

We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900

Open access books available

186,000

International authors and editors

200M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



Evaluating Clinical Effectiveness with CF Registries

Rhonda Szczesniak and Bin Huang

Abstract

Treatment and disease registries have played a vital role in understanding the heterogeneous nature of cystic fibrosis (CF) disease progression. The maturity of so many patient registries and recent national focus on their potential to improve patient-centered outcomes have led to the establishment of guidelines for the conduct of registry data analyses. Despite the insights garnered from utilizing CF patient registries, the analyses are plagued with methodological challenges, such as confounding, missing data, time varying treatment and/or covariates, and treatment-by-selection bias. Nonetheless, these registry studies have been essential for CF clinical effectiveness research. They reflect real-world clinical practice and allow for evaluating patient outcomes in a realistic clinical environment. In this chapter, we reflect on these advancements in registries and study results broadly and specifically in CF. We identify the key statistical challenges with the analysis of CF registry data from start to finish, including design considerations, quality assurance, issues with selection bias, covariate effects, sample size justification and missing data. We describe how these approaches are implemented to answer clinical effectiveness questions and undertake an illustrative example on tobramycin effectiveness and lung function decline.

Keywords: confounding-by-indication bias, instrumental variables, lung function decline, propensity scores, treatment-selection bias

1. Introduction

A registry is “an organized system that uses observational study methods to collect uniform data (clinical or otherwise) to evaluate specified outcomes for a population defined by a particular disease, condition or exposure, and that serves a predetermined scientific, clinical or policy purpose(s)” [1]. Registries and other non-intervention studies are often referred to as *real-world* data to distinguish them from clinical trials or experimental studies.

Treatment and disease registries play a vital role in the advancement of patient-centered outcomes research. These patient registries often include data arising from patient surveillance in observational settings. Numerous epidemiologic studies have used patient registries to characterize disease progression. In more recent years, patient registries have been used for a variety of health-related inquiries, ranging from comparative effectiveness studies to informing clinical decision making at the point of care (see [2], for an example). The maturity of so many patient registries and recent national focuses on their potential to improve patient-centered outcomes have led to the establishment of guidelines for the conduct of registry data analyses [1].

Although these guidelines are recent, the statistical challenges posed in these observational settings were noted decades ago in epidemiology and public health research [3]. Indeed, registry analyses are plagued with methodological challenges, such as confounding, missing data, time varying treatment and/or covariates, and treatment-by-selection bias.

Despite these challenges, registry studies are essential for clinical effectiveness research. They reflect real-world clinical practice and allow for evaluating patient outcomes in a realistic clinical environment. A registry encompasses the general patient population, including those who are severely ill or less likely to adhere with assigned treatment. These patients commonly are excluded from the randomized controlled trials, and are likely to have very different treatment responses. Further, registry study offers the opportunity to examine important factors such as physician's practice behavior, prescription preference and other covariates pertaining to quality of care, which are impossible to assess in an experimental study. Registry studies commonly include long-term observation and therefore can reflect change of treatment practices, in order to provide a timely assessment of emerging research questions. The use of registry data to evaluate outcomes is of mutual benefit to both patients and clinicians, and it facilitates management of patient care, thereby improving the health care system.

1.1 Evaluating the effectiveness of tobramycin on lung function decline

Throughout the chapter, we will refer to an example from a retrospective longitudinal cohort study, which used the Cystic Fibrosis Foundation Patient Registry (CFFPR) to evaluate the clinical effectiveness of a treatment for lung function decline [4]. Cystic fibrosis (CF) is a lethal autosomal disease in which respiratory failure is the primary cause of death. *Pseudomonas aeruginosa* (*Pa*) is a common, chronic pulmonary infection in CF patients. Inhaled tobramycin (hereafter, Tobi) has been shown to improve lung function in CF patients with *Pa* in the clinical trial setting. In this example, it is our objective to evaluate the clinical effectiveness—as opposed to efficacy—of Tobi using the CFFPR. We will refer to this case study, in order to illustrate statistical methods for registry data analysis. The Appendix includes analysis implementation using SAS 9.3 (SAS Institute, Cary, NC).

In this chapter, we focus on the design and statistical analyses of patient registry studies. We begin in Section 2 by describing processes to design a study involving registry data, in accordance with the aforementioned guidelines from Gliklich and colleagues. We follow this section with overviews of inferential analyses methods that can be used in registry study to combat selection bias, missing data, time varying treatment or covariates in Section 3. In Section 4, we describe details of the application to the aforementioned patient registry. We discuss the utility of existing methods and remaining analytic challenges in Section 5. Finally, we provide an appendix in Appendix A with implementation of the statistical analyses in our illustrative application.

2. Design considerations for registry studies

Registries may be organized around conditions or exposures (e.g., a cystic fibrosis registry, stroke registry); a healthcare service (e.g., procedure); or a product (drug or device) and can address questions ranging from treatment effectiveness and safety to the quality of care delivered. Registries vary in complexity from simply recording product use as a requirement for reimbursement to more systematic efforts to collect prospective data on many types of treatment, risk factors, and

clinical events in a defined population. Follow-up can be retrospective, prospective, or a combination of both. The mode and duration of follow-up can range from days (e.g., hospital admission registry) to decades (e.g., orthopedic implant registry). Constructing and maintaining a large registry requires substantial resources, collaborative effort, and often requires a multi-center or inter-institutional agreement, and a governing body that oversees and coordinates all activities. Typically, there are standard guidelines or written procedures in place that help researchers to gain familiarity and/or access to the registry study.

Before utilizing data from any registry, it is imperative to define the research question and develop a study protocol. Clinical or public health questions of interest should be stated as research questions. Each research question should correspond to a testable hypothesis, which may be assessed using an approach fully described in the statistical considerations (this is particularly important for comparative effectiveness studies).

2.1 Selecting a registry and target population

Finding a registry that is appropriate to answer the research question of interest will require us to review preliminary information about each of the prospective registries, particularly regarding the data elements. For example, consider the following two studies. In each study, it is of interest to determine treatment effectiveness for cystic fibrosis (CF) lung disease. The first study utilized the Cystic Fibrosis Foundation Patient Registry (hereafter, CFFPR) [5] to examine the association between ibuprofen and lung function decline [6, 7]. In a subsequent study, Konstan et al. [8] assessed the relationship between a different treatment, dornase alfa, and lung function decline using registry data from the Epidemiologic Study of Cystic Fibrosis (ECSF) [9]. Although both studies examined treatment effectiveness on the same outcome (lung function decline), each study required distinct data elements to answer the research questions regarding treatment effectiveness. The CFFPR includes data collected on ibuprofen usage; however, the ECSF does not include information for this treatment, eliminating this database as an option for the first study. On the other hand, the ECSF has detailed information on pulmonary symptoms (e.g., coughing), which are known predictors of more rapid lung function decline [7] and therefore need to be considered as potential confounders to assessing treatment effectiveness. Although both registries include data elements to measure dornase alfa usage, which are necessary to answer the research question in the second study, the ECSF enabled the authors to consider detailed pulmonary symptoms as potential confounders. If our research question involves a newly diagnosed condition or rare disorder, we may be limited to a single patient registry. In those instances, the research question may need additional refinement.

In the study protocol, we will need to state the specific objectives. The objective of our CF study is to evaluate the effect of tobramycin on lung function decline. Once the objectives are clarified, we consider the most appropriate study design. In registry analyses, the selection of our study design often depends on how the registry was structured. Registries constructed to capture natural histories are often amenable to studies with longitudinal cohort designs. We can identify the population of interest at this point in the study protocol. Acquiring the subset of data which best reflects the population of interest, exposure variables, and primary and secondary endpoints may include some manipulation of the original registry data files. In our CF example, it is of interest to limit our cohort to individuals chronically infected with *Pseudomonas aeruginosa* (*Pa*). We target this population, since our research question is related to the effectiveness of tobramycin, which is a drug recommended for treating CF chronic *Pa* in patients with CF. In our example, we

Characteristics	Treated with tobramycin	Not treated with tobramycin	P-value ^a
Age, mean ± SD (n), y	12.82 ± 4.68 (6451)	12.78 ± 4.59 (6255)	0.84
Male sex, % (n)	47.2% (3046)	53.5% (3346)	<0.0001
FEV ₁ , mean ± SD (n), % predicted	74.46 ± 25.33 (6451)	83.69 ± 22.68 (6255)	<0.0001
Weight-for-age percentile, mean ± SD (n)	30.05 ± 26.08 (6446)	33.92 ± 26.88 (6252)	<0.0001
CF-related diabetes, % (n)	2.3% (150)	1.5% (96)	0.0012
Pancreatic insufficiency, % (n)	95.3% (6145)	94.8% (5932)	0.27
No or state/federal insurance, % (n)	32.3% (2082)	30.5% (1910)	0.0348
Prior hospitalizations ^b	—	—	—
None, % (n)	57.2% (3360)	75.7% (4448)	<0.0001
1, % (n)	23.9% (1401)	16.0% (940)	
2, % (n)	9.3% (546)	4.6% (273)	
3 or more, % (n)	9.6% (566)	3.7% (219)	
Dornase alfa, % (n)	79.3% (5116)	49.4% (3087)	<0.0001

Abbreviations: CF, cystic fibrosis; FEV₁, percentage predicted of forced expiratory volume in 1 s.
^aP-values from Wilcoxon Mann-Whitney or chi-square test.
^bNumber of hospitalizations in the year before baseline.

