Annals of Epidemiology 24 (2014) 593–597

Contents lists available at ScienceDirect

Annals of Epidemiology

journal homepage: www.annalsofepidemiology.org

Original article

Positive predictive value of primary inpatient discharge diagnoses of infection among cancer patients in the Danish National Registry of Patients

Louise Holland-Bill MD [a,*], Hairong Xu MD, PhD [b], Henrik Toft Sørensen MD, PhD, DMSc [a], John Acquavella PhD [b], Claus Sværke MSc [a], Henrik Gammelager MD [a,c], Vera Ehrenstein MPH, DSc [a]

[a] Department of Clinical Epidemiology, Aarhus University Hospital, Aarhus, Denmark
[b] Center for Observational Research, Amgen Inc., Thousand Oaks, CA
[c] Department of Clinical Epidemiology, Aalborg University Hospital, Aalborg, Denmark

ARTICLE INFO

Article history:
Received 19 December 2013
Accepted 20 May 2014
Available online 24 May 2014

Keywords:
Cancer
Infection
International Classification of Disease codes
Predictive value of tests
Pharmacoepidemiology

ABSTRACT

Purpose: Pharmacovigilance studies of cancer treatment frequently monitor infections. Predictive values of algorithms identifying disease depend on prevalence of the disease in the population under study. We therefore estimated the positive predictive value (PPV) of primary inpatient diagnosis of infection among cancer patients in the Danish National Registry of Patients (DNRP).
Methods: The algorithm to identify infections in the DNPR was based on *International Classification of Diseases*, 10th revision (ICD-10) codes. A physician blinded to the type of sampled infection reviewed the medical charts and assessed the presence and type of infection. Using the physician global assessment as gold standard, we computed PPVs with and without requiring agreement on infection type.
Results: We retrieved 266 of 272 medical charts (98%). Presence of infection was confirmed in 261 patients, resulting in an overall PPV of 98% (95% confidence interval, 96%–99%). When requiring agreement on infection type, overall PPV was 77%. For skin infections, pneumonia, and sepsis, PPVs were 79%, 93% and 84%, respectively. For these infections, we additionally calculated PPVs using evidence-based criteria as the gold standard. PPV was similar for pneumonia, but lower for skin infections and sepsis.
Conclusions: The Danish National Registry of Patients is suitable for monitoring infections requiring hospitalization among cancer patients.

Background

Infections frequently complicate cancer and cancer-related therapies and are therefore often examined in pharmacovigilance studies of cancer treatments. Medical databases are commonly used to monitor adverse drug reactions occurring in routine clinical practice [1]. The Danish National Registry of Patients (DNRP) is a potentially valuable data source for monitoring such events. It covers all in-hospital stays, outpatient clinic visits, and emergency contacts in the setting of Denmark's universal health care [2]. However, as the DNRP was not primarily developed to support specific studies or to monitor specific events, the validity of adverse events recorded in the DNRP requires assessment [3].

Previous studies evaluating the positive predictive value (PPV) of codes for infections in the *International Classification of Diseases*, ninth revision (ICD-9) using administrative data have found highly variable results, ranging from 1.3% to 100% depending on type of infection and population examined [4]. In Denmark, two studies estimated PPVs of infections recorded in the DNRP based on the *International Classification of Diseases*, 10th revision (ICD-10) [5,6]. First, a study from 1998 found a low PPV of codes for bacteremia (septicemia) compared against a database of all positive blood cultures from a clinical microbiology department and medical chart review [5]. In contrast, among patients hospitalized for pneumonia in 1994–2003, a high PPV was found for ICD-10 codes of pneumonia [6]. Predictive values of algorithms for identifying disease depend on the prevalence of the disease in the population under study and must therefore be established for each population of interest [7]. To date, no study has examined the validity of discharge diagnoses recorded in the DNRP registries for identifying infectious conditions among cancer patients.

We estimated the PPV of the ICD-10 codes used to record primary inpatient discharge diagnoses of infection in the DNRP among patients with a history of cancer. We estimated PPVs both

* Corresponding author. Department of Clinical Epidemiology, Aarhus University Hospital, Olof Palmes Allé 43-45, 8200 Aarhus N, Denmark. Tel.: +45 871 68254; fax: +45 871 67215.
E-mail address: louise.bill@dce.au.dk (L. Holland-Bill).

1047-2797/$ – see front matter © 2014 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.annepidem.2014.05.011

regardless of infection type and requiring agreement by type of infection. Furthermore, we estimated the PPV of algorithms to identify three infections of special interest: skin infection, pneumonia, and sepsis. The study was performed to validate the use of hospital discharge diagnoses in future pharmacoepidemiologic studies with a primary focus on hospitalized infections endpoint in patients with solid malignant tumors [4,8].

