

Incidence Estimation in Post-ICU Populations: Challenges and Possible Solutions When Using Claims Data

Inzidenzschätzung nach Entlassung von der Intensivstation: Herausforderungen und mögliche Lösungen bei der Verwendung von GKV-Routinedaten



Authors

Magdalena Brandl¹, Christian Apfelbacher^{1,2,3}, Annette Weiß¹, Susanne Brandstetter^{1,4}, Sebastian Edgar Baumeister^{5,6,7}

Affiliations

- 1 Institut für Epidemiologie und Präventivmedizin, Medizinische Soziologie, Universität Regensburg, Regensburg, Germany
- 2 Institut für Sozialmedizin und Gesundheitsökonomie, Otto von Guericke Universität Magdeburg, Magdeburg, Germany
- 3 Lee Kong Chian School of Medicine, Family Medicine and Primary Care, Singapore, Singapore
- 4 Klinik und Poliklinik für Kinder- und Jugendmedizin (KUNO-Kliniken), Universität Regensburg, Regensburg, Germany
- 5 LMU München, Lehrstuhl für Epidemiologie am UNIKA-T, Augsburg, Germany
- 6 Helmholtz Zentrum München Deutsches Forschungszentrum für Umwelt und Gesundheit, Independent Research Group Clinical Epidemiology, Neuherberg, Germany
- 7 Institute for Community Medicine, Greifswald, Universitätsmedizin Greifswald, Germany

Key words

incidence estimation, claims data, competing risks, Post-intensive care syndrome

Schlüsselwörter

competing risks, Post-intensive care syndrome, Inzidenzschätzung, Sekundärdaten

Bibliography

DOI <https://doi.org/10.1055/a-1082-0777>

Online-Publikation: 28.1.2020

Gesundheitswesen 2020; 82 (Suppl. 2): S101–S107

© Georg Thieme Verlag KG Stuttgart · New York

ISSN 0949-7013

Correspondence

Magdalena Brandl
Institut für Epidemiologie und Präventivmedizin,
Medizinische Soziologie, Universität Regensburg,
Dr.-Gessler-Str. 17
93051 Regensburg
Germany
magdalena.brandl@ukr.de

ABSTRACT

Background New or worsening cognitive, physical and/or mental health impairments after acute care for critical illness are referred to as “post-intensive care syndrome” (PICS). Little is known about the incidence of its components, since it is challenging to recruit patients after intensive care unit (ICU) treatment for observational studies. Claims data are particularly suited to achieve incidence estimates in difficult-to-recruit groups. However, some limitations remain when using claims data for empirical research on the outcome of ICU treatment. The objective of this article is to describe three challenges and possible solutions for the estimation of the incidence of PICS based on claims data

Methodological challenges: The presence of competing risk by death, investigating a syndrome and dealing with interval censoring First, in (post) ICU populations the assumption of independence between the event of interest (diagnosis of PICS component) and the competing event (death) is violated. Competing risk is an event whose occurrence precludes the event of interest to be observed, and in ICU populations, death is a frequent secondary event. Methods that estimate incidence in the presence of competing risks are well-established but have not been applied to the scenario described above. Second, PICS is a complex syndrome and represented by various ICD-10 (International Classification of Diseases, 10th Revision) disease codes. The operationalization of this syndrome (case identification) and the validation of cases are particularly challenging.

Third, another major challenge is that the exact date of the event of interest is not available in claims data. It is only known that the event occurred within a certain interval. This feature is called interval censoring. Recently, methods have been developed that address informative censoring due to competing risks in the presence of interval censoring. We will discuss how these methods could be used to tackle the problem when estimating PICS components. Alternatively, it could be possible to assign an exact date for each diagnosis by combining the diagnosis with the exact date of prescriptions of the respective medicines and/or medical services.

Conclusion Estimating incidence in post-ICU populations entails various methodological issues when using claims data. Investigators need to be aware of the presence of competing risks. The application of internal validation criteria to operationalize the event of interest is crucial to achieve reliable incidence estimates. The problem of interval censoring can be solved either by statistical methods or by combining information from different sources.

ZUSAMMENFASSUNG

Hintergrund Neue oder sich verschlechternde kognitive, körperliche und/oder psychische Folgeerkrankungen nach der Behandlung kritischer Erkrankungen auf der Intensivstation (ICU) werden als „Post-Intensiv-Syndrom“ (PICS) bezeichnet. Die Rekrutierung von ehemaligen Intensivpatienten für Beobachtungsstudien gestaltet sich oftmals als schwierig, weshalb kaum Informationen zur Inzidenz einzelner Komponenten dieses Syndroms vorliegen. (GKV-)Routinedaten sind daher besonders gut geeignet, um Inzidenzschätzungen an solchen schwer zu rekrutierenden Gruppen durchzuführen. Allerdings gibt es einige methodische Herausforderungen, die adressiert werden müssen, wenn man empirische Forschung zu Folgen von Intensivaufenthalten auf Basis von Routinedaten durchführen möchte. Ziel dieser Arbeit ist es, 3 wesentliche Herausforderungen und mögliche Lösungen für die Schätzung der Inzidenz von PICS auf Grundlage von GKV-Routinedaten zu beschreiben.