Table 1.
Descriptive analysis of CF registry variables.

determine chronic *Pa* status for each patient by examining the number of recorded *Pa* infections throughout the calendar year. Our primary endpoint is the mean change in FEV₁% predicted over a 2-year period. We selected additional exposure variables of interest, which are known predictors of change in FEV₁% (see **Table 1**).

2.2 Data elements and quality assurance

For many different types of research, particularly comparative effectiveness research or research involving children and/or rare disease conditions, no single institution has a large enough patient population to perform a proper study. This, along with the growing infrastructures of electronic medical records, has led to an increased effort to create distributed research networks. The widespread adoption of electronic health records (EHRs) has enabled them to become a main source for registry data, capable of capturing the necessary elements as part of routine clinical care, and the ever-changing clinical practices.

The number of data elements and scope of collection often increase over the life of the registry. Well-maintained registries typically include data dictionaries, but verifying data quality specific to our study is essential. In our CF example, we had to calculate specific variables for analysis. Understanding how the data have been collected over time and to what extent (e.g., every clinical encounter) will help determine the appropriate subset of data to extract from the registry. For example, the CFFPR data are collected at every clinical encounter and hospitalization, as well as on an annual basis, on each patient and provided to the CF Foundation. Using descriptive statistics, such as the 5-number summary, mean and standard deviation for each variable, and histograms or boxplots can highlight data discrepancies in continuous variables. Similarly, computing the frequency and percentage of each category in a nominal or ordinal variable may identify variables with questionable

entries. Furthermore, summary statistics stratified by calendar year can inform selection of an optimal time frame from natural history registries. In our example, CF-related diabetes, a known predictor of lung function decline that should be included in the analysis, was not collected in earlier calendar years in the CFFPR.

Access to most registries requires approval by a local institutional review board (IRB) prior to data release, and this approval is often necessary to have results of the study peer-reviewed and published. In our experience, developing a protocol that is in accordance with the aforementioned guidelines is sufficient for the IRB review. Although registries rarely contain patient names or medical record numbers, they often include clinical encounter and/or discharge dates. Having this type of protected health information in the data often requires IRB approval.

3. Statistical considerations for comparative effectiveness using registry studies

Statistical analyses in the registry data setting are subject to the statistical challenges previously described for analyses of observational studies [10]. Registries are often established for the purpose of evaluating the effects of interventions. The statistical analysis plan should include appropriate methods to test each hypothesis, methods to address biases and confounding arising from various sources, and sample size/power considerations.

3.1 Selection bias

Regardless of the research question, a registry study will likely be plagued with numerous sources of bias. Selection bias, although inevitable, is typically the most concerning. This type of bias distorts the results for the association of interest and may yield misleading results. Failure to sample from the correct target population and loss to follow-up due to death or some other event are types of selection bias.

A pervasive type of selection bias is confounding by indication, arising from nonrandomized treatment assignment that is often related to the patient's risk to experience poor outcomes. This treatment-by-selection bias creates distinctions between the risk profiles of treated and comparator groups and may violate statistical assumptions in our analyses. In our CF example, treatment selection bias may be more pronounced because the drug in question should only be prescribed to individuals with CF who have a specific chronic infection. Narrowing the cohort to "sicker" individuals can intensify the aforementioned risk profile imbalance between Tobii and non-Tobii groups.

Statistical methods to combat treatment selection bias have been applied in previous studies. Approaches to adjust for treatment selection bias include multivariable regression, propensity score methods, matching and instrumental variables analysis. Stukel et al. [11] applied each of these four approaches to examine the association between cardiac catheterization and long-term acute myocardial infarction mortality. The authors found that the results differed according to the choice of statistical approach. Next, we describe and outline each approach in the context of our CF example.

3.2 Statistical analyses of comparative effectiveness utilized for registry data analysis

3.2.1 Multivariable regression

In the absence of randomization, intervention and comparator groups may exhibit large differences with respect to observed covariates recorded in the

registry. This approach, sometimes referred to as covariate adjustment, attempts to account for such differences that may distort estimates of intervention effects (Figure 1). Most biomedical studies employ ordinal least squares (OLS) regression to adjust the association between the treatment indicator variable (T_i) and outcome variable (Y) for measured confounders (X_1, \dots, X_K). The OLS regression model for each subject ($i = 1, \dots, l$) specifies

$$y_i = \beta_0 + \beta_1 X_1 + \dots + \beta_K X_K + \gamma T_i + u_i \tag{1}$$

where β_0 is the parameter for the model intercept and μ_i is an error term. Each of the model parameters β_1, \dots, β_K correspond to the association between the measured confounder and outcome variable. The parameter for treatment effect is γ ; we denote its OLS estimate as $\hat{\gamma}$. OLS estimation requires that the error term (u) is not correlated with the measured confounders (X_1, \dots, X_K) or the treatment (T). Therefore, the only effect of T on outcome variable (Y) is the direct effect estimated as $\hat{\gamma}$. The challenge of utilizing multivariable regression model for comparative effectiveness is that we must appropriate account for necessary set of confounders. Failure to fully account for necessary confounder may lead to bias estimate of treatment effect.

3.2.2 Propensity score regression

The propensity score (PS) is a summary balancing score indicating the likelihood for a patient to receive the active treatment ($T_i = 1$) using observed set of confounders (X_1, \dots, X_K), represented in Figure 1. It is a balance score, because by conditioning on the propensity score, one could achieve independence between the treatment assignment and confounders; therefore, propensity scores help to achieve quasi-experiment design for natural occurring treatment assignment in a registry study. The PS can be estimated through a logistic regression modeling

$$\text{logit } P(T_i = 1) = \beta_0 + \beta_1 X_1 + \dots + \beta_K X_K \tag{2}$$

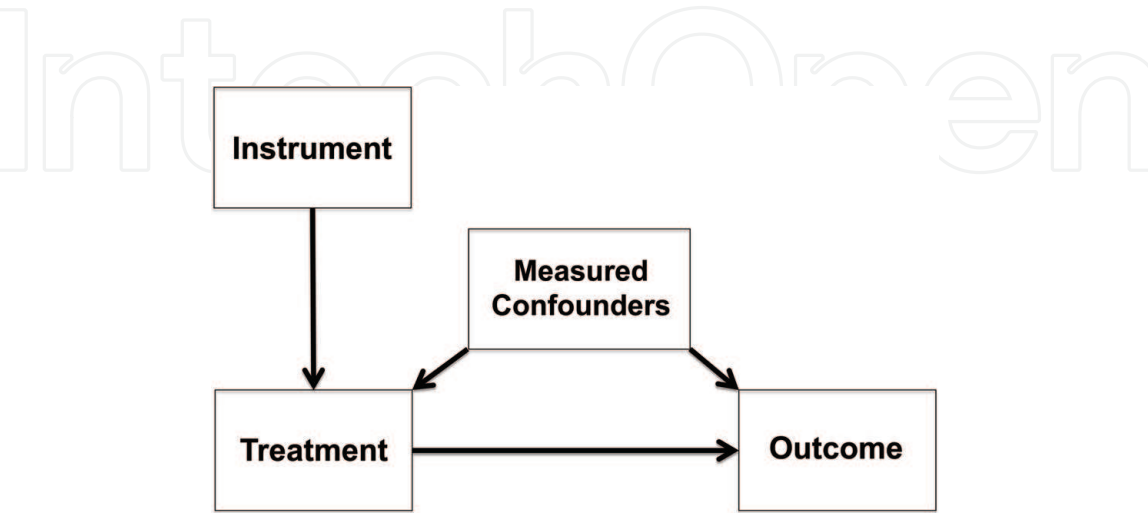


Figure 1. Causal diagram. The multivariable regression in Model (1) examines the treatment-outcome association, after adjustment for measured confounders. The propensity score methods outlined in Model (2) use the measured confounders to balance the treatment groups (exposure). The IV regression from Model (3) examines the treatment-outcome association, to the extent that the exposure is associated with the instrument. The instrument should not be related the measured confounders; therefore no arrow is drawn for this relationship.

