pretty high (e.g., for breast and prostate cancers). Lung cancer had its unique pattern characterized by a concave shape of survival curve with very sharp decrease of survival during the 1st year.

Using the Cox proportional hazard model where the outcome is the risk of death and the covariate is the indicator of recovery/remission and the age is used as a cofactor, allows us to evaluate the effect of recovery/remission and to test whether recovered individuals belong to a healthier or sicker subcohort. Table 2 shows the death hazard ratios of occurrence (versus non-occurrence) of recovery/long-term remission with recovery times 1, 2, or 3 years. For all diseases the hazard rates are significantly lower than one. This means that in all cases the risk of death is significantly lower for the cohort of recovered individuals. It means that the patients with long periods without the appearance of new records with ICD-9 of the given disease belong to the healthier subcohort. The results are similar for the two datasets used in the analyses and for different recovery times.

The survival of recovered individuals was also compared to the survival of the general population to understand the extent in which the recovered patients are sicker than the general population of the same age. The Cox proportional hazard model with age as follow-up variable was applied for both cohorts (for the SEER-M, cohorts represented 5% of total population including both cancer cases and non-cancer control), and the effect of the indicator variable shown the recovery from the given disease was estimated in a series of unidimensional analyses for $\tau_d=1, 2,$ and 3 years. The results for the death hazard ratio were presented in Table 3. For all diseases the hazard ratios are greater than one. This allows us to conclude that time periods without disease-specific records longer than 3 years is required to reach the health state of the general population. Another observation from Table 3 is that for majority of diseases (except breast and prostate cancers, melanoma, hip fracture) hazard ratios for $\tau_d=3$ years are lower than those for $\tau_d=1$ year. Therefore, the cohorts formed with requirement $\tau_d=3$ years are healthier than those for $\tau_d=1$ year. Results for the NLTCs-M data and for total SEER-M population are similar.

For all studied diseases, males had lower survival than females, while recovery/remission rates did not differ for the two genders (the hazard ratios differed by less than several percents). The results for the hazard ratios of being males versus females are presented in Supplementary Table 1.

Discussion

The presented results on recovery or long-term remission of aging-related diseases and survival were obtained from analyses of the two large datasets: SEER-M and NLTCs-M. These results are representative at the national level—specifically, for the U.S. elderly population (aged 65 years and older). At our best knowledge, a majority of published studies, including those based on NLTCs-M data, predominantly focused on the prevalence

of disability rather than of occurrence of a long-term remission of chronic diseases or on estimation of transition rates between disability states and their time trends among the older and oldest adults (Manton, 1988; Manton et al., 1993; Schoeni et al., 2008; Stallard, 2011; Zimmer and House, 2003). Compared with the predictors of disease onset, there were few predictors of recovery (Clark et al., 1998). The mathematical multistate models that explicitly allow for transitions between unhealthy and recovery states has been discussed since the early 1980s (Mathers and Robine, 1997; Putter et al., 2007), however, the estimates of the model parameters from nationally representative datasets and comprehensive analysis of properties of recovery estimates have not been performed. The difficulties in studies on recovery could be mainly due to the fact that the data on disease onset and recovery are sparse. Therefore, the obtained results based on the two nationally representative large datasets are valuable.

We followed and generalized the definition used by Yashin et al. (Yashin et al., 2010) for analysis of time trend in recovery from stroke. In our study a similar definition was also applied to ten other diseases. Recovery within this approach signifies nonappearance of the same diagnosis coded by the respective ICD-9-CM code. For example, for myocardial infarction such a definition means the absence of i) another myocardial infarction and ii) complications directly related to myocardial infarction (when myocardial infarction ICD code appears as secondary code in Medicare records) during the period τ_d . Within this definition we find the median no-recovery (i.e., time period when 50% of patients are recovered) equaling 0.086 year. A stricter definition of recovery can be used with the additional requirement of non-occurring complications such as angina pectoris and heart failure without myocardial infarction codes. In this case the estimate of median non-recovery is larger: 0.408 year. Yashin et al. (Yashin et al., 2010) demonstrated that the estimates of time trend in recovery rates are robust: the results did not depend on different definitions of recovery event and were not sensitive to definitions of disease onset. For example, the authors considered several alternative strategies for definition of the onset of stroke (e.g., all Medicare sources used, keeping not only primary diagnoses, no requirement of the second record with the disease-specific code, and other) and found that positive trends in the recovery rate from stroke took place in all cases independent of the definition of both incidence and recovery rates. However further sensitivity analyses of the results is still required. For example, in our analysis, deaths that occurred before recovery were treated as a censoring event independent of recovery. This is a strong assumption, and since available data (i.e., time to event data that are used in the study) does not allow us to identify rates under the assumption of dependent competing risks (Tsiatis, 1975), the sensitivity analysis based on model assumptions needs to be performed.