Methodische Herausforderungen Konkurrierendes Risiko durch Versterben, die Untersuchung eines Syndroms und der Umgang mit Intervallzensurierung.

Erstens wird in (Post-)ICU-Populationen die Annahme der Unabhängigkeit zwischen dem interessierenden Ereignis (Diagnose der PICS-Komponente) und dem konkurrierenden Ereignis (Tod) verletzt. Das konkurrierende Risiko (competing risk) ist ein Ereignis, dessen Auftreten das zu beobachtende Ereignis von Interesse ausschließt. Das konkurrierende Ereignis bleibt häufig unberücksichtigt, aber in ICU-Populationen ist das Versterben ein häufiges sekundäres Ereignis. Methoden zur Schätzung der Inzidenz bei konkurrierenden Risiken sind gut etabliert, wurden aber nicht auf das oben beschriebene Szenario angewendet.

Zweitens ist PICS ein komplexes Syndrom und kann durch verschiedene ICD-10-Codes (Internationale Klassifikation der Krankheiten, 10. Revision) abgebildet werden. Die Operationalisierung dieses Syndroms (Fallidentifizierung) und die Validierung von Fall-Definitionen sind besonders herausfordernd.

Die dritte große Herausforderung ist die fehlende Dokumentation des genauen Datums des interessierenden Ereignisses in den GKV-Routinedaten. Die vorhandenen Daten geben nur Aufschluss darüber, ob das Ereignis innerhalb eines bestimmten Zeitraums eingetreten ist. Diese Eigenschaft wird als Intervallzensurierung bezeichnet. Seit kurzem gibt es Methoden, die eine informative Zensurierung aufgrund konkurrierender Risiken bei gleichzeitiger Intervallzensurierung berücksichtigen. Eine weitere Alternative besteht darin die Zuordnung eines genauen Datums zu jeder PICS-Diagnose mithilfe von Verwaltungsdaten zu kombinieren, denn für die Verschreibung eines entsprechenden Medikamentes und/oder medizinischer Leistungen ist ein genaues Datum hinterlegt.

Schlussfolgerung Die Inzidenzschätzung in Post-ICU-Populationen bringt verschiedene methodische Probleme bei der Verwendung von GKV-Routinedaten mit sich. Wissenschaftlern muss das Vorhandensein von konkurrierenden Risiken bewusst sein. Die Anwendung interner Validierungskriterien zur Operationalisierung des interessierenden Ereignisses ist entscheidend, um zuverlässige Inzidenzschätzungen zu erhalten. Das Problem der Intervallzensurierung kann entweder durch statistische Methoden oder durch die Kombination von Informationen aus anderen Quellen gelöst werden.

Introduction

Treatment during intensive care has become more effective over the past century. This resulted in a growing population of survivors of critical care treatment [1]. Consequently, the number of patients suffering from long-term impairments associated with intensive care unit (ICU) treatment is increasing [2, 3]. In 2012, these “new or worsening impairments in physical, cognitive, or mental health status arising after critical illness and persisting beyond acute care hospitalization” were summarized as post-intensive care syndrome (PICS) [4]. Approximately half of patients after ICU treatment suffer from some component of PICS [5]. For instance, a systematic

review of posttraumatic stress disorder (PTSD) in ICU survivors reported a prevalence of clinically diagnosed PTSD of 19% [6]. Despite the growing awareness of PICS [7], epidemiological data are scarce and inconsistent [8], especially in Germany.

Valid estimation of the incidence of components of PICS is challenging, as it might be difficult to recruit and interview critically ill patients, especially a large and representative sample. Consequently, healthier ICU survivors will be overrepresented in the resulting data. In addition, there will be a large attrition because of the high mortality among ICU survivors [9]. Claims data from statutory health insurances (SHI) are well suited to overcome, at least in part,

some of these limitations [10]. The most valuable features of health claims data are the availability of information about a large number of persons at little cost to the researcher and the absence of response and recall bias. Another briefly mentioned potential bias is selection bias. Incidence estimates derived from one sickness fund might not be generalizable due to structural differences between funds [11]. One approach to address this selection bias is using auxiliary data to post-stratify incidence estimates [12]. However, this strategy only corrects for differences in the age, sex and regional distributions but not for lifestyle and morbidity related factors that are related to the variables of interest (ICU stay, PICS) but that are not available in official statistics [11].

