Practice of Epidemiology

Comparing New-User Cohort Designs: The Example of Proton Pump Inhibitor Effectiveness in Idiopathic Pulmonary Fibrosis

Tanja Tran and Samy Suissa*

* Correspondence to Dr. Samy Suissa, Centre for Clinical Epidemiology, Lady Davis Institute for Medical Research, Jewish General Hospital, 3755 Côte-Sainte-Catherine, H-461, Montreal, Québec H3T 1E2, Canada (e-mail: samy.suissa@mcgill.ca).

Initially submitted March 26, 2020; accepted for publication October 27, 2020.

The prevalent new-user cohort design is useful for assessing the effectiveness of a medication in the absence of an active comparator. Alternative approaches, particularly in the presence of informative censoring, include a variant of this design based on never users of the study drug and the marginal structural Cox model approach. We compared these approaches in assessing the effectiveness of proton pump inhibitors (PPIs) in reducing mortality among patients with idiopathic pulmonary fibrosis (IPF) using a cohort of IPF patients identified in the United Kingdom's Clinical Practice Research Datalink and diagnosed between 2003 and 2016. The cohort included 2,944 IPF patients, 1,916 of whom initiated use of PPIs during follow-up. There were 2,136 deaths (mortality rate = 25.8 per 100 person-years). Using the conventional prevalent new-user design, we found a hazard ratio for death associated with PPI use compared with nonuse of 1.07 (95% confidence interval (CI): 0.94, 1.22). The variant of the prevalent new-user design comparing PPI users with never users found a hazard ratio of 0.82 (95% CI: 0.73, 0.91), while the marginal structural Cox model found a hazard ratio of 1.08 (95% CI: 0.85, 1.38). The marginal structural model and the conventional prevalent new-user design, both accounting for informative censoring, produced similar results. However, the prevalent new-user design variant based on never users introduced selection bias and should be avoided.

cohort studies; comparative effectiveness research; idiopathic pulmonary fibrosis; informative censoring; pharmacoepidemiology; proton pump inhibitors

Abbreviations: CI, confidence interval; CPRD, Clinical Practice Research Datalink; HES, Hospital Episode Statistics; IPCW, inverse probability of censoring weights; IPF, idiopathic pulmonary fibrosis; IPTW, inverse probability of treatment weights; MSCM, marginal structural Cox model; ONS, Office for National Statistics; PPI, proton pump inhibitor; TCPS, time-conditional propensity score.

Idiopathic pulmonary fibrosis (IPF) is a chronic lung disease associated with a poor prognosis; median survival is 2–5 years after diagnosis (1–3). Because of a potential effect of acid reflux on the progression of IPF, proton pump inhibitors (PPIs) are used to treat this disease and have been given conditional recommendations in international IPF treatment guidelines (4). These recommendations were based on several observational studies that assessed the effectiveness of PPIs among patients with IPF, with particular interest in survival (5, 6). Subsequent studies continued to evaluate the effectiveness of PPIs in IPF (7–13). However, several of these studies were affected by immortal time bias or presented

other limitations, such as a small sample size or a short study period (14).

In a recent study, Tran et al. (15) used a large population-based database and a prevalent new-user cohort design that allowed them to specifically circumvent these limitations while assessing the effectiveness of a drug in the absence of an active comparator. The prevalent new-user cohort design originally received its name because of its ability to compare new users of a medication of interest with prevalent users of an older established drug. For example, patients who switched from the comparison drug to the drug of interest and patients who were prevalent users of the comparison

drug at a similar stage of disease could be included in the study cohort (16). This design, which can be applied to compare new users with nonusers, paired each new user of a PPI with a time-matched nonuser who had a physician visit without a PPI at that time and had a similar value for the time-conditional propensity score (TCPS) (16). The nonuser can receive a PPI prescription later during follow-up, at which point the patient is censored. However, with high incidence of PPI use during follow-up (65%), the nonuser comparison group introduces a significant amount of potential informative censoring, which can be corrected by means of inverse probability of censoring weighting under a set of assumptions. To our knowledge, this conventional approach has yet to be compared with other methods applicable to this context.

In this study, we compared the conventional prevalent new-user approach with alternative approaches in the context of estimating the association between PPIs and mortality and hospitalization in patients with IPF. It was first compared with a variation of this design that would select comparators only among never users of PPIs, specifically to avoid censoring among nonusers due to PPI initiation. Second, it was compared with the conventional marginal structural Cox model (MSCM) approach, which uses the entire cohort to create comparable treatment groups at different time points in follow-up, balanced with respect to time-dependent confounders and differential censoring (17, 18).

METHODS

Study population

In this study, we used data from the United Kingdom's Clinical Practice Research Datalink (CPRD) GOLD database, linked to the Hospital Episode Statistics (HES) and Office for National Statistics (ONS) databases. CPRD GOLD is a large primary-care database which contains electronic medical record data on more than 17 million people from more than 680 general practices in the United Kingdom. Participating general practitioners have been trained to record medical information, such as demographic data, data on lifestyle factors (including smoking and alcohol use), and medical diagnoses, using the Read code classification (19). Medications prescribed by general practitioners are coded on the basis of the United Kingdom Prescription Pricing Authority dictionary (20). CPRD data have been previously validated and shown to be of high quality (21). The HES database records information on all inpatient and outpatient hospital admissions, including primary and secondary diagnoses, using *International Classification of Diseases, Tenth Revision*, codes and on hospital procedures using the Office of Population Censuses and Surveys Classification of Interventions and Procedures, Version 4 (22). The ONS database contains electronic death certificates, including the underlying cause of death recorded using *International Classification of Diseases, Tenth Revision*, codes. The HES and ONS databases can only be linked to general practices in England, which represent approximately 75% of all practices in England (21).

