

Adjusting for time-varying confounding in the subdistribution analysis of a competing risk

Maarten Bekaert · Stijn Vansteelandt ·
Karl Mertens

Received: 18 December 2008 / Accepted: 21 September 2009 / Published online: 10 October 2009
© Springer Science+Business Media, LLC 2009

Abstract Despite decades of research in the medical literature, assessment of the attributable mortality due to nosocomial infections in the intensive care unit (ICU) remains controversial, with different studies describing effect estimates ranging from being neutral to extremely risk increasing. Interpretation of study results is further hindered by inappropriate adjustment (a) for censoring of the survival time by discharge from the ICU, and (b) for time-dependent confounders on the causal path from infection to mortality. In previous work (Vansteelandt et al. *Biostatistics* 10:46–59), we have accommodated this through inverse probability of treatment and censoring weighting. Because censoring due to discharge from the ICU is so intimately connected with a patient's health condition, the ensuing inverse weighting analyses suffer from influential weights and rely heavily on the assumption that one has measured all common risk factors of ICU discharge and mortality. In this paper, we consider ICU discharge as a competing risk in the sense that we aim to infer the risk of 'ICU mortality' over time that would be observed if nosocomial infections could be prevented for the entire study population. For this purpose we develop marginal structural subdistribution hazard models with accompanying estimation methods. In contrast to subdistribution hazard models with time-varying covariates, the proposed approach (a) can accommodate high-dimensional confounders, (b) avoids regression adjustment for post-infection measurements and thereby so-called collider-stratification bias, and (c) results in a well-defined model for the cumulative incidence function. The methods are used to quantify the causal effect of nosocomial pneumonia on ICU mortality

M. Bekaert · S. Vansteelandt (✉)
Department of Applied Mathematics and Computer Science, Ghent University, Ghent, Belgium
e-mail: Stijn.Vansteelandt@ugent.be

K. Mertens
Epidemiology Unit, Scientific Institute of Public Health, Brussels, Belgium

using data from the National Surveillance Study of Nosocomial Infections in ICU's (Belgium).

Keywords Causal inference · Competing risk · ICU · Inverse weighting · Nosocomial infection · Time-dependent confounding

1 Introduction

Nosocomial (i.e., hospital-acquired) infections are highly prevalent in intensive care unit (ICU) patients because of their poor health conditions and because of the prevalent use of invasive treatments (e.g., mechanical ventilation) that make it more difficult to conquer hostile bacteria. They form a major public health problem in the Western world. For instance, in Belgium (which has a population size of about 10.5 million people), it is roughly estimated that nosocomial infections contribute to the mortality of 2600 patients per year, to 700,000 extra hospital days and to a total annual cost of 400 million EUR ([Vrijens et al. 2009](#)). There is a longstanding interest in precise knowledge of the effect of these infections on mortality ([Davis 2006](#); [Heyland et al. 1999](#); [Safdar et al. 2005](#)). This stems from the desire to better understand their severity as well as the importance of a timely treatment. This is additionally motivated by the need to justify the high costs that these infections impose on the society as a result of prolonged hospital stay and extra hospitalisation costs.

The question whether nosocomial infections are severely harmful, remains a subtle one for several reasons. First, infection must be viewed in combination with antibiotic treatment, which has beneficial effects if applied appropriately ([Kollef 2003](#)) but may have harmful effects otherwise (e.g., due to antibiotic resistance). Second, patients who acquire these infections and subsequently die, may have been so severely ill that they would have died even without the infection. In view of this, it will be crucial to have the ability to adjust for the evolution in disease severity over time.

Despite the wide literature on the effect on mortality of acquiring a nosocomial infection in the ICU, assessment of this effect remains controversial, with different studies describing effect estimates ranging from being neutral to extremely risk increasing ([Bercault and Boulain 2001](#); [Fagon et al. 1993](#); [Girou et al. 1998](#); [Heyland et al. 1999](#); [Papazian et al. 1996](#); [Timsit et al. 1996](#)). Interpretation of study results is further hindered by time-dependent bias (i.e., bias due to ignoring the time-dependent nature of the data ([Beyersmann et al. 2008](#); [Van Walraven et al. 2004](#))) and inappropriate adjustment for time-dependent confounding and for informative censoring due to discharge from the ICU ([Vansteelandt et al. 2009](#)).

Informative censoring ([Kalbfleisch and Prentice 2002](#)) results from the decision to discharge a patient from the ICU being closely related to his/her health status. Patients who get discharged from the ICU on a given day since admission are therefore not comparable with those who stay in the ICU on that day. If a rich collection of common risk factors of mortality and ICU discharge were measured, one could in principle accommodate this using inverse probability of censoring weighted estimators ([Keiding et al. 1999](#); [Robins and Rotnitzky 1992](#); [Vansteelandt et al. 2009](#)). However, the resulting analyses rely heavily on the assumption that all such risk factors have been measured

and suffer from influential weights because censoring tends to be strongly associated with mortality in this setting (Vansteelandt et al. 2009). In this paper, we will therefore take a different focus than Vansteelandt et al. (2009) and estimate the impact of acquiring infection in the ICU on the Cumulative Incidence Function (CIF) of ICU death (Fine and Gray 1999; Pepe and Mori 1993; Prentice et al. 1978). This is the probability of dying within the ICU before a given time, as a function of time. Focussing on this estimand not only rules out the censoring problem (since ICU death by a given time is observable for all patients so that censoring becomes purely administrative due to end of follow-up); ICU mortality is also of primary interest to ICU physicians, is frequently used in this domain (see e.g. Resche-Rigon et al. (2006), Vallés et al. (2007), and Wolkewitz et al. (2009)) and would be representative of overall mortality if the hazard of death were relatively small after ICU discharge. A similar focus is taken by e.g. Beyersmann et al. (2006), Wolkewitz et al. (2009), and Wolkewitz et al. (2008) who consider discharge from the ICU and ICU mortality as competing risks to assess the impact of infection on ICU mortality and length of stay in the ICU.

Given our focus on the CIF of ICU death, a natural approach would be to consider the infection status as a time-dependent covariate in a model for the subdistribution hazard (i.e. the hazard function corresponding to the CIF) (Barnett and Graves 2008; Beyersmann and Schumacher 2008; Wolkewitz et al. 2008). This approach acknowledges that patients need to survive and stay within the ICU long enough before acquiring infection (Beyersmann et al. 2008; Van Walraven et al. 2004). However, because patients who do and do not acquire infection at a given time are not comparable in terms of severity of illness, observed mortality differences between these groups cannot be fully attributed to an infection effect. Including measures of disease severity as additional time-varying covariates does not accommodate this. First, it inevitably raises difficulties as to how to define these covariates, which were only measured during the ICU stay, after discharge (i.e., when the competing event has occurred) (Beyersmann and Schumacher 2008; Latouche et al. 2005). Second, because a patient's health status is affected by prior infection, several of these covariates are intermediate on the causal path from early infection to mortality, and therefore should not be adjusted for via standard regression methods (Hernán et al. 2000). This is not only because standard regression would remove indirect infection effects mediated through severity of illness, but also because it would induce a so-called collider-stratification bias (Greenland 2003). Third, subdistribution hazards adjusted for time-dependent covariates do not result in a model for the CIF anymore (Beyersmann and Schumacher 2008), just like standard survival Cox models with time-dependent covariates yield no model for the survival function (Andersen et al. 1993). This complicates the interpretation of the infection effect by the fact that a hazard ratio, unlike a comparison of the CIFs, does not lend itself to a simple causal interpretation. This is so because, even if infected and uninfected patients were exchangeable (Hernán and Robins 2006) upon ICU admission, the risk sets of surviving infected and uninfected patients may not stay exchangeable over time (Hernán et al. 2004).

In view of these limitations, we will develop marginal structural models (Robins et al. 2000) for the counterfactual subdistribution hazard with accompanying estimation methods. Using these models, we will be able to infer the risk of ICU death over time that would be observed if nosocomial infections could be prevented for

the entire study population. The proposed framework for causal effect estimation (a) can accommodate high-dimensional confounders, (b) avoids regression adjustment for post-infection measurements, and (c) results in a model for the CIF. The methodology is developed for complete data, but may be extended to right-censored data using additional inverse probability of censoring weights (Fine and Gray 1999; Robins and Rotnitzky 1992).

The paper is organized as follows. In Sect. 2 we start by introducing the main concepts and assumptions. In Sect. 3, we develop inference under a nonparametric model for the counterfactual cumulative incidence function. This is then extended in Sect. 4 to semiparametric discrete-time marginal structural models for the subdistribution hazard. Both estimation strategies are used in Sects. 3.3 and 4.2, respectively, to quantify the effect of nosocomial pneumonia on ICU mortality using data from the National Surveillance Study of Nosocomial Infections in ICU's (Belgium). We end in Sect. 5 with a discussion on the assumptions and the relative advantages and limitations of the proposed methodology.

2 Notation, definitions and assumptions

2.1 Observed data

For each patient, we observe the time T from ICU admission to death or ICU discharge, whichever comes first and the event type ϵ , which indicates 1 for the event of interest (ICU death) and 2 for the competing event (ICU discharge). In addition, for each patient, we observe a discrete-time counting process $\{A_t, 0 \leq t \leq T\}$ until death or ICU discharge, whichever comes first, where t denotes days since ICU admission. Here, A_t indicates 1 for ICU acquired infection at or prior to time t and 0 otherwise. Finally, for each patient, we observe a collection of baseline confounders L_0 and time-dependent confounders L_t on each day t , which will be listed in Sect. 3.3. Our interest focuses on the Cumulative Incidence Function (CIF) of ICU death: $P(T \leq t, \epsilon = 1)$.