For PICS, the most valuable feature of claims data is that claims data delivers information on critically ill patients where the patient can hardly be interviewed because of his/her condition. However, other limitations remain when using claims data for empirical research on the outcome of ICU treatment. To illustrate the methodological challenges and potential solutions described in this paper, we give an example of a study that sought to estimate the incidence of pneumonia in older inhabitants of nursing homes in Germany using claims data [13]. Patients were considered right censored due to end of insurance or death, and standard methodology (Kaplan-Meier-Estimation) was used to estimate the incidence of pneumonia. Thus, it was assumed that the risk was equal in censored and uncensored individuals over time. However, the cumulative incidence of pneumonia would be underestimated, if the risk of pneumonia were higher for censored than for non-censored observations over time, and the study population under risk would include a progressively greater proportion of lower risk subjects.

In summary, if censoring is informative the standard methods can provide biased estimates. Because of the high mortality in post-ICU populations, PICS is subject to informative censoring and the event death would be seen as competing risk in this case. A competing risk is an event that either hinders the observation of the event of interest or modifies the chance that this event occurs. ICU survivors have a 1-year mortality post ICU ranging from 16 to 44 % [14]. Even 5-year post ICU mortality is still higher (32 %) than in persons who were discharged from hospital but had no ICU stay (22 %) [9]. Therefore, it is not possible to apply standard methods like the Kaplan-Meier (KM) method [15] to calculate the incidence of PICS in ICU survivors.

For this reason, we provide an explanation why methods accounting for the presence of competing risks, like the cumulative incidence function, should be used instead of classical methods like the KM estimator. We also show that the application of these advanced methods entails further challenges, namely possible bias when the wrong case definition is used and/or when information on exact dates is missing. Therefore, the objective of this article is to describe three challenges and possible solutions for the estimation of the incidence of PICS based on claims data.

Methodological challenges when investigating a syndrome in a cohort of ICU survivors in claims data

Challenge I: Right censoring and the competing risk “death”

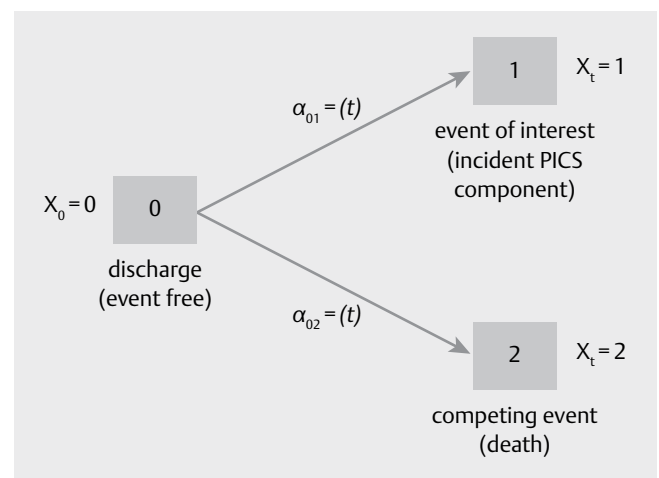
Incidence is measured either as incidence proportion (number of new cases during a specific period of time / number of patients under risk during that time period) or incidence rate (number of new cases

during a specific period of time / total number of person time) [16]. These measures differ in their type of denominators, which describe the population under risk. The incidence proportion uses a period of time during which all persons at risk were observed, whereas for the incidence rate, persons at risk were observed for different time periods and the denominator describes the sum of the time periods that each person was at risk and observed (person-time). But, sometimes, not every person has been followed for the full study period. This condition is called right censoring. Right censoring occurs, for instance, when the study ends prematurely before the event of interest occurs (administrative censoring). For example, when a person switched the insurer during the study period (loss to follow-up), or when a competing event occurs (e. g. death) [17].

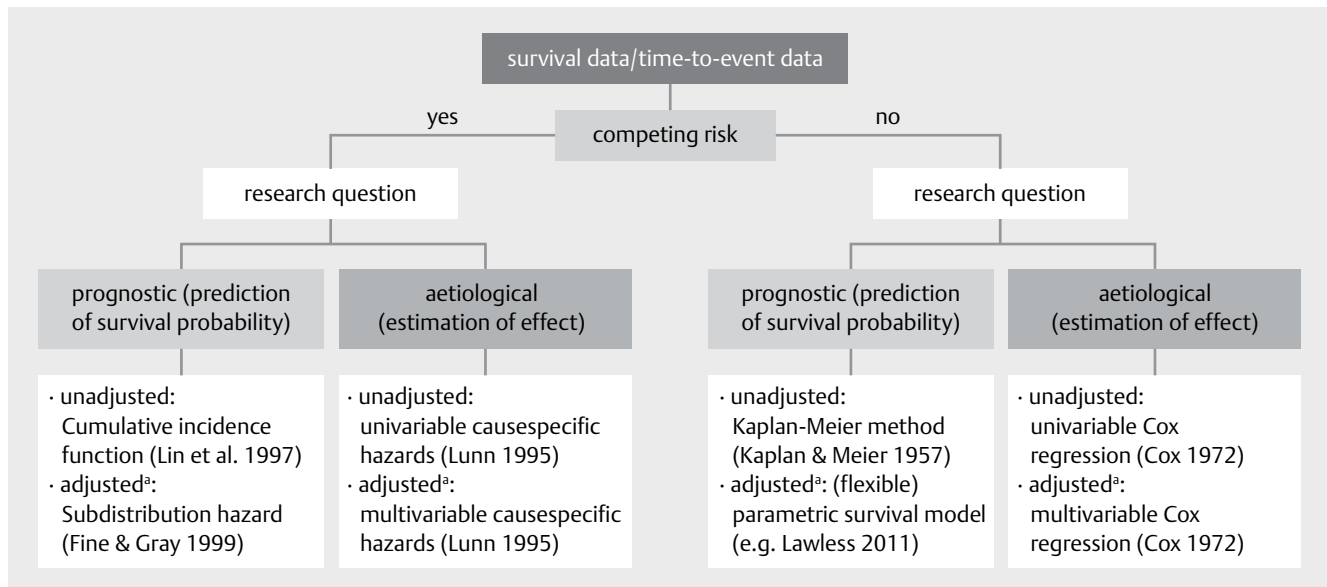
If censored observations have the same probability of the event after censoring as those remaining under risk, it is called noninformative right censoring. In populations with high mortality, this independence between censoring and time-to-event distributions cannot be assumed [18]. This type of censoring is then called informative right censoring. The risk of bias is even larger in populations with persons at high risk of experiencing a competing event.

