Journal of Clinical Epidemiology 105 (2019) 68−79

Journal of Clinical Epidemiology

ORIGINAL ARTICLE

Cohort studies were found to be frequently biased by missing disease information due to death

Nadine Binder[a,*], Anette Blümle[b], James Balmford[a], Edith Motschall[c], Patrick Oeller[b], Martin Schumacher[c]

[a]*Institute for Prevention and Cancer Epidemiology, Faculty of Medicine and Medical Center, University of Freiburg, Elsässerstr. 2, 79110 Freiburg, Germany*
[b]*Institute for Evidence in Medicine, Faculty of Medicine and Medical Center, University of Freiburg, Breisacher Str. 153, 79110 Freiburg, Germany*
[c]*Institute of Medical Biometry and Statistics, Faculty of Medicine and Medical Center, University of Freiburg, Stefan-Meier-Str. 26, 79104 Freiburg, Germany*

Accepted 7 September 2018; Published online 22 September 2018

Abstract

Objectives: In epidemiologic cohort studies with missing disease information due to death (MDID), conventional analyses right-censoring death cases at the last observation or at death may yield significant bias in relative risk and hazard ratio estimates. The aim of this study was to investigate susceptibility to this bias and assess its potential direction and magnitude.

Study Design and Setting: Literature review of selected epidemiologic, geriatric, and environmental journals in 2011−2012 and simulation study of various conventional approaches to handling missing disease data. A study was considered susceptible to MDID bias if disease information was collected at follow-up visits only, and a conventional analysis was performed on the data.

Results: Of 125 identified studies, 58 (46.4%, 95% confidence interval [CI]: 37.7−55.1%) were classified as susceptible to MDID bias, of which six (10.3%, 95% CI: 2.5−18.2%) attempted to address this in sensitivity analyses. The simulation revealed that depending on the analytic strategy for handling missing disease data, the potential exists for significant under- or over-estimation of risk factor effect estimates.

Conclusion: Awareness of MDID bias is important as more adequate analysis methods exist permitting an unbiased analysis. Recommendations for better reporting and analysis of MDID are provided.

Keywords: Cohort studies; Epidemiological biases; Illness-death model; Missing disease information due to death; Regression models; Time-to-event

1. Introduction

In epidemiologic or clinical cohort studies investigating disease incidence or effects of potential factors on the risk of developing a certain (chronic) disease of interest, disease information, that is, the disease status and onset time, are often collected at discrete follow-up visits. This may be assessed by questionnaire in which individuals are asked about the onset time (retrospectively exact), or it may be assessed via an examination, in which case the disease onset time is typically interval-censored.

The problem of missing disease information due to death (MDID), in which disease information is missing for individuals who die between visits, can affect any long-term or aging follow-up cohort study, where disease information is only collected at discrete time points, and death is likely to occur. The underlying model for data with MDID is the illness-death multistate model (Fig. 1). Each cohort subject starts in the initial state and may move into the other states as indicated by the arrows, that is, may become diseased (outcome of interest), may die after contracting the disease, may die without having been diseased, or may stay in the initial state until the end of the study. If being diseased increases the risk of dying (compared with not being diseased) and the disease event is not captured before death, a conventional approach ‘‘censoring’’ the death cases at the last visit observed disease-free (CensVisit) underestimates the disease incidence [1], as does ‘‘censoring’’ at baseline or at the time of death. Not only will some disease cases be missing (numerator of incidence estimate) but also the inherent problem with the three ‘‘censoring’’ schemes is that the resulting risk sets (denominator of incidence estimate) are no longer representative of the sample. ‘‘Censoring’’ in these approaches is not predictable from the past but conditional on the future, that is, on

Conflicts of interest: none.

* Corresponding author. Tel.: +49-761-27077330; fax: +49-761-27077340.
E-mail address: nadine.binder@uniklinik-freiburg.de (N. Binder).

https://doi.org/10.1016/j.jclinepi.2018.09.010
0895-4356/

What is new?

Key findings

- Many epidemiologic cohort studies in which participants can die with or without the disease of interest are susceptible or potentially susceptible to missing disease information due to death (MDID) bias. The most commonly used methods for handling MDID are to exclude death cases from the analysis or to "censor" death cases at the time of death.
- Simulated data revealed that both these conventional approaches are likely to result in significant bias with regard to risk factor effect estimates on disease risk. Where the risk factor affects the risk of dying after being diseased, risk factor estimates on disease risk are generally biased downward (i.e., underestimating the risk factor effect).

What this adds to what was known?

- This article draws attention to the existence of MDID bias and shows that it is likely to be prevalent in studies of aging cohorts. It demonstrates that conventional analysis approaches to cohort data are inadequate in datasets susceptible to this bias.

What is the implication and what should change now?

- Authors should clearly define how the primary outcome of interest was ascertained, specify what procedures were in place to establish the primary outcome in death cases, and summarize characteristics of observed death cases.
- Authors should use analysis approaches based on the illness-death multistate model, which permit the calculation of unbiased risk factor estimates and are readily available. The approaches should also be referenced in reporting guidelines for cohort studies.

being dead and hence does not fulfill the condition of independent censoring [2,3], hence we put this term in quotation marks. This bias also affects the corresponding hazard ratio estimates, for example, from a Cox regression [4,5], and may then directly affect meta-analyses of such studies [6]. The extent and direction of the bias have been found to depend on risk factor–related differential mortality, that is, differing risks for dying between those diseased and those disease-free, because of different baseline death hazards or a nonzero risk factor effect on time to death. We note that this problem is as relevant to estimates of associations as it is to causal effect estimates.

To more adequately estimate disease incidence or effects of potential risk factors on disease for data with MDID,

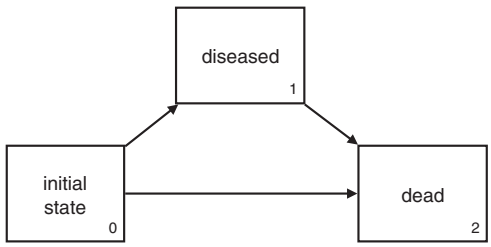

Fig. 1. Three-state illness-death model with an initial state 0, a diseased state of interest 1, and the death state 2. The arrows illustrate the potential transitions between the states.

statistical approaches based on the illness-death model have been proposed: a parametric approach [7], a semiparametric penalized likelihood approach [1], and an imputation-based approach [8]. These analytic approaches, investigated in detail in Binder et al. [9], require time-to-death information for all study participants and event history information from at least two follow-up visits so that disease during follow-up can be observed for some participants who are observed to die afterward. There is evidence that these approaches reduce the extent of MDID bias in comparison to the "CensVisit" approach [5,9].

The present article is divided into two substudies. First, we conducted a systematic literature review to determine the proportion of published cohort studies potentially susceptible to bias arising from the conventional analysis of data with MDID. The literature review also examined the extent to which investigators take steps to quantify or overcome the problem, for example, by reporting the number of death cases or using medical records to retrieve disease cases in deceased subjects. Second, based on the findings of the literature review, we performed a small simulation study to characterize the direction and magnitude of the potential bias in relative risk estimates. We note that comparable investigations have been performed for other time-related biases: time-dependent bias, introduced in estimates if time-dependent variables that change values after study initiation but were not measured at baseline are analyzed as fixed variables [10], and competing risks bias, introduced in Kaplan–Meier estimates if patients experiencing a competing risk are censored at the time of death in a time-to-disease incidence analysis [11,12].