2.2 Counterfactual subdistribution hazard and cumulative incidence function

Unlike a competing risk analysis with time-dependent infection status (Beyersmann and Schumacher 2008; Wolkewitz et al. 2008), we will evaluate the CIF for the entire population under different *hypothetical or potential infection paths* to guarantee that we compare the same patient population each time. This requires that we define for each patient a counterfactual event time $T_{\bar{a}}$ (Robins 1986; Rubin 1978), which represents the time until ICU death or ICU discharge, whichever comes first, which an ICU patient would—possibly contrary to fact—have had under the given hypothetical infection path $\bar{a}$. Such potential infection path $\bar{a} = (a_1, a_2, \dots, a_E)$ equals a vector of elements a_t , for t going from day 1 (i.e. admission) in the ICU to a fixed end-of-follow-up time E , representing a hypothetical regime in which *all* ICU patients would be uninfected in the ICU during all days t with $a_t = 0$ and infected in the ICU on all remaining days t' with $a_{t'} = 1$. Correspondingly, we define $\{\epsilon_{\bar{a}}(t), t \geq 0\}$ to

be a discrete-time stochastic process expressing the counterfactual event status under infection path $\bar{a}$ on each day t . This indicates 1 (or 2) on day t if the patient would have died in (been discharged from) the ICU under that potential infection path, and 0 if no event would have occurred by that day. In addition, we define $\epsilon_{\bar{a}} \equiv \epsilon_{\bar{a}}(T_{\bar{a}})$. For an event of type k ($k = 1, 2$), we thus conceptualize the corresponding counterfactual subdistribution event time $\theta_{\bar{a}k} \equiv \inf\{t \in [0, \infty[: \epsilon_{\bar{a}}(t) = k\}$. For $k = 1$ ($k = 2$), this is equal to $T_{\bar{a}} = \inf\{t \in [0, \infty[: \epsilon_{\bar{a}}(t) \neq 0\}$ for patients who die within (get discharged from) the ICU under the given infection path $\bar{a}$ and ∞ for patients who get discharged from (die within) the ICU under that infection path. We further define the counterfactual CIF corresponding to event type k as $F_{\bar{a}k}(t) \equiv P(T_{\bar{a}} \leq t, \epsilon_{\bar{a}} = k) = P(\theta_{\bar{a}k} \leq t)$, for all $t \geq 0$. This is the probability that, under infection path $\bar{a}$, an event of type k occurs at or before time t .

Our focus in this article will be on inferring the CIF under monotone infection paths in which $a_t \geq a_s$ for $t > s$. This will enable us to infer the effect of acquiring infection on a given day, disregarding the possible impact of the duration of infection. For instance, we can define the causal effect of acquiring infection on day t since admission in the ICU on the cumulative incidence of ICU death as a comparison of 2 counterfactual CIFs under different monotone infection paths $(a_1, \dots, a_{t-1}, a_t, \dots, a_E) = (0, \dots, 0, 1, \dots, 1)$ and $(a_1, \dots, a_E) = (0, \dots, 0)$, indicating regimes where *all* patients acquire infection on day t since admission versus never in the ICU, respectively. By considering monotone infection paths only, we will be able to ignore information on the duration of infection, which is often imprecisely measured, and will thus average over patients with different duration of infection. Alternatively, the causal effect of infection can be reported in terms of the population attributable fraction of ICU mortality related to infection, which we define at each time t as

$$\frac{\hat{F}_1(t) - \hat{F}_{01}(t)}{\hat{F}_1(t)}, \quad (1)$$

where $\hat{F}_1(t)$ is the estimated observed CIF of ICU mortality at time t and $\hat{F}_{01}(t)$ is the estimated counterfactual CIF of ICU mortality at time t under the no infection path. This expresses what percentage of the observed ICU deaths by time t could be avoided by preventing infection. A related comparison of observed and counterfactual infection-free responses was considered in [Schulgen and Schumacher \(1996\)](#) to estimate the prolongation of stay attributable to nosocomial infections. We believe that (1) links more closely to the originally concept of attributable risk than other proposed measures ([Schumacher et al. 2007](#)), in particular because at each time it compares the same population twice.

2.3 Assumptions

To be able to identify the counterfactual CIF under all infection paths, the following 2 assumptions must be made (see Sect. 5 for a further discussion of the assumptions). First, we will make the so-called consistency assumption which links the counter-

factual data to the observed data by assuming that $T_{\bar{a}}$ and $\epsilon_{\bar{a}}$ equal the observed event time T and status ϵ for patients whose observed infection data are *compatible with* (i.e. would have been observed under) the given infection path $\bar{a}$ at that time (see also Sect. 3.1). To appreciate this assumption, note that our definition of ‘infection path’ assumes the existence of an intervention whereby infection is prevented for all patients during a given period of time and acquired at a given later time t (provided they are still in the ICU). The consistency assumption then states that this intervention is non-invasive in the sense that it has no effect other than causing infection at and not before time t (VanderWeele and Vansteelandt 2009). In particular, it assumes that for patients who would naturally acquire (prevent) infection on day t , the same event time and status would be observed under that intervention as was naturally observed. In this sense, the population attributable fraction (1) reflects a ‘pure’ infection effect that would be realized if patients would somehow ‘naturally’ avoid infections. Second, we will make the no unmeasured confounders assumption that, for each time t , among uninfected patients who have been in the ICU with the same covariate history $\bar{L}_{t-1} \equiv (L_0, \dots, L_{t-1})$ up to time $t - 1$, those who acquire infection at time t are exchangeable with those who do not in the sense that $\epsilon_{\bar{a}}(t) \prod A_t | \bar{A}_{t-1}, \epsilon_{\bar{a}}(t - 1) = 0, \bar{L}_{t-1}$ for all t and all infections paths $\bar{a}$; here, for any vector W_t , $\bar{W}_t$ refers to the history $(W_1, \dots, W_t)$ and $\underline{W}_t$ refers to the future $(W_t, \dots, W_E)$. Here we assume that, as was the case in our data analysis (see Sect. 3.3), the infection status was measured on day t for all patients who did not die or get discharged before day t .

3 Estimation under a nonparametric model for the counterfactual cumulative incidence function

3.1 Compatible infection paths

To infer the counterfactual CIF under a given infection path $\bar{a}$, we must select those patients whose observed data are compatible with it: i.e. those for whom we would have observed the same event time and status had such infection path been implemented. As an illustration, we will show how this can be done for the infection path $\bar{a}(0) = (0, 0, 0, \dots, 0)$ under which patients stay infection-free and for the infection path $\bar{a}(5) = (0, 0, 0, 0, 1, \dots, 1)$ under which patients acquire infection 5 days after admission. Consider the following 4 hypothetical patients in Fig. 1. Patient 1 stays in the ICU without infection for at least 10 days, patients 2 and 3 acquire infection in the ICU on day 5 and 4 respectively, and finally patient 4 gets discharged from the ICU without infection on day 3. The days on which a patient is compatible with the hypothetical infection paths $\bar{a}(0)$ and $\bar{a}(5)$ are indicated with a larger font size in Fig. 1a and b, respectively, as we now explain.

No infection: Figure 1a illustrates that patients who have not acquired infection and are still in the ICU by time t are compatible with $\bar{a}(0)$ at time t . For instance, at $t = 3$, all patients in Fig. 1a are compatible with $\bar{a}(0)$; at $t = 4$, patients 1 and 2, unlike patient 3, are compatible with $\bar{a}(0)$. In addition, patients who get discharged or die without

	t	1	2	3	4	5	6	7	8	9	10	...
	$\mathbf{a_t(0)}$	0	0	0	0	0	0	0	0	0	0	...
Patient 1	A_t	0	0	0	0	0	0	0	0	0	0	...
Patient 2	A_t	0	0	0	0	1	1	1	1	1	1	...
Patient 3	A_t	0	0	0	1	1	1	1	1	1	1	...
Patient 4	A_t	0	0	0								

A

	t	1	2	3	4	5	6	7	8	9	10	...
	$\mathbf{a_t(5)}$	0	0	0	0	1	1	1	1	1	1	...
Patient 1	A_t	0	0	0	0	0	0	0	0	0	0	...
Patient 2	A_t	0	0	0	0	1	1	1	1	1	1	...
Patient 3	A_t	0	0	0	1	1	1	1	1	1	1	...
Patient 4	A_t	0	0	0								

B

Fig. 1 Selection of patients compatible with infection paths $\bar{a}(0)$ (a) and $\bar{a}(5)$ (b)

infection before time t (e.g. patient 4) are compatible with $\bar{a}(0)$ at time t because their same observed data would have been obtained under a regime in which infection is prevented for all patients.

Acquiring infection on day 5: Figure 1b illustrates that on all days $t < 5$, this infection path coincides with the no infection path and thus the same patients are compatible with both $\bar{a}(5)$ and $\bar{a}(0)$ on those days. For instance, patient 1 in Fig. 1b is compatible with $\bar{a}(5)$ on all days $t < 5$; patient 3 is compatible with $\bar{a}(5)$ until the day before acquiring infection and obviously, patients who acquire infection on day 5 (i.e. patient 2) are compatible with $\bar{a}(5)$ on all days t . In addition, patients who get discharged or die without infection before day 5 (i.e. patient 4), are compatible with $\bar{a}(5)$ on day t because they would also have been discharged/have died prior to day 5 under a regime in which infection is acquired on day 5 (and not before) by all patients who are still in the ICU on that day.

Note from the previous examples that a patient's observed data can be compatible with (and thus carry information about) multiple infection paths. For instance, a patient who is discharged without infection 20 days after admission is compatible with all infection paths in which infection is acquired after day 20. When the end-of-follow-up time E is 30 days, this patient thus has 10 compatible infection paths (see Fig. 2). In general, a patient who gets discharged without infection at a given time t is compatible with all infection paths $\bar{a}(t + s)$ for $s > 0$.

3.2 Estimation under a given infection path

In this section, we discuss the estimation of the counterfactual CIF for a given infection path $\bar{a}$. Define $d_{\bar{a}kt}$ to be the number of events of type k that occur at time t for the

Because patients who are discharged or have died before time s stay compatible with infection path $\bar{a}$ when they were compatible with this path at the time of discharge/death, no further selection is made within that subgroup and thus we define $P(A_{is} = a_s | \epsilon_i(s-1), \bar{A}_{i,s-1}, \bar{L}_{i,s-1})$ to equal 1 when $\epsilon(s-1) \neq 0$. When $\epsilon(s-1) = 0$, this probability can be estimated using a pooled logistic regression model.