Base cohort

The base cohort included patients with a first diagnosis of IPF between January 1, 2003, and December 31, 2016, who had at least 1 year of medical history data prior to diagnosis and were aged ≥ 40 years at diagnosis. Patients were only included in the base cohort if their medical record was linked to the HES and ONS databases. Base cohort entry was defined as the date of IPF diagnosis.

Conventional prevalent new-user cohort design with nonusers

Within the base cohort, all patients with a first PPI prescription after IPF diagnosis were identified (new PPI users). All first prescriptions were ordered chronologically, and time-based exposure sets were created (± 1 month). Potential reference patients had a record of a physician visit within the corresponding exposure set, with no PPI prescription. PPI users and reference patients were matched 1:1 without replacement on TCPS (23) using conditional logistic regression to estimate their propensity to receive a PPI after IPF diagnosis. Patients who could not be matched were excluded from the analysis. PPI users were considered exposed throughout follow-up. Matched reference patients could be prescribed a PPI later during follow-up and were censored at that point. This introduced the potential for informative censoring, and we addressed it by using inverse probability of censoring weighting in the analysis. Entry into the study cohort was defined as the date of the first PPI prescription for PPI users and the date of the corresponding physician visit for nonusers. To assess the impact of sampling in this design, we matched PPI users 1:1 to reference patients using sampling with replacement in an additional analysis.

Prevalent new-user cohort design variant with never users

We used the prevalent new-user cohort design within the base cohort of newly diagnosed IPF patients, allowing only never users of PPIs to be selected as potential matches for PPI users in order to avoid the issue of informative censoring by analyzing only uncensored nonusers (24). Thus, PPI users and never users were matched 1:1 without replacement on the basis of time-based exposure sets, using TCPSs estimated by conditional logistic regression (16, 23). Patients who could not be matched were excluded from the analysis. Throughout follow-up, PPI users were considered exposed and never users were considered unexposed. Cohort entry was defined as the day of the first PPI prescription for new PPI users and the date of the corresponding physician visit for never users.

Conventional MSCM

The conventional approach to assessment of time-varying exposure included all patients from the base cohort. Patients entered the study cohort on the day of their IPF diagnosis.

Follow-up was measured in monthly time units, allowing time-varying covariates to be updated in each person-month. Patients who were exposed to PPIs during follow-up were considered unexposed until their first PPI prescription and exposed from then on. Stabilized inverse probability of treatment weights (IPTW) and inverse probability of censoring weights (IPCW) were applied to create a pseudopopulation that was balanced with regard to measured treatment selection and censoring selection factors that affect the outcome (time-dependent confounders and differential censoring) (17). To obtain IPTW for each person-month in which a patient had not yet started using a PPI, we estimated the probabilities that a patient would receive his or her own observed PPI prescription in a certain month, given the patient's medical history up to that month. The weights, which are the inverse of the probability of observed exposure history, were multiplied over time (17). This approach, using time-updated covariate information collected up to PPI initiation, is similar to the TCPS used in the prevalent new-user design, where the probability of initiating PPI use is calculated on the basis of information collected prior to each exposure set defined by a PPI prescription. The probability of censoring in each person-month was similarly estimated by calculating the conditional probability of remaining uncensored up to a certain month and included PPI treatment as an additional regressor. Two mechanisms for censoring were considered in the calculation of the IPCW. First, we considered administrative censoring due to the ending of the study or loss to follow-up in the CPRD. Second, we also estimated the probability of remaining uncensored due to PPI initiation to account for informative censoring. The resulting weight from this estimation is comparable to using the IPCW to account for informative censoring due to PPI initiation among nonusers in the conventional prevalent new-user cohort design. The weights obtained from the inverse probabilities of treatment and censoring were then multiplied to calculate the stabilized weight for each person-month.

Outcome

The primary study outcome was death from any cause identified in the CPRD or ONS. Other outcomes of interest were respiratory deaths identified in the ONS database and a composite outcome of respiratory related-hospitalization identified in the HES database or death from any cause. Patients were followed to the earliest of the following events: the occurrence of one of the study outcomes, death, lung transplantation, the end of registration with the general practice, or the end of the study period (May 31, 2017).

Covariates

Baseline covariates were identified during the year prior to study cohort entry. Time-varying covariates were identified in each person-month during follow-up in the MSCM. The same covariates were included in the TCPS model of the prevalent new-user cohort design and the model for calculating the IPTW in the MSCM. These variables were age, sex, body mass index (weight (kg)/height (m)²), smoking history, alcohol-related disorders (based on diagnoses such as alco-

holism, alcoholic cirrhosis of the liver, alcoholic hepatitis, other alcohol-related disorders, or consumption of more than 15 units of alcohol (>150 mL or >120 g of ethanol) per week), race/ethnicity, prior hospitalizations, comorbidity, and medication use. Comorbid conditions included asthma, chronic obstructive pulmonary disease, gastroesophageal reflux disease, a history of Nissen fundoplication, cardiovascular disease, diabetes, cancer, renal disease, and depression. Medications evaluated included oral and inhaled corticosteroids, angiotensin-converting enzyme inhibitors, β -blockers, anticoagulants, diuretics, statins, nonsteroidal antiinflammatory drugs, H₂ blockers, and previous PPI use prior to IPF diagnosis. The models used to estimate the probability of censoring due to PPI initiation in the conventional prevalent new-user cohort design and the MSCM included the same covariates.