2. Literature review

2.1. Methods

2.1.1. Study selection and inclusion criteria

We implemented a search strategy (see Supplementary Material Section 1) to identify (prospective) cohort studies with follow-up and time-to-event or dichotomous outcome (i.e., collecting events over a period of time) published in 2011 and 2012 in six highly cited journals within three categories: (1) general epidemiologic; (2) geriatric, focusing

on aging populations where death is a predominant natural competitor; and (3) environmental, where we first noticed the problem of MDID [13]. The journals, displayed in Table 1, are indexed in the electronic database MEDLINE (Ovid) and have a high impact factor in their category (median IF in 2011: 4.9).

Studies solely based on register data, (nested) case–control, cross-sectional, case-crossover studies, and studies not examining the association of one or more factors with a disease-type outcome were excluded. We further excluded methods articles not presenting any study results, systematic reviews, and meta-analyses. We allowed, however, for retrospective or prevalent disease event collection based on a prospectively planned cohort study. We allowed any factors to be primary risk factor variables or covariates. We did not restrict the total study period, length of time between follow-ups, number of follow-up visits, or age at study commencement and also allowed for retrospective or prevalent disease event collection (e.g., where a new research question is proposed after initiation of the cohort).

The title and abstract of all retrieved studies were screened (by A.B. or P.O.) to identify obvious exclusions, according to the predefined criteria stated previously. If we were unable to determine eligibility based on title and abstract, the full text of the article was examined.

Table 1. Description of 125 identified illness-death-type cohort studies

Characteristic	Illness-death-type (n = 125)	Susceptible to missing disease information due to death bias (n = 58)
Journal		
American Journal of Epidemiology	45	19 (42.2%)
International Journal of Epidemiology	10	3 (30.0%)
Journal of the American Geriatrics Society	35	19 (54.3%)
Journal of Gerontology, Series A	20	14 (70.0%)
Environmental Health Perspectives	8	2 (25.0%)
Journal of Occupational and Environmental Medicine	7	1 (14.3%)
Year		
2011	71	33 (46.5%)
2012	54	25 (46.3%)
Investigated outcome type		
Time-to-event	94	34 (36.2%)
Dichotomous	31	24 (77.4%)

2.1.2. Data extraction and assessment of criteria

Predefined study characteristics of eligible studies as reported in the full text were recorded (A.B. or P.O.; cross-check N.B.). For this, we modified the item bank for assessment of risk of bias and precision for observational studies of interventions or exposures [14]. We collected study data relevant to estimating the prevalence of bias in risk factor estimates (see Supplementary Material, Section 2). We classified a study of illness-death type (i.e., able to be modeled in terms of the illness-death multistate model) if the primary outcome was a nonterminal event, and death was prespecified as an outcome in the methods section or was otherwise mentioned somewhere in the study article, irrespective of whether the actual number of observed death cases was reported.

One author (N.B.) evaluated eligible illness-death type studies for susceptibility to MDID bias. A study was classified as susceptible if information on the primary disease outcome was collected only at follow-up visits, and a conventional regression analysis was performed. We classified studies as not susceptible to MDID bias if investigators reported having performed disease retrieval to reduce the rate of MDID, for example, by consulting death certificates or contacting hospitals if possible or if a combined endpoint was investigated. We classified susceptibility to MDID bias as unclear if the study accessed an established registry (independently) collecting the events of interest, but if no further information on the registry follow-up scheme was reported.

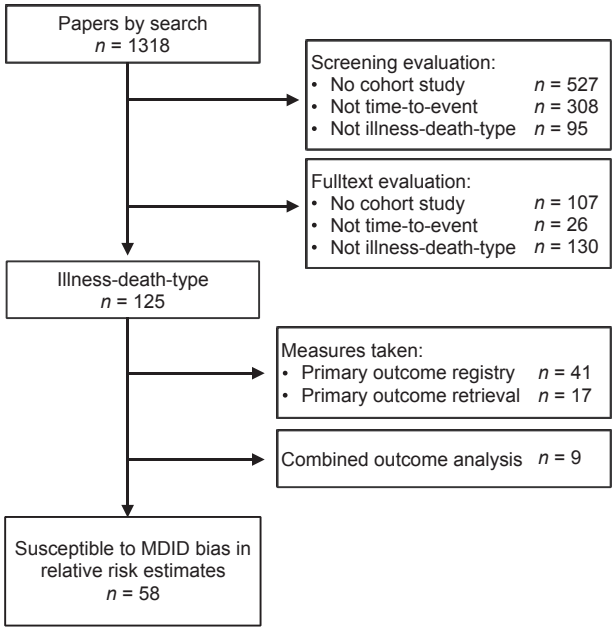

Fig. 2. Flow chart of study classification.

Table 2. Identified illness-death-type cohort studies susceptible to MDID bias in relative risk and/or hazard ratio estimates

Reference	Primary outcome		Risk factor(s)	Age	Follow-up interval	Death handling (sensitivity)	Missing data report
	Type	Ascertainment					
Makris et al. [15]	Restricting back pain	Questionnaire		≥70	1 mo	CensDeath	No
Oh-Park et al. [16]	Fear of falling	Questionnaire	Several risk factors	≥70	3 mo	CensVisit	No
Ottenbacher et al. [17]	Hospital readmission	Questionnaire	Stroke after postacute inpatient rehabilitation	>50	3 mo	Excluded	No
Quinlan et al. [18]	Postoperative functional decline	Questionnaire	Delirium after noncardiac surgery	≥60	3 mo (once)	Excluded (description)	Yes
Viccaro et al. [19]	Multiple geriatric outcomes (ADL, hospitalization)	Assessments	Timed Up and Go, gait speed	≥65	3/6 mo	Excluded	Yes
Bowling et al. [20]	ADL decline	Questionnaire	Chronic kidney disease	77.4 (mean)	6 mo	Excluded	No
Margolis et al. [21]	Hypertension in postmenopausal women	Questionnaire	Levels of 25(OH)D and changes in blood pressure	50–79	6 mo	CensDeath	No
Salanitro et al. [22]	Hospitalizations and emergency department visits	Questionnaire (phone)	Cumulative symptom burden	>65	6 mo	CensDeath (Excluded)	No
Clay et al. [23]	Surgical and nonsurgical overnight hospital admissions	Questionnaire (phone)	Race	65–106	6 mo	CensDeath	No
Houston et al. [24]	Mobility disability and activities of daily living disability	Questionnaire (phone)	25-hydroxyvitamin D	77–100	6 mo	CensDeath	Yes
Verghese et al. [25]	Frailty, disability	Examinations and by phone interview	Mobility stress test	≥70	1 yr/3 mo	CensDeath	No
Cesari et al. [26]	Mobility disability, severe mobility disability, mortality	Examinations and by phone interview	Inflammation, oxidative damage, and platelet activation	70–79	1 yr/6 mo	CensDeath	No
Thorpe et al. [27]	Mobility limitation	Assessments	Race, socioeconomic status	70–79	1 yr/6 mo	Excluded (description)	Yes
Sanders et al. [28]	Incident disability	Assessments	Modified Physiologic Index	70–79	1 yr/6 mo	CensDeath	No
Sikkes et al. [29]	Dementia	Assessments	Instrumental ADL	≥55	1 yr	Excluded	No
James et al. [30]	Disability	Assessments	Level of social activity	55–101	1 yr	Excluded	No
Lee et al. [31]	Connective-tissue disease	Questionnaire	Breast implants	57.0 (mean)	1 yr	CensVisit	No
Shah et al. [32]	Severe performance-based mobility disability	Examinations	Musculoskeletal pain	73.9 (mean)	1 yr	Excluded	No
Shah et al. [33]	Constriction of life space	Questionnaire	Driving status	76.1 (mean)	1 yr	Excluded	No
Potvin et al. [34]	Cognitive impairment	Assessments	Anxiety and depression	≥65	1 yr (once)	Excluded	No
Valenzuela et al. [35]	Dementia, mortality	Examinations	Cognitive lifestyle factors	≥65	1–2 yr	CensDeath	No
Gregory et al. [36]	Preclinical mobility disability	Questionnaire	Educational attainment	70–79	1.5 yr	CensDeath	No