where $\text{logit}^{-1}(p) = \log \frac{p}{1-p}$, and the propensity score is estimated by $\hat{\pi}_i = \text{logit}^{-1}(\beta_0 + \beta_1 X_1 + \dots + \beta_K X_K)$. There are several propensity score approaches: propensity score adjustment, stratified analyses by the quintiles of propensity score, propensity score sub classification matches treated and untreated patients on their propensity score sub-classes (often by percentiles), and inverse weighting of propensity scores. The first approach includes propensity score directly in the regression equation as a covariate to obtain adjusted treatment effect,

$$y_i = \beta_0 + \beta_1 \pi_1 + \gamma T + u_i \quad (3)$$

The second and the third approaches often categorize patients into five groups using propensity score quintiles. The stratified analyses will perform the regression model of $y_i = \beta_{0,k} + \gamma_k T_i$ for $k = 1, \dots, 5$, and estimate the treatment effect by $\gamma = \sum_{k=1 \dots 5} \gamma_k / 5$. The PS sub-classification matched analyses will be matching the Tobi and non-Tobi patients on their propensity score groups, then perform analyses for matched pairs. The propensity matching could also be performed on a finer grouping, for example, using 10 groups, or fine matching where a Tobi patient finding matching non-Tobi patient(s) through a distance measure. The method of inverse PS weighting assigns higher (lower) weight to patients who has lower (higher) propensity of receiving Tobi, where the weight is defined as $w_i = \frac{T_i}{\hat{f}_i} + \frac{1-T_i}{1-\hat{f}_i}$. The intuition behind the weighting approach came from the survey sampling method, and through inverse weighting, one could align the Tobi and non-Tobi patients to have comparable distribution of the confounders. There are advantages and disadvantage of each propensity score methods. Comparisons of these methods can be found in an excellent review paper by Austin and Mandani [12] and the references therein. Different methods are available for deriving propensity score. Other than the logistic regression, one could use more flexible classification and regression tree [13], boosted logistic regression [14], and covariate balancing propensity score method [15]. When applying PS approaches, it is important to check PS balance between the two treatment groups. Patients who have extremely high or low PS values that are not compatible with values from any patients in the other treatment group should be excluded from the PS analyses. The balancing check can be presented in graphic presentation, usually presenting the absolute standardized mean difference (SMD).

3.2.3 Instrumental variables (IV) analysis

One of our primary analysis goals in the registry setting is to identify potential sources of confounding and make the appropriate adjustments in our statistical analysis. Failure to identify sources of measured confounding results in residual confounding. This type of unaddressed confounding goes into the error term, u , which was introduced in Model (1).

Inferential results can also be impacted by what is known as unmeasured confounding. McClellan et al. [16] propose a technique known as instrumental variables (IV) to combat both measured and unmeasured confounding. We introduce the following notation for IV regression. From Model (1), recall that the variables (Y_i, T_i, X_i) correspond to data from the i^{th} patient in the registry, and we assume that there is no correlation between the treatment variable, T_i , and the error term, u_i . This correlation is present when patients receive treatment based on unmeasured characteristics. Let R_i represent an instrument. Consider the following example of a randomized controlled trial. If R_i represents random assignment to

treatment, it is the ideal instrument. By construction, it is related to outcome only through treatment assignment [17].

In the typical clinical setting, a provider does not flip a coin to determine whether she will prescribe her patient treatment A, as opposed to some alternative. By construction, *real-world* data contained in registries represent non-random assignment to treatment. Instead, we identify a variable—“an instrument”—that is related to the outcome only through treatment. The variable R_i is a valid instrument, provided the following assumptions are met:

- i. R_i is associated with the treatment variable or exposure of interest, T_i ;
- ii. R_i is not directly associated with the outcome, Y_i ; R_i is only associated with Y_i through the treatment variable, T_i

Fortunately, assumption (i) is testable by performing least-squares regression of the proposed IV on the treatment variable and measured confounders:

$$T_i = b_0 + \lambda R_i + b_1 X_1 + \dots + b_K X_K + \delta_i, \quad (4)$$

where b_0 is the intercept; b_1, \dots, b_K are the parameters corresponding to the aforementioned measured confounders, X_1, \dots, X_K ; λ is the parameter estimate of the association between the treatment variable, T_i , and the IV, R_i . The magnitude of this association is a measure of the strength of the instrument [17]. Higher magnitude corresponds to greater strength. Let T_i be the resulting prediction of the treatment value, obtained from Model (4). This association is illustrated in **Figure 1** by the arrow moving downward from the instrument to exposure.

We continue this approach, often referred to as two-stage least squares regression, by substituting T_i from Model (4) into the multiple linear regression defined in Model (1):

$$y_i = \beta'_0 + \beta'_1 X_1 + \dots + \beta'_K X_K + \gamma' \hat{T}_i + \eta_i \quad (5)$$

In this regression, the same method of estimation is used; however, we use distinct notation because parameter estimates and residual error will differ from Model (1). Finally, we use the estimate of γ' from Model (5) for our interpretations of treatment effect on the outcome. This estimate corresponds to the association in **Figure 1** from treatment to outcome. Note that it is the same path as the multiple linear regression, but the treatment effect has been “instrumented.” Assumption (ii) cannot be formally tested, but can be explained in the context of the registry analysis at-hand. We provide this type of explanation in our illustrative application. Sensitivity analyses are imperative to determine the robustness of the IV. We recommend analyzing the data in subgroups to understand how these groups may drive heterogeneous treatment effects.

3.3 Time varying treatment/exposure and covariate

Incorporating time-varying treatment and/or covariate effects is a pervasive issue in registry data analyses. The fundamental challenge arising from the change in treatment and covariates over time often results from a patient’s responses and/or experiences with the previous treatment assignment. Thus, simply including the time varying treatment or covariate in such cases could induce bias in estimating treatment effect. Special attention is needed to address this issue when analyzing

registry data. Relatively few statistical approaches are available to assess time-varying treatment effects or intermediate outcomes. Hogan and Lancaster [18] proposed inverse probability weighting and instrumental variables as time-varying treatment approaches; another population-based approach is the G-computation formula [19].

3.4 Sample size justification

Completing this process implies that we have carefully considered the hypothesis test and analysis variables, ultimately arriving at a statistical model that will rigorously address the research question. Sample size assessments will differ according to the statistical approach proposed to test the hypothesis, and should incorporate previously established public health or clinical information.

If the statistical approach entails adjustment for confounding and other sources of bias, the sample size calculation is often straightforward. Suppose we plan to test the significance of the treatment effect, γ' , previously defined in Model (1), and we have already identified measured confounders (i.e., covariates) that should be included in the model, referred to as X_1, \dots, X_K . Our null hypothesis corresponds to $\gamma = 0$, while our alternative hypothesis corresponds to $\gamma \neq 0$. Testing this hypothesis corresponds to determining sample size/power for a multiple linear regression model [20].

We now reconsider the importance of sample size justification for analyses involving a large registry. Statistical significance depends on the sample size and is typically declared if the P value obtained from the test statistic falls below a predetermined threshold (e.g., 0.05). This type of significance may be reached in any study, provided that the sample size is large enough; therefore, in addition to this mathematical criterion, we recommend specifying conditions that must be met to achieve practical (public health or clinical) significance within the context of the research question. In biomedical studies, these criteria can often be defined by determining the minimal clinically important difference (MCID). This technique was originally proposed for clinical trials [21] but has spawned several other approaches [22] to determine the MCID. Once we incorporate the MCID into our null and alternative hypothesis statements, we can perform the sample size calculation that corresponds to our proposed inferential analysis.

3.5 Missing data mechanisms and missing data modeling

Missing data can occur in the registry setting for a variety of reasons. Simply put, a missing data point is an observation that should have been recorded; however, for some reason, it was not recorded. It is our desire, as analysts, to understand the reason for this “missingness.” In this section, we outline practical analytic approaches to identify potential sources attributable to missing data and methods to combat the resulting bias. We begin with a brief description of the three fundamental missing data mechanisms. For an elegant mathematical treatment of the distinctions among the mechanisms, we refer the reader to the original work by Rubin [23].

3.5.1 Missing completely at random (MCAR)

If the registry data are MCAR, then the reason for missingness is not related to the data that we were able to observe or to the data that we were not able to observe.

We now consider the CF example. MCAR could correspond to the following. The probability of a lung function observation (the outcome variable) being missing from the registry does not depend on any of the observed data (e.g., patient's age) or any of the unobserved data (e.g., having lower lung function does not alter the risk of the observation being missing). Our analysis results from this subset of data will be no different (aside from larger standard errors) than if we had been able to perform the analysis on the entire dataset.

3.5.2 Missing at random (MAR)

This assumption is more relaxed than MCAR but still has specific requirements. For MAR to hold, the missingness cannot be related to unobserved data, given what we have been able to observe. In other words, the missingness can depend upon data that we have already observed (i.e., data entries that were recorded in the registry). Referring again to our CF example, the probability of a lung function observation being missing does not depend upon the actual lung function value, provided that we have the other covariate data. In this case, missingness can depend upon characteristics that have been recorded in the CFFPR (e.g., gender).