Methods

Setting

The Danish National Health Service provides tax-supported medical care for the entire Danish population, guaranteeing unfettered access to public hospitals free of charge [9]. Since 1968, the Danish Civil Registration System has assigned a unique 10-digit Civil Personal Registration (CPR) number to each person born in or immigrating to Denmark. This number, which encodes date of birth and sex, is used in all public records and allows unambiguous individual-level linkage between all Danish medical databases and virtually complete follow-up of patients receiving care from the Danish National Health Service. The Danish Civil Registration System is updated daily and tracks residence, migration, and vital status of all Danish residents [9]. The DNRP contains records of all inpatient stays to Danish hospitals since 1977, and all contacts with emergency departments and outpatient clinics since 1995. Each record in the DNRP contains the patient's CPR number, dates of admission and discharge, one primary diagnosis, reflecting the main reason for the hospital contact, and up to 19 secondary diagnoses, coded according to the ICD-10 from 1994 onward [2,10]. Diagnoses are assigned by the attending physician at the time of hospital discharge. Data are initially recorded in the hospitals' patient administrative system and then electronically transferred to the DNPR. Reporting to the DNPR is mandatory, and DNPR data provide the basis for financial reimbursement to hospitals from the government [9].

Study population

We used the DNRP to identify all patients with an inpatient hospital stay at Aalborg University Hospital, who had a primary discharge diagnosis of preselected infectious conditions between January 1, 2006 and December 31, 2010, inclusive. Aalborg University Hospital is the main provider of secondary care—including cancer care—to residents of the North Denmark Region with a population of 580,000 inhabitants [11]. We restricted the study population to patients with a history of a solid malignancy (ICD-10 codes C00–C79, excluding nonmelanoma skin cancer C44) within 5 years before the admission with an infection. For each patient, we considered only the first inpatient hospital stay with a primary diagnosis of infection during the study period. Sampling was restricted to patients older than 18 years at the time of cancer diagnosis.

We planned to sample at least 250 patients with a diagnosis of any infection, with additional sampling allowed to ensure at least 40 observations of each of the potential infections of special interest: pneumonia, skin infection, and sepsis. The ICD-10 codes used for the case-ascertainment algorithms for infections are listed in Appendix A.

Medical chart review

We used medical chart review to obtain information on presence and type of infections [3]. We used each patient's CPR number to link between registry data and the medical chart [9]. The identified charts were abstracted by one of two study physicians using a standard Medical Chart Abstraction Form (Appendix B). The form was used to collect information on evidence-based criteria for the infections of special interest, as outlined in the validation study in a US Veterans Affair hospital setting by Schneeweiss et al. [12] based on published guidelines. Based on the information recorded in the chart, the reviewing physician assessed the presence and type of infection and documented it in the abstraction form as a physician global assessment (PGA) (Appendix B). If charts were unavailable, this was noted and reported. The reviewers were blinded to type of infection and to the evidence-based diagnostic criteria for given infections. The evidence-based diagnostic criteria are described in Appendix C. The abstraction form was tested for comprehensibility, validity of questions, and reproducibility of information extracted using a random sample of eight medical charts. Possible sources of error and discrepancy were identified and corrected in the abstraction form before initiation of chart review.

Statistical analysis

We described patients with available charts in terms of sampled ICD-10 codes, sex, age at diagnosis of infection, length of hospital stay, and comorbidity level (using the Charlson Comorbidity Index score excluding the patient's cancer diagnosis) [13]. We extracted data on comorbidities from the DNRP for up to 5 years before the infection-related admission.

The overall PPV was calculated as the proportion of patients for whom a primary discharge diagnosis of infection was confirmed by medical record review using the PGA as the gold standard, with and without requiring agreement by type of infection. When agreement on type of infection was required, we applied both specific and broad definition, as listed in Table 1 .

For the infections of special interest, we additionally computed the PPV using as the gold standard information on evidence-based criteria abstracted from the chart. All PPV estimates were reported with 95% confidence intervals (CI) using Jeffrey's method for binomial proportions.

The overall sample size was targeted to include 250 randomly selected patients with a primary discharge diagnosis of infection to allow estimating 95% CI limits within ±6%. It was planned to include a minimum of 40 patients for each specific infectious condition of interest (pneumonia, skin infection, and sepsis) to allow estimating 95% CI limits within ±14%, which is considered ample to assess the PPV for a single infection.

The study was approved by the Danish Data Protection Agency (record numbers 2010-41-5171, 2012-41-0045).

Results

We sampled 272 potential cases of infection associated with inpatient stays, including at least 40 potential cases of the infections of special interest and could retrieve medical charts for 266 patients (98%). Table 1shows the distribution of the sampled ICD-10 codes among the 266 patients with available medical charts. Median age of the patients was 67 years (quartiles: 59–76 years), and 118 of them were women (44%). Median length of hospital stay was 6 days (quartiles: 3–11 days), and more than half of the patients had no comorbidity after excluding cancer diagnosis (Table 2). According to the information recorded in medical charts, one-third of the patients had a history of an earlier infection and 18% of the patients had undergone a surgery (Table 2).

The presence of any infection, regardless of type, was confirmed by PGA in 261 of the 266 patients, yielding an overall PPV of 98% (95% CI: 96%–99%; Table 3). The remaining five patients had ICD-10 codes for pneumonia (two patients), cellulitis (one patient), intestinal infection (one patient), and urinary tract infection (one patient).