(*Continued*)

Table 2. Continued

Reference	Primary outcome		Risk factor(s)	Age	Follow-up interval	Death handling (sensitivity)	Missing data report
	Type	Ascertainment					
Cozier et al. [37]	Sarcoidosis	Questionnaire; subsequent physician	Exogenous and endogenous estrogen	21–69	2 yr	CensDeath	No
Wise et al. [38]	Leiomyoma	Questionnaire	Hair relaxer	21–69	2 yr	CensDeath	No
Palmer et al. [39]	Breast cancer	Questionnaire; subsequent validation	Individual and neighborhood socioeconomic status	21–69	2 yr	CensDeath	No
Stuebe et al. [40]	Maternal hypertension	Questionnaire	Duration of lactation	25–42	2 yr	CensDeath	No
Abdullah et al. [41]	Type 2 diabetes	Examinations	Obese-years	28–62	2 yr	CensDeath	No
Fung et al. [42]	Postmenopausal breast cancer	Questionnaire; subsequent validation	Dietary approaches to stop hypertension	30–55	2 yr	CensDeath	No
Lucas et al. [43]	Clinical depression	Questionnaire	Physical activity and time spent watching television	30–55	2 yr	CensDeath	No
Song et al. [44]	Skin cancer	Questionnaire; subsequent physician	Smoking	30–75	2 yr	CensDeath	No
Puett et al. [45]	Type 2 diabetes	Questionnaire	Particulate matter and distance to roadways	30–75	2 yr	CensDeath	No
Chen et al. [46]	Proteinuria	Examinations	Arsenic exposure from drinking water	37 (mean)	2 yr	CensDeath	No
Dublin et al. [47]	Dementia, Alzheimer's disease	Examinations	Atrial fibrillation	≥65	2 yr	CensDeath	Yes
Li et al. [48]	Psoriasis	Questionnaire	Smoking	No restriction	2 yr	CensDeath	Yes
Engel et al. [49]	Osteoporotic fractures	Questionnaire	Menopausal hormone therapy	53.8 (mean)	2 yr	Excluded	No
Davydow et al. [50]	Cognitive impairment in survivors of severe sepsis	Phone interview	Presepsis depressive symptoms	≥50	2 yr	Excluded (AllDiseased)	Yes
Himes et al. [51]	Risk of falling; injury or greater ADL disability due to falling	Questionnaire	Obesity	≥65	2 yr	Excluded (description)	Yes
Dong et al. [52]	Chronic diseases	Questionnaire	Occupation and the aging process	>50	2 yr	Excluded	Yes
Spira et al. [53]	Incident instrumental ADL impairment	Questionnaire	Sleep quality	70–94	2.5 yr	Excluded	No
Krishnan et al. [54]	Type 2 diabetes	Examinations	Hyperuricemia as a marker	18–30	2–5 yr	Not reported	Yes
Kvaskoff et al. [55]	Melanoma	Questionnaire	Endogenous hormonal factors	40–65	2–3 yr	CensDeath	No
Di Nisio et al. [56]	Venous thromboembolism	Assessments	Obesity and poor muscle strength	≥60	3 yr	Excluded	No
Viljanen et al. [57]	Incident walking difficulty	Questionnaire	Fear of falling and coexisting sensory difficulties	63–76	3 yr (once)	Excluded	No
Gureje et al. [58]	Dementia	Assessments	Social, health, lifestyle	≥65	3 yr	Excluded	Yes

(*Continued*)

Table 2. Continued

Reference	Primary outcome		Risk factor(s)	Age	Follow-up interval	Death handling (sensitivity)	Missing data report
	Type	Ascertainment			(once)	(description)	
Liao et al. [59]	Functional disability	Examinations	Healthy behaviors	≥60	3–4 yr	AllDiseased	Yes
Smith et al. [60]	Type 2 diabetes	Assessments	Socioeconomic position	34.0 (mean)	4 yr	Excluded	No
Peron et al. [61]	Urinary incontinence	Questionnaire	Antihypertensive agents	70–79	4 yr (once)	Excluded	No
Feng et al. [62]	Cognitive and instrumental ADL decline	Assessments	Glomerular filtration rate or chronic kidney disease	≥55	4 yr (once)	Excluded	Yes
Gast et al. [63]	Risk of diabetes	Self-report or examinations	Changes in biologic risk factors	20–59	5 yr	CensDeath	No
Spira et al. [64]	Nursing home/ personal care home	Questionnaire	Sleep	83.0 (mean)	5 yr (once)	Excluded	No
Sasaki et al. [65]	Dementia	Assessments	Chronic kidney disease	≥65	5 yr (once)	Excluded	Yes
Annweiler et al. [66]	Dementia	Examination	Vitamin D	≥75	7 yr (once)	Excluded	No
Tsubota-Utsugi et al. [67]	Higher-level functional capacity	Assessments		≥60	7 yr (once)	Excluded (as primary outcome)	Yes
Seow et al. [68]	Skin lesion recovery	Assessments	Reduced arsenic exposures	34.2 (mean)	8 yr (once)	Excluded (description)	No
Gao et al. [69]	Parkinson disease	Questionnaire	Durations of daytime napping or nighttime sleeping	50–71	10 yr (once)	Excluded	Yes
Curhan et al. [70]	Hearing loss	Questionnaire	Analgesic use	31–48	14 yr (once)	CensDeath	No
Adams et al. [71]	Age-related macular degeneration	Assessments	Body composition	40–69	14 yr (once)	Excluded	No
Adams et al. [72]	Age-related macular degeneration	Examination	Alcohol use	40–69	15 yr (once)	Excluded	Yes

Abbreviations: ADL, Activities of daily living; MDID, missing disease information due to death.

Note: The terms in the ''Death handling'' column are described in the Introduction; briefly, ''AllDiseased'' refers to analyses in which all death cases were assumed to have been diseased before death, with disease onset time at the midpoint between last visit observed disease-free and time of death, ''CensDeath'' refers to ''censoring'' at the time of death, and ''CensVisit'' refers to ''censoring'' at the last follow-up visit observed disease-free. In this column, any sensitivity analyses undertaken are noted in brackets; these refer to the analysis of death cases only. ''Follow-up interval'' indicates the interval between visits. ''Missing data report'' refers to acknowledgment in the article that missing disease information due to death was a potential limitation of the analysis, irrespective of the analysis conducted.

We estimate the prevalence of MDID bias as the number of studies identified as susceptible to this bias as a proportion of all illness-death type studies. Confidence intervals (CIs) for proportions were calculated using the binomial distribution.