3.5.3 Missing not at random (MNAR)

We are more likely to encounter this mechanism in registry data, compared to the other mechanisms. If data are MNAR, then the missingness is related to unobserved data (unlike MAR). The missing observation follows a different distribution than the observed data, regardless of whether the two types of data have other characteristics that are the same. Despite the fact that we have registry data, the data that we are able to observe are not representative of the entire population. Within the CFFPR example, consider the longitudinal data. According to CF Foundation guidelines, patients are supposed to have at least one pulmonary function test per quarter [5]. Suppose there is a subset of patients who do not have lung function data recorded at every clinical encounter. There are many plausible explanations for why these data are missing. For an individual patient, there may be a lack of interest in managing his disease progression, or it could be an entry error. In general, we may lack relations to observed values or those relations may be irrelevant.

In practice, we do not have the information necessary to declare the reason for the missingness. Even thoughtfully developed, well-maintained registries will have missing data; therefore, sensitivity analyses are needed as part of the statistical considerations. As a preliminary step, we recommend creating an indicator (dummy) variable to indicate whether the observation is missing (=1) or otherwise (=0). Regress this dichotomous variable on the other variables to determine whether the missing indicator is associated with observed characteristics. If no association is found, we may conclude that the data are MCAR; however, we still encourage caution when making the MCAR assumption for statistical models using registry data. Although small sample size may produce this result, it is not a likely culprit in settings with large data sources. It is possible that the extent of the missingness may be too low (e.g., 5% of observations are missing) to substantially alter results, but having a low proportion of missing observations is also unlikely in a registry setting. If there is a significant association from our preliminary regression with the indicator variable, then we can rule out the MCAR assumption and more intently investigate the MAR and MNAR assumptions.

We can further examine the MAR assumption by checking for variables that are often missing simultaneously or other potential patterns of missingness. Whenever possible, we recommend performing the analysis under the MAR assumption. The two most common approaches under this mechanism are direct modeling and multiple imputation. Direct modeling implies that we will consider all available data points in our parameter estimation. This method is sometimes referred to as “available case analysis” [24]. In other words, the analysis will not exclude the records of any individual subject who has at least one observed entry. There is a second approach, multiple imputation [25], which has gained favor among analysts with the expansion of computing resources. To perform this approach, several data points for each missing data point are generated, resulting in several distinct datasets. We employ our proposed statistical model separately on each dataset and obtain parameter estimates. The estimates are combined to produce an aggregate estimate. The aggregate estimate and standard error are used to make interpretations of results. This technique is available in many software packages (e.g., SAS `proc mi`, `proc mianalyze`).

Unfortunately, there is no way to know whether the data are MAR or MNAR. Previous work by experts in the analysis of missing data has shown that any model we develop under the MNAR assumption will have an equivalent MAR counterpart [26]. Developing an MNAR model requires technical steps that are beyond the scope of our current chapter. Dmitrienko et al. [27] provide an applied approach to investigating MNAR assumptions in the context of sensitivity analyses. Although their text focuses on analyses for data from clinical trials, their approach and accompanying SAS implementation may be adapted to registry data analyses.

3.6 Interpretations of registry data analyses

To simplify interpretation and improve accuracy of the results, sources of potential confounding (measured or unmeasured) should be considered as much in advance as possible. Propensity score regression offers an effective method to further balance the treatment and non-treatment groups. Like multivariable regression, this approach accounts for treatment selection bias [28] only for measured confounders (e.g., measured comorbidities and severity of illness). The propensity score could utilize measured confounders to remove treatment-selection bias. However, when there are unmeasured confounders that determine treatment-selection bias, the propensity-score approach will be limited. In analyzing registry data, IV analyses should be considered when unmeasured confounders are suspected.

Although the IV analysis is a powerful approach, this method has some noteworthy constraints. Large sample size is essential for performing IV analysis, but this issue may not be a challenge in the registry setting. The IV must only affect treatment assignment and have no direct association with outcome. If these assumptions are satisfied, then the IV analysis will yield a consistent estimate of the average causal effect [29]. Assumption (i) is directly testable, but making a heuristic argument for assumption (ii) is a common approach. See Kahn et al. [30] for an example. A weak IV will produce larger standard errors and may lead to incorrect inferential results. This approach is ideal in the presence of small/moderate confounding but becomes less reliable in the presence of large confounding. Admittedly, this is a limitation of the IV analysis in the registry setting. On the other hand, an appropriate IV minimizes the potential impacts of measured and unmeasured confounding [31].

Sensitivity analyses should be performed to examine potential impacts of missing data and particular subgroups that may drive inferential results. Analyses corresponding to the missing at random assumption should be explored in the registry setting. Subgroup analyses are essential to identify heterogeneous treatment effects, particularly in the IV analysis. These sensitivity analyses should be performed regardless of the statistical model that we choose to employ.

4. Illustrative application

4.1 Data summary and descriptive analysis

The CFFPR contains data on individuals receiving care from any CF center in the United States, which has been accredited by the CF Foundation. Like many registries, we underwent an application process to receive the data. The CFFPR data that we received were in separate databases. We used the following two databases. The encounter-level database had one record per patient, per clinical encounter. The annual-level database contained one record per patient, per year. We merged these data to extract the information necessary to determine whether there is a significant association between the use of inhaled tobramycin and lung function in individuals with CF who are chronically infected with *Pa*. Our primary outcome, lung function, was defined as mean change in FEV₁% predicted (FEV₁). In this application, we study short-term effectiveness of inhaled tobramycin, in order to facilitate use of instrumental variables, which still pose several challenges in longitudinal settings with multiple data points and time-varying exposures [17].

We considered the following restrictions to target the study cohort of interest. We requested CFFPR data ranging from January 1, 1998 to December 31, 2009, in order to capture the time at which inhaled tobramycin (Tobi) was recorded in the registry on a consistent basis. We did not consider study records with individuals <6 years of age, due to limitations of modality to measure lung function in young children. We limited the maximum age to 21 years, in an effort to focus on first occurrence of chronic *Pa*. We identified the first chronic *Pa* infection for each individual by examining all *Pa* culture results available in the encounter-level data. Patients recorded as having a positive *Pa* culture more than 50% of time in a given year were considered as eligible for the study. This was determined by using the *Pa* culture (indicator) variable available in the CFFPR. We took the first year that the patient had chronic *Pa* infection as the baseline year. In an effort to keep our study data to one record per patient, we only considered the first chronic *Pa* infection for each patient. Patients who also had another infection at the same time, *Burkholderia cepacia* complex, were not considered as part of the study cohort, because of previously established criteria [32]. An indicator variable for patient-level tobramycin use was defined as receiving inhaled tobramycin within 6 months of initial chronic *Pa*. Baseline FEV₁ was defined as the closest FEV₁ measurement recorded within 6 months after initial chronic *Pa* record. Follow-up FEV₁ was defined as the closest recorded FEV₁ within 1.5–2.5 years of the baseline FEV₁. Patients who did not have a recorded FEV₁ measurement within 6 months after meeting criteria for chronic *Pa* infection were excluded. The outcome variable, decline in FEV₁, was calculated as the difference between follow-up and baseline FEV₁ for each patient. A negative value implies that FEV₁ declined over the 2-year period; a positive value indicates that FEV₁ increased over the 2-year period. **Figure 2** illustrates steps to determining the analysis cohort and resulting sample size.