In the Appendix, we show that (3) is a consistent estimator of the counterfactual cumulative incidence at time t under infection path $\bar{a}$ when, in addition to the assumptions of Sect. 2.3, $P(A_t = a_t | \epsilon(t-1) = 0, \bar{A}_{t-1}, \bar{L}_{t-1})$ is correctly specified at each time t . In the Appendix, we further show that for given t , a conservative, asymptotic $(1 - \alpha)100\%$ confidence interval for $F_{\bar{a}k}(t)$ can be obtained as

$$\text{expit} \left[\text{logit} \left\{ \hat{F}_{\bar{a}k}(t) \right\} \pm z_{\alpha/2} \sqrt{\frac{1}{n} \frac{\widehat{\text{Var}}(U_{i\bar{a}k})}{\hat{F}_{\bar{a}k}(t)\{1 - \hat{F}_{\bar{a}k}(t)\}}} \right], \quad (5)$$

where $z_{\alpha/2}$ is the $(1 - \alpha/2)100\%$ percentile of the standard normal distribution, $U_{i\bar{a}k}$ is the weighted residual

$$U_{i\bar{a}k} \equiv W_{i\bar{a}t} \left\{ I(\epsilon_i(t) = k) - \hat{F}_{\bar{a}k}(t) \right\} \quad (6)$$

and $\widehat{\text{Var}}(U_{i\bar{a}k})$ refers to the sample variance of $U_{i\bar{a}k}$.

3.3 Data analysis

We use the proposed techniques for the analysis of the National ICU Surveillance Study in Belgium (Suetens et al. 1999, 2003; Vansteelandt et al. 2009). All ICU's in Belgian hospitals were invited to participate in this surveillance study on a voluntary basis. For all patients admitted to the ICU, data were recorded on personal characteristics, reasons for ICU admission, baseline health status, and daily indicators of received invasive treatments and acquired infections in the ICU. Nosocomial pneumonia (NP) was defined as pneumonia acquired by patients after the second day of ICU stay, to exclude infections that were in incubation upon enrollment in the ICU. The third day of stay in ICU will therefore be the starting point for our analysis, thus excluding patients who stayed less than 3 days. We restrict the analysis to surveillance data collected for the years 2002 and 2003 in three of the largest hospitals, which have accurate daily measurements of received invasive treatments and acquired infections. A total of 4288 ICU patients were analysed. Of the 360 (8.4%) patients who acquired NP in the ICU and stayed more than 2 days, 75 (20.8%) died in the ICU, as compared to 360 (9.2%) deaths among the 3928 patients who remained NP-free in the ICU; of these 360 (infected) patients, 245 or 68.2% (328 or 91.3%) acquired infection during the first week (first two weeks) after admission. Among patients who stayed more than 2 days in the ICU, the median length of stay in the ICU was 4 days (IQR 3, 95th percentile 14) for those without an infection history and 18 days (IQR 19, 95th percentile 30) for the remaining patients.

Table 1 Number of patients who died at or before time t and the number of patients compatible with $\bar{a}(0)$ and $\bar{a}(5)$, before and after weighting

t	Naïve				IPTW			
	$n_{\bar{a}(0)}$	$d_{\bar{a}(0)}$	$n_{\bar{a}(5)}$	$d_{\bar{a}(5)}$	$n_{\bar{a}(0)}$	$d_{\bar{a}(0)}$	$n_{\bar{a}(5)}$	$d_{\bar{a}(5)}$
3	4228	86	4228	86	4290	87	4290	88
4	4167	144	4167	144	4283	148	4283	148
5	4118	176	2482	144	4290	182	4135	148
6	4075	208	2482	144	4289	221	4135	148
7	4043	226	2482	144	4278	243	4135	148
8	4024	242	2482	145	4277	264	4135	328
9	4008	257	2482	146	4280	283	4135	356
10	3996	272	2482	146	4285	303	4135	356
11	3983	286	2482	147	4279	322	4135	417
12	3976	296	2482	148	4282	336	4135	424
13	3967	307	2482	149	4280	351	4135	439
14	3960	311	2482	150	4276	358	4135	461
15	3952	319	2482	150	4273	372	4135	461
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
30	3928	360	2482	152	4284	443	4135	543

n_a and d_a refer to the denominator and numerator, respectively, of the naïve estimator (2) and the IPTW estimator (3)

To estimate the CIF of ICU death under the no infection path $\bar{a}(0)$, we select those patients for whom we would have observed the same event time and status had such infection path been implemented (see Fig. 1a for an illustration). The number of patients $n_{\bar{a}(0)t}$ whose observed data are compatible with $\bar{a}(0)$ at each time t , along with the number $d_{\bar{a}(0)t}$ of those who died at or before time t are given in the second and third column of Table 1. Similar numbers are reported in columns 4 and 5 for the infection path $\bar{a}(5)$. Note that $n_{\bar{a}(0)t}$ diminishes each time with the number of patients who acquire infection at that time. Patients who get discharged without infection by a given time t stay compatible with this path until the end-of-study time E , which was chosen to equal 30 days in our analyses. Note also that, by construction, identical results are obtained for the infection paths $\bar{a}(0)$ and $\bar{a}(5)$ up to day 5. From day 5 onwards, the number of patients whose data are compatible with $\bar{a}(5)$ stays constant until the end-of-study time as it equals the number of patients who acquire infection on day 5 together with those who died or were discharged without infection prior to day 5.

The naïve estimator (2) corresponding to $\bar{a}(0)$ is obtained as the ratio of $d_{\bar{a}(0)t}$ over $n_{\bar{a}(0)t}$. The resulting estimates are displayed in Fig. 3a and suggest roughly a 1% absolute reduction in ICU mortality at 30 days if infection were avoided for all patients. These estimates are biased because they ignore confounding. We therefore calculate the IPTW estimator (3). This requires estimating the probability of acquiring infection

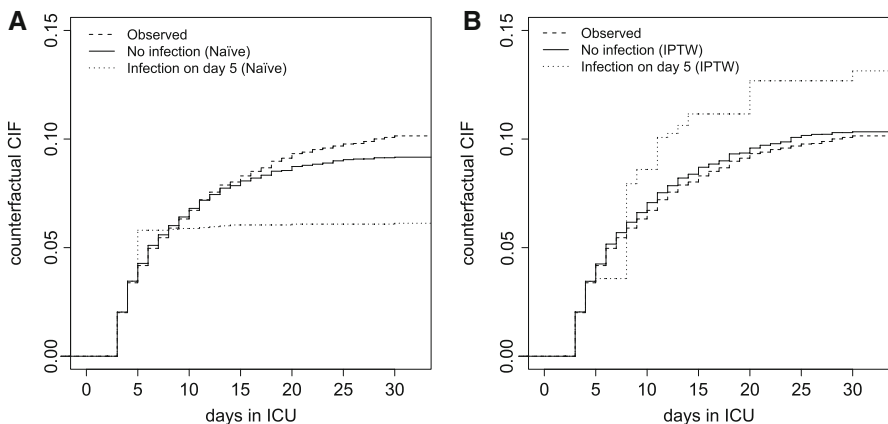
at each time t . We have built a pooled logistic regression model for this purpose, following the strategy explained in Vansteelandt et al. (2009). The probability of acquiring infection was associated with the presence of infection (other than NP) upon ICU admission and the following invasive treatments, received the day before infection (i.e., at time $t - 1$): mechanical ventilation, parenteral feeding and feeding through naso- or oro- intestinal tube, tracheotomy, oral intubation, the presence/absence of tracheotomy intubation, surgery and a central vascular catheter; and antibiotic treatment received two days before infection (i.e., at time $t - 2$). We did not adjust for antibiotic treatment the day before infection to exclude the possibility that this was already an effect of a latent infection. Additionally there was an association with gender and study center, evidence for effect modification of the SAPS II score by type of admission, multiple trauma and surgery, and evidence for effect modification of the effect of antibiotic treatment and parenteral feeding by study center. Finally, we used regression splines for the time effect. Using the predicted values from the final model, we calculated the probability of each patient being compatible with the chosen infection path up to time t , given baseline variables and time-dependent variables $\bar{L}_{t-1}$. The calculation of the resulting weights $W_{i\bar{a}t}$ is illustrated in Table 2 for three effectively observed patients who were discharged from the ICU at 5, 9 and 3 days after admission, respectively. In particular, the weight $W_{i\bar{a}t}$ is obtained by taking the reciprocal of the product of the probabilities π_{2is} over all time points $s = 3, \dots, t$ (the remaining columns of Table 2 will be explained in Sect. 4.2). Note the sharp increase in the weights when patients become infected. It follows that an analysis of most infection paths, such as $\bar{a}(5)$, may be much influenced by extreme weights so that less reliable inferences are typically obtained under these infection paths. For instance, the weights of patients who are compatible with $\bar{a}(5)$ had a median and mean of 1.01 and 1.59, an interquartile range and standard deviation of 0.03 and 9.47 and 1% and 99% percentiles of 1 and 9.93 (min. 1, max. 387.4). In contrast, inference for the infection-free path $\bar{a}(0)$ tends not to suffer much from influential weights because most observed patient-days correspond to no infection. For instance, the weights of patients who are compatible with $\bar{a}(0)$ had a median and mean of 1.02 and 1.08, an interquartile range and standard deviation of 0.05 and 0.22 and 1% and 99% percentiles of 1 and 1.93 (min. 1, max. 9.64). This is useful because the most relevant intervention is one where infection would be prevented for all patients and thus our most substantive interest lies in $\bar{a}(0)$.