2.2. Results

Our search strategy resulted in 1,318 articles (Supplementary Table 1), of which 930 were excluded during title and abstract screening. A further 263 articles were excluded after full-text evaluation (Fig. 2). This left 125 illness-death-type cohort studies with follow-up determining the association of one or more risk factor variables with a primary disease-type outcome. Most measured a time-to-event outcome (n = 94; 75.2%; 95% CI: 67.8–82.7%); the others measured a dichotomous outcome (i.e., did not account for any temporal information). None of the identified studies used an analysis approach recommended by Binder et al. [9]. Rather, all used a conventional approach (i.e., primarily Cox regression for time-to-event outcomes and logistic regression for dichotomous outcomes). Nine studies were classified as not susceptible to MDID bias, as they analyzed a combined endpoint (Fig. 2). Of the remaining 116 studies, 17 were classified as not susceptible to MDID bias as a result of having performed disease retrieval for death cases. In 41 studies, susceptibility to MDID bias was unclear, as the authors reported access to a disease registry but provided no

Table 3. Bias resulting from conventional Cox analyses of data with MDID in simulated data

	Assumptions						Ref. full cohort		CensVisit
Scenarios	h_{01}	h_{02}	h_{12}	β_{01}	β_{02}	β_{12}	ME	95% CI	ME
1	0.1	0.1	0.1	0	−0.5	−0.5	0	−0.09, 0.1	0
2	0.1	0.1	0.1	0.5	0.5	0	0.5	0.40, 0.60	0.59
3	0.1	0.1	0.1	0.5	0	0.5	0.5	0.41, 0.60	0.41
4	0.1	0.1	0.1	0.5	0.5	−0.5	0.5	0.40, 0.60	0.64
5	0.1	0.1	0.1	−0.5	0	−0.5	−0.5	−0.61, −0.39	−0.44
6	0.3	0.1	0.1	0	−0.5	−0.5	0	−0.08, 0.08	0
7	0.3	0.1	0.1	0.5	0.5	0	0.5	0.42, 0.59	0.57
8	0.3	0.1	0.1	0.5	0	0.5	0.5	0.42, 0.59	0.43
9	0.3	0.1	0.1	0.5	0.5	−0.5	0.5	0.42, 0.59	0.62
10	0.3	0.1	0.1	−0.5	0	−0.5	−0.5	−0.58, −0.41	−0.45
11	0.1	0.3	0.1	0	−0.5	−0.5	0	−0.13, 0.12	−0.12
12	0.1	0.3	0.1	0.5	0.5	0	0.5	0.35, 0.64	0.75
13	0.1	0.3	0.1	0.5	0	0.5	0.5	0.37, 0.63	0.39
14	0.1	0.3	0.1	0.5	0.5	−0.5	0.5	0.35, 0.64	0.81
15	0.1	0.3	0.1	−0.5	0	−0.5	−0.5	−0.66, −0.35	−0.43
16	0.1	0.1	0.3	0	−0.5	−0.5	0	−0.09, 0.1	0.11
17	0.1	0.1	0.3	0.5	0.5	0	0.5	0.40, 0.60	0.6
18	0.1	0.1	0.3	0.5	0	0.5	0.5	0.41, 0.60	0.27
19	0.1	0.1	0.3	0.5	0.5	−0.5	0.5	0.40, 0.60	0.76
20	0.1	0.1	0.3	−0.5	0	−0.5	−0.5	−0.61, −0.39	−0.35

Abbreviations: AllDiseased, All death cases assumed to have been diseased before death, with disease onset time at the midpoint between last visit observed disease-free and time of death; CensDeath, ''censored'' at the time of death; CensVisit, ''censored'' at last visit observed disease-free; CI, confidence interval; MDID, missing disease information due to death; ME, mean estimated regression coefficient on disease; Ref., Reference cohort.

Mean estimated regression coefficient on disease of 20 scenarios assuming different baseline hazards and effects of a binary risk factor on these hazards (Assumptions) are summarized and compared with estimates derived from fully observed cohort data (reference cohort) and a multistate model analysis assuming Weibull-type hazards.

further information. The remaining 58 studies (46.4%, 95% CI: 37.7–55.1%) were classified as susceptible to MDID bias. The proportion was the same in both years (Table 1). Studies investigating a dichotomous endpoint (n = 24; 77.4%, 95% CI: 62.7–92.1%) were more often identified as susceptible as those with a time-to-event endpoint (n = 34; 36.2%, 95% CI: 26.5–45.9%).

Table 2 describes the characteristics of the 58 susceptible studies identified as to MDID bias, sorted by follow-up interval which ranged from 1 month to 15 years (25.8% [n = 15] had only one visit). Contrary to expectations, the most common means of handling MDID data were to exclude death cases from the analysis (Excluded; 29 studies) and ''censoring'' death cases at the time of death (CensDeath; 25 studies). Only two studies ''censored'' cases at their last disease-free visit (CensVisit). Finally, one study assumed all death cases were diseased before dying (AllDiseased), whereas in one study, the method was not reported. Six studies (10.3%, 95% CI: 2.5–18.2%) attempted to address potential MDID in sensitivity analyses by describing the baseline characteristics of those who died (designated in Table 2 as ''description''), including the number of deaths as a primary outcome (''as primary outcome'') or excluding death cases as defined previously (Excluded).

Overall, only 17 studies acknowledged the fact that MDID was a potential limitation of the analysis. This was higher in studies which excluded death cases (n = 11; 46%, 95% CI: 25.9–65.8%) than otherwise (n = 6; 18%, 95% CI: 4.8–30.4%).

3. Simulation study

3.1. Methods

The literature review identified four conventional approaches to analyzing data with MDID. Event history data for six exemplary individuals with continuous follow-up and two discrete follow-up visits, demonstrating the effect of these four ways of handling MDID in time-to-disease analyses, are provided in Section 3 of the Supplementary Material.

To determine the direction and magnitude of the bias resulting from the use of these approaches under a range of assumptions, we first simulated N = 1,000 complete illness-death reference datasets (without any additional