Statistical analyses

Prevalent new-user cohort design. The statistical analyses for both sampling strategies using the prevalent new-user cohort design were the same. However, the conventional prevalent new-user design with nonusers additionally accounted for informative censoring due to PPI initiation among nonusers by using IPCW, estimated by logistic regression (25). The Cox proportional hazards model was used to estimate crude and adjusted hazard ratios and 95% confidence intervals for all-cause mortality, respiratory-disease-related mortality, and the combined outcome of respiratory-disease-related hospitalization or death associated with PPI use as compared with either no use or never use, using the intention-to-treat approach. The models adjusted for potential confounders measured at cohort entry, including age, sex, smoking history, cardiovascular disease, and previous hospitalization. In the weighted analyses, in order to take within-subject correlation into account, we calculated 95% confidence intervals for the marginal hazard ratios based on robust standard errors (26). Kaplan-Meier curves for all-cause mortality were estimated, weighted in the conventional approach to take informative censoring into account (27). We also estimated the risk difference for each outcome at 3 years after cohort entry in the prevalent new-user cohort design (28).

Conventional MSCM. We used pooled logistic regression models to estimate stabilized IPTW and IPCW (17, 18). We fitted the models for the numerator of the weights using baseline covariates only and the models for the denominator of the weights using fixed and time-updated covariates. A time-dependent Cox model weighted by the resulting stabilized weights was then used to estimate the marginal hazard ratio adjusted for potential confounders measured at cohort entry, including age, sex, smoking history, cardiovascular disease, and previous hospitalization. In order to take within-subject correlation into account, we calculated 95% confidence intervals for the marginal hazard ratios based on robust standard errors (26). We calculated weighted Kaplan-Meier curves for all-cause mortality (27). We also estimated risk differences at 3 years after cohort entry (28).

All analyses were conducted using SAS, version 9.4 (SAS Institute, Inc., Cary, North Carolina). The study was approved by the Independent Scientific Advisory Committee of the CPRD and the Ethics Committee of Jewish General Hospital (Montreal, Quebec, Canada).

RESULTS

The base cohort included 2,944 patients newly diagnosed with IPF during 2003–2016, with a mean follow-up of 2.1 years (interquartile range, 0.8–4.0; maximum, 13.6 years).

Overall, 1,916 (65%) patients were prescribed a PPI at some point after their IPF diagnosis (mean time to PPI initiation = 7.8 months; range, 0–136.9 months) (see Web Figure 1, available at <https://doi.org/10.1093/aje/kwaa242>). In total, there were 8,277 person-years of follow-up, of which 4,685 person-years involved exposure to PPIs. There were 2,136 (73%) deaths from any cause during follow-up. The overall all-cause mortality rate was 25.8 per 100 person-years (95% confidence interval (CI): 24.7, 26.9). Median survival in the full cohort was 2.8 years.

Figure 1 shows the formation of the base cohort, which was the study cohort for the MSCM, and the 2 matched

Figure 1. Selection of 3 cohorts of patients with idiopathic pulmonary fibrosis (IPF) from the Clinical Practice Research Datalink (CPRD), United Kingdom, 2003–2016. HES, Hospital Episode Statistics; ONS, Office for National Statistics; PPI, proton pump inhibitor.

Table 1. Baseline Characteristics of Patients With Idiopathic Pulmonary Fibrosis at Diagnosis (Base Cohort) and by Exposure at Study Cohort Entry, United Kingdom, 2003–2016

Covariate	Study Cohort									
	Base Cohort (n = 2,944)		Sampling From Nonusers				Sampling From Never Users			
			PPI Use (n = 1,852)		No Use of PPIs (n = 1,852)		PPI Use (n = 1,017)		Never Use of PPIs (n = 1,017)	
	No.	%	No.	%	No.	%	No.	%	No.	%
Demographic and lifestyle factors										
Age at diagnosis, years ^a	75.8 (9.5)		75.4 (9.4)		75.6 (9.5)		76.2 (9.5)		76.7 (9.6)	
Male sex	1,884	64.0	1,150	62.1	1,209	65.3	636	62.5	685	67.4
Duration of IPF, years ^a	0.5 (0.9)		0.5 (0.9)		0.5 (0.9)		0.5 (0.9)		0.5 (0.9)	
Body mass index ^b										
<25	832	28.3	549	29.6	531	28.7	310	30.5	326	32.1
25–29	1,060	36.0	663	35.8	692	37.4	341	33.5	357	35.1
≥30	702	23.9	466	25.2	439	23.7	257	25.3	201	19.8
Unknown	350	11.9	174	9.4	190	10.3	109	10.7	133	13.1
Smoking status										
Former smoker	1,403	47.7	866	46.8	873	47.1	466	45.8	466	45.8
Never smoker	1,049	35.6	674	36.4	657	35.5	369	36.3	347	34.1
Current smoker	453	15.4	299	16.1	309	16.7	174	17.1	191	18.8
Unknown	39	1.3	13	0.7	13	0.7	8	0.8	13	1.3
Alcohol-related disorder	831	28.3	537	29.0	551	29.8	290	28.5	290	28.5
Race/ethnicity										
White	2,683	91.1	1,715	92.6	1,698	91.7	932	91.6	904	88.9
Other	83	2.8	59	3.2	52	2.8	28	2.8	19	1.9
Unknown	178	6.1	78	4.2	102	5.5	57	5.6	94	9.2
Medical history prior to cohort entry										
Hospitalization ^c	1,424	48.4	1,066	57.6	982	53.0	547	53.8	508	50.0
Asthma	586	19.9	410	22.1	379	20.5	196	19.3	183	18.0
COPD	797	27.1	595	32.1	576	31.1	310	30.5	308	30.3
GERD	551	18.7	481	26.0	348	18.8	117	11.5	86	8.5
Arrhythmia	421	14.3	297	16.0	300	16.2	159	15.6	147	14.5
Heart failure	586	19.9	410	22.1	398	21.5	233	22.9	224	22.0
Hypertension	1,594	54.1	1,076	58.1	1,042	56.3	575	56.5	530	52.1
Pulmonary hypertension	54	1.8	50	2.7	44	2.4	33	3.2	25	2.5
Myocardial infarction	370	12.6	277	15.0	239	12.9	126	12.4	114	11.2
Stroke	229	7.8	149	8.1	144	7.8	91	9.0	99	9.7
Diabetes mellitus	430	14.6	286	15.4	255	13.8	136	13.4	139	13.7
Cancer	471	16.0	341	18.4	322	17.4	210	20.7	170	16.7
Lung cancer	25	0.9	31	1.7	23	1.2	16	1.6	15	1.5
Renal disease	717	24.4	511	27.6	493	26.6	276	27.1	260	25.6
Depression	429	14.6	305	16.5	281	15.2	140	13.8	124	12.2