In our population of ICU survivors, informative right censoring due to death is a major issue as the mortality remains high even many years after ICU [9]. In general, competing risks are present when an individual can experience more than one type of event and the occurrence of one type of event prevents any other event from ever happening [19]. If the primary outcome is PICS, death from other causes serves as the competing risk because death before occurrence of the PICS precludes the latter event. Therefore, death must be considered as competing event to the occurrence of PICS components because persons who died were probably sicker and had an even higher risk to develop PICS.

In general, the term ‘time-to-event data’ is used in the medical literature to refer to data in which the outcome denotes time to the occurrence of an event of interest. The KM method for estimating survival functions and the Cox proportional hazards model for estimating the relative effect of covariates in the hazard of occurrence of the event are commonly used for the analysis of survival data (► Fig. 2). Time-to-event data are generally characterized by cen-



► Fig. 1 Competing risks multistate model. Competing risks process with cause-specific hazards α_{0j} , $j = 1, 2$.



► **Fig. 2** Overview of methods to use for time-to-event analysis in presence and absence of competing risks for prognostic or aetiological research questions. ^aregression adjustment for confounding

soring, i. e. the timing of the event of censored observations is unknown. The KM method requires that censoring does not affect potential failure times (= time to event). It assumes that censoring is independent or noninformative [20]. However, this assumption is often not met when failure can occur for more than one reason.

Competing risk is a special case of multistate models in which each of the different events are absorbing states (► **Fig. 1**) [21]. Although a competing risk analysis may include several types of competing events, for the sake of simplicity, we focus on only one type of competing event in this article. There are two competing absorbing states, which represent the possible event types. Occurrence of an event is modeled by a transition from the initial state (discharge from ICU) into the corresponding event state.

Standard methods for time-to-event data typically censor subjects when a competing risk occurs, and assume that competing risks are absent. For example, when the primary outcome is time to post-traumatic stress disorder, the primary outcome is censored when a study participant dies. However, this might violate the assumption of noninformative censoring because deceased subjects cannot be adequately represented by uncensored subjects.

When there are no competing risks, the KM method estimates the survival function and its complement provides an estimate of the cumulative incidence of an event over time. However, when the KM estimator is applied in the presence of competing risk then there is an upward bias to be expected in the incidence. Instead, when estimating the incidence of an outcome in the presence of competing risks, suitable estimators need to be chosen. Applying competing risks theory allows us to connect the two initially described measures of incidence.

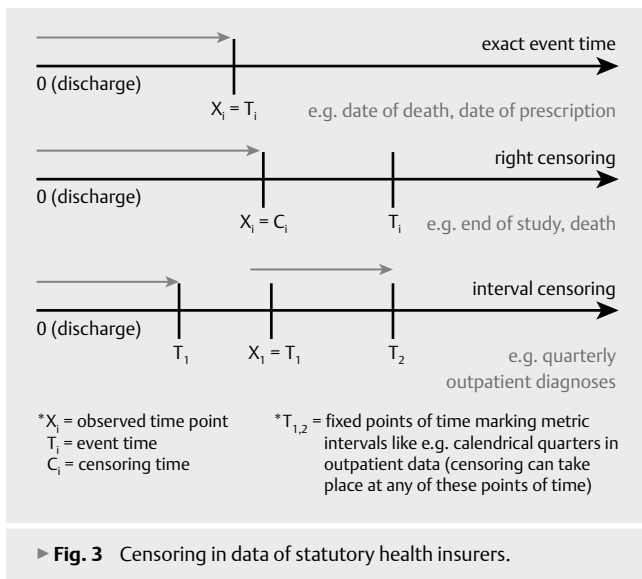
Solution I: Applying competing risks theory