Using the calculated weights, we can now predict the number of patients who would have died within the ICU by a given time had infection been prevented for all. This is given by the numerator of (3) and is displayed as $d_{\bar{a}(0)}$ in the 7th column of Table 1. The corresponding overall number of patients (i.e. the denominator of (3)) is shown as $n_{\bar{a}(0)}$ in column 6. It approximates the overall sample size (i.e., 4288) because the idea behind inverse weighting is to infer how the data would have looked like had all patients received the same intervention (e.g. been uninfected). The last 2 columns of Table 1 show similar results for the infection path $\bar{a}(5)$. The IPTW estimators (3) corresponding to $\bar{a}(0)$ and $\bar{a}(5)$ are again obtained as $d_{\bar{a}(0)}/n_{\bar{a}(0)}$ and $d_{\bar{a}(5)}/n_{\bar{a}(5)}$, respectively, and displayed in Fig. 3b along with the observed CIF. Not surprisingly, the CIF corresponding to the infection-free path is now higher than before because it acknowledges that patients who avoid infection during their ICU stay are relatively

Table 2 Estimated weights for three patients who are discharged on days 5, 9 and 3, respectively, corresponding to different infection histories

i	t	A_{it}	π_{1it}	π_{2it}	π_{3it}	$W_{i\bar{a}t}$	$h_{t\bar{a}_tk}(L_{i0})$	$h_{t\bar{a}_tk}(L_{i0})W_{i\bar{a}t}$
1	3	0	0.98	0.99	0	1.01	0.98	0.99
1	4	1	0.02	0.01	0.36	100	0.03	3
1	5	1	1	1	0.29	100	0.03	3
2	3	0	0.98	0.99	0	1.01	0.98	0.99
2	4	0	0.98	0.99	0.36	1.02	0.97	0.99
2	5	1	0.02	0.02	0.29	50	0.05	2.5
2	6	1	1	1	0.23	50	0.05	2.5
2	7	1	1	1	0.18	50	0.05	2.5
2	8	1	1	1	0.15	50	0.05	2.5
2	9	1	1	1	0.12	50	0.05	2.5
3	3	0	0.98	0.99	0	1.01	0.98	1

$$\pi_{1it} = P(A_{it} = a_{it} | \epsilon_i(t-1) = 0, A_{i,t-1} = 0, L_{i0}), \pi_{2it} = P(A_{it} = a_{it} | \epsilon_i(t-1), A_{i,t-1} = 0, \bar{L}_{i,t-1}), \\ \pi_{3it} = P(\epsilon_i(t-1) > 0 | \epsilon_i(t-2) = 0, \bar{A}_{i,t-1}, L_{i0})$$

**Fig. 3** Naïve estimate (a) and IPTW estimate (b) of the counterfactual CIF for ICU death under infection paths $\bar{a}(0)$ and $\bar{a}(5)$, together with the observed CIF

more healthy. The IPTW estimate of the CIF under infection path $\bar{a}(5)$ (Fig. 3b) is now more realistic than the naïve one, although very imprecise (see Fig. 4b). The fact that the IPTW estimate of the CIF under $\bar{a}(0)$ suggests slightly higher ICU death rates than in the observed population, might be counterintuitive, although there is little that can be said about this comparison given the imprecision of these estimates. The more efficient inferences developed in the next section will shed further light on this.

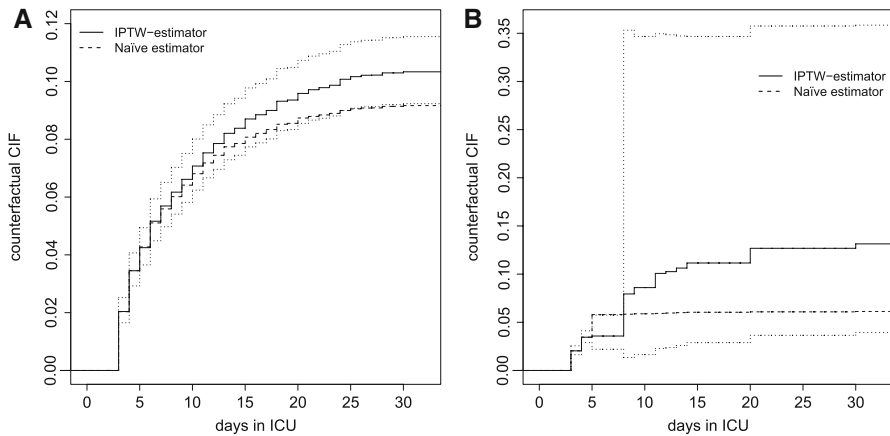


Fig. 4 IPTW estimate (with 95% pointwise confidence intervals) and naïve estimate of the counterfactual CIF for ICU death under infection paths $\bar{a}(0)$ (a) and $\bar{a}(5)$ (b)

4 Estimation under a semiparametric discrete-time marginal structural subdistribution hazard model

4.1 Model and estimation

The previous IPTW estimator is useful to infer the CIF of death that would be realized if all patients were uniformly exposed to the same infection path, but does not allow for direct comparison of CIFs corresponding to different chosen infection paths. It thus yields no effect estimates. Furthermore, in finite samples, the IPTW estimator is not guaranteed to yield a non-decreasing counterfactual CIF with time and can be very inefficient (see e.g. Fig. 4b). In view of this, we will develop models and accompanying inference for the counterfactual subdistribution hazard corresponding to event type k , which we define conditional on baseline covariates L_0 as:

$$\lambda_{\bar{a}k}(t|L_0) = P(T_{\bar{a}} = t, \epsilon_{\bar{a}} = k | T_{\bar{a}} \geq t \cup (T_{\bar{a}} < t \cap \epsilon_{\bar{a}} \neq k), L_0).$$

For $k = 1$, this represents the risk for a patient to die within the ICU on day t under infection path $\bar{a}$, given that he/she did not die in the ICU before that day. Note that the risk set at each time t is composed of those who did not experience the event of interest (i.e., who did not die) by time t , thus including those who experienced a competing event (i.e., who got discharged) by time t . We can now define a discrete-time marginal structural subdistribution hazard model for event type k to be a model for the counterfactual discrete-time hazard, possibly conditional on baseline covariates L_0 . For example, with $\delta_t \equiv \sum_{s=1}^{t-1} a_s$ denoting the number of days a patient would have been infected prior to time t under infection path $\bar{a}$, one might postulate that

$$\text{logit} \{ \lambda_{\bar{a}k}(t|L_0) \} = \omega(t) + \beta_1 a_t + \beta_2 \delta_t + \beta_3 L_0, \quad (7)$$

where $\omega(t)$ represents an unknown function of time which will later be modelled using regression splines and where β_1, β_2 and β_3 are unknown finite-dimensional parameters. When $k = 1$ and, as usual, the counterfactual hazard of death is relatively small at any given time, then $\exp(\beta_1 a_t + \beta_2 \delta_t)$ approximates the causal subdistribution rate ratio of ICU death at time t due to infection path $\bar{a}$ (as compared to the no infection path).

If, for each patient in the sample, we had the ability to measure $\epsilon_{\bar{a}}$ for each potential infection path $\bar{a}$, then we could fit model (7) by solving the following estimating equation

$$0 = \sum_{i=1}^n \sum_{t=1}^E \sum_{\bar{a}_t} U_{i\bar{a}_t,k} (h_{i\bar{a}_t,k}; \beta), \quad (8)$$

where $\beta = (\beta_1 \ \beta_2 \ \beta_3)'$ and

$$U_{i\bar{a}_t,k} (h_{i\bar{a}_t,k}; \beta) \equiv I(\epsilon_{i\bar{a}}(t-1) \neq k) h_{i\bar{a}_t,k}(L_{i0}) \Gamma \\ \times [I(\epsilon_{\bar{a}}(t) = k) - \text{expit}\{\omega(t) + \beta_1 a_t + \beta_2 \delta_t + \beta_3 L_{i0}\}].$$

Here, $h_{i\bar{a}_t,k}(L_{i0})$ is an arbitrary scalar function of $(t, \bar{a}_t, L_{i0})$, e.g. the constant 1, and Γ is the vector of predictors in the discrete-time marginal structural subdistribution hazard model, e.g. $(1 \ t \ a_t \ \delta_t \ L_{i0})'$ if $\omega(t)$ is chosen to be a linear function of time (i.e., $\omega(t) = \beta_0 + \beta_4 t$ with β_0 and β_4 unknown). The above estimating equation is not feasible because $\{\epsilon_{\bar{a}}(t), t = 1, \dots, E\}$ is not measured for each patient under each potential infection path. Thinking along the lines explained in Sect. 3, we will therefore select for each considered infection path, those patients whose observed data are compatible with that path and use inverse probability weighting to account for the selective nature of those subgroups. This is realized by solving the following equations,

$$0 = \sum_{i=1}^n \sum_{t=1}^E \sum_{\bar{a}_t} U_{i\bar{a}_t,k} (h_{i\bar{a}_t,k}; \beta) W_{i\bar{a}t} \quad (9)$$

which only involve observed data.

The choice of $h_{i\bar{a}_t,k}(L_{i0})$ can be very important in the analysis as it links directly to the efficiency of the estimator $\hat{\beta}$ for β obtained by solving this equation. In the Appendix, we show that for a single infection path $\bar{a}$, a locally efficient estimator within the above class is obtained by choosing

$$h_{i\bar{a}_t,k}(L_0) = \left[\sum_{d=1}^t P(\epsilon(d) \neq 0 | \epsilon(d-1) = 0, \bar{A}_d = \bar{a}_d, L_0) \right. \\ \left. \times \frac{\prod_{s=1}^d P(\epsilon(s-1) = 0 | \epsilon(s-2) = 0, \bar{A}_{s-1} = \bar{a}_{s-1}, L_0)}{\prod_{s=1}^d P(A_s = a_s | \epsilon(s-1) = 0, \bar{A}_{s-1} = \bar{a}_{s-1}, L_0)} \right]^{-1}$$

where, for each t , we define $P(\epsilon(d) \neq 0 | \epsilon(d-1) = 0, \bar{A}_d = \bar{a}_d, L_0)$ equal to 1 if $d = t$ and equal to 0 if $d = 0$. Here, the efficiency is local in the sense that it is only attained when infection and discharge/death are independent of the considered time-varying confounders, conditional on the considered baseline confounders. The index function $h_{i\bar{a}_t,k}(L_{i0})$ can be interpreted as the reciprocal of a prediction of the considered subject's inverse weight $W_{i\bar{a}_t}$ in Eq. 9 on the basis of time, the considered infection path and baseline confounders. As such, using this index function has the effect of stabilizing the weights in the sense of making patients who get assigned small probabilities $P(A_t = a_t | \epsilon(t-1) = 0, \bar{A}_{t-1}, \bar{L}_{t-1})$, less influential in the analysis. In line with others (Hernán et al. 2000), we will refer to $W_{i\bar{a}_t}$ as an unstable weight and to $h_{i\bar{a}_t,k}(L_{i0})W_{i\bar{a}_t}$ as a stable weight.