Table continues

cohorts based on the prevalent new-user cohort design. Table 1 gives an overview of the baseline characteristics in the base cohort, the matched cohorts sampling from non-users or never users only.

Conventional prevalent new-user cohort design with nonusers

The study cohort included 1,852 new PPI users matched to 1,852 nonusers, after IPF diagnosis. The mean length of

Table 1. Continued

Covariate	Study Cohort									
	Base Cohort (n = 2,944)		Sampling From Nonusers				Sampling From Never Users			
			PPI Use (n = 1,852)		No Use of PPIs (n = 1,852)		PPI Use (n = 1,017)		Never Use of PPIs (n = 1,017)	
	No.	%	No.	%	No.	%	No.	%	No.	%
Medications prescribed in the year prior to cohort entry										
Inhaled corticosteroids	368	12.5	240	13.0	230	12.4	133	13.1	120	11.8
Oral corticosteroids	645	21.9	675	36.5	587	31.7	535	52.6	293	28.8
Azathioprine	61	2.1	66	3.6	53	2.9	33	3.2	28	2.8
ACE inhibitors	895	30.4	575	31.1	556	30.0	299	29.4	310	30.5
Angiotensin II receptor blockers ^d	12	0.4	8	0.4	7	0.4	NA	<0.5	NA	<0.5
β-blockers	760	25.8	533	28.8	481	26.0	280	27.5	235	23.1
Diuretics	1,334	45.3	863	46.6	826	44.6	474	46.6	462	45.4
Anticoagulants	302	10.3	199	10.8	218	11.8	117	11.5	126	12.4
Antiplatelet agents	1,202	40.8	859	46.4	790	42.7	420	41.3	353	34.7
Statins	1,275	43.3	914	49.4	853	46.1	426	41.9	379	37.3
NSAIDs	366	12.4	241	13.0	214	11.6	119	11.7	72	7.1
H ₂ blockers	166	5.6	117	6.3	111	6.0	84	8.3	60	5.9
PPI use in the year prior to IPF diagnosis	1,318	44.8	1,153	62.3	840	45.4	442	43.5	148	14.6

Abbreviations: ACE, angiotensin-converting enzyme; COPD, chronic obstructive pulmonary disease; GERD, gastroesophageal reflux disease; IPF, idiopathic pulmonary fibrosis; NA, not available; NSAID, nonsteroidal antiinflammatory drug; PPI, proton pump inhibitor.

^a Values are expressed as mean (standard deviation).

^b Weight (kg)/height (m)².

^c During the year prior to cohort entry.

^d NA indicates that data were suppressed to protect patient confidentiality, since the cell size was <5.

follow-up was 2.4 years (median, 1.7 years) for PPI users and 1.1 years (median, 0.3 years) for nonusers. There were 1,703 deaths from any cause during follow-up (range, 0–13.4 years). The overall all-cause mortality rate in this cohort was 26.7 per 100 person-years, and median survival was 2.8 years. The weighted Kaplan-Meier curve for all-cause mortality among PPI users compared with nonusers is displayed in Figure 2A (Web Figure 2A). Overall, there was no association between all-cause mortality (hazard ratio (HR) = 1.07, 95% CI: 0.94, 1.22), respiratory-disease-related mortality (HR = 1.10, 95% CI: 0.94, 1.28), or the composite outcome of respiratory-disease-related hospitalization or death (HR = 1.01, 95% CI: 0.91, 1.12) and PPI use as compared with no use (Table 2). There was also no reduction in absolute risk for these outcomes at 3 years associated with PPI use compared with no use. The risk differences were −2.3% (95% CI: −7.0, 2.1), −2.7% (95% CI: −5.9, 1.7), and −0.4% (95% CI: −5.1, 3.4), respectively. Sampling with replacement resulted in a study cohort of 1,905 PPI users matched to 1,905 nonusers. The hazard ratio for

all-cause mortality associated with PPI use versus nonuse was 1.22 (95% CI: 1.03, 1.45).