The most important measures for competing risk data are the cumulative incidence function (CIF) [22, 23] and the cause-specific hazard rate [24] (► **Fig. 2**). The CIF is used for the estimation of the

occurrence of an event (PICS) while accounting for competing risks (death). The CIF for each competing risk k gives the probability, as a function of time, that an event occurs in the presence of the other competing risks. The CIF is defined as $P(T \leq t | \text{failure from cause } k)$, for T equal to the time to first failure from cause k , and denotes the probability of event k before time t and before occurrence of a different type of event. Unlike using the complement of the KM estimator, the sum of the CIFs will equal the CIF estimate of the incidence of the composite outcome (PICS or death), defined as any of the event types. Therefore, the CIF is not biased upward like the KM, as the estimator is lowered by the occurrence of the competing event (dead patients can no more experience PICS). This results from the fact that the conditional probability of the event of interest at time t is multiplied by the cumulative survival of any event. The cause-specific hazard is the instantaneous risk of failure from a specified cause given that no failure from any cause has yet occurred and derived using

$$\lim_{\Delta t \rightarrow 0} \frac{p(t \leq T < t + \Delta t, \text{ failure from cause } k | T \geq t)}{\Delta t}$$

The CIF, and its extensions to the multivariable setting, are suitable in situations where we are interested in incidence estimation and for prognostic reasons (e. g., what is an individual's probability of experiencing a PICS within 12 month after discharge?, ► **Fig. 2**). Cause-specific hazards are more appropriate for addressing aetiological questions (e. g., is ICU associated with the rate of occurrence of a diagnosis?) [25, 26]. The CIF can be estimated in R (cmprsk), SAS (%CIF), and Stata (stcompet). The cause-specific hazard can be estimated by treating events due to competing causes as censored observations in a standard Cox model. The k cause-specific hazards can be derived by fitting k separate models by stacking the events (having k rows per individual) and fitting a Cox model stratified by cause [27]. Covariates can be easily incorporated into a



cause-specific Cox model. Multivariable models for the CIF can be fit on the subhazard scale, where the subhazard is the CIF transformed to the hazard scale, using a Fine-Gray model [28]. Subdistribution hazard models are available in R (cmprsk), SAS (PHREG), and Stata (stcrreg). These methods are described in detail elsewhere [29, 30], for simplicity reasons they are beyond the scope of this paper. ► **Fig. 2** provides a decision tree on how to select a suitable time-to-event method for individual patient data (IPD).

When analyses of prospective individual level claims data are not feasible (e.g. due to cost restraints or data protection issues), an alternative strategy is to estimate the incidence based on age-specific cross-sectional prevalence and mortality data [31, 32]. One of these incidence estimation techniques makes use of the theoretical relationship between age-specific prevalence, incidence and mortality data to estimate the incidence from current status data to solve an ordinary differential equation. Details on the method are provided in [33].

No matter which approach is chosen, information about the type of event and the exact time to the occurrence of this event must be available. In order to get this information two further challenges emerge in claims data: first, how to operationalize the event of interest (challenge II) and second, how to define the time to event (challenge III).

Challenge II: Operationalization of PICS in claims data

Operationalization of the event of interest (case identification) in claims data requires two steps: In a first step, cases must be comprehensively defined based on the disease under study. A 3-digit ICD-10 (International Classification of Diseases, 10th Revision) code is usually used to specify the diagnosis. To complete the case definition, the possible sectors (e.g. inpatient, outpatient, or incapacity for work data) in which these codes were documented must be defined. In a second step, we have to validate the case definition. Physicians in Germany code outpatient diagnoses as “ruled out”, “asymptomatic”, “suspected”, and “confirmed”. In most cases, it is more reliable to use “confirmed” outpatient diagnoses. Inpatient diagnoses are divided into a (main) discharge diagnosis and sec-

ondary diagnoses. Internal validation is strongly recommended for diagnoses from both sectors [34].

Applying the criterion of “at least two quarters” (M2Q), is probably one of the best-known validation approaches to reflect persistence of diagnosis. This criterion requires a confirmative diagnosis in at least one further quarter or a second diagnosis by another physician in the same quarter [35, 36].

Both, the case definition and the validation, are particularly challenging in an ICU survivor population for several reasons. First, no agreement on a list of ICD-codes exists to define the PICS and, additionally, the diseases included in the PICS syndrome are not generally agreed by experts [7]. Second, all health care sectors are of interest to assess the incidence of the diseases after ICU stay. Many patients are admitted to a rehabilitative care after discharge from ICU. In terms of rehabilitation stays, limited information is available in the data from the SHI. Third, internal validation strategies, which are based on measures of persistence, might not be the most suitable approach for a patient group that has a high lethality. Thus, applying M2Q to sicker patients who are censored because they died will result in downwardly biased incidence estimates.