In principle, the third summation in (8) runs over all possible infection paths that are considered to be relevant to the investigator. In practice, depending on the chosen discrete-time marginal structural subdistribution hazard model, the estimating functions $U_{i\bar{a}_t,k}(h_{i\bar{a}_t,k}; \beta)$ for a given subject at a given time t may be identical with probability 1 between a number of infection paths. For instance, suppose that β_2 were set to zero in model (7). Then all infection paths with $a_t = 1$ (or $a_t = 0$) yield the same estimating function value $U_{i\bar{a}_t,k}(h_{i\bar{a}_t,k}; \beta)$ at time t , regardless of the infection history $\bar{a}_{t-1}$ or future $\bar{a}_{t+1} \equiv (a_{t+1}, \dots, a_E)$. In view of this, we recommend to sum at each time t only over those infection paths in the estimating Eq. 9 that yield unique estimating function values (with probability 1) at that time. This is inspired by the fact that the efficient score for the parameters indexing a conditional mean model counts perfectly correlated observations only once (Chamberlain 1987). The resulting analysis continues to yield a consistent (and likely more efficient) estimator because it follows from the Appendix that each term in the summation (9) has mean zero when evaluated at the truth.

A conservative estimate for the asymptotic variance of the estimator $\hat{\beta}$ which solves Eq. 9 can be obtained via the sandwich estimator

$$\frac{1}{n} E^{-1} \left(\frac{\partial U_i(\beta)}{\partial \beta} \right) \text{Var} \{U_i(\beta)\} E^{-1} \left(\frac{\partial U_i(\beta)}{\partial \beta} \right)',$$

where we define

$$U_i(\beta) = \sum_{t=1}^E \sum_{\bar{a}_t} U_{i\bar{a}_t,k}(h_{i\bar{a}_t,k}; \beta) W_{i\bar{a}_t}$$

and where population expectations and variances can be replaced with sample analogs and β with $\hat{\beta}$. The counterfactual CIF corresponding to infection path $\bar{a}$ among patients with baseline covariate $L_0 = l_0$ can be estimated under the discrete-time marginal structural subdistribution hazard model (7) as

$$\hat{F}_{\bar{a}k}(t) = 1 - \prod_{s \leq t} \{1 - \lambda_{\bar{a}k}(s | l_0)\}.$$

This counterfactual CIF is always well-defined because the marginal structural subdistribution hazard model (7) only involves baseline covariates and no time-varying covariates. Using the Delta method, a corresponding conservative, asymptotic $100(1 - \alpha)\%$ confidence interval for given t is obtained as

$$1 - \exp \left[\log \{1 - \hat{F}_{\bar{a}k}(t)\} \pm z_{\alpha/2} \sigma_{\bar{a}k}(t) \right]$$

where

$$\sigma_{\bar{a}k}^2(t) = \left(\sum_{s \leq t} \frac{\lambda'_{\bar{a}k}(s|l_0)}{1 - \lambda_{\bar{a}k}(s|l_0)} \right) \text{Var}(\hat{\beta}) \left(\sum_{s \leq t} \frac{\lambda'_{\bar{a}k}(s|l_0)}{1 - \lambda_{\bar{a}k}(s|l_0)} \right)'$$

and where $\lambda'_{\bar{a}k}(s|l_0)$ is the first order derivative of $\lambda_{\bar{a}k}(s|l_0)$ w.r.t. β .

4.2 Data analysis

The stable weights $h_{t\bar{a},k}(L_{i0})W_{i\bar{a}t}$ were obtained by estimating $h_{t\bar{a},k}(L_{i0})$ based on two pooled logistic regression models for discharge/death and infection in function of time. In our analysis, the model for discharge/death contained the infection history and regression splines for the time effect, but (for computational simplicity) no baseline covariates. Table 2 illustrates the calculation of the stable weights for three effectively observed patients. Note indeed that the resulting weights in the last column of Table 2 are more stable than the unstable weights $W_{i\bar{a}t}$ that were previously considered. It can be shown that the unstable weights stay constant after discharge and that the stable weights for a given infection path $\bar{a}$ remain constant at those times t where $a_t = 1$. In general, the stable weights continue to change with time after discharge because they involve the probability of getting discharged at each time.

Parameter estimates of the infection effect were obtained by solving the weighted estimating Eq. 9. This required extending the dataset by including all compatible infection paths, as illustrated in Fig. 2. We estimate the time-dependent intercept $\omega(t)$ in model (7) using natural cubic splines with five knots. The analysis was adjusted for baseline covariates (ICU center, SAPSII score on admission, infection upon admission, oral intubation on admission and type of admission), which were also used in the infection models in the calculation of the stable weights. To check the stability of the weights for a given patient at a given time t we calculated that patient's overall weight at that time as the sum of the weights over all compatible infection paths. The resulting stable weights had a median and mean of 2.37 and 2.74, an interquartile range and standard deviation of 2.57 and 1.97 and 1% and 99% percentiles of 0.41 and 7.35 (min. 0.16, max. 53.4). In contrast, the unstable weights had a median and mean of 13.08 and 15.71, an interquartile range and standard deviation of 15.91 and 25.19 and 1% and 99% percentiles of 1 and 90.36 (min. 1, max. 529.81).

The results for model (7) are given in the first two rows of Table 3. Note the large difference between the estimates obtained with and without inverse weighting. This underscores the importance of confounding adjustment. The IPTW estimates for β_1

Table 3 The naïve parameter estimates for model (7) including (the most important) baseline covariates, along with the IPTW estimates using unstable and stable weights

	Naïve			IPTW (unstable)			IPTW (stable)		
	Estimate	SE	OR	Estimate	SE	OR	Estimate	SE	OR
β_1	-3.27 ^a	0.21	0.038	-0.33	0.35	0.72	-0.49	0.28	0.61
β_2	0.12 ^a	0.022	1.13	0.095 ^a	0.035	1.10	0.095 ^a	0.024	1.10
β_2				0.087 ^a	0.032	1.09	0.059 ^a	0.021	1.06

^a Significant at the 5% significance level

suggest a drop in the hazard of death during the first 4–5 days after acquiring infection (which could possibly be explained by extra care after infection or delayed decision of treatment interruption), but the evidence for this is weak. We hence fitted a reduced model by setting $\beta_1 = 0$ and found a 6.1% (95%CI 1.8–10.0) increase in the hazard of ICU death per additional day since acquiring infection (see the third row of Table 3). This reflects the increased risk of ICU death which is partly attributable to a direct effect of infection on death, and partly to the increase in length of stay after infection, which implies a greater chance of death in the ICU. Figure 5 shows the estimated counterfactual CIFs obtained using the naïve and IPTW estimates from the final models (but now no longer conditioning on baseline covariates). They give similar results as obtained in Sect. 3 using the nonparametric estimators, but a drastic increase in efficiency for the infection paths. In particular, the stable IPTW estimator suggests a roughly 3% increase in ICU mortality due to infection on day 5 as compared to no infection. Figure 6 allows for evaluating how much the CIF of ICU death would change if all infections could be avoided. It suggests a population attributable fraction of 30-day ICU mortality related to infection of 6.4%, indicating that 6.4% of the observed ICU deaths at 30 days could have been avoided by preventing infection.

5 Discussion

We have developed a general framework for estimating the causal effect of a time-varying exposure on a survival outcome in the presence of competing risks. The validity of the approach relies on the following four assumptions. First, it assumes correctly specified models for the probability of infection at each time in function of the history of measured time-varying confounders. Second, it assumes that there are no patients who, on the basis of their time-varying confounder history, are precluded from avoiding or acquiring infection. While we believe the second assumption was realistic in our data analysis, in general, there may be patients who have a small daily risk of infection on the basis of their time-varying confounder history. This can lead to patients having large weights in the analysis who may thus become influential. In future work, we will try to decrease the reliance on inverse weights by extracting additional information from models for the subdistribution hazard of ICU death in function of time-varying

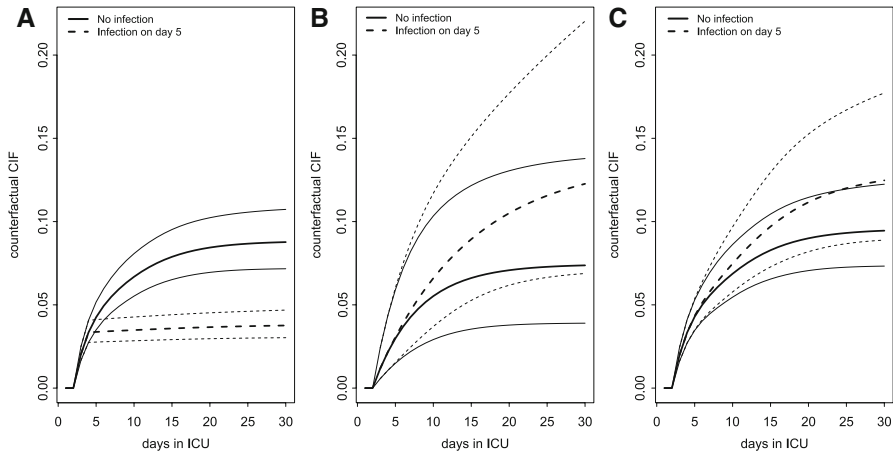


Fig. 5 Counterfactual CIF of ICU death for infection paths $\bar{a}(0)$ (solid) and $\bar{a}(5)$ (dashed), along with 95% pointwise confidence intervals, as based on the naïve estimator (a), the unstable IPTW estimator (b) and the stable IPTW estimator (c)

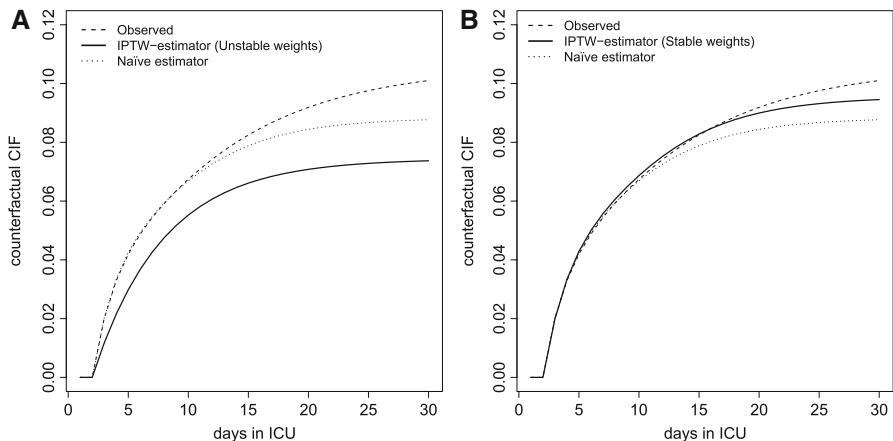


Fig. 6 Counterfactual CIF of ICU death for infection paths $\bar{a}(0)$, as based on the naïve estimator, the unstable IPTW estimator (a) and the stable IPTW estimator (b) together with the observed CIF of death

measurements, along the lines of [Goetgeluk et al. \(2008\)](#). This might lower the impact of both model misspecification and of influential weights.