We identified potential confounders by looking at previous literature (see [6], for example). These variables, measured in the CFFPR, included gender, baseline

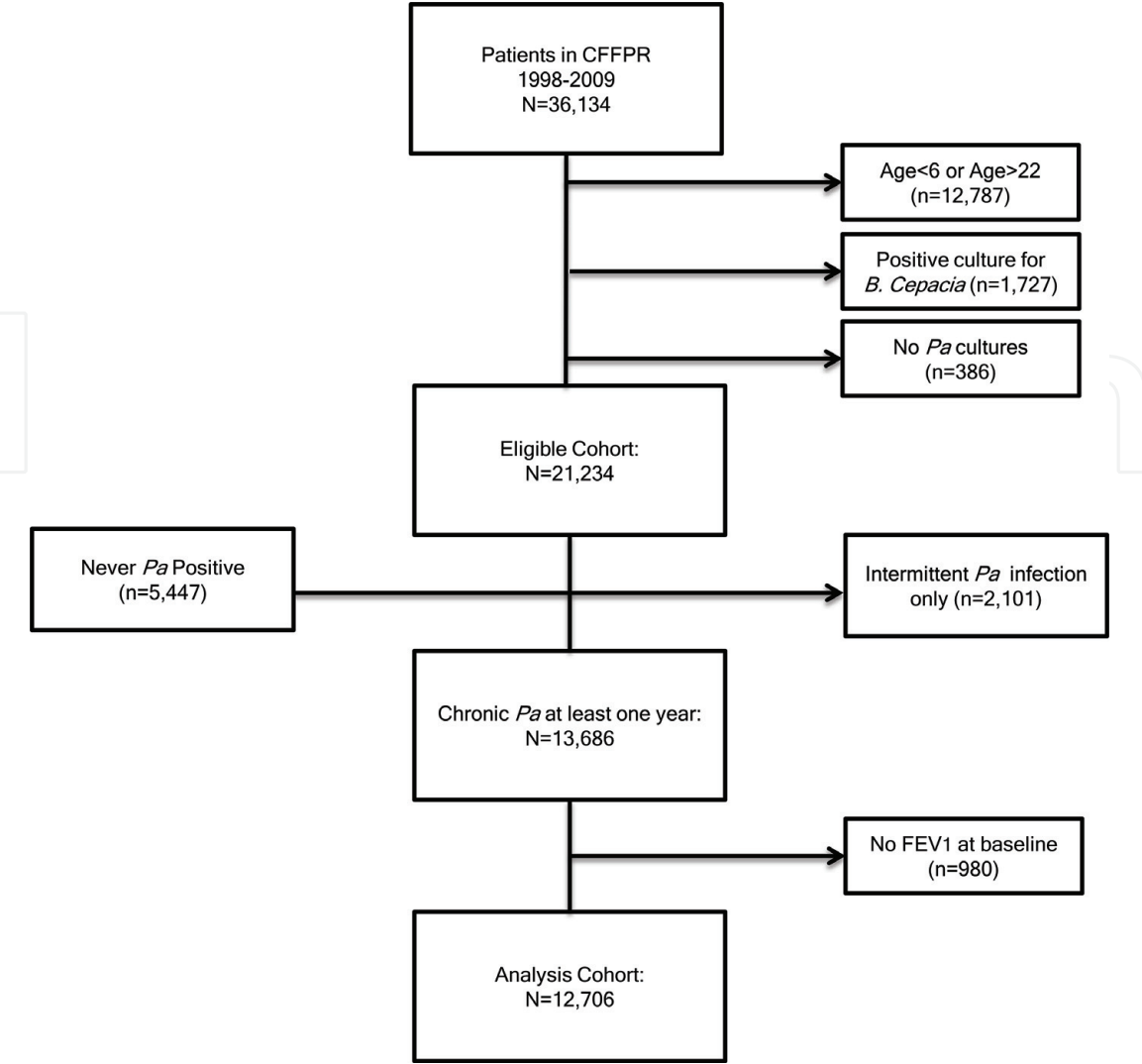


Figure 2.
Diagram of study population in the illustrative CF example, showing inclusion and exclusion steps to obtain an analysis cohort from the registry. CFFPR, Cystic Fibrosis Foundation Patient Registry; Pa, Pseudomonas aeruginosa.

measurements for age, FEV₁, weight-for-age percentile, insurance coverage, CF-related diabetes (with or without fasting hyperglycemia), dornase alfa use, pancreatic insufficiency (defined as taking pancreatic enzymes) and number of hospitalizations in the preceding year. We can compare Tobi and non-Tobi groups with respect to each of these variables using basic inferential testing (i.e., nonparametric test for continuous variables and Chi-square test for categorical variables). Results of the descriptive analysis are presented in **Table 1**. Our descriptive analysis reveals that Tobi and non-Tobi groups differed by several demographic and clinical characteristics. We note that the groups did not differ according to age or being pancreatic insufficient. Next, we utilize the aforementioned statistical models to test this association.

4.2 Multiple linear regression

We use Model (1) to test the association between lung function and tobramycin use, adjusting for potential confounders as covariates, represented as X_1, \dots, X_K . **Table 2** shows the results of the multiple linear regression, which suggest that the treated group experienced greater mean decline in FEV₁% predicted than the

	Type of model	
	Multiple linear regression ^a	Propensity score regression ^b
Covariates	Coefficient (SE), (P-value)	Coefficient (SE), (P-value)
Patient tobramycin use		
Treated	−1.74 (0.31) (<0.0001)	−1.71 (0.30) (<0.0001)
Not treated	0	0
Age	−0.87 (0.04) (<0.0001)	−0.86 (0.04) (<0.0001)
Baseline FEV ₁	−0.27 (0.01) (<0.0001)	−0.27 (0.01) (<0.0001)
Sex		
Female	−1.16 (0.30) (<0.0001)	−1.15 (0.31) (0.0002)
Male	0	0
Weight-for-age percentile	0.06 (0.01) (<0.0001)	0.05 (0.01) (<0.0001)
CF-related diabetes		
Yes	2.06 (1.44) (0.15)	2.19 (1.36) (0.11)
No	0	0
Pancreatic insufficiency	—	—
Yes	0.52 (0.83) (0.54)	0.44 (0.83) (0.60)
No	0	0
Insurance	—	—
None or state/federal	−1.66 (0.34) (<0.0001)	−1.66 (0.34) (<0.0001)
Other	0	0
Baseline hospitalizations ⁺	—	—
None	5.05 (0.70) (<0.0001)	4.63 (0.69) (<0.0001)
1	2.74 (0.74)	2.26 (0.74)
2	0.40 (0.87)	0.37 (0.87)
3 or more	0	0
Dornase alfa use	—	—
Yes	−0.46 (0.39) (0.25)	−0.38 (0.40) (0.34)
No	0	0
Abbreviations: CF, cystic fibrosis; FEV ₁ , percentage predicted of forced expiratory volume in 1 s.		
^a For each categorical variable in the first-stage model, the coefficient is the difference in patient tobramycin use between the indicated category and the reference category (labeled as coefficient = 0). For each continuous variable, it is the change in patient tobramycin use when the variable is increased by 1 unit. A negative value implies decreased patient tobramycin use.		
^b Predicted treatment obtained in Stage 1 serves as propensity score in Stage 2. For each categorical variable in the second-stage model, the coefficient is the difference in FEV ₁ decline between the indicated group and the reference group (labeled as coefficient = 0).		
⁺ significant at 2-sided p value < 0.05		
For each continuous variable, it is the change in FEV ₁ when the variable is increased by 1 unit. A negative value implies greater FEV ₁ decline.		

Table 2.
Multiple linear regression and propensity score method to predict lung function decline.

untreated group. Although most covariates were statistically significant at $P < 0.05$, we found that CF-related diabetes, pancreatic insufficiency, and dornase alfa use were not significant predictors of outcome.

4.3 Propensity score method

The patient characteristics at the baseline, which are known to impact FEV outcomes, are considered into the multivariable logistic regression model (Eq. (2)) for estimating propensity scores. **Figure 3** presented the histograms of propensity score for the Tobis treated and not-treated patient groups, showing different but overlapping propensity scores between the two groups. Propensity scores are grouped into five groups by quintiles. The distribution of propensity scores are compared between the Tobis treated and not treated patients within each of the five PS categories; as one could see from **Figure 4**, within each quintile categories, the two patient groups present comparable patterns in their likelihood of receiving

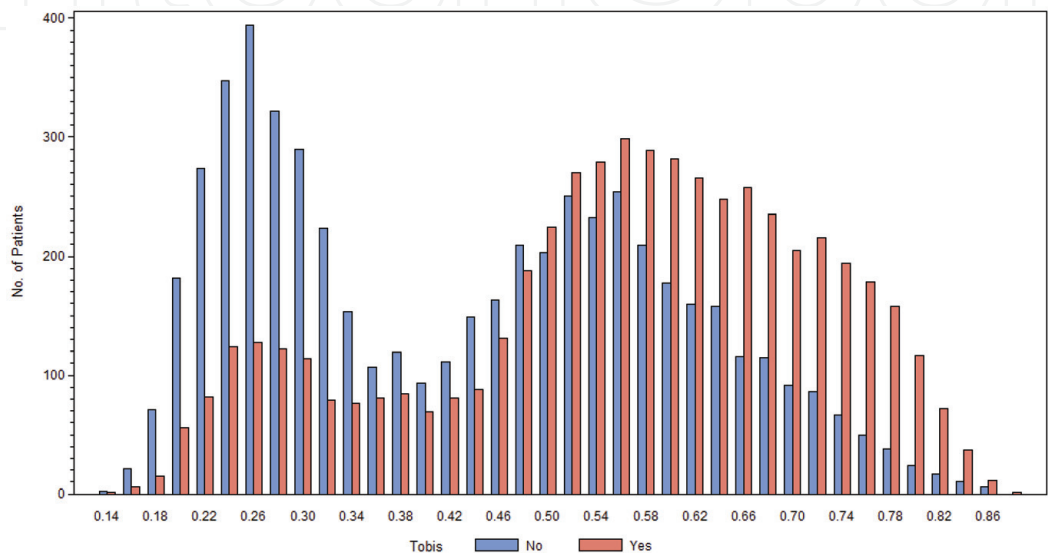


Figure 3. Histogram of the propensity score distributions by Tobis use (red) and not-group groups (blue). Related the measured confounders; therefore no arrow is drawn for this relationship.