Solution II: Reporting single components of PICS that were validated with measures of congruence

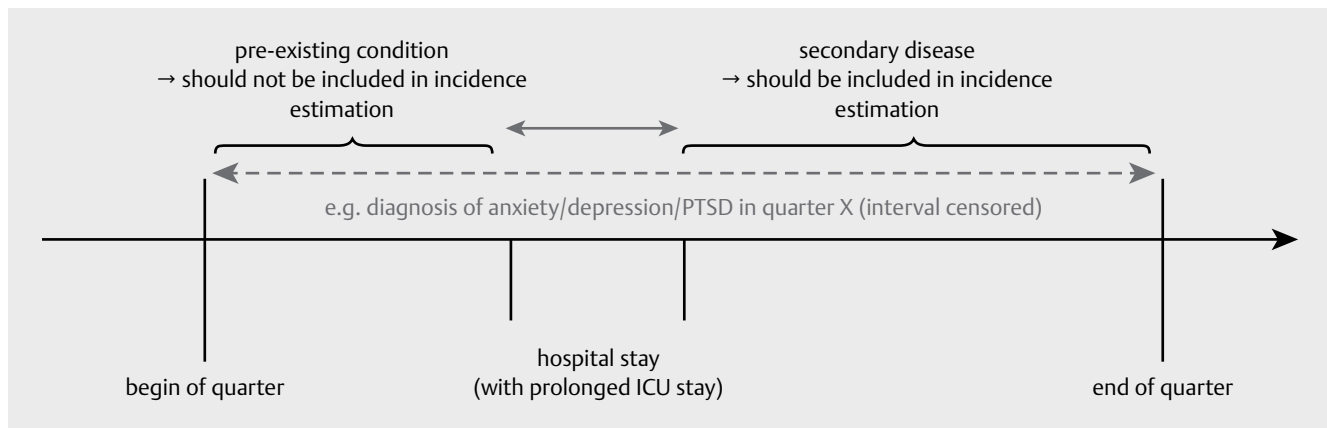
Until now, there is neither consensus on the definition of the PICS nor is there a list of ICD-10 codes available to define specific components of the PICS. In this situation, we propose to apply a case-identifying algorithm that is based on ICD-10 codes, which reflect the PICS and capture all three domains covered by PICS: physical, mental and cognitive impairments [4]. The current selection of ICD-10 codes, which represent PICS components, should be based on literature [4, 37, 38] and be informed by expert consultation with intensive care clinicians.

Because M2Q like other criteria of persistence are not applicable, the criteria of congruence can be used to validate (confirmed) diagnoses. As confirmed diagnoses alone can include false positive diagnoses [34], the second approach is to use information on prescriptions (e.g. ATC (Anatomical Therapeutic Chemical) Classification System codes or EBM-Codes (Einheitlicher Bewertungsmaßstab, meaning outpatient physician’s service codes of the German outpatient fee-for-service reimbursement system)) as a measure of congruence. For instance, a diagnosis of dementia can be assumed when a patient is diagnosed with one of the respective ICD-10 codes in the inpatient or outpatient sector and when a prescription of antedementiva ATC code is also reported.

In summary, depending on the respective goal, two possible ways are plausible to define cases: we can apply more sensitive, but less restrictive selection criteria to ensure full coverage of true positive cases or we use less sensitive, but more restrictive criteria to avoid the inclusion of false positives. Previous studies countered this issue by providing results with different case definitions in sensitivity analyses [35, 39].

Challenge III: Interval Censoring

The data of SHI are also affected by interval censoring (► **Fig. 3**). Interval censoring describes that the exact time of diagnosis is unknown, it is only known to lie within an interval. In outpatient claims data, which are forwarded on a quarterly basis no exact date for



► **Fig. 4** Content related consequences of interval censoring.

outpatient diagnoses is known. Likewise, for inpatient diagnoses only the date of admission and discharge from the hospital is known. More generally, data of this type typically arise in studies where follow up is done at fixed intervals. In the context of interval censoring, many researchers use imputation techniques, especially mid-point or right-point imputation (i. e., replacing an interval censored observation by its right endpoint), and then apply standard techniques for right censored data [40]. However, these approaches may not be appropriate. Imputation approaches propose an estimate for an unknown distribution, and obtain biased estimates and standard errors, especially if the intervals are wide, by falsely assuming that the time of event is equal to the right-point or mid-point of the time interval [41]. Special techniques for interval-censored data should therefore be preferred, as will be seen later in this section. Alternatively, the underlying continuous durations can be conceptualized as a latent variable with observed disjoint (discrete) time intervals [42]. A related methodological issue arises when the time of the study entry was date of hospital discharge, and a patient's hospital stay took place in one quarter. In this case, it is unknown whether an outpatient diagnosis documented per quarter was coded before or after the hospital stay (see ► **Fig. 4**).

