Journal of Clinical Epidemiology 70 (2016) 206–213

A standardized approach to estimating survival statistics for population-based cystic fibrosis registry cohorts

Jenna Sykes[a,*], Sanja Stanojevic[b,c], Christopher H. Goss[d], Bradley S. Quon[e], Bruce C. Marshall[f], Kristofer Petren[f], Josh Ostrenga[f], Aliza Fink[f], Alexander Elbert[f], Anne L. Stephenson[a,g]

[a]*Department of Respirology, St. Michael's Hospital, 30 Bond Street, 6th Floor, Bond Wing, Toronto, Ontario, Canada M5B 1W8*
[b]*Division of Respiratory Medicine, Department of Pediatrics, The Hospital for Sick Children, 555 University Avenue, Toronto, Ontario, Canada M5G 1X8*
[c]*Institute of Health Policy, Management and Evaluation, University of Toronto, 155 College Street, Toronto, Ontario, Canada M5T 3M6*
[d]*Division of Pulmonary and Critical Care Medicine, Department of Medicine and Pediatrics, University of Washington Medical Center, 1959 N.E. Pacific, Seattle, WA, USA 98195-6522*
[e]*Department of Medicine, University of British Columbia, 2775 Laurel Street, Vancouver, British Columbia, Canada V5Z 1M9*
[f]*Cystic Fibrosis Foundation, 6931 Arlington Road, Bethesda, MD, USA 20814*
[g]*Keenan Research Centre, Li Ka Shing Knowledge Institute of St Michael's Hospital, 209 Victoria Street, Toronto, Ontario, Canada M5B 1T8*

Accepted 31 August 2015; Published online 3 October 2015

Abstract

Objectives: Our objective was to quantify the effect of different statistical techniques, inclusion/exclusion criteria, and missing data on the predicted median survival age.

Study Design and Setting: Using the Canadian cystic fibrosis registry (CCFR), the median age of survival was calculated using both the Cox proportional hazards (PH) and the life-table methods. Through simulations, we examined how the median age of survival would change when: (1) patients were excluded, (2) death dates were inaccurate, (3) patients were lost to follow-up, (4) entire years with no clinic visits were excluded even if the patient had a visit in subsequent years, and (5) censoring patients at their date of transplant. Simulations were run assuming 5–35% of data were affected by each scenario.

Results: Over the period 2009–2013, there were 4,666 individuals in the CCFR with 240 deaths. The observed median age of survival calculated by the Cox PH method was 50.9 [95% confidence interval (CI): 47.4, 54.3] and 50.5 from the life-table method (95% CI: 47.5, 53.5). Censoring patients at their transplant date overestimated the median age of survival by 7.2 years (58.1; 95% CI: 53.3, 64.7). Simulations determined that by missing just 15% of deaths, the median age of survival can be overestimated by 3.5 years (54.4; 95% CI: 54.2, 56.1), and having 25% of patients lost to follow-up can underestimate the median age of survival by 3.3 years (47.6; 95% CI: 46.8, 47.7).

Conclusion: We present several recommendations to assist national cystic fibrosis registries in calculating and reporting the median age of survival in a standardized fashion. It is imperative to state the statistical method used as well as the proportion lost to follow-up and the treatment of missing data and transplanted patients. Registries must be diligent in their data collection as incomplete data can lead to overestimation and underestimation of survival.

Keywords: Cystic fibrosis; Survival; Life tables; Cox proportional hazards; Registry; International comparisons

1. Introduction

National cystic fibrosis (CF) data registries are powerful resources for describing patient characteristics, understanding epidemiologic trends, and predicting life expectancy in the CF population. By analyzing these registry data, several countries are able to calculate the median age of survival, the age past which 50% of the population is expected to live assuming the mortality rate and age distribution are held constant. In 2012, the median age of survival varied between countries: 49.7 years in Canada [95% confidence interval (CI): 46.1, 52.2] [1], 43.5 years in the United Kingdom (UK) (95% CI: 37.6, 49.9) [2], and 41.1 years in the United States (95% CI: 37.4, 43.1) [3]. Understanding the reasons behind these international differences in survival may provide valuable information for the CF community; however, the lack of a standardized approach for

Funding: This research was jointly supported by the Canadian Institutes of Health Research grant 119333 and the US Cystic Fibrosis Foundation.