Table 3 *(continued)*

CensVisit	CensDeath		Exclude		AllDiseased		Multistate model	
95% CI	ME	95% CI	ME	95% CI	ME	95% CI	ME	95% CI
−0.12, 0.12	0.07	−0.05, 0.19	−0.06	−0.19, 0.06	−0.23	−0.32, −0.14	0	−0.14, 0.13
0.47, 0.72	0.49	0.37, 0.62	0.64	0.52, 0.77	0.5	0.41, 0.58	0.5	0.37, 0.64
0.29, 0.53	0.37	0.25, 0.50	0.42	0.29, 0.53	0.26	0.17, 0.34	0.5	0.37, 0.64
0.52, 0.77	0.56	0.44, 0.69	0.68	0.56, 0.81	0.5	0.41, 0.58	0.5	0.36, 0.64
−0.58, −0.3	−0.43	−0.57, −0.29	−0.44	−0.58, −0.30	−0.2	−0.29, −0.11	−0.5	−0.65, −0.34
−0.09, 0.09	0.09	0, 0.17	−0.05	−0.14, 0.04	−0.12	−0.2, −0.05	0	−0.1, 0.1
0.48, 0.67	0.45	0.36, 0.55	0.56	0.47, 0.66	0.49	0.41, 0.56	0.5	0.4, 0.61
0.33, 0.52	0.3	0.21, 0.39	0.42	0.33, 0.52	0.36	0.28, 0.43	0.5	0.39, 0.6
0.52, 0.71	0.56	0.47, 0.65	0.58	0.49, 0.67	0.49	0.41, 0.57	0.5	0.4, 0.6
−0.55, −0.35	−0.4	−0.51, −0.31	−0.43	−0.54, −0.34	−0.32	−0.40, −0.24	−0.5	−0.62, −0.39
−0.27, 0.03	0.07	−0.08, 0.21	−0.28	−0.44, −0.13	−0.37	−0.45, −0.3	0	−0.17, 0.15
0.59, 0.91	0.48	0.32, 0.64	0.78	0.62, 0.95	0.49	0.41, 0.57	0.49	0.29, 0.68
0.24, 0.53	0.36	0.20, 0.50	0.33	0.18, 0.48	0.11	0.04, 0.19	0.5	0.33, 0.67
0.65, 0.96	0.56	0.40, 0.72	0.81	0.65, 0.97	0.49	0.42, 0.57	0.48	0.29, 0.66
−0.61, −0.25	−0.42	−0.61, −0.25	−0.37	−0.55, −0.19	−0.08	−0.16, 0.00	−0.49	−0.7, −0.3
−0.03, 0.25	0.18	0.05, 0.33	0.03	−0.11, 0.17	−0.24	−0.33, −0.15	0	−0.17, 0.17
0.46, 0.75	0.48	0.34, 0.62	0.68	0.53, 0.83	0.5	0.41, 0.58	0.5	0.34, 0.67
0.12, 0.43	0.2	0.05, 0.35	0.33	0.18, 0.47	0.26	0.18, 0.34	0.5	0.32, 0.68
0.62, 0.91	0.67	0.53, 0.81	0.8	0.65, 0.95	0.5	0.41, 0.58	0.5	0.34, 0.67
−0.51, −0.19	−0.31	−0.47, −0.15	−0.37	−0.52, −0.21	−0.2	−0.30, −0.11	−0.5	−0.68, −0.32

right-censoring or truncation), each consisting of $n = 3{,}000$ individuals. We then artificially induced MDID by setting two discrete prespecified visits in follow-up time and changed the complete data as if we had collected the event information only at these visits. For ease of illustration, we assumed the baseline hazards h_{01} (disease), h_{02} (death without prior disease), h_{12} (death after disease) to be constant and specified them to follow a Cox proportional hazards model, that is, $h_{ij}(Z) = h_{ij}\exp(\beta_{ij}Z)$ with $i, j \in \{01, 02, 12\}$. We consider four different assumptions for the underlying baseline hazards: (1) all equal (scenarios 1−5 in Table 3); (2) the underlying (baseline) risk for disease is most pronounced (scenarios 6−10); (4) the underlying risk for death without disease is most pronounced (scenarios 11−15); and (4) the underlying risk for illness-death (death after disease) is most pronounced (scenarios 16−20). A binary risk factor $Z \in \{0, 1\}$ with prevalence set to $P(Z = 1) = 0.5$ is assumed to have differing effects β_{01}, β_{02}, β_{12} on each of the hazards. Differential mortality was introduced by the risk factor only (scenarios 1−10) or both from the risk factor and the baseline hazards (scenarios 11−20).

For each simulated data set, we fit Cox models (using the R function coxph) to the reference data and the corresponding MDID data. The mean estimated regression coefficients on disease for each of the four conventional analysis strategies are summarized and compared with estimates derived from the reference cohort for each scenario. For comparison, we also use a multistate model approach based on the illness-death model to the simulated data with artificially induced MDID—an approach recommended for analyzing MDID data in which it is assumed that baseline hazards are parametrized by a Weibull distribution [9]. This analysis was implemented using the R package SmoothHazard.

Five additional scenarios, in which we assume the death hazards to be increasing over time, are also presented in Section 4 of the Supplementary Material. These scenarios provide an illustration that the findings hold irrespective of the shape of the hazard (see also the study by Binder et al. [9]).

Further information on the simulation procedure is available in the Supplementary Material or in the study by Binder and Schumacher [4].

3.2. Results

Table 3 displays the simulation scenarios and results. For each simulation, the risk factor effect estimate is for disease risk (β_{01}). The findings for "CensVisit" replicate earlier studies [5,9]. A risk factor effect estimate is biased upward if the factor increases the risk of death without prior disease. If the factor affects risk of death with prior disease, irrespective of the direction (i.e., increasing or decreasing) and the magnitude of the baseline hazards, the estimated effect on disease risk is less pronounced. No bias is found whenever there is no differential mortality (scenarios 1 and 6), in contrast to scenarios 11 and 16 in which the baseline death hazards differ. For "CensDeath," there is little evidence of bias where the risk factor increases risk of death without prior disease, with the exception of scenario 19 in which the risk factor effect estimate for disease is significantly biased upward. However, risk factor estimates are biased downward if the factor increases risk of death with prior disease to a greater extent than for "CensVisit." The bias for "Excluded" is in all instances similar to that of "CensVisit." Where the risk of death without disease is greater than the risk of death with disease (scenarios 11−15), the bias is slightly more pronounced. Similar results are expected for a logistic regression analysis, in which death cases are typically excluded. For "AllDiseased," risk factor effect estimates on risk of disease are unbiased if the risk factor effect is the same for disease onset and for death without prior disease, irrespective of the magnitude of baseline hazards. Otherwise, the simulations suggest that the estimates are generally and substantially biased downward if the risk factor has a nonnegative effect on disease risk and biased upward otherwise (scenarios 5, 10, 15, and 20).

As expected, the analytic approach based on the Weibull multistate model [9] resulted in unbiased estimates of the risk factor effect under all simulated scenarios. The CIs were marginally wider than in the complete reference data, but some loss of precision is expected, given the existence of cases with missing disease information.

A brief discussion of two of the articles identified in the literature review [22,59], with reference to the simulated data, is presented in the Supplementary Material Section 4. The intention is not to criticize these studies but to provide the interested reader with real-world examples of how the conventional analysis of MDID data can potentially lead to biased findings.

4. Discussion

Susceptibility to MDID bias resulting from the conventional regression analysis of cohort data appears to be common in the literature. In a systematic investigation of six highly cited journals in epidemiology, geriatrics, and environmental research, we found that typical statistical approaches to handling MDID involved "censoring" of death cases based on known information or the exclusion of death cases altogether. In total, 46% of all interval cohorts of illness-death type were considered susceptible to MDID bias across the 2 years that we examined (2011 and 2012), whereas in a further 33% of cohorts, susceptibility to MDID bias was unclear. Susceptibility to MDID bias was particularly high in the two geriatric journals, reflecting the fact that MDID is more relevant in aging cohorts.

Our simulation study showed that the use of "conventional" approaches to handling MDID can overestimate or underestimate the actual risk factor effect on the disease of interest and may even change its sign. The existence and magnitude of this bias depend on the size of and the difference between the two baseline death hazards (with or without prior disease) and the risk factor effect on these death hazards. It was beyond the scope of this study to determine whether susceptibility to MDID bias in the identified studies translated to actual bias or whether study conclusions were affected although it is clear that the "conventional" statistical approaches used fail to adequately account for any potential MDID bias. This is difficult to ascertain because (1) the bias results from both the study design (discrete follow-up and disease ascertainment) and the improper analysis of incomplete data; (2) it differs in direction and size depending on exposure-related differential mortality, making detailed subject-matter knowledge necessary; and (3) information on the number of death cases was often not reported. In addition, it depends on the frequency of and time between follow-up visits, as infrequent follow-ups increase the likelihood of MDID.