Solution III: Approximation of the exact date

A conceivable solution might be to associate an existing date with a diagnosis and use this date as the diagnostic date and apply standard techniques for right-censored data. For this purpose, we could use the exact date of an outpatient physician visit or use the date of a PIC-specific procedure that was coded using the uniform evaluation scale (EBM, "Einheitlicher Bewertungsmaßstab"), an exact prescription date of a medication or of a non-pharmaceutical therapy or technical aid in the data of the SHI. In case of less disease specific medications (such as pain killers) the date of medication cannot be used to approximate the date of diagnosis. In this scenario, dates of physician visits can be considered. Dates attached to physician visits are not as specific as prescription dates, as patients may have several encounters with an outpatient care provider during one quarter. In this case, we consider the first and last visit of the patient during a quarter and code the middle of this period to further limit the period of diagnoses. Certain outcomes

could possibly be indicated by a physician visit with a specific EBM code. If so, these outcomes' incidence could be measured by the exact date of the physician visit with the delivered EBM-code and not only by the time frame spanned by multiple physician visits. There are new statistical methods allowing to model competing risk data in combination with interval censoring [43]. These techniques can provide more exact dates for a relevant part of the sample. However, the allocation of the diagnosis within the quarter of the hospital stay is still not satisfactorily solved.

Conclusion

A better understanding of patients' health after ICU is urgently needed. Claims data are a valuable data source for advancing the scientific knowledge about PICS. However, various methodological issues have to be accounted for when analyzing health claims data of post-ICU populations. Investigators need to be aware of the presence of competing risks when performing time-to-event analysis in studies with highly morbid patient groups. The internal validation of diagnoses in health claims data, and techniques to impute or approximate exact dates of diagnoses are crucial to achieve reliable estimates of frequency. Incorrectly treating competing events as censored events may have practical implications for patient care or health care planning.

To minimize bias when estimating incidences, investigators can use the cumulative incidence function or cause-specific hazards to estimate the incidence of PICS or other diseases in highly morbid patient groups subject to competing risk.

Funding

The project on which this publication is based was funded by the innovation fund of the Joint Federal Committee (Gemeinsamer Bundesausschuss, G-BA). Funding code: 01VSF16056.

Conflict of Interest

The authors declare that they have no conflict of interest.