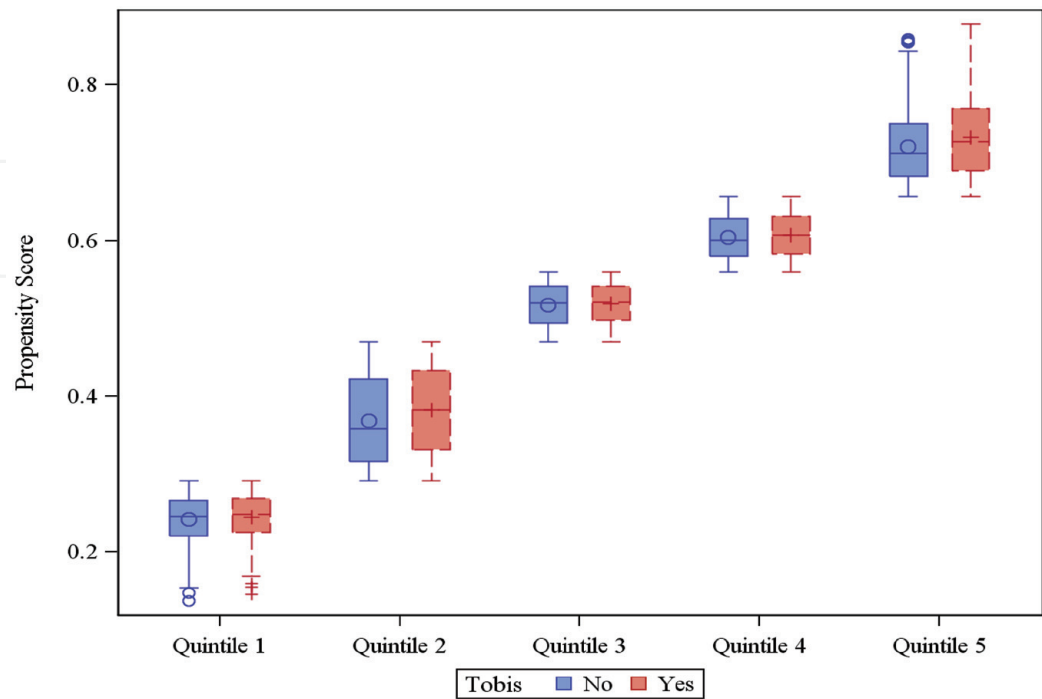


Figure 4. Box-Whisker plots of the distribution of propensity scores by Tobis use (red) and not use (blue) groups stratified by the quintiles.

Characteristics	Tobi	Level	Before PS matching		After PS matching		After IPW weighting	
			Mean	P-value	Mean	P-value	Mean	P-value
Sex	Treated	Male	47.8%	<0.01	49.1%	0.44	50.8%	0.91
	Not treated		54.0%		48.2%		50.9%	
FEV ₁ % predicted	Treated		76.38	<0.01	81.68	0.86	81.26	0.85
	Not treated		85.25		81.75		81.17	
Age	Treated		12.10	0.51	11.98	0.73	12.01	0.94
	Not treated		12.05		12.01		12.01	
Weight-for-age percentile	Treated		30.24	<0.01	30.34	0.80	32.33	0.79
	Not treated		33.78		30.18		32.19	
CF-related diabetes	Treated	Yes	1.5%	<0.01	1.3%	0.53	1.2%	0.49
	Not treated		1.0%		1.2%		1.4%	
Pancreatic insufficiency, % (n)	Treated	Yes	96.0%	0.94	96.6%	0.85	96.5%	0.72
	Not treated		96.0%		96.7%		96.6%	
No or state/federal insurance	Treated	None or state/federal	30.7%	0.53	30.2%	0.62	30.7%	0.81
	Not treated		30.2%		30.7%		30.5%	
Prior hospitalizations	Treated	None	58.5%	<0.01	69.1%	0.12	67.4%	0.95
		1	23.8%		18.9%		19.8%	
		2	9.4%		6.3%		6.9%	
		3 or more	8.4%		5.7%		5.9%	
	Not treated	None	75.9%		70.2%		67.4%	
		1	16.3%		19.5%		19.7%	
		2	4.6%		5.8%		6.9%	
		3 or more	3.3%		4.6%		6.0%	
Dornase alfa	Treated	Yes	77.6%	<0.01	68.7%	0.24	63.2%	0.88
	Not treated		49.3%		67.4%		63.3%	

Abbreviations: CF, cystic fibrosis; FEV₁, percentage predicted of forced expiratory volume in 1 s; PS, propensity score. Calculations for standardized differences are described in Section 4.3.

Table 3.
Standardized difference (T-val) between Tobi treated and untreated patients.

Tobi. To check for propensity score balance, we compared the Tobi treated and not treated patients on their baseline covariates, the standardized differences between the treated and not treated groups are presented in **Table 3**. The results show that there is a significant difference between the treated and not treated patients groups according to their gender, baseline FEV₁, CF-related diabetes, pancreatic insufficiency, insurance status, prior hospitalization and dornase alfa use. After matching patients on their PS categories, as well as after adjusting by inverse propensity score weighting, we are able to achieve balance between the Tobi treated and not treated groups. Subsequently, we proceed with the propensity score analyses using the inverse propensity score weighted approach. The results are presented in **Table 4**, which can be contrasted with the results from the multivariable regression analyses in **Table 2**. The results from these two approaches are very similar; both are

	Stage 1 (predicts patient tobramycin use) ^a	Stage 2 (predicts change in lung function) ^b
Covariates	Coefficient (SE), (P-value)	Coefficient (SE), (P-value)
Patient tobramycin use	—	—
Treated	—	2.55 (1.22), (0.0366)
Not treated	—	0
Age	−0.013 (0.003), (0.0002)	−0.86 (0.04), (<0.0001)
Baseline FEV ₁	−0.010 (0.001), (<0.0001)	−0.27 (0.01), (<0.0001)
Sex	—	—
Female	0.112 (0.027), (<0.0001)	−1.23 (0.30), (<0.0001)
Male	0	0
Weight-for-age percentile	−0.000 (0.001), (0.74)	0.06 (0.01), (<0.0001)
CF-related diabetes	—	—
Yes	0.112 (1.27), (0.38)	1.93 (1.44), (0.18)
No	0	0
Pancreatic insufficiency	—	—
Yes	0.064 (0.074), (0.39)	0.52 (0.83), (0.54)
No	0	0
Insurance	—	—
None or State/Federal	−0.128 (0.030), (<0.0001)	−1.58 (0.34), (<0.0001)
Other	0	0
Baseline hospitalizations ⁺	—	—
None	−0.598 (0.064), (<0.0001)	5.44 (0.69), (<0.0001)
1	−0.251 (0.068)	2.89 (0.74)
2	−0.148 (0.080)	0.48 (0.87)
3 or more	0	0
Dornase alfa use	—	—
Yes	0.224 (0.036), (<0.0001)	0.28 (0.40), (0.48)
No	0	0

Abbreviations: CF, cystic fibrosis; FEV₁, percentage predicted of forced expiratory volume in 1 s.
*Each model is adjusted for measured confounders by including each listed variable as a covariate. For each categorical variable, the coefficient is the difference in FEV₁ decline between the indicated category and the reference category (labeled as coefficient = 0). For each continuous variable, it is the change in FEV₁ decline when the variable is increased by 1 unit. A negative value implies greater FEV₁ decline.
^aMultivariable analysis with standard adjustment for confounding by including characteristics as covariates.
^bMultivariable analysis weighted using propensity scores.

Table 4.
Instrumental variable analysis to predict lung function decline*.

suggesting negative Tobi treatment effect on the improvement of FEV. The results from randomized clinical trials, however, all suggest a positive Tobi treatment effect. Such differences might be explained by unmeasured confounding that is

related to treatment selection bias but not recorded in the registry. We further proceed with IV analyses to examine the Tobi treatment effect.

4.4 Instrumental variables analysis

It is possible that the discrepancy between the previously described registry analysis and clinical trial findings of the treatment effect are due to unmeasured confounding. It is common in observational settings to encounter confounding by indication bias that is not recorded in registries. In this application, we selected a preference-based instrument, center-level prescribing patterns, to combat this bias. The CFFPR includes more than 240 centers. For each center, we calculated the tobramycin-prescribing rate during the time frame of the study. This rate was calculated as the number of times the center prescribed tobramycin to the patient when eligible divided by the total number of times the center should have prescribed tobramycin. We considered a patient to be eligible for the treatment once he met the CFF guidelines for its use.

