The Cystic Fibrosis Foundation Patient Registry

Design and Methods of a National Observational Disease Registry

Emily A. Knapp[1], Aliza K. Fink[1], Christopher H. Goss[2], Ase Sewall[3], Josh Ostrenga[1], Christopher Dowd[1], Alexander Elbert[1], Kristofer M. Petren[1], and Bruce C. Marshall[1]

[1]Cystic Fibrosis Foundation, Bethesda, Maryland; [2]Department of Medicine, University of Washington, Seattle, Washington; and [3]Sewall, Inc., Bethesda, Maryland

ORCID ID: 0000-0003-3953-9874 (E.A.K.).

Abstract

Rationale: The Cystic Fibrosis Foundation Patient Registry (CFFPR) is an ongoing patient registry study that collects longitudinal demographic, clinical, and treatment information about persons with cystic fibrosis (CF) in the United States. CF is a life-shortening genetic disorder that occurs in approximately 1 in 3,500 births in the United States. High-quality observational data is important for clinical research, quality improvement, and clinical management.

Objectives: To describe the data collection, patient population, and key limitations of the CFFPR.

Methods: Inclusion criteria for the CFFPR include diagnosis with CF or a CFTR-associated disorder, care at an accredited care center program, and provision of informed consent. Data from clinic visits and hospitalizations are collected through a secure website. Loss to follow-up and generalizability were examined using several methods. The accuracy of CFFPR data was evaluated with an audit of 2012 CFFPR data compared to the medical record.

Measurements and Main Results: Since 1986, the CFFPR contains the records of 48,463 individuals with CF. Participation among individuals seen at accredited care centers is high, and loss to follow-up is low. An audit of 2012 CFFPR data suggests that the CFFPR contains 95% of clinic visits and 90% of hospitalizations found in the medical record for these patients, and nearly all of the audited fields were highly accurate.

Conclusions: Registries such as the CFFPR are important tools for research, clinical care, and tracking incidence, mortality and population trends.

Keywords: registries; cystic fibrosis

(Received in original form November 25, 2015; accepted in final form April 12, 2016)

The Cystic Fibrosis Foundation Patient Registry is funded by the Cystic Fibrosis Foundation.

Author Contributions: E.A.K., A.K.F., A.S., J.O., C.D., A.E., K.M.P., and B.C.M.: contributed to the management and data quality of the Cystic Fibrosis Foundation Patient Registry (CFFPR); C.D.: managed the audit of CFFPR data; and E.A.K., A.K.F., and J.O.: designed and performed data analyses. All authors were involved with the analysis of the data in the CFFPR, and all authors critically reviewed the final and earlier versions of the manuscript.

Correspondence and requests for reprints should be addressed to Aliza K. Fink, D.Sc., Cystic Fibrosis Foundation, 6931 Arlington Road, Bethesda, MD 20814. E-mail: afink@cff.org

This article has an online supplement, which is accessible from this issue's table of contents at www.atsjournals.org

Ann Am Thorac Soc Vol 13, No 7, pp 1173–1179, Jul 2016

DOI: 10.1513/AnnalsATS.201511-781OC
Internet address: www.atsjournals.org

Cystic fibrosis (CF) is a genetic disease that affects multiple organs in the body. It is the most common life-shortening recessive genetic disorder among white persons and occurs in approximately 1 in 3,500 births in the United States (1). Mutations in the cystic fibrosis transmembrane conductance regulator (*CFTR*) gene lead to malfunctioning or absent CFTR protein, resulting in impaired mucosal clearance mechanisms that lead to recurring infections and obstruction in the lungs. As of 2014, the median predicted survival had increased to almost 40 years of age for individuals in the CF Foundation Patient Registry (CFFPR) from approximately 28 years in the late 1980s (2).

The CFFPR was established in the 1960s to collect information on patient demographics and survival (3). Regular updates have been made to comply with changes in regulations, improvements in data collection, and advances in new technology (Figure 1) (4). Today, the CFFPR is an important tool for research; clinical care; and tracking incidence, survival, and population trends (5, 6).