Further limitations should be noted. First, the literature review focused on interval cohorts following individuals at the same time interval for all and thus excluded clinic-based cohorts driven by the individual's health care needs [73], which may similarly be affected by MDID. More generally, our strict inclusion criteria may have missed studies not reporting death as an observed outcome but where death cases were likely present because of the baseline age or follow-up duration. Second, unclear reporting of how disease was ascertained or how death was handled in analyses or our incorrect interpretation of this may have resulted in misclassification of studies as susceptible to MDID bias. Third, we did not consider studies with a continuous (nonbinary) primary outcome (e.g., body mass index, quality of life), which are likely to be affected by so-called "truncation by death" [74−78]. We found many studies of this type in the environmental journals, which may explain the lower prevalence of studies susceptible to MDID bias in this category. "Truncation by death" is conceptually different to MDID, as it refers to the problem that repeated continuous or ordinal outcome values (e.g., quality of life) are "cut off" for those who are dead and therefore rendered as "undefined" and not unobserved as with MDID. Therefore, excluding studies with repeated continuous outcomes should not have affected our findings.

Finally, we note that MDID bias affecting the risk factor of interest is also likely to be influenced by the presence of additional covariates, which we did not consider in the simulation study. The intention was only to investigate whether MDID bias existed for each of the conventional analysis approaches we identified and estimate its direction and magnitude across a limited set of scenarios. Modeling MDID bias in the presence of different covariates and comparing the findings to analyses of crude models would be worth pursuing in future research.

We close by making the following recommendations for improving the reporting and analysis of data with MDID. MDID is not yet explicitly addressed in general epidemiologic reporting guidelines, for example, STROBE [79,80], although it was recently included in a reporting guideline for dementia research [81]. Consequently, we have not noticed a significant increase in awareness of the problem in the time that has elapsed since this analysis was carried out. In the absence of formal guidelines, we recommend that authors clearly define how the primary outcome of interest was ascertained and what procedures were in place to establish the primary outcome in death cases. Where access to a disease registry is reported, authors should indicate whether this routinely resulted in the collection of prior disease status among deceased participants. Second, authors should summarize characteristics of observed death cases (e.g., total number, number of those with observed prior primary outcome, and relative risk estimates of how the risk factor of interest affects the time to overall death). Finally, we recommend that authors discuss whether and how their statistical methods have accounted for observed MDID. Where possible, analytic approaches based on the illness-death multistate model [9] should be used, at least in sensitivity analyses.

Acknowledgments

The authors thank Professor Niels Keiding for stimulating discussions on the ‘‘censoring’’ schemes in conventional analysis approaches not fulfilling the condition of independent censoring.

Funding: This work was funded by the German Research Foundation (grant number DFG SCHU 756/12-1). The funding agreement ensured the authors’ independence in designing the study, collecting, analyzing, and interpreting the data, writing of the article, and in the decision to submit the article for publication.

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jclinepi.2018.09.010.

References

[1] Joly P, Commenges D, Helmer C, Letenneur L. A penalized likelihood approach for an illness–death model with interval-censored data: application to age-specific incidence of dementia. Biostatistics 2002;3(3):433–43.

[2] Andersen PK, Keiding N. Interpretability and importance of functionals in competing risks and multistate models. Stat Med 2012; 31:1074–88.

[3] Lesko CR, Edwards JK, Cole SR, Moore RD, Lau B. When to censor? Am J Epidemiol 2018;187:623–32.

[4] Binder N, Schumacher M. Missing information caused by death leads to bias in relative risk estimates. J Clin Epidemiol 2014;67:1111–20.

[5] Leffondré K, Touraine C, Helmer C, Joly P. Interval-censored time-to-event and competing risk with death: is the illness-death model more accurate than the cox model? Int J Epidemiol 2013;42: 1177–86.

[6] Binder N, Manderscheid L, Schumacher M. The combined association of alcohol consumption with dementia risk is likely biased due to lacking account of death cases. Eur J Epidemiol 2017;32(7):627–9.

[7] Touraine C, Gerds TA, Joly P. SmoothHazard: An R package for fitting regression models to interval-censored observations of illness-death models. J Stat Softw 2017;79(7):1–22.

[8] Yu B, Saczynski JS, Launer L. Multiple imputation for estimating the risk of developing dementia and its impact on survival. Biom J 2010;52(5):616–27.

[9] Binder N, Herrnböck AS, Schumacher M. Estimating hazard ratios in cohort data with missing disease information due to death. Biom J 2017;59(2):251–69.

[10] van Walraven C, Davis D, Forster AJ, Wells GA. Time-dependent bias was common in survival analyses published in leading clinical journals. J Clin Epidemiol 2004;57:672–82.

[11] van Walraven C, McAlister FA. Competing risk bias was common in Kaplan-Meier risk estimates published in prominent medical journals. J Clin Epidemiol 2016;69:170–173.e8.

[12] van Walraven C, Hawken S. Competing risk bias in Kaplan-Meier risk estimates can be corrected. J Clin Epidemiol 2016;70:101–5.

[13] Krämer U, Herder C, Sugiri D, Strassburger K, Schikowski T, Ranft U, et al. Traffic-related air pollution and incident type 2 diabetes: results from the SALIA cohort study. Environ Health Perspect 2010;118(9):1273–9.

[14] Viswanathan M, Berkman ND. Development of the RTI item bank on risk of bias and precision of observational studies. J Clin Epidemiol 2012;65:163–78.

[15] Makris UE, Fraenkel L, Han L, Leo-Summers L, Gill TM. Epidemiology of restricting back pain in community-living older persons. J Am Geriatr Soc 2011;59(4):610–4.

[16] Oh-Park M, Xue X, Holtzer R, Verghese J. Transient versus persistent fear of falling in community-dwelling older adults: incidence and risk factors. J Am Geriatr Soc 2011;59(7):1225–31.

[17] Ottenbacher KJ, Graham JE, Ottenbacher AJ, Lee J, Al Snih S, Karmarkar A, et al. Hospital readmission in persons with stroke following postacute inpatient rehabilitation. J Gerontol A Biol Sci Med Sci 2012;67(8):875–81.

[18] Quinlan N, Rudolph JL. Postoperative delirium and functional decline after noncardiac surgery. J Am Geriatr Soc 2011;59(Suppl 2):S301–4.

[19] Viccaro LJ, Perera S, Studenski SA. Is timed up and go better than gait speed in predicting health, function, and falls in older adults? J Am Geriatr Soc 2011;59(5):887–92.

[20] Bowling CB, Sawyer P, Campbell RC, Ahmed A, Allman RM. Impact of chronic kidney disease on activities of daily living in community-dwelling older adults. J Gerontol A Biol Sci Med Sci 2011;66(6):689–94.

[21] Margolis KL, Martin LW, Ray RM, Kerby TJ, Allison MA, Curb JD, et al. Women’s health initiative investigators. A prospective study of serum 25-hydroxyvitamin D levels, blood pressure, and incident hypertension in postmenopausal women. Am J Epidemiol 2012;175:22–32.