We had to determine whether the IV met the previously mentioned criteria to be a valid instrument. We began by performing the first-stage analysis outlined in Model (4). We include all potential confounders as explanatory variables, and we include the IV. The response variable in this equation is the tobramycin use. The first-stage results are presented in **Table 4** and reflect what we found in the exploratory analysis from **Table 1**. The IV included in this regression was a highly significant predictor of tobramycin use. The corresponding t -statistic was 28.2, $P < 0.0001$. These results indicate that we have met assumption (i) for center-level prescribing to be a valid instrument. We also note that **Table 4** shows that dornase alfa use is strongly associated with tobramycin use. We will revisit this finding in sensitivity analysis of our instrument. We performed the multiple linear regression specified in Model (5) to determine the association between tobramycin and lung function decline. This regression accounts for observed patient characteristics and provides an instrumented version of tobramycin use. The last column in **Table 4** shows that tobramycin was associated with less FEV₁ decline, suggesting the existence of a positive treatment effect.

Assumption (ii) is not directly testable, but we examine it through sensitivity analyses of heterogeneous treatment effects. These effects may be caused by confounding from other medication use or differences in quality of care received across centers. We performed three different types of sensitivity analyses. First, we extracted quality of care markers through the CF Foundation Annual Report (1) and calculated them for each center. We correlated each marker with our IV and found no significant association. Second, we used subgroup analyses to determine the impact of dornase alfa use on tobramycin effectiveness. We divided the cohort into two distinct groups according to whether they reportedly used dornase alfa. We performed the IV analysis separately on each group. The two sets of results were similar with regard to first- and second-stage analyses. Third, we performed a secondary analysis of patients with *B. cepacia*. Although these patients are traditionally excluded from clinical trials and other effectiveness assessments because of their significantly poorer outcome, they often receive tobramycin in clinical practice. The first-stage analysis of this cohort was similar to the primary results; however, their second-stage analysis showed no significant treatment effect.

4.5 Concluding remarks

Registry data plays an increasingly important role in health care research. Appropriate design and careful statistical approaches to the analyses of registry data

are essential. In this chapter, we have described a step-by-step approach to formulating and implementing a registry data analysis. Understanding the research question, selecting the appropriate data source and identifying potential sources of bias are necessary before beginning to construct an analytic plan. The statistical considerations should include data quality assessments and descriptive analyses, and it is critically important to address selection bias due to both measured and unmeasured confounding. This is because selection bias is ubiquitous; failure to adequately address selection bias will lead to biased conclusions. Multivariable regression has been the primary means to combat selection bias. While this technique can help to minimize differences between groups, it is limited to relatively fewer covariates in the adjustment process. Propensity scores, which correspond to the probability of treatment assignment given pre-treatment characteristics, provide a way to summarize multiple covariates into a single score for each individual. Therefore, this approach is capable of handling a large dimension of confounders, which is particularly useful in registry studies when confounders are measured. Another advantage of PS is that it allows one to check between the treatment groups when conditioning on propensity score whether the confounding factors is balanced out. However, when important confounders are not measured, the PS method is limited. One solution is to perform sensitivity analyses by evaluating how estimated treatment effectiveness might change if there exists an unmeasured confounder with varying levels of prevalence. Such practice will allow one to gauge the impact of unmeasured confounders to the treatment effect.

In this example, the likelihood of tobramycin use depends on unmeasured characteristics at the patient, family or care level. The adjustment of unmeasured confounding that is possible through IV analysis may have led to more intuitive conclusions regarding treatment effect. Since CF care is organized by care center, it was reasonable to examine the validity of a preference-based instrument to combat treatment-selection bias. Thorough sensitivity analyses are necessary to examine the robustness of the IV. We limit our illustrative application to a single instrument. It is possible to include multiple instruments and gain more formal properties to testing assumption (ii).

5. Conclusions

When designing and analyzing registry data, it is critically important to address biases and confounding that are inherent in this type of study. Although we have focused, in this chapter, on describing methods for controlling selection biases, registry data are often subject to other types of biases related to measurement and miss-classification error, immortal time bias, loss to follow up, and missing data. We encourage use of sensitivity analyses to understand the impacts of these potential biases to the study conclusions. There are rich literature sources and several guidelines for design and analysis of registry data. In addition to the literature referenced in this chapter, a very useful resource is the recent report on standards in the conduct of registry studies for patient centered outcomes research and the references therein [33].

In addressing selection bias, most often, treatment effects are examined using multiple linear regression with measured confounders included as covariates [34]. Increasingly, PS methods are employed. However, existing statistical methods to address unmeasured confounding may be underutilized in registry settings. The models that we have presented are by no means exhaustive. There is room to develop more methodology, particularly to combat time-varying treatment effects and utilize time-varying instruments [12]. It is possible that preference-based

instruments will provide a feasible approach to interrogating registries [14]. Admittedly, there are some situations, such as the IV regression specified in Model (3), where the sample size/power analysis calculation is not straightforward. There are approaches to simulate power for this model, but additional assumptions are necessary. Furthermore, in most controlled studies, we can follow up with subjects who drop out. We rarely have this capability in registry settings, which further limits our ability to diagnose the missing data mechanism.

Acknowledgements

We are grateful to the Cystic Fibrosis Foundation Patient Registry Committee for dispensing the data utilized in the illustrative application. We thank Laurie Kahill, M.S., for information regarding the process of center-specific reporting for this registry. **Tables 1, 2 and 4** reprinted with permission of the American Thoracic Society. Copyright © 2014 American Thoracic Society [4].

Conflict of interest

The authors have no relevant conflicts of interest to report.

A. Appendix

Below, we present code from SAS 9.3 (SAS Institute, Cary, NC) to implement the statistical analyses for the application in Section 5.4. See Leslie and Ghomrawi [35] for additional details on the implementation of instrumental variables regression using the QLIM procedure in SAS.

*/*For each implementation below, we begin with *analysis_data*, which is the cleaned version of the registry data with all coded variables necessary for analyses. The variable *Tobi* is the indicator variable for whether the subject received tobramycin; *dfev1* refers to the outcome variable (change in FEV₁% predicted). First, we examine the initial difference between the treated and untreated groups.*/**

```
title 'Unadjusted Analysis';
proc ttest data=analysis_data;
class Tobi;
var dfev1;
run;
```

*/*The code below performs a multivariable linear regression to determine the association between tobramycin and change in lung function, with adjustment for the previously described measured confounders. The variables below correspond to sex (*gender*), baseline measures of age (*age*), FEV₁% predicted (*base_fev1*), weight-for-age percentile (*wt pct*), insurance coverage (*inscat*), CF-related diabetes (*cfrd*), dornase alfa (*dnase*), pancreatic insufficiency (*pancr*), and number of hospitalizations in year prior to baseline year (*numhosp*), categorized as 0, 1, 2, 3 or more*/**

```
title 'Model (1): Traditional Regression';
proc glm data=analysis_data;
class Tobi inscat cfrd dnase pancr numhosp gender;
model dfev1=Tobi base_fev1 wt pct age inscat cfrd dnase pancr numhosp gender/
cl solution;
lsmeans Tobi/pdiff cl;
run;
```

```

/*Next, we implement the propensity score regression model previously
described. First, we use logistic regression to estimate propensity scores for each
subject.*/
title 'Model (2): Propensity Score Regression';
proc logistic data=analysis_data;
class inscat cfrd dnase pancr numhosp gender;
model Tob1=base_fev1 wtpct age inscat cfrd dnase pancr numhosp gender/
link=logit;
output out=props pred=ps;
run;
/*We use the commands below to assign a subject-specific weight that corre-
sponds to his or her propensity score from the logistic regression above. Since the
propensity score, denoted ps below, corresponds to predicted probability of receiv-
ing the treatment, each subject who received the treatment will have weight  $1/ps$ ,
while each subject who did not receive the treatment will have weight  $1/(1-ps)$ . The
resulting dataset, props2, will consist of the analysis_data, propensity scores that
were previously created and stored in props, and the ps_weight corresponding to
each subject's weighting derived from the propensity score.*/
data props2;
set props;
if Tob1=1 then ps_weight=1/ps;
if Tob1=0 then ps_weight=1/(1-ps);
run;
/*We now implement the weighted multivariable regression. The commands are
similar to our previous regression, except for our use here of the weight statement.
By using this statement, we request computation of weighted means and variance
estimates that are inversely proportional to the corresponding sum of weights.*/
proc glm data=props2;
class Tob1 inscat cfrd dnase pancr numhosp gender;
model dfev1=Tob1 base_fev1 wtpct age inscat cfrd dnase pancr numhosp gender/
cl solution;
lsmeans Tob1/pdiff cl;
weight ps_weight;
run;
/*Finally, we present commands for the instrumental variables regression. The
first model statement performs the first-stage regression of the treatment indicator
Tob1 on the instrument (cid_iv) and all measured confounders. The result is a probit
model with predicted probabilities of tobramycin use for each subject. The second
model statement performs multiple linear regression with the instrumented version
of the tobramycin variable from the first model statement.
title 'Model (3): Instrumental Variables Regression';
proc qlim data=analysis_data;
class inscat cfrd dnase pancr numhosp gender;
model Tob1=cid_iv base_fev1 wtpct age inscat cfrd dnase pancr numhosp gender
/discrete;
model dfev1=base_fev1 wtpct age inscat cfrd dnase pancr numhosp gender /
select(Tob1=1);
output out=Tob1 prob proball predicted residual;
run;