Both previous assumptions are (at least to some degree) testable on the basis of the observed data. The next two assumptions are inherently untestable, so that their plausibility must be judged against subject-matter knowledge. First, the proposed approach assumes that at each time, all time-varying confounders for the association between infection and ICU mortality have been measured. While this assumption might be approximately justified for our analysis, it is likely not entirely met because we could only adjust for the applied invasive treatments, which should be regarded as a surrogates for the actual disease severity. In future work, we will apply the proposed

methods to the Outcomerea (French multicenter ICU database) and the COSARA (ICU database at the Ghent University Hospital, Belgium) database, which carry very extensive information about the patients' severity of illness. Second, the approach additionally assumes that confounders measured at time $t - 1$, i.e. L_{t-1} , are a cause, rather than a consequence of infection at time t . This assumption may not be satisfied when infection is in incubation several days prior to its detection.

An important advantage of our approach is that it separates the adjustment for time-varying confounders from the final model for the counterfactual CIF. This implies that this approach can technically handle high-dimensional confounders without imposing a high computational burden and even when confounders are themselves influenced by the infection. In addition, because the model involves no time-varying covariates, the corresponding CIF is well defined and can be displayed for the overall patient population, even with adjustment for high-dimensional confounders, without the need to restrict the CIF to patient subgroups. A further advantage is that, because infection paths are well defined upon ICU admission, our approach does not require imputing or stopping the covariate process (including infection status) for patients who experience the competing event (Beyersmann and Schumacher 2008; Latouche et al. 2005). A final advantage is that it yields fairly easily interpretable results. For instance, the results can be expressed in terms of population attributable fractions comparing the observed population with how that same population would have looked like if infection were prevented for all patients. However, caution must be exercised because the results state the impact of hospital-acquired infections on ICU mortality rather than overall mortality. This implies that the magnitude of the infection effect is partly influenced by e.g. the decision to keep infected patients longer in the ICU (which implies a greater chance of observing ICU death) or to send terminally ill patients home. Care must also be exercised in noting that late infections are forced to have a small effect under our model because most patients will have left the ICU by that time and thus not be affected. This explains the discrepancy with the results of Vansteelandt et al. (2009), who estimate the effect of acquiring infections while still being present in the ICU.

We did not discuss how to deal with censored event times as this was not an issue in our data analysis. In the presence of the usual right-censoring, our method can be extended using additional inverse probability of censoring weights along the lines of Fine and Gray (Fine and Gray 1999; Robins and Rotnitzky 1992). Finally, we have chosen a development which adopts the principles underlying marginal structural models (Robins et al. 2000). Alternatively, we could have considered a development based on structural nested failure time models, as in Keiding et al. (1999) and Schulgen and Schumacher (1996). We have chosen not to do so because the survival time of critically ill patients can be greatly influenced by the decision of doctors and family to keep patients artificially alive, so that 30-day ICU mortality tends to be a more relevant endpoint to physicians. In addition, it is so far unclear how to account for censoring due to ICU discharge in this approach, without running into the aforementioned difficulties caused by inverse probability of censoring weighting.

In conclusion, we hope that the proposed methods will shed new light on addressing issues of time-dependent confounding in survival analyses with competing risks and,

furthermore, will help improve our understanding of the role of nosocomial infections in ICU mortality.

Acknowledgements We would like to thank Martin Schumacher, Jan Beyersmann, Martin Wolkowitz and Christine Gall (University of Freiburg, Germany) for organizing a workshop on hospital epidemiology in February 2008, out of which followed stimulating discussions that prompted the writing of this paper. The authors are also grateful to Johan Decruyenaere, Dominique Benoit (Ghent University, Belgium) and Jean-Francois Timsit (Centre de recherche Inserm, France) for helpful discussions and to two referees for very helpful comments. We acknowledge support from IAP research network grant nr. P06/03 from the Belgian government (Belgian Science Policy) and the Institute for the Promotion of Innovation through Science and Technology in Flanders (IWT-Vlaanderen).

Appendix

We drop the subject index i for notational simplicity.

Derivation of expression (5) To derive expression (5), note that $\hat{F}_{\bar{a}k}(t)$ is such that $\sum_{i=1}^n U_{i\bar{a}_t k} = 0$, with $U_{i\bar{a}_t k}$ defined in (6). It then follows using standard arguments for M-estimators that, under weak regularity conditions and for given t , the asymptotic variance of $\hat{G}_{\bar{a}k}(t) \equiv \text{logit} \left\{ \hat{F}_{\bar{a}k}(t) \right\}$ equals

$$\text{Var}(\hat{G}_{\bar{a}k}(t)) = \frac{1}{n} E^{-2} \left(\frac{\partial U_{i\bar{a}_t k}}{\partial G_{\bar{a}k}(t)} \right) \text{Var}(U_{i\bar{a}_t k}),$$

where

$$E \left(\frac{\partial U_{i\bar{a}_t k}}{\partial G_{\bar{a}k}(t)} \right) = F_{\bar{a}k}(t) \{1 - F_{\bar{a}k}(t)\}.$$

The result (5) is now immediate. It yields a conservative $(1 - \alpha)100\%$ confidence interval by the fact that it does not acknowledge estimation of the unknown population probabilities appearing in the inverse weights (van der Laan and Robins 2003).

Unbiasedness of estimating Eq. 9 For simplicity of notation, we write $U_{i\bar{a}_t}(\beta) \equiv U_{i\bar{a}_t k}(h_{i\bar{a}_t k}; \beta)$ for a given choice of $h_{i\bar{a}_t k}(L_{i0})$. Using the law of iterated expectations, the population expectation of (9), evaluated at β , equals

$$\begin{aligned} & E \left[\sum_{t=1}^E \sum_{\bar{a}_t} E \left(U_{i\bar{a}_t}(\beta) \prod_{s=1}^t \frac{I(A_s = a_s) I(\epsilon(s-1) = 0) + I(\epsilon(s-1) \neq 0)}{P(A_s = a_s | \epsilon(s-1), \bar{A}_{s-1}, \bar{L}_{s-1})} \middle| \bar{A}_{t-1}, \right. \right. \\ & \quad \left. \left. \bar{L}_{t-1}, \overline{\epsilon_{\bar{a}}(t)} \right) \right] \\ & = E \left[\sum_{t=1}^E \sum_{\bar{a}_t} U_{i\bar{a}_t}(\beta) \prod_{s=1}^{t-1} \frac{I(A_s = a_s) I(\epsilon(s-1) = 0) + I(\epsilon(s-1) \neq 0)}{P(A_s = a_s | \epsilon(s-1), \bar{A}_{s-1}, \bar{L}_{s-1})} \right] \end{aligned}$$

$$\times E \left(\frac{I(A_t = a_t)I(\epsilon(t-1) = 0) + I(\epsilon(t-1) \neq 0)}{P(A_t = a_t | \epsilon(t-1), \bar{A}_{t-1}, \bar{L}_{t-1})} \mid \bar{A}_{t-1}, \bar{L}_{t-1}, \overline{\epsilon_{\bar{a}}(t)} \right) \Bigg]. \quad (10)$$

Here, the expectation in the last line can be rewritten as

$$\begin{aligned} & \frac{P(A_t = a_t | \overline{\epsilon_{\bar{a}}(t)}, \bar{A}_{t-1}, \bar{L}_{t-1}) I(\epsilon(t-1) = 0)}{P(A_t = a_t | \epsilon(t-1), \bar{A}_{t-1}, \bar{L}_{t-1})} \\ & + \frac{I(\epsilon(t-1) \neq 0)}{P(A_t = a_t | \epsilon(t-1), \bar{A}_{t-1}, \bar{L}_{t-1})}. \end{aligned} \quad (11)$$

The first term in (11) equals $I(\epsilon(t-1) = 0)$ under the no unmeasured confounders assumption. The second term in (11) equals $I(\epsilon(t-1) \neq 0)$ by the fact that we defined $P(A_t = a_t | \overline{\epsilon_{\bar{a}}(t)}, \bar{A}_{t-1}, \bar{L}_{t-1}) = P(A_t = a_t | \epsilon(t-1), \bar{A}_{t-1}, \bar{L}_{t-1})$ to equal 1 for subjects with $\epsilon(t-1) \neq 0$. Equation 10 thus simplifies to

$$E \left[\sum_{t=1}^T \sum_{\bar{a}_t} U_{t\bar{a}_t}(\beta) \prod_{s=1}^{t-1} \frac{I(A_s = a_s)I(\epsilon(s-1) = 0) + I(\epsilon(s-1) \neq 0)}{P(A_s = a_s | \epsilon(s-1), \bar{A}_{s-1}, \bar{L}_{s-1})} \right]$$

Repeating these arguments another $t-1$ times, we obtain $E \left\{ \sum_{t=1}^T \sum_{\bar{a}_t} U_{t\bar{a}_t}(\beta) \right\}$ which has mean zero by the fact that $U_{t\bar{a}_t}(\beta)$ is an unbiased estimating function for each t and $\bar{a}_t$. A similar proof can be developed, starting from estimating function (6) to show that (3) is a consistent estimator of the CIF under infection path $\bar{a}$ at time t .