* Corresponding author. Tel.: 416-864-5409; fax: 416-864-5651.
E-mail address: sykesj@smh.ca (J. Sykes).

http://dx.doi.org/10.1016/j.jclinepi.2015.08.026
0895-4356/

What is new?

Key findings

- The Cox proportional hazards and life-table method yield nearly identical results when data processing techniques are the same.
- Losing patients to follow-up has the potential to underestimate the median survival age, whereas missing death dates will overestimate the median survival age.
- Censoring patients at transplant results in an overestimation of the median survival age.

What this adds to what was known?

- The magnitude of the bias on predicting cystic fibrosis (CF) survival when data are missing or patients were lost to follow-up is now quantified.
- Recommendations are provided to help reduce bias in calculating the median age of survival and to improve transparency in reporting survival estimates.

What is the implication and what should change now?

- Researchers should report the sample size, the statistical method used, the proportion lost to follow-up, treatment of missing data, and whether transplanted patients were censored or not when reporting the median age of survival from data registries to acknowledge the magnitude of under/overestimation in survival.
- National CF registries should standardize the reporting of median survival to aid in international comparison.

calculating and reporting median survival makes direct comparison between countries challenging.

The statistical technique for calculating the median age of survival differs between countries with the United States and UK applying the period life-table method and Canada adopting the period Cox proportional hazards (PH) method [4]. International comparisons may still be limited even when the same statistical approach is used. Jackson et al. [5] analyzed data from the United States (US) and Republic of Ireland (RoI) CF registries using the same statistical method and found that the estimates of median predicted survival varied considerably from year to year, particularly in the RoI where the number of annual deaths is small. To reduce year-to-year variability in survival, a longer time window can be used to smooth out estimates, particularly when death numbers are few. For this reason, a 5-year time window is used by Canada, the US, and the UK when calculating the median age of survival [4]. Despite best efforts, registries often suffer from missing data. Variables essential for survival calculations (e.g., date of diagnosis, date of death) may be difficult to accurately obtain, and patients may become lost to follow-up. Healthier and nonadherent patients may defer regular clinic appointments, creating missing intervening years in the registry, which would affect estimates when the median age of survival is calculated using the life-table method. The distance of CF centers may also influence the degree of follow-up. Furthermore, a recent study by Nick et al. [6] showed that although approximately 10% of deaths occur over the age of 45 years, the US CF registry only captured 45.9% of deaths in this age group when compared with Centre for Disease Control Mortality statistics, indicating that registries may be underestimating the number of deaths among older patients. Inaccurate data within the registry will impact survival estimates for any population; however, there is no literature to date which quantifies the impact of these biases on survival statistics in CF.

A standardized approach to calculating and reporting median age of survival for CF registries is needed for comparisons across countries. In this study, we aim to (1) compare the impact of different statistical methodologies, specifically, the life-table and the Cox PH methods, on survival estimates in the CF population; (2) investigate the impact of different methodological and data processing approaches on median age of survival when (a) patients were excluded, (b) death dates were inaccurate, (c) patients were lost to follow-up, (d) entire years with no clinic visits were excluded even if the patient had a visit in subsequent years, and (e) censoring patients at their date of transplant; and (3) develop a unified standardized approach for calculating the median age of survival that can be used in all CF registries.

2. Methods

2.1. Data source

We used the Canadian CF registry (CCFR). Individuals followed in the registry between January 1, 2009, and December 31, 2013, were included in these analyses. The CCFR was first developed in 1970 and incorporates patients from 42 accredited centers across the country. CF Canada provides funding to centers contingent on submitting data to the registry. In addition, there are incentives for CF patients to attend CF centers as many medications are covered by provincial drug plans if dispensed from a CF pharmacy. It is estimated that less than 1% of the Canadian CF population have declined consent to have their data captured in the registry. For these reasons, it is felt that the CCFR captures essentially all CF patients in Canada. The CCFR collects demographic and clinical information annually, and these data undergo routine validation checks to minimize