The estimates in Table 2 show that “recovered” patients (those who did not have medical care for one, two or three years, depending on the corresponding scenario analyzed) had a better survival. Therefore they represent i) a “healthier” subgroup of patients who started feeling well enough to discontinue use of medical services or ii) a subgroup of the patients who at the moment of diagnosis likely had a functional disorder rather than a serious disease but the similarity of the symptoms lead to “overweighting” of diagnosis. Also the a subcohort of “recovered” individuals can have a fraction of patients who stopped visiting medical specialists, probably, due to their disappointment in treatment results (instead, they could turning to chiropractors, homeopaths, or herbal medicine) or patients who started experiencing difficulties with transportation to reach the medical facilities (note, that this subcohort of patients did not have a better survival). However this subcohort is not distinguishable from the healthier (and much larger) fraction of recovered individuals.

We also found (see Table 3) that recovered individuals had lower survival than general population. Among studied diseases, lung cancer, hip fracture, and nephritis had the lowest

relative survival (characterized by highest hazard ratios presented in Table 3), while interruption of visiting medical specialists, and ACHD, prostate cancer and skin melanoma had the highest relative survival. This can be related to the characteristics of these diseases (e.g., earlier diagnosis and higher success of the treatment for melanoma and prostate cancer compared with other cancers, relatively good “repairing capacity” of cardiovascular system after myocardial infarction) and/or higher comorbidities in certain groups of the patients such as those with kidney disease and hip fracture. The differences in the sex-specific relative survivals for ACHD are in agreement with findings that ACHD symptoms are more likely to be unrecognized in females (Lerner and Kannel, 1986) and that women, especially those older than 65 years old, delays longer than do men before seeking medical treatment for symptoms of MI (Lefler and Bondy, 2004). In addition, that could be due to their higher depression score which is also associated with increased risk of coronary mortality after myocardial infarction (Welin et al., 2000). Thus, the selected group of diagnosed patients is still heterogeneous; e.g., they may include a fraction of false positive diagnoses or be due to comorbid conditions. Further investigation of groups of patients stratified over characteristics of diagnosed cancers is required. For example, we evaluated the Charlson comorbidity index (Charlson et al., 1987, Quan et al., 2006) at the month of diagnosis and investigated its effects for all considered diseases. We found that i) recovery is less probable for patients with higher comorbidity: HR of the Charlson index from 0.84 (breast cancer) to 0.97 (hip fracture) per index unit, ii) as expected, the Charlson comorbidity index influences survival with HR=1.07–1.17 (except lung cancer HR=1.01 and asthma 1.24), and iii) comorbidity modifies the effect of recovery (HRs of the interaction term are in the range of 1.02–1.08). At the next stages of (disease-specific) analyses the detailed consideration of related comorbidities and analysis of their effects on recovery rates and subsequent survival would be useful and performed in separate studies. The examples of expected effects that are observed in clinical practice of a gerontologist could be the following: i) mortality among the elderly patients with peptic ulcer is higher among those with diabetes mellitus compared with the patients without diabetes (Thomsen et al., 2006); an impairment of ulcer healing could be, at least in part, due to reduction in gastric microcirculation with impaired effects of gastric vascular endothelial growth factor and gastric tumor necrosis factor alpha among those with diabetes (Baraka et al., 2010); and ii) the preexisting renal disease, even mild, has been shown to be a major risk factor for delaying the recovery period and developing of cardiovascular complications such as heart failure, left ventricular systolic dysfunction, or both, after a myocardial infarction (Anavekar et al., 2004).

A close similarity in patterns of the dynamics of recovery/remission and survival over a 5-year period observed for different diseases could be due, at least in part, to the similarity of these disease courses. For example, ACHD and asthma both have periods of acute attacks followed by periods with lighter symptoms or symptom free. Ulcer, hip fracture and skin melanoma all could be completely treated during a relatively short period, without a recurrence. Breast and prostate carcinomas are both hormonal dependent and have an intensive screening strategy. Stroke (with moderate and severe symptoms), nephritis/nephrosis, and colon cancer have progradiently progressive courses leading to severe impairments of organ functions including the whole organism’s systemic disorders.

It is well known, that the average lifespan in males is shorter than in females. Our results provide specific details on which aging-related diseases contribute to these gender-specific disparities. While recovery or long-term remission rates did not differ in males and females, survival rates of many diseases (such as ulcer, asthma, nephritis/nephrosis, lung cancer, and skin melanoma) were much lower in males. So, we could suppose that gender-specific disparities are minor in “healthier” populations but are pronounced among “sicker” populations.

