Journal of Clinical Epidemiology 82 (2017) 103–111

Journal of Clinical Epidemiology

Time-to-event methodology improved statistical evaluation in register-based health services research

Tobias Bluhmki[a,*], Peter Bramlage[b], Michael Volk[c], Matthias Kaltheuner[d,e], Thomas Danne[f], Wolfgang Rathmann[g], Jan Beyersmann[a]

[a]*Institute of Statistics, Ulm University, Helmholtzstrasse 20, 89081 Ulm, Germany*
[b]*Institute for Pharmacology and Preventive Medicine, Menzelstrasse 21, 15831 Mahlow, Germany*
[c]*axaris software & systeme GmbH, Max-Eyth-Weg 2, 89160 Dornstadt, Germany*
[d]*Gemeinschaftspraxis Kaltheuner—v. Boxberg, Kalkstraße 117, 51377 Leverkusen, Germany*
[e]*winDiab GmbH, Kehler Strasse 24, 40468 Düsseldorf, Germany*
[f]*Children's and Youth Hospital "Auf der Bult", Janusz-Korczak-Allee 12, 30173 Hannover, Germany*
[g]*Institute for Biometrics and Epidemiology, German Diabetes Center (DDZ) Düsseldorf, Auf'm Hennekamp 65, 40225 Düsseldorf, Germany*

Accepted 4 November 2016; Published online 12 November 2016

Abstract

Objectives: Complex longitudinal sampling and the observational structure of patient registers in health services research are associated with methodological challenges regarding data management and statistical evaluation. We exemplify common pitfalls and want to stimulate discussions on the design, development, and deployment of future longitudinal patient registers and register-based studies.

Study Design and Setting: For illustrative purposes, we use data from the prospective, observational, German DIabetes Versorgungs-Evaluation register. One aim was to explore predictors for the initiation of a basal insulin supported therapy in patients with type 2 diabetes initially prescribed to glucose-lowering drugs alone.

Results: Major challenges are missing mortality information, time-dependent outcomes, delayed study entries, different follow-up times, and competing events. We show that time-to-event methodology is a valuable tool for improved statistical evaluation of register data and should be preferred to simple case–control approaches.

Conclusion: Patient registers provide rich data sources for health services research. Analyses are accompanied with the trade-off between data availability, clinical plausibility, and statistical feasibility. Cox' proportional hazards model allows for the evaluation of the outcome-specific hazards, but prediction of outcome probabilities is compromised by missing mortality information.

Keywords: Diabetes mellitus; Registers; Health services research; Mortality; Longitudinal study; Survival analysis

1. Introduction

During the last decades, diabetes has developed into a major public health concern [1]. An improved understanding of the disease and treatment strategies are traditionally established by large, randomized, multicenter, controlled clinical trials [2–5]. However, noninterventional real-world settings, long follow-up times, and partly rare outcomes cannot be addressed in such trials due to cost-effectiveness and practicability.

Thus, population-based and administrative registers containing "routine data" have become more and more popular in epidemiological diabetes research [6–13]. Such structural health registers are rich databases with a great potential for modern epidemiology [7,14]. They provide standardized and systematically collected information about a large population over long periods of time including patient's antidiabetic as well as concomitant drug use, changes in treatment strategies, comorbidities, and selected lifestyle conditions. Rapid hard- and software-based developments extremely simplified data collection, exchange, and management.

Conflict of interest: None.

Funding: The register received funding from Sanofi Aventis Deutschland GmbH. T.B. and J.B. acknowledged support by Diabetes Agenda 2010 gGmbH.

* Corresponding author. Tel.: +49-(0)731-50-33104; fax: +49-(0)731-50-33110.

E-mail address: tobias.bluhmki@uni-ulm.de (T. Bluhmki).

http://dx.doi.org/10.1016/j.jclinepi.2016.11.001
0895-4356/

What is new?

Key findings

- Methodological challenges in health services research include complex timing of events.
- We exemplify missing mortality information, time-dependent outcomes, delayed study entries, different follow-up times, and competing events on the basis of a diabetes register-based study.

What this adds to what was known?

- We show that time-to-event methodology improves statistical evaluation and should be preferred to simple case−control approaches.

What is the implication what should change now?

- Our presentation should contribute to the evaluation of existing register-based data and may influence the planning and design of future registers and studies.

However, the complexity and observational nature of such data make the definition of the population underlying the statistical analysis challenging. This is an omnipresent problem in register-based epidemiological research and requires advanced scientific expertise [15]. Registers do often not completely meet corresponding study objectives in terms of availability, accuracy, reliability, and completeness of data [7,16−18]. In addition, individual information can hardly be controlled by researchers. This data structure displays aspects that are possibly more challenging than encountered in diabetes randomized clinical trials. In the latter, time-to-event techniques are well established, although, for example, proper handling of competing events is still an open issue in the field [19]. The more involved observational structure of registers may have led to an overemphasis on simpler statistical techniques. Our aim is to both show that advanced time-to-event techniques are well suited to reflect the data structure and to also point out open issues.

In spite of growing interest in the evaluation of diabetes registers, specific methodological literature is hardly available [6,15,19−23]. The present article intends to contribute to this current field of research by discussing in detail major administrative as well as statistical challenges illustrated on a recently published register-based cohort study in patients with type 2 diabetes [24]. We exemplify the determination of the population, the timescale, and outcomes to be analyzed as well as the connection of these choices to determining the baseline covariates of interest. We further draw a comparison between time-to-event methodology and standard case−control approaches. Database was the German DIabetes Versorgungs-Evaluation (DIVE) register [24−27].

2. Materials and methods

2.1. Study example

The DIVE data collection system is a prospective, observational, multicenter diabetes register. Enrollment of patients diagnosed either with type 1 or 2 diabetes mellitus independent of the progression of the disease started during the year 2011 with a trace-back of data to January 1, 2011. From there on, the register prospectively tracks participants in up to 200 countrywide doctor's surgeries. DIVE is ongoing and currently includes more than 135,000 patients. Written informed consent is required for participation. The study protocol received ethical approval from the Hannover Medical School Ethics Committee and was carried out in accordance with the Declaration of Helsinki [26]. Data collection is accomplished by the professional diabetes documentation software "DPV2 Diamax" (axaris software & systems GmbH, Dornstadt, Germany). Data are either transferred from the practice administration systems or directly entered into the software by physicians. To ensure accuracy, data input is linked with software-based plausibility checks. The aim of the DIVE register is to establish a national diabetes register of standardized high-quality data and to provide access to health services research.

For illustrative purposes, the present article takes up data from a recent publication retrospectively analyzing the DIVE register [24]. One aim was to explore prognostic factors associated with the initiation of a basal-supported oral therapy (BOT) in patients with type 2 diabetes, which added long-acting basal insulin to a first-line oral antidiabetic drug (OAD), for instance, metformin. The observational period was defined as the time between January 1, 2011, and October 31, 2014.

2.2. Structure of the DIVE registry

The DIVE register consists of three main components: First, "master data" contain demographic characteristics such as diabetes type, gender, age, and date of diabetes diagnosis. Second, "medication data" prospectively provide information on start and end dates as well as dosages of prescribed antidiabetic treatment strategies and concomitant therapies. Third, "course data" include time-varying parameters such as weight, blood parameters (glycated hemoglobin, fasting/plasma blood glucose level), smoking status, and comorbidities (for instance, hypoglycemic and macrovascular/microvascular incidences). Drug information as well as course data are only collected when physicians actually see the patients.