Prevalent new-user cohort design variant with never users

There were 1,028 never users and 1,916 PPI users in the base cohort of 2,944 patients diagnosed with IPF. After time- and TCPS-matching, the study cohort included 1,017 PPI users matched to 1,017 never users. The mean length of follow-up was 1.6 years (median, 0.9 years) for never users and 2.2 years (median, 1.4 years) for PPI users. There were 1,353 (67%) deaths from any cause during follow-up (range, 0–13.4 years). The overall all-cause mortality rate was 34.9 per 100 person-years (95% CI: 33.0, 36.7). Median survival in this cohort was 1.9 years. Figure 2B shows the Kaplan-Meier curve for all-cause mortality among PPI users compared with never users (Web Figure 2B).

PPI users had lower risks of all-cause mortality (HR = 0.82, 95% CI: 0.73, 0.91), respiratory-disease-related mortality

Table 2. Crude and Adjusted Hazard Ratios for the Association Between Use of a Proton Pump Inhibitor After Diagnosis of Idiopathic Pulmonary Fibrosis (Versus No Use) and Hospitalization and Mortality Outcomes, Obtained Using 3 Different Study Designs^a, United Kingdom, 2003–2016

Study Design and Exposure	No. of Patients	No. of Events	No. of Person-Years	IR per 100 Person-Years	95% CI for IR	Crude HR	Adjusted HR	95% CI
<i>All-Cause Mortality</i>								
Conventional prevalent new-user cohort design								
PPI use	1,852	1,221	4,390	27.8	26.2, 29.4	1.20	1.07 ^b	0.94, 1.22 ^c
No use	1,852	482	1,978	24.3	22.2, 26.5	1.00	1.00	Referent
Prevalent new-user cohort design variant with never users								
PPI use	1,017	684	2,231	30.7	28.4, 33.0	0.79	0.82 ^d	0.73, 0.91
Never use	1,017	669	1,647	40.6	37.5, 43.7	1.00	1.00	Referent
Marginal structural Cox model								
PPI use	1,916 ^e	1,376	4,685	29.4	27.9, 31.0	1.37	1.08 ^f	0.85, 1.38 ^c
No use	2,821 ^e	760	3,592	21.2	19.7, 22.7	1.00	1.00	Referent
<i>Respiratory-Disease-Related Mortality</i>								
Conventional prevalent new-user cohort design								
PPI use	1,852	821	4,390	18.7	17.4, 20.0	1.28	1.10 ^b	0.94, 1.28 ^c
No use	1,852	304	1,978	15.4	13.6, 17.1	1.00	1.00	Referent
Prevalent new-user cohort design variant with never users								
PPI use	1,017	461	2,231	20.7	18.8, 22.6	0.82	0.85 ^d	0.74, 0.98
Never use	1,017	436	1,647	26.5	24.0, 29.0	1.00	1.00	Referent
Marginal structural Cox model								
PPI use	1,916 ^e	932	4,685	19.9	18.6, 21.2	1.44	1.00 ^f	0.73, 1.36 ^c
No use	2,821 ^e	491	3,592	13.7	12.5, 14.9	1.00	1.00	Referent
<i>Respiratory-Disease-Related Hospitalization or Death</i>								
Conventional prevalent new-user cohort design								
PPI use	1,852	1,358	3,532	38.4	36.4, 40.5	1.15	1.01 ^b	0.91, 1.12 ^c
No use	1,852	641	1,713	37.4	34.5, 40.3	1.00	1.00	Referent
Prevalent new-user cohort design variant with never users								
PPI use	1,017	748	1,807	41.4	38.4, 44.4	0.85	0.86 ^d	0.77, 0.95
Never use	1,017	712	1,385	51.4	47.6, 55.2	1.00	1.00	Referent
Marginal structural Cox model								
PPI use	1,916 ^e	1,319	3,479	37.9	35.9, 40.0	1.27	1.03 ^f	0.83, 1.27 ^c
No use	2,821 ^e	977	3,155	31.0	29.0, 32.9	1.00	1.00	Referent

Abbreviations: CI, confidence interval; HR, hazard ratio; IPF, idiopathic pulmonary fibrosis; IR, incidence rate; PPI, proton pump inhibitor.

^a The conventional prevalent new-user cohort design with sampling from all nonusers, the prevalent new-user cohort design variant with sampling from never users only, and the marginal structural Cox model.

^b After matching on time-conditional propensity score, results were further adjusted for age, sex, smoking, history of hospitalization in the year prior to cohort entry, and concomitant cardiovascular disease and weighted by inverse probability of censoring weights.

^c 95% CIs were calculated on the basis of robust standard errors.

^d After matching on time-conditional propensity score, results were further adjusted for age, sex, smoking, history of hospitalization in the year prior to cohort entry, and concomitant cardiovascular disease.

^e Number of patients who were PPI users or nonusers at some point during follow-up. Patients in the marginal structural model could contribute person-time to both comparison groups.

^f Weighted by means of inverse probability of treatment weights and inverse probability of censoring weights and adjusted for age at IPF diagnosis, sex, history of smoking, history of hospitalization in the year prior to IPF diagnosis, and cardiovascular disease at the time of IPF diagnosis.

Figure 2. Kaplan-Meier curves for all-cause mortality comparing proton pump inhibitor (PPI) use (solid line) with a comparison group (dashed line) using 3 different approaches among patients diagnosed with idiopathic pulmonary fibrosis, United Kingdom, 2003–2016. A) The conventional prevalent new-user cohort design (weighted by inverse probability of censoring weights); B) the prevalent new-user cohort design with never users; and C) the marginal structural Cox model (weighted by means of inverse probability of treatment weights and inverse probability of censoring weights).