Limitations of the approach are basically due to our focus on developing a common view on disease onsets and disease recovery/remission. Such an approach is natural for the first stage of analysis of complex phenomena in public health. It is certainly useful for investigating demographic health related aspects of the U.S. population of older adults, but this approach is not quite conclusive for clinical aspects of recovery. However, it could be generalized to become clinically related by considering detailed disease-specific definitions of onsets and recovery (e.g., by extending the set of codes). Note that the SEER-Medicare data does not have information about changes in individual functional status since the disease onset. Therefore the analyses of the effects of individual functional status, including physical, mental, cognitive, and social functions, should involve other data (e.g., the NLTCs-Medicare data). Only age, time after diagnoses, and sex of individuals were used in the Cox regression model for evaluation of recovery hazard ratios. The population can remain heterogeneous after adjusting over these variables. Among other variables that can be used for adjusting are characteristics of cancer diagnoses, comorbidities, cost of diagnosis (estimated as the total cost during the month of diagnosis) as an approximate measure of disease severity, and race. Further sensitivity studies could also take into account the effects of disease severity, insurance status, social-economic status, as well as the MSA or geographic region of the country.

In summary, a new approach is developed for quantitative analyses of individuals with recovery/long-term remission after onset of chronic diseases using Medicare data. The effects of heterogeneity in recovery rates from common geriatric diseases were evaluated for the U.S. elderly population at the national level. To our knowledge, such type of analysis has never been done before. The approach opens new opportunities for developing predictive models with the time dependent covariates representing health status. Such models could be further used to better quantify the contribution of age-related diseases to healthy life expectancy and to improve forecasts of health and mortality. The results of such analysis can be important for prognosis of medical expenditures for the elderly, because people living with chronic conditions are the groups of the population which requires expensive long-term medical services.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

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.

Reference

- Akushevich I, Kravchenko J, Akushevich L, Ukraintseva S, Arbeev K, Yashin AI. Cancer Risk and Behavioral Factors, Comorbidities, and Functional Status in the US Elderly Population. *ISRN Oncology* 2011. 2011a:415790.
- Akushevich I, Kravchenko J, Akushevich L, Ukraintseva S, Arbeev K, Yashin AI. Medical Cost Trajectories and Onsets of Cancer and NonCancer Diseases in US Elderly Population. *Comput Math Methods Med* 2011. 2011b:857892.
- Akushevich I, Kravchenko J, Ukraintseva S, Arbeev K, Yashin AI. Circulatory diseases in the U.S. elderly in the linked National Long Term Care Survey-Medicare database: population-based analysis of incidence, comorbidity, and disability. *Research on Aging in press*. 2012a

- Akushevich I, Kravchenko J, Ukraintseva S, Arbeev K, Yashin AI. Age patterns of incidence of geriatric disease in the u.s. Elderly population: medicarebased analysis. *J Am Geriatr Soc.* 2012b; 60:323–327. [PubMed: 22283485]
- Anavekar NS, McMurray JJV, Velazquez EJ, Solomon SD, Kober L, Rouleau JL, White HD, Nordlander R, Maggioni A, Dickstein K. Relation between renal dysfunction and cardiovascular outcomes after myocardial infarction. *New England Journal of Medicine.* 2004; 351:1285–1295. [PubMed: 15385655]
- Baraka AM, Guemei A, Gawad HA. Role of modulation of vascular endothelial growth factor and tumor necrosis factor-alpha in gastric ulcer healing in diabetic rats. *Biochemical pharmacology.* 2010; 79:1634–1639. [PubMed: 20144589]
- Campion EW. The oldest old. *N Engl J Med.* 1994; 330:1819–1820. [PubMed: 8190160]
- Christensen K, Doblhammer G, Rau R, Vaupel JW. Ageing populations: the challenges ahead. *Lancet.* 2009; 374:1196–1208. [PubMed: 19801098]
- Clark DO, Stump TE, Wolinsky FD. Predictors of onset of and recovery from mobility difficulty among adults aged 51–61 years. *Am J Epidemiol.* 1998; 148:63–71. [PubMed: 9663405]
- Costa DL. Changing chronic disease rates and long-term declines in functional limitation among older men. *Demography.* 2002; 39:119–137. [PubMed: 11852833]
- Crimmins EM. Trends in the health of the elderly. *Annu Rev Public Health.* 2004; 25:79–98. [PubMed: 15015913]
- Crimmins EM, Hayward MD, Hagedorn A, Saito Y, Brouard N. Change in disability-free life expectancy for Americans 70-years-old and older. *Demography.* 2009; 46:627–646. [PubMed: 19771948]
- Freedman VA, Crimmins E, Schoeni RF, Spillman BC, Aykan H, Kramarow E, Land K, Lubitz J, Manton K, Martin LG, Shinberg D, Waidmann T. Resolving inconsistencies in trends in old-age disability: report from a technical working group. *Demography.* 2004; 41:417–441. [PubMed: 15461008]
- Goff D Jr, Brass L, Braun L, Croft J, Flesch J, Fowkes F, Hong Y, Howard V, Huston S, Jencks S. Essential features of a surveillance system to support the prevention and management of heart disease and stroke: a scientific statement from the American Heart Association Councils on Epidemiology and Prevention, Stroke, and Cardiovascular Nursing and the Interdisciplinary Working Groups on Quality of Care and Outcomes Research and Atherosclerotic Peripheral Vascular Disease. *Circulation.* 2007; 115:127–155. [PubMed: 17179025]
- Lefler LL, Bondy KN. Women's delay in seeking treatment with myocardial infarction: a meta-synthesis. *Journal of Cardiovascular Nursing.* 2004; 19:251–268. [PubMed: 15326981]
- Lerner DJ, Kannel WB. Patterns of coronary heart disease morbidity and mortality in the sexes: a 26-year follow-up of the Framingham population. *American heart journal.* 1986; 111:383–390. [PubMed: 3946178]
- Manton KG. A longitudinal study of functional change and mortality in the United States. *J Gerontol.* 1988; 43:S153–S161. [PubMed: 2971088]
- Manton KG, Corder LS, Stallard E. Estimates of change in chronic disability and institutional incidence and prevalence rates in the U.S. elderly population from the 1982, 1984, and 1989 National Long Term Care Survey. *J Gerontol.* 1993; 48:S153–S166. [PubMed: 8315240]
- Manton KG, Gu X, Lamb VL. Change in chronic disability from 1982 to 2004/2005 as measured by long-term changes in function and health in the US elderly population. *Proceedings of the National Academy of Sciences of the United States of America.* 2006; 103:18374–18379. [PubMed: 17101963]
- Mathers CD, Robine J-M. How good is Sullivan's method for monitoring changes in population health expectancies? *J Epidemiol Community Health.* 1997; 51:80–86. [PubMed: 9135793]
- Parker MG, Thorslund M. Health trends in the elderly population: getting better and getting worse. *Gerontologist.* 2007; 47:150–158. [PubMed: 17440120]
- Perenboom RJ, Van Hertem LM, Boshuizen HC, Van Den Bos GA. Trends in disability-free life expectancy. *Disabil Rehabil.* 2004; 26:377–386. [PubMed: 15204474]
- Putter H, Fiocco M, Geskus R. Tutorial in biostatistics: competing risks and multi-state models. *Statistics in Medicine.* 2007; 26:2389–2430. [PubMed: 17031868]