[22] Salanitro AH, Hovater M, Hearld KR, Roth DL, Sawyer P, Locher JL, et al. Symptom burden predicts hospitalization

independent of comorbidity in community-dwelling older adults. J Am Geriatr Soc 2012;60(9):1632−7.
[23] Clay OJ, Roth DL, Safford MM, Sawyer PL, Allman RM. Predictors of overnight hospital admission in older African American and Caucasian Medicare beneficiaries. J Gerontol A Biol Sci Med Sci 2011;66(8):910−6.
[24] Houston DK, Tooze JA, Davis CC, Chaves PHM, Hirsch CH, Robbins JA, et al. Serum 25-hydroxyvitamin D and physical function in older adults: the cardiovascular health study all stars. J Am Geriatr Soc 2011;59(10):1793−801.
[25] Verghese J, Holtzer R, Lipton RB, Wang C. Mobility stress test approach to predicting frailty, disability, and mortality in high-functioning older adults. J Am Geriatr Soc 2012;60(10):1901−5.
[26] Cesari M, Kritchevsky SB, Nicklas B, Kanaya AM, Patrignani P, Tacconelli S, et al. Oxidative damage, platelet activation, and inflammation to predict mobility disability and mortality in older persons: results from the health aging and body composition study. J Gerontol A Biol Sci Med Sci 2012;67(6):671−6.
[27] Thorpe RJ, Koster A, Kritchevsky SB, Newman AB, Harris T, Ayonayon HN, et al. Race, socioeconomic resources, and late-life mobility and decline: findings from the health, aging, and body composition study. J Gerontol A Biol Sci Med Sci 2011;66A(10):1114−23.
[28] Sanders JL, Boudreau RM, Penninx BW, Simonsick EM, Kritchevsky SB, Satterfield S, et al. Association of a modified physiologic index with mortality and incident disability: the health, aging, and body composition study. J Gerontol A Biol Sci Med Sci 2012;67(12):1439−46.
[29] Sikkes SAM, Visser PJ, Knol DL, de Lange-de Klerk ESM, Tsolaki M, Frisoni GB, et al. Do instrumental activities of daily living predict dementia at 1- and 2-year follow-up? Findings from the development of screening guidelines and diagnostic criteria for predementia Alzheimer's disease study. J Am Geriatr Soc 2011;59(12):2273−81.
[30] James BD, Boyle PA, Buchman AS, Bennett DA. Relation of late-life social activity with incident disability among community-dwelling older adults. J Gerontol A Biol Sci Med Sci 2011;66(4):467−73.
[31] Lee I-MM, Cook NR, Shadick NA, Pereira E, Buring JE. Prospective cohort study of breast implants and the risk of connective-tissue diseases. Int J Epidemiol 2011;40:230−8.
[32] Shah RC, Buchman AS, Boyle PA, Leurgans SE, Wilson RS, Andersson GB, et al. Musculoskeletal pain is associated with incident mobility disability in community-dwelling elders. J Gerontol A Biol Sci Med Sci 2011;66(1):82−8.
[33] Shah RC, Maitra K, Barnes LL, James BD, Leurgans S, Bennett DA. Relation of driving status to incident life space constriction in community-dwelling older persons: a prospective cohort study. J Gerontol A Biol Sci Med Sci 2012;67(9):984−9.
[34] Potvin O, Forget H, Grenier S, Prville M, Hudon C. Anxiety, depression, and 1-year incident cognitive impairment in community-dwelling older adults. J Am Geriatr Soc 2011;59(8):1421−8.
[35] Valenzuela M, Brayne C, Sachdev P, Wilcock G, Matthews F. Cognitive lifestyle and long-term risk of dementia and survival after diagnosis in a multicenter population-based cohort. Am J Epidemiol 2011;173:1004−12.
[36] Gregory PC, Szanton SL, Xue Q-LL, Tian J, Thorpe RJ, Fried LP. Education predicts incidence of preclinical mobility disability in initially high-functioning older women. The women's health and aging study II. J Gerontol A Biol Sci Med Sci 2011;66(5):577−81.
[37] Cozier YC, Berman JS, Palmer JR, Boggs DA, Wise LA, Rosenberg L. Reproductive and hormonal factors in relation to incidence of sarcoidosis in US Black women: the black women's health study. Am J Epidemiol 2012;176:635−41.
[38] Wise LA, Palmer JR, Reich D, Cozier YC, Rosenberg L. Hair relaxer use and risk of uterine leiomyomata in African-American women. Am J Epidemiol 2012;175:432−40.
[39] Palmer JR, Boggs DA, Wise LA, Adams-Campbell LL, Rosenberg L. Individual and neighborhood socioeconomic status in relation to breast cancer incidence in African-American women. Am J Epidemiol 2012;176:1141−6.
[40] Stuebe AM, Schwarz EB, Grewen K, Rich-Edwards JW, Michels KB, Foster EM, et al. Duration of lactation and incidence of maternal hypertension: a longitudinal cohort study. Am J Epidemiol 2011;174:1147−58.
[41] Abdullah A, Wolfe R, Mannan H, Stoelwinder JU, Stevenson C, Peeters A. Epidemiologic merit of obese-years, the combination of degree and duration of obesity. Am J Epidemiol 2012;176:99−107.
[42] Fung TT, Hu FB, Hankinson SE, Willett WC, Holmes MD. Low-carbohydrate diets, dietary approaches to stop hypertension-style diets, and the risk of postmenopausal breast cancer. Am J Epidemiol 2011;174:652−60.
[43] Lucas M, Mekary R, Pan A, Mirzaei F, O'Reilly EJ, Willett WC, et al. Relation between clinical depression risk and physical activity and time spent watching television in older women: a 10-year prospective follow-up study. Am J Epidemiol 2011;174:1017−27.
[44] Song F, Qureshi AA, Gao X, Li T, Han J. Smoking and risk of skin cancer: a prospective analysis and a meta-analysis. Int J Epidemiol 2012;41:1694−705.
[45] Puett RC, Hart JE, Schwartz J, Hu FB, Liese AD, Laden F. Are particulate matter exposures associated with risk of type 2 diabetes? Environ Health Perspect 2011;119(3):384−9.
[46] Chen Y, Parvez F, Liu M, Pesola GR, Gamble MV, Slavkovich V, et al. Association between arsenic exposure from drinking water and proteinuria: results from the health effects of arsenic longitudinal study. Int J Epidemiol 2011;40:828−35.
[47] Dublin S, Anderson ML, Haneuse SJ, Heckbert SR, Crane PK, Breitner JCS, et al. Atrial fibrillation and risk of dementia: a prospective cohort study. J Am Geriatr Soc 2011;59(8):1369−75.
[48] Li W, Han J, Choi HK, Qureshi AA. Smoking and risk of incident psoriasis among women and men in the United States: a combined analysis. Am J Epidemiol 2012;175:402−13.
[49] Engel P, Fabre A, Fournier A, Mesrine S, Boutron-Ruault MCC, Clavel-Chapelon F. Risk of osteoporotic fractures after discontinuation of menopausal hormone therapy: results from the E3N cohort. Am J Epidemiol 2011;174:12−21.
[50] Davydow DS, Hough CL, Langa KM, Iwashyna TJ. Presepsis depressive symptoms are associated with incident cognitive impairment in survivors of severe sepsis: a prospective cohort study of older Americans. J Am Geriatr Soc 2012;60(12):2290−6.
[51] Himes CL, Reynolds SL. Effect of obesity on falls, injury, and disability. J Am Geriatr Soc 2012;60(1):124−9.
[52] Dong XS, Wang X, Daw C, Ringen K. Chronic diseases and functional limitations among older construction workers in the United States: a 10-year follow-up study. J Occup Environ Med 2011;53(4):372−80.
[53] Spira AP, Covinsky K, Rebok GW, Punjabi NM, Stone KL, Hillier TA, et al. Poor sleep quality and functional decline in older women. J Am Geriatr Soc 2012;60(6):1092−8.
[54] Krishnan E, Pandya BJ, Chung L, Hariri A, Dabbous O. Hyperuricemia in young adults and risk of insulin resistance, prediabetes, and diabetes: a 15-year follow-up study. Am J Epidemiol 2012;176:108−16.
[55] Kvaskoff M, Bijon A, Mesrine S, Boutron-Ruault MCC, Clavel-Chapelon F. Cutaneous melanoma and endogenous hormonal factors: a large French prospective study. Am J Epidemiol 2011;173:1192−202.
[56] Di Nisio M, Di Iorio A, Porreca E, Abate M, Ferrante N, Bandinelli S, et al. Obesity, poor muscle strength, and venous thromboembolism in older persons: the InCHIANTI study. J Gerontol A Biol Sci Med Sci 2011;66(3):320−5.
[57] Viljanen A, Kulmala J, Rantakokko M, Koskenvuo M, Kaprio J, Rantanen T. Fear of falling and coexisting sensory difficulties as