References

- [1] Wunsch H, Guerra C, Barnato AE et al. Three-year outcomes for Medicare beneficiaries who survive intensive care. *Jama* 2010; 303: 849–856
- [2] Adhikari NKJ, Fowler RA, Bhagwanjee S et al. Critical care and the global burden of critical illness in adults. *Lancet* 2010; 376: 1339–1346
- [3] Vincent J-L, Marshall JC, Namendys-Silva SA et al. Assessment of the worldwide burden of critical illness: The Intensive Care Over Nations (ICON) audit. *Lancet Respir Med* 2014; 2: 380–386
- [4] Needham DM, Davidson J, Cohen H et al. Improving long-term outcomes after discharge from intensive care unit: Report from a stakeholders' conference. *Crit Care Med* 2012; 40: 502–509
- [5] Marra A, Pandharipande PP, Girard TD et al. Co-Occurrence of post-intensive care syndrome problems among 406 survivors of critical illness. *Crit Care Med* 2018; 46: 1393–1401
- [6] Davydow DS, Gifford JM, Desai SV et al. Posttraumatic stress disorder in general intensive care unit survivors: A systematic review. *Gen Hosp Psychiatry* 2008; 30: 421–434
- [7] Dodoo-Schittko F, Brandstetter S, Apfelbacher C et al. Folgen kritischer Erkrankung und mögliche Interventionen. *Anästhesiol Intensivmed Notfallmed Schmerzther* 2017; 52: 137–144
- [8] Harvey MA, Davidson JE. Postintensive Care Syndrome: Right Care, Right Now... and Later. *Crit Care Med* 2016; 44: 381–385
- [9] Lone NI, Gillies MA, Haddow C et al. Five-Year mortality and hospital costs associated with surviving intensive care. *Am J Respir Crit Care Med* 2016; 194: 198–208
- [10] Swart E, Ihle P, Gothe H et al. Routinedaten im Gesundheitswesen. Verlag Hans Huber; Hogrefe AG: 2014
- [11] Hoffmann F, Koller D. Different regions, differently insured populations? Socio-demographic and health-related differences between insurance funds. *Gesundheitswesen* 2017; 79: e1–e9
- [12] Smith TM.. Post-stratification. *J Royal Stat Soc* 1991; 40: 315–323
- [13] Fassmer AM, Spreckelsen O, Hoffmann F. Incidence of pneumonia in nursing home residents in Germany: Results of a claims data analysis. *Epidemiol Infect* 2018; 146: 1123–1129
- [14] Brinkman S, Bakhshi-Raiez F, Abu-Hanna A et al. Determinants of mortality after hospital discharge in ICU patients: literature review and Dutch cohort study. *Crit Care Med* 2013; 41: 1237–1251
- [15] Kleinbaum DG, Klein M. *Survival Analysis*. Springer; 2010
- [16] Beyersmann J, Gastmeier P, Schumacher M. Incidence in ICU populations: How to measure and report it? *Intensive Care Med* 2014; 40: 871–876
- [17] Putter H, Fiocco M, Geskus RB. Tutorial in biostatistics: Competing risks and multi-state models. *Stat Med* 2007; 26: 2389–2430
- [18] Andersen PK, Geskus RB, de Witte T et al. Competing risks in epidemiology: Possibilities and pitfalls. *Int J Epidemiol* 2012; 41: 861–870
- [19] Gichangi A, Vach W. The analysis of competing risks data: A guided tour. *Stat Med*. 2005; 132: 1–41
- [20] Kalbfleisch J, Prentice R. *Competing risks and multistate models. The statistical analysis of failure time data* 2002; 247: 277
- [21] Andersen PK, Abildstrom SZ, Rosthøj S. Competing risks as a multi-state model. *Statistical methods in medical research* 2002; 11: 203–215
- [22] Lin D. Non-Parametric inference for cumulative incidence functions in competing risks studies. *Stat Med* 1997; 16: 901–910
- [23] Pintilie M. *Competing risks: A practical perspective*. John Wiley & Sons; 2006
- [24] Lunn M, McNeil D. Applying Cox regression to competing risks. *Biometrics* 1995; 524–532
- [25] Latouche A, Allignol A, Beyersmann J et al. A competing risks analysis should report results on all cause-specific hazards and cumulative incidence functions. *Journal of Clinical Epidemiology* 2013; 66: 648–653
- [26] Lau B, Cole SR, Gange SJ. Competing risk regression models for epidemiologic data. *Am J Epidemiol* 2009; 170: 244–256
- [27] Lunn M, McNeil D. Applying Cox regression to competing risks. *Biometrics* 1995; 51: 524–532
- [28] Fine JP, Gray RJ. A Proportional Hazards Model for the Subdistribution of a Competing Risk. *J Am Stat Assoc* 1999; 94: 496–509
- [29] Verduijn M, Grootendorst DC, Dekker FW et al. The analysis of competing events like cause-specific mortality – beware of the Kaplan-Meier method. *Nephrology, dialysis, transplantation: Official Publication of the European Dialysis and Transplant Association - European Renal Association* 2011; 26: 56–61
- [30] Austin PC, Lee DS, Fine JP. Introduction to the Analysis of Survival Data in the Presence of Competing Risks. *Circulation* 2016; 133: 601–609
- [31] Brinks R, Hoyer A, Weber S et al. Age-specific and sex-specific incidence of systemic lupus erythematosus: An estimate from cross-sectional claims data of 2.3 million people in the German statutory health insurance 2002. *Lupus Sci Med* 2016; 3: e000181
- [32] Sewerin P, Brinks R, Schneider M et al. Prevalence and incidence of psoriasis and psoriatic arthritis. *Ann Rheum Dis* 2019; 78: 286–287
- [33] Landwehr S, Brinks R. A comparative study of prevalence-based incidence estimation techniques with application to dementia data in Germany. *Stat Med* 2016; 35: 768–781
- [34] Hoffmann F, Andersohn F, Giersiepen K et al. [Validation of secondary data. Strengths and limitations]. *Bundesgesundheitsblatt, Gesundheitsforschung, Gesundheitsschutz* 2008; 51: 1118–1126
- [35] Nerius M, Fink A, Doblhammer G. Parkinson's disease in Germany: prevalence and incidence based on health claims data. *Acta Neurol Scand* 2017; 136: 386–392
- [36] Schubert I, Ihle P, Koster I. [Internal confirmation of diagnoses in routine statutory health insurance data: concept with examples and case definitions]. *Gesundheitswesen* 2010; 72: 316–322
- [37] Bramesfeld A, Grobe TG, Schwartz FW. Who is diagnosed as suffering from depression in the German statutory health care system? An analysis of health insurance data. *Eur J Epidemiol* 2007; 22: 397–403
- [38] Ziegler U, Doblhammer G. [Prevalence and incidence of dementia in Germany – a study based on data from the public sick funds in 2002]. *Gesundheitswesen* 2009; 71: 281–290
- [39] Schmedt N, Khil L, Berger K et al. Incidence of Multiple Sclerosis in Germany: A Cohort Study Applying Different Case Definitions Based on Claims Data. *Neuroepidemiology* 2017; 49: 91–98
- [40] Collett D. *Modelling survival data in medical research*: Chapman and Hall/CRC. 2015;
- [41] Leung KM, Elashoff RM, Afifi AA. Censoring issues in survival analysis. *Annu Rev Public Health* 1997; 18: 83–104
- [42] Singer JD, Willett JB. *Applied longitudinal data analysis: Modeling change and event occurrence*. Oxford university press; 2003
- [43] Bakoyannis G, Yu M, Yiannoutsos CT. Semiparametric regression on cumulative incidence function with interval-censored competing risks data. *Stat Med* 2017; 36: 3683–3707