- Reed CE. The natural history of asthma. *J Allergy Clin Immunol.* 2006; 118:543–548. quiz 549–50. [PubMed: 16950268]
- Rhodes HL, Thomas P, Sporik R, Holgate ST, Cogswell JJ. A birth cohort study of subjects at risk of atopy: twenty-two-year follow-up of wheeze and atopic status. *Am J Respir Crit Care Med.* 2002; 165:176–180. [PubMed: 11790650]
- Schoeni RF, Freedman VA, Martin LG. Why is late-life disability declining? *Milbank Q.* 2008; 86:47–89. [PubMed: 18307477]
- Schoeni RF, Freedman VA, Wallace RB. Persistent, consistent, widespread, and robust? Another look at recent trends in old-age disability. *J Gerontol B Psychol Sci Soc Sci.* 2001; 56:S206–S218. [PubMed: 11445613]
- Slavin RG, Haselkorn T, Lee JH, Zheng B, Deniz Y, Wenzel SE. Asthma in older adults: observations from the epidemiology and natural history of asthma: outcomes and treatment regimens (TENOR) study. *Ann Allergy Asthma Immunol.* 2006; 96:406–414. [PubMed: 16597074]
- Stallard E. Estimates of the Incidence, Prevalence, Duration, Intensity, and Cost of Chronic Disability among the US Elderly. *North American Actuarial Journal.* 2011; 15:32–58.
- Taussig LM, Wright AL, Holberg CJ, Halonen M, Morgan WJ, Martinez FD. Tucson Children's Respiratory Study: 1980 to present. *J Allergy Clin Immunol.* 2003; 111:661–675. quiz 676. [PubMed: 12704342]
- Thomsen RW, Riis A, Christensen S, Nørgaard M, Sørensen HT. Diabetes and 30-Day Mortality From Peptic Ulcer Bleeding and Perforation A Danish population-based cohort study. *Diabetes Care.* 2006; 29:805–810. [PubMed: 16567819]
- Tsiatis A. Nonidentifiability Aspect of Problem of Competing Risks. *Proceedings of the National Academy of Sciences of the United States of America.* 1975; 72:20–22. [PubMed: 1054494]
- Waidmann TA, Liu K. Disability trends among elderly persons and implications for the future. *J Gerontol B Psychol Sci Soc Sci.* 2000; 55:S298–S307. [PubMed: 10985301]
- Welin C, Lappas G, Wilhelmsen L. Independent importance of psychosocial factors for prognosis after myocardial infarction. *Journal of Internal medicine.* 2000; 247:629–639. [PubMed: 10886484]
- Yashin A, Akushevich I, Ukraintseva S, Akushevich L, Arbeev K, Kulminski A. Trends in survival and recovery from stroke: evidence from the National Long-Term Care Survey/Medicare data. *Stroke.* 2010; 41:563–565. [PubMed: 20093634]
- Zimmer Z, House JS. Education, income, and functional limitation transitions among American adults: contrasting onset and progression. *Int J Epidemiol.* 2003; 32:1089–1097. [PubMed: 14681281]