Stable weights For simplicity of notation and without loss of generality, we let $k = 1$ and drop this index in all further expressions. To derive stabilized weights, we focus on 1 infection path $\bar{a}$ and calculate the efficient index function $h_{t\bar{a}_t}(L_0)$ in the estimating function

$$\sum_{t=1}^T h_{t\bar{a}_t}(L_0) V_{t\bar{a}_t} W_{t\bar{a}_t}, \quad (12)$$

where we define, for notational convenience,

$$\begin{aligned} V_{t\bar{a}_t} & \equiv \{I(\epsilon_{\bar{a}}(t) = 1) - \text{expit}(\beta' X)\} I(\epsilon_{\bar{a}}(t-1) \neq 1) \\ W_{t\bar{a}_t} & \equiv \prod_{s=1}^t \frac{I(A_s = a_s)I(\epsilon(s-1) = 0) + I(\epsilon(s-1) \neq 0)}{P(A_s = a_s | \epsilon(s-1), \bar{A}_{s-1}, \bar{L}_{s-1})}, \end{aligned}$$

where X is, for instance, equal to $(1 \ t \ a_t \ \delta_t \ L_0)'$. It follows from standard semiparametric theory (Chamberlain 1987) that the efficient index function (in the sense of leading to an estimator of β with minimal variance among all estimators obtained by

solving an estimating equation based on estimating functions of the form (12) with $P(A_s = a_s | \epsilon(s-1), \bar{A}_{s-1}, \bar{L}_{s-1})$ known for $s = 1, \dots, E$ is given by

$$h_{i\bar{a}_t}(L_0) = E \left(\frac{\partial}{\partial \beta} V_{i\bar{a}_t} W_{i\bar{a}_t} | L_0 \right) \text{Var}^{-1} (V_{i\bar{a}_t} W_{i\bar{a}_t} | L_0). \quad (13)$$

Define $\epsilon_{i\bar{a}} \equiv (I(\epsilon_{\bar{a}}(1) \neq 1), \dots, (\epsilon_{\bar{a}}(t) \neq 1))$. Using similar arguments as in the previous appendix, one can show that $E \left(\frac{\partial}{\partial \beta} V_{i\bar{a}_t} W_{i\bar{a}_t} | L_0 \right)$ equals $X \text{expit}(\beta' X) \{1 - \text{expit}(\beta' X)\} \times P(\epsilon_{\bar{a}}(t-1) \neq 1 | L_0)$. Using the law of iterated variances, the variance in (13) can be calculated as

$$E \{ \text{Var} (V_{i\bar{a}_t} W_{i\bar{a}_t} | \epsilon_{i\bar{a}}, L_0) | L_0 \} + \text{Var} \{ E (V_{i\bar{a}_t} W_{i\bar{a}_t} | \epsilon_{i\bar{a}}, L_0) | L_0 \}. \quad (14)$$

The second term in (14) equals

$$\text{expit}(\beta' X) \{1 - \text{expit}(\beta' X)\} P(\epsilon_{\bar{a}}(t-1) \neq 1 | L_0), \text{ because } E (W_{i\bar{a}_t} | \epsilon_{i\bar{a}}, L_0) = 1.$$

The first term in (14) equals

$$\begin{aligned} & E \left\{ V_{i\bar{a}_t}^2 \text{Var} (W_{i\bar{a}_t} | \epsilon_{i\bar{a}}, L_0) | L_0 \right\} \\ &= E \left[V_{i\bar{a}_t}^2 \left\{ E (W_{i\bar{a}_t}^2 | \epsilon_{i\bar{a}}, L_0) - E (W_{i\bar{a}_t} | \epsilon_{i\bar{a}}, L_0)^2 \right\} | L_0 \right]. \end{aligned}$$

We first calculate $E (W_{i\bar{a}_t}^2 | \bar{A}_{t-1}, \bar{L}_{t-1}, \epsilon_{i\bar{a}})$ as

$$\begin{aligned} & E \left[W_{t-1, \bar{a}_{t-1}}^2 \frac{I(A_t = a_t) I(\epsilon(t-1) = 0) + I(\epsilon(t-1) \neq 0)}{P(A_t = a_t | \epsilon(t-1), \bar{A}_{t-1}, \bar{L}_{t-1})^2} | \bar{A}_{t-1}, \bar{L}_{t-1}, \epsilon_{i\bar{a}} \right] \\ &= E \left\{ W_{t-1, \bar{a}_{t-1}}^2 \left[\frac{I(\epsilon(t-1) = 0)}{P(A_t = a_t | \epsilon(t-1) = 0, \bar{A}_{t-1} = \bar{a}_{t-1}, \bar{L}_{t-1})} \right. \right. \\ &\quad \left. \left. + I(\epsilon(t-1) \neq 0) \right] | \bar{A}_{t-1}, \bar{L}_{t-1}, \epsilon_{i\bar{a}} \right\}, \end{aligned}$$

where we use the no unmeasured confounders assumption along with the simplifying assumption that $A_t \perp\!\!\!\perp \bar{L}_{t-1} | \epsilon(t-1) = 0, \bar{A}_{t-1}, L_0$ (note that violation of this assumption may affect the efficiency, but not the asymptotic unbiasedness of the estimator because each choice of $h_{i\bar{a}_t}(L_0)$ yields a consistent estimator). Starting from here, we next calculate $E (W_{i\bar{a}_t}^2 | \bar{A}_{t-1}, \bar{L}_{t-2}, \epsilon_{i\bar{a}})$ as

$$\begin{aligned}
& E \left(W_{t-1, \bar{a}_{t-1}}^2 \left[\frac{I(\epsilon(t-1) = 0)}{P(A_t = a_t | \epsilon(t-1) = 0, \bar{A}_{t-1} = \bar{a}_{t-1}, \bar{L}_0)^2} \right. \right. \\
& \quad \left. \left. + I(\epsilon(t-1) \neq 0) \right] | \bar{A}_{t-1}, \bar{L}_{t-1}, \epsilon_{t\bar{a}} \right) \\
&= E \left\{ W_{t-1, \bar{a}_{t-1}}^2 \left(I(\epsilon(t-2) = 0) \left\{ \frac{P(\epsilon(t-1) = 0 | \epsilon(t-2) = 0, \bar{A}_{t-1} = \bar{a}_{t-1}, L_0)}{P(A_t = a_t | \epsilon(t-1) = 0, \bar{A}_{t-1} = \bar{a}_{t-1}, \bar{L}_0)} \right. \right. \right. \\
& \quad \left. \left. + P(\epsilon(t-1) \neq 0 | \epsilon(t-2) = 0, \bar{A}_{t-1} = \bar{a}_{t-1}, L_0) \right\} \right. \\
& \quad \left. \left. + I(\epsilon(t-2) \neq 0) \right) | \bar{A}_{t-1}, \bar{L}_{t-1}, \epsilon_{t\bar{a}} \right\}
\end{aligned}$$

where we now make the simplifying assumption that $P(\epsilon(t-1) | \epsilon(t-2) = 0, \bar{A}_{t-1} = \bar{a}_{t-1}, \bar{L}_{t-1}, \epsilon_{t\bar{a}}) = P(\epsilon(t-1) | \epsilon(t-2) = 0, \bar{A}_{t-1} = \bar{a}_{t-1}, L_0)$ (note again that violation of this assumption may affect the efficiency, but not the asymptotic unbiasedness of the estimator). Starting from here, we then calculate $E(W_{t\bar{a}_t}^2 | \bar{A}_{t-2}, \bar{L}_{t-2}, \epsilon_{t\bar{a}})$ as

$$\begin{aligned}
& E \left\{ W_{t-2, \bar{a}_{t-2}}^2 \left[I(\epsilon(t-2) \neq 0) \right. \right. \\
& \quad \left. \left. + \frac{I(A_{t-1} = a_{t-1}) I(\epsilon(t-2) = 0)}{P(A_{t-1} = a_{t-1} | \epsilon(t-2) = 0, \bar{A}_{t-2} = \bar{a}_{t-2}, \bar{L}_{t-2})^2} \right. \right. \\
& \quad \left. \left. \times \left\{ \frac{P(\epsilon(t-1) = 0 | \epsilon(t-2) = 0, \bar{A}_{t-1} = \bar{a}_{t-1}, L_0)}{P(A_t = a_t | \epsilon(t-1) = 0, \bar{A}_{t-1} = \bar{a}_{t-1}, \bar{L}_0)} \right. \right. \right. \\
& \quad \left. \left. \left. + P(\epsilon(t-1) \neq 0 | \epsilon(t-2) = 0, \bar{A}_{t-1} = \bar{a}_{t-1}, L_0) \right\} \right] | \bar{A}_{t-2}, \bar{L}_{t-2}, \epsilon_{t\bar{a}} \right\} \\
&= E \left\{ W_{t-2, \bar{a}_{t-2}}^2 \left[I(\epsilon(t-2) \neq 0) \right. \right. \\
& \quad \left. \left. + \frac{I(\epsilon(t-2) = 0)}{P(A_{t-1} = a_{t-1} | \epsilon(t-2) = 0, \bar{A}_{t-2} = \bar{a}_{t-2}, \bar{L}_{t-2})} \right. \right. \\
& \quad \left. \left. \times \left\{ \frac{P(\epsilon(t-1) = 0 | \epsilon(t-2) = 0, \bar{A}_{t-1} = \bar{a}_{t-1}, L_0)}{P(A_t = a_t | \epsilon(t-1) = 0, \bar{A}_{t-1} = \bar{a}_{t-1}, \bar{L}_0)} \right. \right. \right. \\
& \quad \left. \left. \left. + P(\epsilon(t-1) \neq 0 | \epsilon(t-2) = 0, \bar{A}_{t-1} = \bar{a}_{t-1}, L_0) \right\} \right] \right. \\
& \quad \left. | \bar{A}_{t-2}, \bar{L}_{t-2}, \epsilon_{t\bar{a}} \right\},
\end{aligned}$$