(HR = 0.85, 95% CI: 0.74, 0.98), and respiratory-disease-related hospitalization or death combined (HR = 0.86, 95% CI: 0.77, 0.95) than never users (Table 2). Risk differences were 7.3% (95% CI: 3.0, 11.0), 5.2% (95% CI: 0.4, 9.9), and 5.4% (95% CI: 1.8, 9.4), respectively.

Conventional MSCM

The analysis using the conventional MSCM included all 2,944 patients diagnosed with IPF, including 1,916 who were prescribed a PPI at some point. There were 1,376 deaths among PPI users out of the 2,136 deaths that occurred during follow-up. The mortality rate during exposed follow-up time was 29.4 per 100 person-years (95% CI: 27.9, 31.0). During unexposed follow-up time, it was 21.2 per 100 person-years (95% CI: 19.7, 22.7). Median survival from IPF diagnosis was 2.2 years in the PPI user group and 3.5 years in the nonuser group. Figure 2C displays the weighted Kaplan-Meier curves by treatment group (Web Figure 2C).

Using the MSCM approach and after adjustment for potential confounders, the hazard ratio for all-cause mortality associated with PPI use was 1.08 (95% CI: 0.85, 1.38) compared with nonuse (Table 2). The hazard ratio for respiratory-disease-related mortality was 1.00 (95% CI: 0.73, 1.36), and for the combined outcome of respiratory-disease-related hospitalization or death it was 1.03 (95% CI: 0.83, 1.27). There was no reduction in absolute risk for these outcomes associated with PPI use versus no use at 3 years. The risk differences were −2.6% (95% CI: −11.2, 5.1), −0.0% (95% CI: −9.0, 7.7), and −1.1% (95% CI: −8.8, 6.6), respectively.

Figure 3 shows the hazard ratios for the 3 outcomes estimated under the 3 approaches.

DISCUSSION

We explored alternative approaches to assess the effectiveness of treating IPF with PPIs, compared with nonuse. The unique methodological challenge in this study was the large proportion of the cohort (65%) exposed to PPIs at some point during follow-up, which introduced censoring. Consequently, the conventional prevalent new-user cohort design with nonusers as the comparison group had to account for informative censoring, and it found no association between PPI use and mortality (HR = 1.07). In contrast, a variant of the prevalent new-user cohort design that avoided censoring by sampling the comparison group exclusively from never users found a significantly lower risk of mortality with PPIs (HR = 0.82). In comparison, the conventional MSCM approach yielded an estimate similar to that of the conventional prevalent new-user cohort design (HR = 1.08). Below, we discuss how the 3 approaches led to different results and consider their feasibility in addressing informative censoring.

The variant of the prevalent new-user cohort design based on sampling exclusively from never users of PPIs as the comparator group produced a hazard ratio of 0.82 (95% CI: 0.73, 0.91) for all-cause mortality associated with PPI use. While use of this unexposed comparator group avoided the issue of informative censoring, which had to be adjusted for in the conventional prevalent new-user cohort design, it introduced selection bias by conditioning on future exposure information (29, 30). These never users probably included either patients who were healthy survivors and did not ever require a prescription for PPIs or patients who were too sick to receive a prescription and died early. Indeed, never users

Figure 3. Hazard ratios (HRs) for respiratory-disease-related hospitalization and mortality associated with proton pump inhibitor use (compared with no use) obtained using 3 different study designs in a cohort of patients with idiopathic pulmonary fibrosis, United Kingdom, 2003–2016. Bars, 95% confidence intervals (CIs).

were more likely to die shortly after their IPF diagnosis compared with PPI users (Web Figure 3). Therefore, by choosing never users as the comparator group and thus primarily selecting patients with poor health, the mortality rate in the comparator group was inflated, and PPI use appeared to be beneficial. Even though choosing never use as the comparison group seems to be a quick fix to avoid informative censoring, it will introduce selection bias at the study design level, which cannot be corrected using data analytical methods.

The MSCM approach is useful in the presence of time-varying confounding due to the time-varying nature of exposure and the absence of an active comparator in this study (17). For example, patients who receive a PPI prescription are likely to be different from a nonuser with regard to their health status at the same point during follow-up, which may also affect their likelihood of death. The MSCM creates a reweighted pseudo-patient population that would lead to balanced treatment groups of PPI new users and nonusers with regard to time-varying predictors of treatment and includes inverse probability of censoring weights to account for informative censoring due to PPI initiation (17).

This analysis showed the similarity between the MSCM approach and the conventional prevalent new-user cohort design. Indeed, both approaches estimate the probability of initiating treatment based on time-varying patient characteristics up until the first PPI prescription and account for informative censoring using IPCW. It is thus not surprising that the resulting hazard ratios for all-cause mortality associated with PPI use were similar between the MSCM (HR = 1.08, 95% CI: 0.85, 1.38) and the prevalent new-user

approach (HR = 1.07, 95% CI: 0.94, 1.22). However, while the point estimates were similar, the confidence interval for the MSCM was wider. Although both methods adjusted for informative censoring, this 91% wider confidence interval on the log scale may have been due to the use of inverse probability weighting by the MSCM to adjust for confounding compared with the 1:1 matching used in the conventional prevalent new-user approach. Such weighting for confounding can decrease precision in the presence of extreme weights. Truncation of these weights could lead to more precise effect estimates but could also affect the control of confounding (Web Table 1) (18, 31). More research is needed to evaluate the impact of these approaches on the precision of these estimates.