Research Highlights

- Medicare data is a powerful source for analyzing recovery and survival
- Most patients who stopped visiting doctors comprise a healthier subcohort
- Medicare-based data provide reliable estimates valid for the US elderly population
- Recovery rates are consistently evaluated using different Medicare-based data
- Survival of recovered patients is lower than survival of the general population

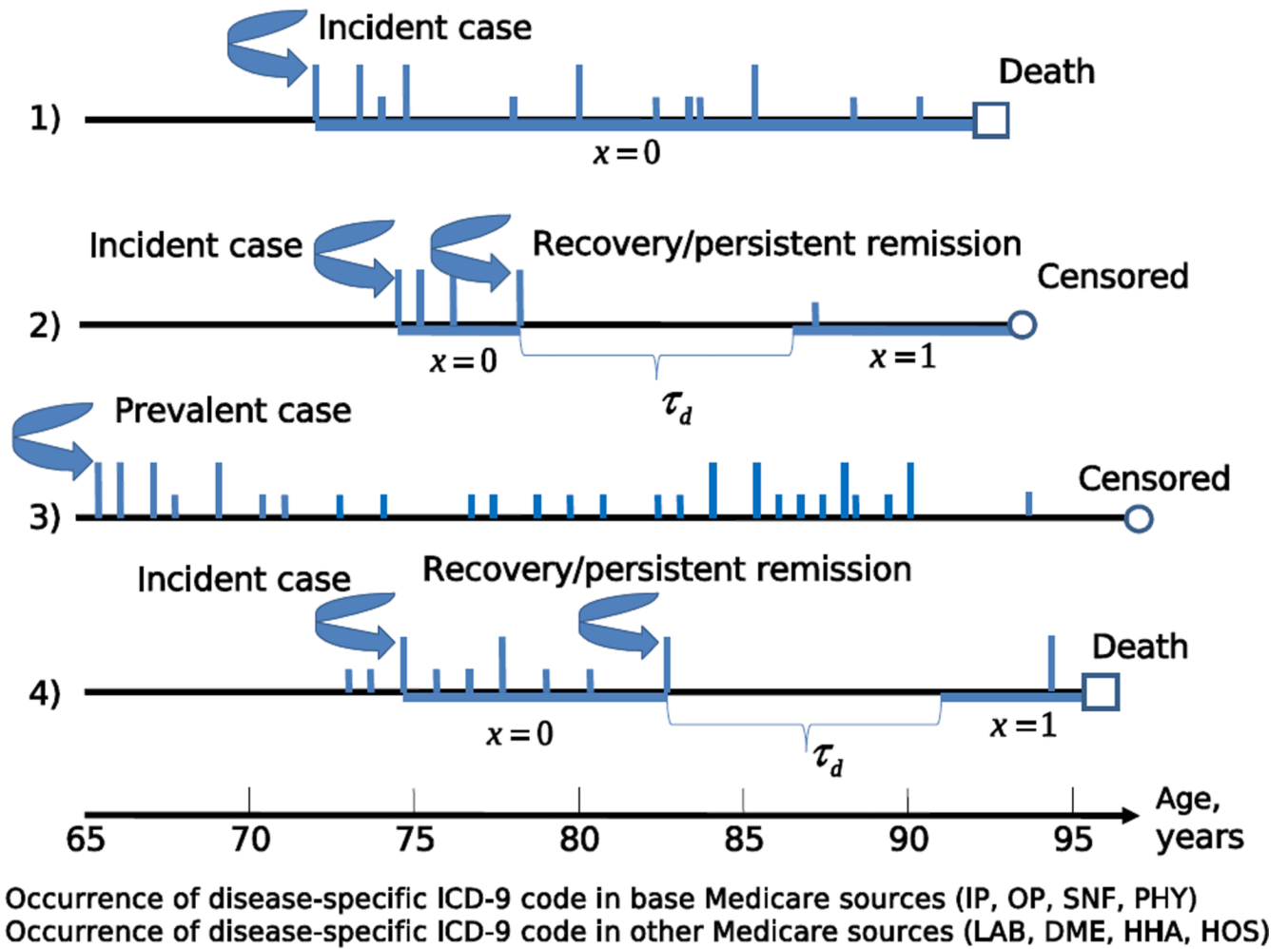


Figure 1.

Scheme with four hypothetical time lines (or individual disease-specific trajectories) illustrating our definitions. The date at onset is detected for the first individual at age 72 years. There is no recovery detected for this individual because there is no time period, τ_d without occurrence of ICD-9 codes for the given disease between dates at onset and death. Period of individual follow-up is marked by thick horizontal line. Since recovery was not detected during the follow-up period the covariate x (meaning “binary indicator of recovery after diagnosis”) equals to zero. Both incidence and recovery are detected for the second individual at ages 74 and 77 years. Two time periods (marked by thick lines) contribute to the partial likelihood for the proportional hazard model of recovery effect on survival with the covariate $x = 0$ before recovery and $x = 1$ after recovery. Time period τ_d just after recovery is excluded from the follow-up when estimating the effect of recovery on survival, because this period was used to identify the recovery event, and therefore survival probability for this period equals 1 exactly. Disease-specific records for the third individuals appeared just after the beginning of his/her follow-up, therefore, this case is considered as a prevalent case and excluded from all analyses. The fourth individual represents another example when both incidence and recovery are detected.