where we use the simplifying assumption that $A_{t-1} \perp\!\!\!\perp \bar{L}_{t-2} | \epsilon(t-2) = 0, \bar{A}_{t-2}, L_0$. Continuing along these lines, we obtain that

$$E\left(W_{t\bar{a}_t}^2 | \epsilon_{t\bar{a}_t}, L_0\right) = \sum_{d=1}^t P(\epsilon(d) \neq 0 | \epsilon(d-1) = 0, \bar{A}_d = \bar{a}_d, L_0) \\ \times \frac{\prod_{s=1}^d P(\epsilon(s-1) = 0 | \epsilon(s-2) = 0, \bar{A}_{s-1} = \bar{a}_{s-1}, L_0)}{\prod_{s=1}^d P(A_s = a_s | \epsilon(s-1) = 0, \bar{A}_{s-1} = \bar{a}_{s-1}, L_0)},$$

where, for each t , we define $P(\epsilon(d) \neq 0 | \epsilon(d-1) = 0, \bar{A}_d = \bar{a}_d, L_0)$ equal to 1 if $d = t$ and equal to 0 if $d = 0$. Because $E(W_{t\bar{a}_t} | \epsilon_{t\bar{a}_t}, L_0) = 1$, it follows that (14) equals

$$\text{Var}(V_{t\bar{a}_t} | L_0) \sum_{d=1}^t P(\epsilon(d) \neq 0 | \epsilon(d-1) = 0, \bar{A}_d = \bar{a}_d, L_0) \\ \times \frac{\prod_{s=1}^d P(\epsilon(s-1) = 0 | \epsilon(s-2) = 0, \bar{A}_{s-1} = \bar{a}_{s-1}, L_0)}{\prod_{s=1}^d P(A_s = a_s | \epsilon(s-1) = 0, \bar{A}_{s-1} = \bar{a}_{s-1}, L_0)}$$

and that $h_{t,\bar{a}_t}(L_0)$ in (13) equals

$$X \left[\sum_{d=1}^t P(\epsilon(d) \neq 0 | \epsilon(d-1) = 0, \bar{A}_d = \bar{a}_d, L_0) \right. \\ \left. \times \frac{\prod_{s=1}^d P(\epsilon(s-1) = 0 | \epsilon(s-2) = 0, \bar{A}_{s-1} = \bar{a}_{s-1}, L_0)}{\prod_{s=1}^d P(A_s = a_s | \epsilon(s-1) = 0, \bar{A}_{s-1} = \bar{a}_{s-1}, L_0)} \right]^{-1}.$$

References

- Andersen P, Borgan O, Gill R, Keiding N (1993) Statistical models based on counting processes, Springer series in statistics New York. Springer, NY
- Barnett A, Graves N (2008) Competing risks models and time-dependent covariates. *Crit Care* 12:143
- Bercault N, Boulain T (2001) Mortality rate attributable to ventilator-associated nosocomial pneumonia in an adult intensive care unit: a prospective case-control study. *Crit Care* 29:2303–2309
- Beyersmann J, Schumacher M (2008) Time-dependent covariates in the proportional subdistribution hazards model for competing risks. *Biostatistics* 9:765–776
- Beyersmann J, Gastmeier P, Grundmann H, Bärwolff S, Geffers C, Behnke M, Rüden H, Schumacher M (2006) Use of multistate models to assess prolongation of intensive care unit stay due to nosocomial infection. *Inf Control Hosp Epidemiol* 27:493–499
- Beyersmann J, Wolkewitz M, Schumacher M (2008) The impact of time-dependent bias in proportional hazards modelling. *Stat Med* 27:6439–6454
- Chamberlain G (1987) Asymptotic efficiency in estimation with conditional moment restrictions. *J Econom* 34:305–334
- Davis K (2006) Ventilator-associated pneumonia: a review. *J Intensive Care Med* 21:211–226
- Fagon J, Chastre J, Hance A, Montravers P, Novara A, Gibert C (1993) Nosocomial pneumonia in ventilated patient: a cohort study evaluating attributable mortality and hospital stay. *Am J Med* 94:281–299
- Fine J, Gray R (1999) A proportional hazards model for the subdistribution of a competing risk. *J Am Stat Assoc* 94:496–509

- Girou E, Stephan F, Novara A, Safar M, Fagon J (1998) Risk factors and outcome of nosocomial infections: results of a matched case-control study of ICU patients. *Am J Respir Crit Care Med* 157:1151–1158
- Goetghebeur S, Vansteelandt S, Goetghebeur E (2008) Estimation of controlled direct effects. *J R Stat Soc Ser B Methodol* 70:1049–1066
- Greenland S (2003) Quantifying biases in causal models: classical confounding vs collider-stratification bias. *Epidemiology* 14:300–306
- Hernán M, Robins J (2006) Estimating causal effects from epidemiological data. *J Epidemiol Commun Health* 60:578–586
- Hernán M, Brumback B, Robins J (2000) Marginal structural models to estimate the causal effect of zidovudine on the survival of HIV-positive men. *Epidemiology* 11:561–570
- Hernán M, Diaz SH, Robins J (2004) A structural approach to selection bias. *Epidemiology* 15:615–625
- Heyland D, Cook D, Griffith L (1999) The attributable morbidity and mortality of ventilator-associated pneumonia in the critically ill patient. *Am J Respir Crit Care Med* 159:1249–1256
- Kalbfleisch J, Prentice R (2002) The statistical analysis of failure time data, 2nd edn. John Wiley & Sons, Hoboken
- Keiding N, Filiberti M, Esbjerg S, Robins J, Jacobsen N (1999) The graft versus leukemia effect after bone marrow transplantation: a case study using structural nested failure time models. *Biometrics* 55:23–28
- Kollef M (2003) The importance of appropriate initial antibiotic therapy for hospital-acquired infections. *Am J Med* 115:582–584
- Latouche A, Porcher R, Chevret S (2005) A note on including time-dependent covariate in regression model for competing risks data. *Biometric J* 47:807–814
- Papazian L, Bregeon F, Thirion X (1996) Effect of ventilator-associated pneumonia on mortality and morbidity. *Am J Respir Crit Care Med* 154:91–97
- Pepe M, Mori M (1993) Kaplan–Meier, marginal or conditional probability curves in summarizing competing risks failure time data? *Stat Med* 12:737–751
- Prentice R, Kalbfleisch J, Peterson A, Flournoy N, Farewell V, Breslow N (1978) The analysis of failure times in the presence of competing risks. *Biometrics* 34:541–554
- Resche-Rigon M, Azoulay E, Chevret S (2006) Evaluating mortality in intensive care units: contribution of competing risks analyses. *Crit Care* 10:R5
- Robins J (1986) A new approach to causal inference in mortality studies with sustained exposure periods—application to control of the healthy worker survivor effect. *Math Model* 7:1393–1512
- Robins J, Rotnitzky A (1992) Recovery of information and adjustment for dependent censoring using surrogate markers. *AIDS Epidemiol Methodol Issue* 29:297–331
- Robins J, Hernán M, Brumback B (2000) Marginal structural models and causal inference in epidemiology. *Epidemiology* 11:550–560
- Rubin D (1978) Bayesian inference for causal effects: the role of randomization. *Ann Stat* 6:34–58
- Safdar N, Dezfulian C, Collard H, Saint S (2005) Clinical and economic consequences of ventilator-associated pneumonia: a systematic review. *Crit Care Med* 33:2184–2193
- Schulgen G, Schumacher M (1996) Estimation of prolongation of hospital stay attributable to nosocomial infections: new approaches based on multistate models. *Lifetime Data Anal* 2:219–240
- Schumacher M, Wangler M, Wolkewitz M, Beyersmann J (2007) Attributable mortality due to nosocomial infections: a simple and useful application of multistate models. *Method Inf Med* 46:595–600
- Suetens C, Jans B, Carsauw H (1999) Nosocomial infections in intensive care: results from the Belgian national surveillance—1996 to 1998. *Arch Public Health* 57:221–231
- Suetens C, Savey A, Lepape A, Morales I, Carlet J, Fabry J (2003) Surveillance of nosocomial infections in intensive care units: towards a consensual approach in Europe. *Réanimation* 12:205–213
- Timsit J, Chevret S, Valcke J (1996) Mortality of nosocomial pneumonia in ventilated patients: influence of diagnostic tools. *Am J Respir Crit Care Med* 154:116–123
- Vallés J, Pobo A, García-Esquivel O, Mariscal D, Real J, Fernández R (2007) Excess ICU mortality attributable to ventilator-associated pneumonia: the role of early vs late onset. *Int Care Med* 33:1363–1368
- van der Laan M, Robins J (2003) Unified methods for censored longitudinal data and causality. Springer, New York
- VanderWeele T, Vansteelandt S (2009) Conceptual issues concerning mediation, interventions and composition. *Stat Interface* (in press)
- Van Walraven C, Davis D, Forster A, Wells G (2004) Time-dependent bias was common in survival analyses published in leading clinical journals. *J Clin Epidemiol* 57:672–682

- Vansteelandt S, Mertens K, Suetens C, Goetghebeur E (2009) Marginal structural models for partial exposure regimes. *Biostatistics* 10:46–59
- Vrijens F, Hulstaert F, Gordts B, De Laet C, Devriese S, Van de Sande S, Huybrechts M, Peeters G (2009) Infecties in België, deel II: impact op Mortaliteit en Kosten. Health Services Research (HSR). Tech rep Federaal Kenniscentrum voor de Gezondheidszorg (KCE), Brussel
- Wolkewitz M, Vonberg R, Grundmann H, Beyersmann J, Gastmeier P, Barwolff S, Geffers C, Behnke M, Ruden H, Schumacher M (2008) Risk factors for the development of nosocomial pneumonia and mortality on intensive care units: application of competing risks models. *Crit Care* R44
- Wolkewitz M, Beyersmann J, Gastmeier P, Schumacher M (2009) Modeling the effect of time-dependent exposure on intensive care unit mortality. *Int Care Med* 35:826–832