Despite the similar findings, it should be noted that the time of cohort entry for the 2 approaches is different and that these 2 approaches estimate slightly different parameters. While they both aim at estimating the average treatment effect of PPIs on survival among patients with IPF, the conventional prevalent new-user design estimates the effect among patients with IPF who were matched, from cohort entry on. This approach answers the question as to whether patients should be treated with PPIs at varying times during the course of their disease. The conventional MSCM estimates the treatment effect within the total cohort from IPF diagnosis on and gives information on whether patients should be treated with PPIs as of the time of IPF diagnosis. Since IPF is a rare disease with limited treatment options and treatment is not necessarily initiated at diagnosis, the matched analysis using the conventional prevalent new-user cohort design gives an answer that would more closely

resemble a clinical trial in this population—namely, patients at different stages of the disease. This estimate is thus useful for treatment decision-making in this population. A reason for the similar estimates in both approaches is that these medications do not seem to be associated with mortality, irrespective of when they are given. Moreover, the study populations and the time of cohort entry in both approaches are quite similar. Most PPI users received their first PPI shortly after IPF diagnosis, with 49% receiving their first prescription within 1 month. Despite the similarities, there may be some advantages of one approach over the other. The cohort formation in the MSCM is less complex than in the prevalent new-user cohort design, but the data analytical techniques are more challenging. However, for the typical consumer of clinical research, the prevalent new-user cohort design may be simpler because of its resemblance to the randomized controlled trial in terms of data presentation, which is not the case for the MSCM. Indeed, for example, Table 1 presents data on the exposed and unexposed groups at cohort entry, making these 2 groups directly comparable with respect to the time-dependent covariates. Outcome events are also directly attributed to the comparison groups, rather than to exposed and unexposed person-time. In addition, this analysis suggests that it may also result in increased precision, though this needs further research for confirmation.

This study had several strengths. It provided alternative strategies for comparative effectiveness studies in the absence of an active comparator and a high degree of informative censoring using real-world data. We used approaches, such as the prevalent new-user cohort design and MSCMs, to minimize confounding. Moreover, we accounted for informative censoring due to the high frequency of PPI initiation in the unexposed group, which had not been an issue in prior applications of the prevalent new-user design where exposure was relatively infrequent.

There were also limitations. First, using inverse probability weighting in the MSCM may be more similar to sampling with replacement. The conventional MSCM approach also uses data from the full cohort, whereas the conventional prevalent new-user cohort design includes only patients who have been matched. This can lead to loss of information and generalizability if the proportion of exposed individuals who are matched is small. In this study, 1,852 (97%) of the patients who received treatment were matched, making bias due to exclusion less likely. In an additional analysis, we sampled with replacement to increase the likelihood of identifying the best possible match for each PPI user and found that the risk of all-cause mortality among PPI users was slightly elevated (HR = 1.22, 95% CI: 1.03, 1.45). This approach included overall fewer patients ($n = 2,350$ (80%)) from the base cohort compared with the sampling approach without replacement ($n = 2,668$ (91%)) and may thus represent a different study population. Future research studies need to further evaluate sampling with replacement in the prevalent new-user cohort design. Second, inverse probability weights, as applied in the MSCM and as control for informative censoring, are only valid if the measured covariates are sufficient to adjust for both confounding and selection bias due to loss of follow-up (32). Relevant infor-

mation may not be available at the time of each risk set in databases with routinely collected data. However, because data in the CRPD are entered by the general practitioner who also prescribes PPIs, we assume that most covariates that may indicate treatment initiation or censoring are captured at that time point. It also assumes that the models for initiation of PPIs and censoring, given the past, are correctly specified. Thus, residual confounding and differential censoring may still be present in the reweighted study populations.

In summary, the prevalent new-user cohort design variant that selected only never users to circumvent informative censoring introduced selection bias by conditioning on future exposure and should be avoided when choosing the comparator group in the prevalent new-user cohort design. The MSCM and the conventional prevalent new-user cohort design produced similar results and can both be used in the absence of an active comparator and when informative censoring is present. The prevalent new-user cohort design may be favored by some because of its simplicity and familiarity in data presentation.

ACKNOWLEDGMENTS

Author affiliations: Department of Epidemiology, Biostatistics, and Occupational Health, Faculty of Medicine and Health Sciences, McGill University, Montreal, Quebec, Canada (Tanja Tran, Samy Suissa); and Centre for Clinical Epidemiology, Lady Davis Institute for Medical Research, Jewish General Hospital, Montreal, Quebec, Canada (Tanja Tran, Samy Suissa).

Access to the data was funded by grants from the Canadian Institutes of Health Research and the Canadian Foundation for Innovation. S.S. is a recipient of the Distinguished James McGill Professorship award.

The funding sources played no role in the design of the study, the analysis of the data, or interpretation of the study results.

S.S. has received research grants from Boehringer Ingelheim (Ingelheim am Rhein, Germany) and Novartis International AG (Basel, Switzerland) and has participated in advisory board meetings or served as a speaker for AstraZeneca AB (Cambridge, United Kingdom), Boehringer Ingelheim, and Novartis. T.T. declares no competing interests.

REFERENCES

1. Nathan SD, Shlobin OA, Weir N, et al. Long-term course and prognosis of idiopathic pulmonary fibrosis in the new millennium. *Chest*. 2011;140(1):221–229.
2. Ley B, Collard HR, King TE Jr. Clinical course and prediction of survival in idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med*. 2011;183(4):431–440.
3. Navaratnam V, Fleming KM, West J, et al. The rising incidence of idiopathic pulmonary fibrosis in the U.K. *Thorax*. 2011;66(6):462–467.