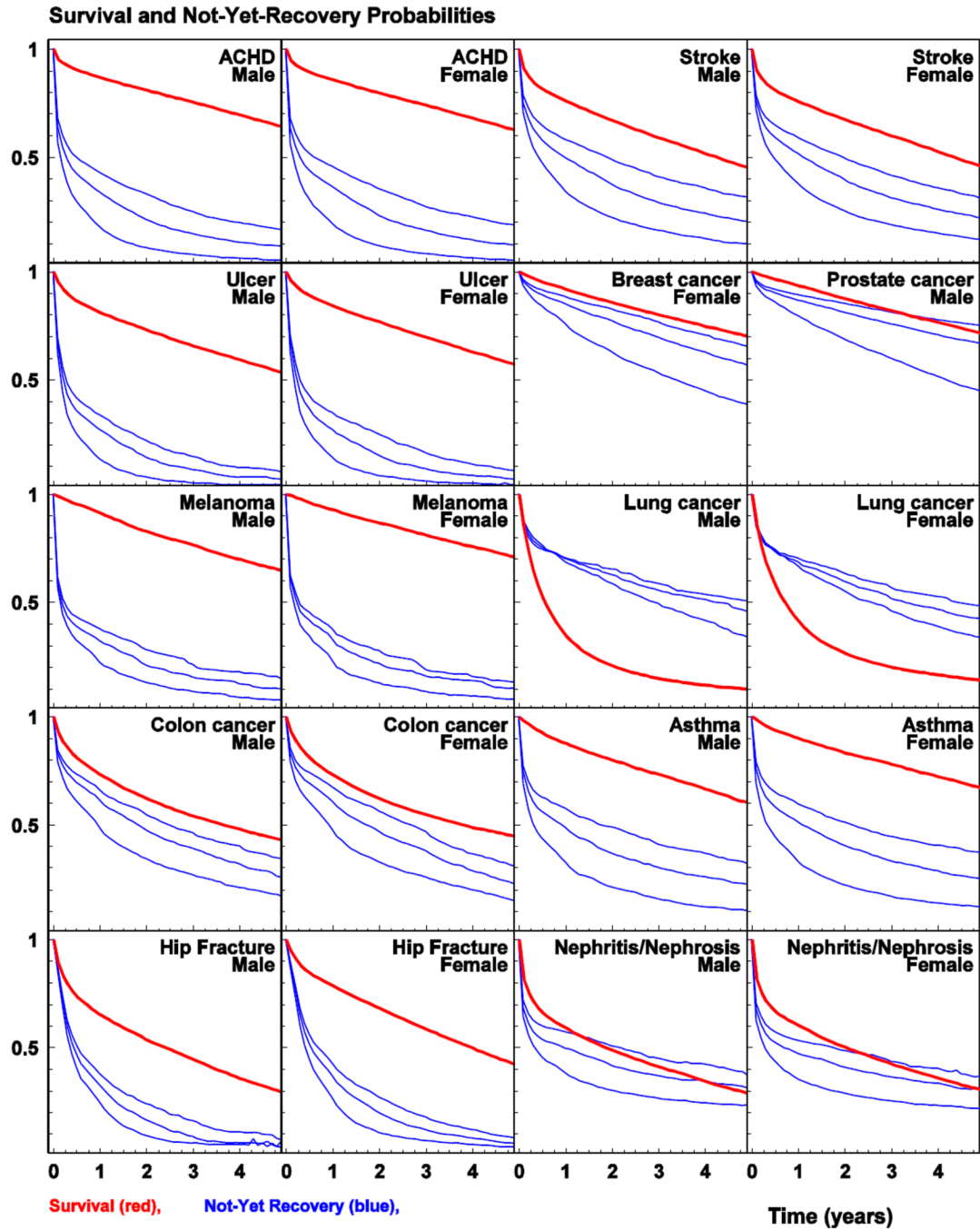


Figure 2.

The survival (thick curve) and not-yet-recovery probabilities (three thin curves corresponding to recovery times of one, two, and three years from the bottom to top respectively) calculated using SEER-M.

Table 1

The characteristics of individuals in SEER-M and NLTCS-M with disease of onset recorded in April 1, 1994 - April 1, 1997 and April 1, 1999 - April 1, 2001: the number of patients, fraction of died during follow-up (f_d), mean follow-up times for died (t_d) and censored (t_c) individuals, and the fractions of individuals removed because their follow-up time was lower than τ_d (i.e., recovery is not possible) for $\tau_d=1, 2$, or 3 years (f_1, f_2 or f_3 , respectively).