```

IntechOpen

Author details

Rhonda Szczesniak^{1,2*} and Bin Huang^{1,2}

1 Division of Biostatistics and Epidemiology, Cincinnati Children's Hospital Medical Center, Cincinnati, Ohio, USA

2 Department of Pediatrics, University of Cincinnati, Cincinnati, Ohio, USA

*Address all correspondence to: rhonda.szczesniak@cchmc.org

IntechOpen

© 2019 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

References

- [1] Gliklich RE, Dreyer NA, editors. Registries for Evaluating Patient Outcomes: A User's Guide. 2nd ed. Rockville, MD: Agency for Healthcare Research and Quality; 2010 (Prepared by Outcome DEcIDE Center [Outcome Sciences, Inc. d/b/a Outcome] under Contract No. HHS290200500351 TO3); AHRQ Publication No.10-EHC049
- [2] McGoon MD, Benza RL, Escribano-Subias P, Jiang X, Miller DP, Peacock AJ, et al. Pulmonary arterial hypertension: Epidemiology and registries. *Journal of the American College of Cardiology*. 2013;**62**(25 Suppl):D51-D59
- [3] Lilienfeld M. Epidemiological methods and inferences in studies of noninfectious diseases PHR review. *Public Health Reports*. 1957;**72**(1):51-60
- [4] VanDyke RD, McPhail GL, Huang B, Fenchel MC, Amin RS, Carle AC, et al. Inhaled tobramycin effectively reduces FEV₁ decline in cystic fibrosis: An instrumental variables analysis. *Annals of the American Thoracic Society*. 2013; **10**(3):205-212
- [5] Cystic Fibrosis Foundation. Patient Registry: 2012 Annual Report; 2013
- [6] Konstan MW, Morgan WJ, Butler SM, Pasta DJ, Craib ML, Silva SJ, et al. Risk factors for rate of decline in forced expiratory volume in one second in children and adolescents with cystic fibrosis. *The Journal of Pediatrics*. 2007; **151**(2):134-139
- [7] Konstan MW, Schluchter MD, Xue W, Davis PB. Clinical use of ibuprofen is associated with slower FEV₁ decline in children with cystic fibrosis. *American Journal of Respiratory and Critical Care Medicine*. 2007;**176**(11):1084-1089
- [8] Konstan MW, Wagener JS, Pasta DJ, Millar SJ, Jacobs JR, Yegin A, et al. Clinical use of dornase alfa is associated with a slower rate of FEV₁ decline in cystic fibrosis. *Pediatric Pulmonology*. 2011;**46**:545-553
- [9] Morgan WJ, Butler SM, Johnson CA, Colin AA, Fitzsimmons SC, Geller DE, et al. Epidemiologic study of cystic fibrosis: Design and implementation of a prospective, multicenter, observational study of patients with cystic fibrosis in the U.S. and Canada. *Pediatric Pulmonology*. 1999;**28**:231-241
- [10] Rosenbaum PR. Design of Observational Studies. New York, NY: Springer Series in Statistics; 2009
- [11] Stukel TA, Fisher ES, Wennberg DE, Alter DA, Gottlieb DJ, Vermeulen MJ. Analysis of observational studies in the presence of treatment selection bias: Effects of invasive cardiac management on AMI survival using propensity score and instrumental variable methods. *JAMA*. 2007;**297**(3):278-285
- [12] Austin PC, Mamdani MM. A comparison of propensity score methods: A case-study estimating the effectiveness of post-AMI statin use. *Statistics in Medicine*. 2006;**25**(12): 2084-2106. DOI: 10.1002/sim.2328
- [13] Lee BK, Lessler J, Stuart EA. Improving propensity score weighting using machine learning. *Statistics in Medicine*. 2009;**29**(3):n/a-n/a. DOI: 10.1002/sim.3782
- [14] McCaffrey DF, Griffin BA, Almirall D, Slaughter ME, Ramchand R, Burgette LF. A tutorial on propensity score estimation for multiple treatments using generalized boosted models. *Statistics in Medicine*. 2013;**32**(19):3388-3414. DOI: 10.1002/sim.5753
- [15] Imai K, Ratkovic M. Covariate balancing propensity score. *Journal of the Royal Statistical Society. Series B*:

Statistical Methodology. 2014;**76**(1): 243-263. DOI: 10.1111/rssb.12027

[16] McClellan M, McNeil BJ, Newhouse JP. Does more intensive treatment of acute myocardial infarction in the elderly reduce mortality? Analysis using instrumental variables. *JAMA*. 1994;**272**: 859-866

[17] Angrist JD, Imbens GW, Rubin DB. Identification of causal effects using instrumental variables. *Journal of the American Statistical Association*. 1996; **91**(434):444-455

[18] Hogan JW, Lancaster T. Instrumental variables and inverse probability weighting for causal inference from longitudinal observational studies. *Statistical Methods in Medical Research*. 2004;**13**: 17-48

[19] Robins JM. A new approach to causal inference in mortality studies with a sustained exposure period—Application to control of the healthy worker survivor effect. *Mathematical Modeling*. 1986;**7**:1393-1512

[20] Hsieh FY, Bloch DA, Larsen MD. A simple method of sample size calculation for linear and logistic regression. *Statistics in Medicine*. 1998; **17**:1623-1634

[21] Jaeschke R, Singer J, Guyatt GH. Measurement of health status. Ascertaining the minimal clinically important difference. *Controlled Clinical Trials*. 1989;**10**(4):407-415

[22] Copay AG, Subach BR, Glassman SD, Polly DW, Schuler TC. Understanding the minimum clinically important difference: A review of concepts and methods. *The Spine Journal*. 2007;**7**(5):541-546

[23] Rubin DB. Inference and missing data. *Biometrika*. 1976;**63**(3):581-592

[24] Kim J-O, Curry J. The treatment of missing data in multivariate analysis. *Sociological Methods & Research*. 1977; **6**(2):215-240

[25] Schafer JL. *Analysis of Incomplete Multivariate Data*. London, England: Chapman and Hall; 1997

[26] Molenberghs G, Kenward MG. *Missing Data in Clinical Studies*. New York, NY: Wiley Statistics in Practice; 2007. p. 286

[27] Dmitrienko A, Molenberghs G, Chuang-Stein C, Offen W. *Analysis of Clinical Trials Using SAS®: A Practical Guide*. Cary, NC: SAS Institute; 2005. p. 353

[28] Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. *Biometrika*. 1983;**70**:41-55

[29] Hernan MA, Robins JM. Instruments for causal inference—An epidemiologist's dream? *Epidemiology*. 2006;**17**(4):360-372

[30] Kahn KL, Tisnado DM, Adams JL, Liu H, Chen WP, Hu FA, et al. Does ambulatory process of care predict health-related quality of life outcomes for patients with chronic disease? *Health Services Research*. 2007;**42**:63-83

[31] Newhouse JP, McClellan M. Econometrics in outcomes research: The use of instrumental variables. *Annual Review of Public Health*. 1998;**19**:17-34

[32] Ramsey BW, Pepe MS, Quan JM, Otto KL, Montgomery AB, Williams-Warren J, et al. Intermittent administration of inhaled tobramycin in patients with cystic fibrosis. *The New England Journal of Medicine*. 1999;**340**: 23-30

[33] PCORI. In: Gliklich RE, Dreyer NA, editors. *Standards in the Conduct of Registry Studies for Patient-Centered*

Outcomes Research. Methodology Committee, Patient-Centered Outcomes Research Institute (PCORI), by Outcome, A Quintiles Company, PCORI Methodology Standards; 2012. Retrieved from: <https://www.pcori.org/sites/default/files/PCORI-Methodology-Standards.pdf>

[34] Ionescu-Ittu R, Abrahamowicz M, Pilote L. Treatment effect estimates varied depending on the definition of the provider prescribing preference-based instrumental variables. *Journal of Clinical Epidemiology*. 2012;**65**:155-162

[35] Leslie RS, Ghomrawi H. The use of propensity scores and instrumental variable methods to adjust for treatment selection bias. *SAS Global Forum*, Paper 366-2008; 2008