4. Raghu G, Rochwerf B, Zhang Y, et al. An official ATS/ERS/JRS/ALAT clinical practice guideline: treatment of idiopathic pulmonary fibrosis. An update of the 2011 clinical practice guideline. *Am J Respir Crit Care Med*. 2015; 192(2):e3–e19.
5. Lee JS, Ryu JH, Elicker BM, et al. Gastroesophageal reflux therapy is associated with longer survival in patients with idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med*. 2011;184(12):1390–1394.
6. Lee JS, Collard HR, Anstrom KJ, et al. Anti-acid treatment and disease progression in idiopathic pulmonary fibrosis: an analysis of data from three randomised controlled trials. *Lancet Respir Med*. 2013;1(5):369–376.
7. Ghebremariam YT, Cooke JP, Gerhart W, et al. Pleiotropic effect of the proton pump inhibitor esomeprazole leading to suppression of lung inflammation and fibrosis. *J Transl Med*. 2015;13:Article 249.
8. Ekström M, Bornefalk-Hermansson A. Cardiovascular and antacid treatment and mortality in oxygen-dependent pulmonary fibrosis: a population-based longitudinal study. *Respirology*. 2016;21(4):705–711.
9. Kreuter M, Wuyts W, Renzoni E, et al. Antacid therapy and disease outcomes in idiopathic pulmonary fibrosis: a pooled analysis. *Lancet Respir Med*. 2016;4(5):381–389.
10. Lee CM, Lee DH, Ahn BK, et al. Protective effect of proton pump inhibitor for survival in patients with gastroesophageal reflux disease and idiopathic pulmonary fibrosis. *J Neurogastroenterol Motil*. 2016;22(3):444–451.
11. Kreuter M, Spagnolo P, Wuyts W, et al. Antacid therapy and disease progression in patients with idiopathic pulmonary fibrosis who received pirfenidone. *Respiration*. 2017;93(6): 415–423.
12. Liu B, Su F, Xu N, et al. Chronic use of anti-reflux therapy improves survival of patients with pulmonary fibrosis. *Int J Clin Exp Med*. 2017;10(3):5805–5810.
13. Jo HE, Corte TJ, Glaspole I, et al. Gastroesophageal reflux and antacid therapy in IPF: analysis from the Australia IPF Registry. *BMC Pulm Med*. 2019;19(1):Article 84.
14. Tran T, Suissa S. The effect of anti-acid therapy on survival in idiopathic pulmonary fibrosis: a methodological review of observational studies. *Eur Respir J*. 2018;51(6): 1800376.
15. Tran T, Assayag D, Ernst P, et al. Effectiveness of proton pump inhibitors in idiopathic pulmonary fibrosis: a population-based cohort study. *Chest*. 2021;159(2):673–682.
16. Suissa S, Moodie EE, Dell’Aniello S. Prevalent new-user cohort designs for comparative drug effect studies by time-conditional propensity scores. *Pharmacoepidemiol Drug Saf*. 2017;26(4):459–468.
17. Hernán MA, Brumback B, Robins JM. Marginal structural models to estimate the causal effect of zidovudine on the survival of HIV-positive men. *Epidemiology*. 2000;11(5): 561–570.
18. Cole SR, Hernán MA. Constructing inverse probability weights for marginal structural models. *Am J Epidemiol*. 2008;168(6):656–664.
19. Chisholm J. The Read clinical classification. *BMJ*. 1990; 300(6732):1092.
20. Walley T, Mantgani A. The UK General Practice Research Database. *Lancet*. 1997;350(9084):1097–1099.
21. Herrett E, Thomas SL, Schoonen WM, et al. Validation and validity of diagnoses in the General Practice Research Database: a systematic review. *Br J Clin Pharmacol*. 2010; 69(1):4–14.
22. National Health Service, United Kingdom Department of Health and Social Care. Clinical classifications. <https://digital.nhs.uk/services/terminology-and-classifications/clinical-classifications>. Updated April 17, 2020. Accessed February 9, 2021.
23. Lu B. Propensity score matching with time-dependent covariates. *Biometrics*. 2005;61(3):721–728.
24. Leung K-M, Elashoff RM, Afifi AA. Censoring issues in survival analysis. *Annu Rev Public Health*. 1997;18(1): 83–104.
25. Robins JM, Finkelstein DM. Correcting for noncompliance and dependent censoring in an AIDS clinical trial with inverse probability of censoring weighted (IPCW) log-rank tests. *Biometrics*. 2000;56(3):779–788.
26. Lin DY, Wei LJ. The robust inference for the Cox proportional hazards model. *J Am Stat Assoc*. 1989;84(408): 1074–1078.
27. Cole SR, Hernán MA. Adjusted survival curves with inverse probability weights. *Comput Methods Programs Biomed*. 2004;75(1):45–49.
28. Altman DG, Andersen PK. Calculating the number needed to treat for trials where the outcome is time to an event. *BMJ*. 1999;319(7223):1492–1495.
29. Hernán MA, Hernandez-Diaz S, Robins JM. A structural approach to selection bias. *Epidemiology*. 2004;15(5): 615–625.
30. Lund JL, Horváth-Puhó E, Komjáthiné Szépligeti S, et al. Conditioning on future exposure to define study cohorts can induce bias: the case of low-dose acetylsalicylic acid and risk of major bleeding. *Clin Epidemiol*. 2017;9:611–626.
31. Karim ME, Gustafson P, Petkau J, et al. Marginal structural Cox models for estimating the association between beta-interferon exposure and disease progression in a multiple sclerosis cohort. *Am J Epidemiol*. 2014;180(2): 160–171.
32. Robins JM, Hernán MA, Brumback B. Marginal structural models and causal inference in epidemiology. *Epidemiology*. 2000;11(5):550–560.