Disease	Sex	SEER-M						NLTCS-M							
		Total	f_d %	t_d year	t_c year	f_1 %	f_2 %	f_3 %	Total	f_d %	t_d year	t_c year	f_1 %	f_2 %	f_3 %
ACHD	M	33621	45	3.4	7.2	14	19	23	509	54	3.1	7.1	17	23	31
	F	25038	47	3.5	7.3	14	19	23	521	49	3.4	7.3	14	21	28
Stroke	M	23148	67	2.9	6.6	24	33	40	379	70	2.6	6.3	27	34	45
	F	23567	65	2.9	6.6	23	31	38	670	68	2.5	6.7	26	36	45
Ulcer	M	9191	58	3.2	7.3	19	26	33	114	56	3.2	7.4	20	23	30
	F	8690	53	3.5	7.5	14	20	26	174	55	3.7	7.2	12	17	26
Breast cancer	F	56851	40	3.6	6.7	9	15	20	247	44	3.6	7.2	8	13	21
Prostate cancer	M	68931	41	3.9	6.6	7	13	18	296	46	3.9	7.2	9	17	22
Melanoma	M	8227	42	3.4	6.3	8	16	21	41	41	2.9	6.6	12	15	22
	F	5495	37	3.5	6.7	8	13	18	30	53	3.4	6	6	15	32
Lung cancer	M	38026	91	1.1	6.2	63	77	83	143	95	1.2	7.3	63	82	87
	F	31652	87	1.2	6.3	56	70	76	134	89	1.3	6.6	54	72	78
Colon cancer	M	20629	65	2.4	6.4	26	37	45	93	64	3.1	5.6	18	33	39
	F	25534	60	2.3	6.6	25	36	43	109	61	2.3	7.7	20	39	48
Asthma	M	6361	47	3.7	6.8	10	16	22	61	53	3.6	6.2	7	16	25
	F	7952	37	4.2	7.2	7	11	15	118	44	3.8	7.6	6	14	21
Hip fracture	M	9245	79	2.3	6.3	34	46	54	187	77	2.3	6.6	29	39	53
	F	20632	72	3.2	6.6	19	29	38	601	70	2.7	6.8	25	36	42
Nephritis/ Nephrosis	M	15404	77	1.8	5.9	42	51	58	215	84	1.7	5.4	43	53	61
	F	10537	76	1.9	5.9	40	49	56	225	81	1.5	5.6	43	53	67

Table 2

Death hazard ratio of occurrence (versus non-occurrence) of recovery/long-term remission with recovery times 1, 2, or 3 years (i.e., time period since the last disease-specific record in Medicare file) using the predictor variable “time after recovery” (per one year). Disease onsets were recorded in April 1, 1994 – April 1, 1997 and April 1, 1999 – April 1, 2001.

Disease	Sex	SEER-M			NLCS-M		
		$\tau_{t=1}$	$\tau_{t=2}$	$\tau_{t=3}$	$\tau_{t=1}$	$\tau_{t=2}$	$\tau_{t=3}$
ACHD	M	0.39	0.45	0.46	0.29	0.53	0.38
	F	0.42	0.43	0.46	0.33	0.45	0.47
Stroke	M	0.50	0.53	0.57	0.53	0.38	0.53
	F	0.47	0.48	0.53	0.60	0.52	0.48
Ulcer	M	0.37	0.35	0.35	0.29*	0.11	0.19
	F	0.34	0.30	0.37	0.19	0.23	0.54*
Breast cancer	F	0.87	0.97*	0.96**	1.16**	0.75**	0.56**
Prostate cancer	M	0.93	0.88	0.89	0.71*	0.70**	1.09**
Melanoma	M	0.28	0.39	0.52	0.70**	0.62**	3.25**
	F	0.29	0.45	0.65	0.38*	0.31*	0.11*
Lung cancer	M	0.39	0.43	0.45	0.20	0.68**	1.18**
	F	0.33	0.41	0.53	0.31	0.38*	0.26*
Colon cancer	M	0.46	0.52	0.51	0.33	0.38*	0.88**
	F	0.48	0.53	0.58	0.41	0.55**	0.55**
Asthma	M	0.59	0.63	0.70	1.26**	0.88**	1.45**
	F	0.68	0.67	0.62	1.97*	0.69**	0.72**
Hip fracture	M	0.44	0.39	0.49	0.42	0.36	0.24
	F	0.44	0.38	0.41	0.51	0.35	0.42
Nephritis/ Nephrosis	M	0.53	0.52	0.55	0.63*	0.92**	0.89**
	F	0.53	0.49	0.53	0.49	0.46	0.38

Note: 1 - $\tau_{t=1}$ is the recovery time (in years)

* 0.05 $p < 0.3$, 0.3 $p > 0.05$

** 0.3 p , and p 0.05 when there is no asterisk.

Table 3

Death hazard ratios (with 95%CI) of patients with occurrence of recovery/long-term remission with recovery time of 1, 2, or 3 years (i.e., time period since the last disease-specific record in Medicare file) and estimated recovery times for which death hazard ratios of recovered patients and the general population are equal.