predictors of mobility decline in older women. J Gerontol A Biol Sci Med Sci 2012;67(11):1230−7.
[58] Gureje O, Ogunniyi A, Kola L, Abiona T. Incidence of and risk factors for dementia in the Ibadan study of aging. J Am Geriatr Soc 2011;59(5):869−74.
[59] Liao W-CC, Li C-RR, Lin Y-CC, Wang CC, Chen YJ, Yen CH, et al. Healthy behaviors and onset of functional disability in older adults: results of a national longitudinal study. J Am Geriatr Soc 2011; 59(2):200−6.
[60] Smith BT, Lynch JW, Fox CS, Harper S, Abrahamowicz M, Almeida ND, et al. Life-course socioeconomic position and type 2 diabetes mellitus: the Framingham Offspring Study. Am J Epidemiol 2011;173:438−47.
[61] Peron EP, Zheng Y, Perera S, Newman AB, Resnick NM, Shorr RI, et al. Antihypertensive drug class use and differential risk of urinary incontinence in community-dwelling older women. J Gerontol A Biol Sci Med Sci 2012;67(12):1373−8.
[62] Feng L, Yap KB, Yeoh LY, Ng TP. Kidney function and cognitive and functional decline in elderly adults: findings from the Singapore longitudinal aging study. J Am Geriatr Soc 2012;60(7):1208−14.
[63] Gast GC, Spijkerman AM, Van der A DL, Jacobs-van der Bruggen MA, Verschuren WM. Five-year changes in biologic risk factors and risk of type 2 diabetes: are attained but not initial risk factor levels of importance? Am J Epidemiol 2012;176:720−5.
[64] Spira AP, Covinsky K, Rebok GW, Stone KL, Redline S, Yaffe K. Objectively measured sleep quality and nursing home placement in older women. J Am Geriatr Soc 2012;60(7):1237−43.
[65] Sasaki Y, Marioni R, Kasai M, Ishii H, Yamaguchi S, Meguro K. Chronic kidney disease: a risk factor for dementia onset: a population-based study. The Osaki-Tajiri project. J Am Geriatr Soc 2011;59(7):1175−81.
[66] Annweiler C, Rolland Y, Schott AM, Blain H, Vellas B, Herrmann FR, et al. Higher vitamin D dietary intake is associated with lower risk of Alzheimer's disease: a 7-year follow-up. J Gerontol A Biol Sci Med Sci 2012;67(11):1205−11.
[67] Tsubota-Utsugi M, Ito-Sato R, Ohkubo T, Kikuya M, Asayama K, Metoki H, et al. Health behaviors as predictors for declines in higher-level functional capacity in older adults: the Ohasama study. J Am Geriatr Soc 2011;59(11):1993−2000.
[68] Seow WJ, Pan WC, Kile ML, Baccarelli AA, Quamruzzaman Q, Rahman M, et al. Arsenic reduction in drinking water and improvement in skin lesions: a follow-up study in Bangladesh. Environ Health Perspect 2012;120(12):1733−8.
[69] Gao J, Huang X, Park Y, Hollenbeck A, Blair A, Schatzkin A, et al. Daytime napping, nighttime sleeping, and Parkinson disease. Am J Epidemiol 2011;173:1032−8.
[70] Curhan SG, Shargorodsky J, Eavey R, Curhan GC. Analgesic use and the risk of hearing loss in women. Am J Epidemiol 2012;176:544−54.
[71] Adams MK, Simpson JA, Aung KZ, Makeyeva GA, Giles GG, English DR, et al. Abdominal obesity and age-related macular degeneration. Am J Epidemiol 2011;173:1246−55.
[72] Adams MKM, Chong EW, Williamson E, Aung KZ, Makeyeva GA, Giles GG, et al. 20/20−Alcohol and age-related macular degeneration: the Melbourne collaborative cohort study. Am J Epidemiol 2012;176:289−98.
[73] Lau B, Gange SJ, Moore RD. Interval and clinical cohort studies: epidemiological issues. AIDS Res Hum Retroviruses 2007;23: 769−76.
[74] Zhang JL, Rubin DB. Estimation of causal effects via principal stratification when some outcomes are truncated by ''death''. J Educ Behav Stat 2003;28(4):353−68.
[75] Kurland BF, Johnson LL, Egleston BL, Diehr PH. Longitudinal data with follow-up truncated by death: match the analysis method to research aims. Stat Sci 2009;24(2):211.
[76] Weuve J, Tchetgen Tchetgen EJ, Glymour MM, Beck TL, Aggarwal NT, Wilson RS, et al. Accounting for bias due to selective attrition: the example of smoking and cognitive decline. Epidemiology 2012;23:119−28.
[77] Tchetgen Tchetgen EJ. Identification and estimation of survivor average causal effects. Stat Med 2014;33:3601−28.
[78] Mayeda ER, Tchetgen Tchetgen EJ, Power MC, Weuve J, Jacqmin-Gadda H, Marden JR, et al. A simulation platform for quantifying survival bias: an application to research on determinants of cognitive decline. Am J Epidemiol 2016;184:378−87.
[79] von Elm E, Altman DG, Egger M, Pocock SJ, Gotzsche PC, Vandenbroucke JP, et al. The strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. Epidemiology 2007;18:800−4.
[80] Vandenbroucke JP, von Elm E, Altman DG, Gotzsche PC, Mulrow CD, Pocock SJ, et al. STROBE initiative. Strengthening the reporting of observational studies in epidemiology (STROBE): explanation and elaboration. Epidemiology 2007;18:805−35.
[81] Weuve J, Proust-Lima C, Power MC, Gross AL, Hofer SM, Thibaut R, et al. Guidelines for reporting methodological challenges and evaluating potential bias in dementia research. Alzheimers Dement 2015;11(9):1098−109.