Disease	Sex	SEER-M			NLTC-S-M		
		$\tau_{t=1}$	$\tau_{t=2}$	$\tau_{t=3}$	$\tau_{t=1}$	$\tau_{t=2}$	$\tau_{t=3}$
AChD	M	1.15(1.10,1.19)	1.13(1.08,1.18)	1.11(1.06,1.17)	1.24(1.07,1.45)	1.26(1.05,1.50)	1.04(0.83,1.30)
	F	1.44(1.39,1.49)	1.42(1.36,1.48)	1.41(1.34,1.48)	1.26(1.10,1.46)	1.20(1.01,1.41)	1.11(0.90,1.37)
Stroke	M	1.76(1.68,1.83)	1.69(1.61,1.79)	1.57(1.47,1.68)	1.61(1.37,1.91)	1.58(1.29,1.94)	1.75(1.37,2.24)
	F	1.83(1.77,1.90)	1.78(1.70,1.86)	1.76(1.67,1.86)	1.85(1.63,2.10)	1.76(1.50,2.07)	1.65(1.32,2.07)
Ulcer	M	1.43(1.34,1.53)	1.38(1.28,1.49)	1.36(1.25,1.49)	1.03(0.77,1.39)	0.90(0.63,1.29)	0.95(0.65,1.41)
	F	1.51(1.43,1.60)	1.47(1.38,1.57)	1.47(1.37,1.58)	1.32(1.05,1.65)	1.32(1.02,1.72)	1.43(1.07,1.93)
Breast cancer	F	1.42(1.31,1.54)	1.57(1.41,1.75)	1.62(1.41,1.85)	1.53(1.20,1.95)	1.25(0.88,1.77)	1.35(0.83,2.18)
Prostate cancer	M	1.22(1.13,1.32)	1.21(1.09,1.34)	1.31(1.15,1.50)	0.88(0.66,1.16)	0.81(0.53,1.25)	0.85(0.49,1.45)
Melanoma	M	1.09(0.94,1.26)	1.06(0.89,1.26)	1.19(0.98,1.45)	0.89(0.51,1.54)	0.91(0.48,1.71)	1.41(0.70,2.84)
	F	1.29(1.07,1.54)	1.24(1.00,1.54)	1.42(1.13,1.80)	1.81(1.05,3.12)	1.26(0.56,2.83)	0.65(0.17,2.45)
Lung cancer	M	2.44(2.11,2.82)	2.26(1.86,2.73)	2.03(1.58,2.61)	2.37(1.45,3.86)	2.21(1.21,4.01)	1.52(0.57,4.11)
	F	2.69(2.33,3.11)	2.31(1.91,2.79)	2.36(1.86,3.00)	3.02(2.00,4.55)	2.12(1.16,3.90)	1.35(0.48,3.82)
Colon cancer	M	1.59(1.43,1.77)	1.44(1.26,1.65)	1.41(1.19,1.67)	1.47(1.01,2.13)	1.88(1.25,2.82)	1.89(1.16,3.09)
	F	1.54(1.40,1.70)	1.40(1.24,1.58)	1.36(1.17,1.58)	1.14(0.78,1.68)	0.81(0.48,1.38)	0.81(0.42,1.53)
Asthma	M	1.43(1.30,1.56)	1.42(1.28,1.59)	1.43(1.25,1.63)	1.67(1.14,2.46)	1.83(1.16,2.91)	1.83(1.01,3.31)
	F	1.37(1.28,1.47)	1.36(1.25,1.48)	1.36(1.23,1.50)	1.70(1.31,2.21)	1.40(0.99,1.96)	1.11(0.70,1.75)
Hip fracture	M	2.02(1.90,2.16)	1.99(1.85,2.15)	2.03(1.86,2.23)	1.80(1.44,2.24)	1.92(1.47,2.50)	1.19(0.79,1.79)
	F	1.77(1.72,1.83)	1.78(1.71,1.84)	1.83(1.75,1.91)	1.36(1.19,1.54)	1.24(1.06,1.45)	1.23(1.02,1.49)
Nephritis/ Nephrosis	M	2.06(1.93,2.20)	1.89(1.74,2.05)	1.91(1.72,2.11)	3.25(2.56,4.12)	3.83(2.89,5.09)	3.76(2.52,5.61)
	F	2.56(2.41,2.72)	2.24(2.07,2.42)	2.06(1.86,2.28)	2.81(2.25,3.52)	2.49(1.79,3.47)	1.67(0.95,2.93)

Note: $\bar{1}$ - τ_{tj} is the recovery time (in years), non-significant results for estimated τ_{tj} are not shown (marked by n/a)