Researchers may apply to a registry oversight committee for use of the data (7). Data are used for both retrospective analyses and prospective studies, with data collection performed via registry-embedded case report forms. Additionally, the CF Foundation collaborates with pharmaceutical

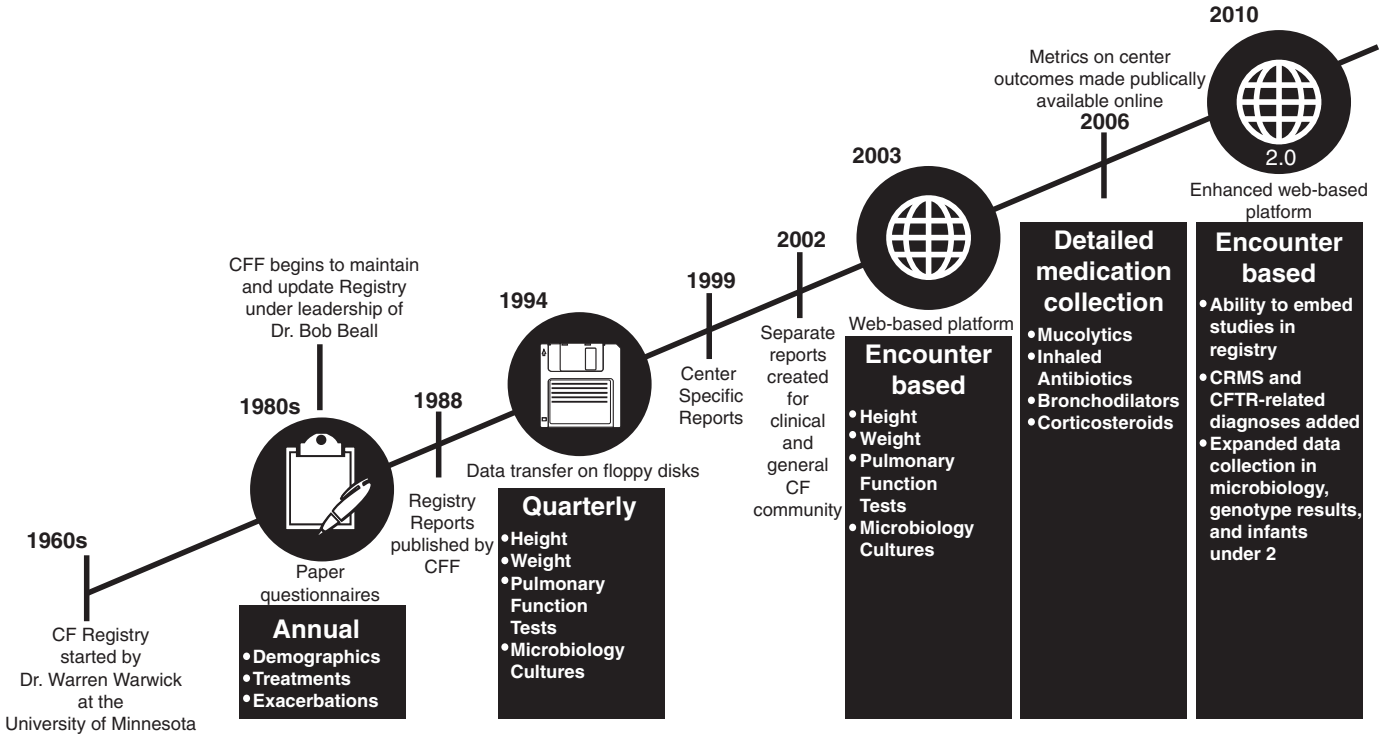

Figure 1. Timeline of key events in Cystic Fibrosis Foundation Patient Registry history. The Cystic Fibrosis Foundation Patient Registry was established in the 1960s and has continually evolved to keep pace with changes in technology and regulations, as well as improvements in the treatment of cystic fibrosis (CF). CFF = Cystic Fibrosis Foundation; CFTR = cystic fibrosis transmembrane conductance regulator; CRMS = CFTR-related metabolic syndrome.

sponsors to leverage the CFFPR to fulfill postapproval regulatory requirements.

The CFFPR supports the work of clinical care teams in their interactions with patients and families. Specifically, it is used by clinicians for previsit planning and patient care management. CFFPR data are used to examine center level variation in treatment and care practices and to establish and evaluate quality improvement initiatives (6, 8). Since 2006, key metrics on individual care centers have been released annually, adjusted for the case mix of each center (9). Last, the CFFPR provides annual estimates of incidence, prevalence, and mortality of CF and its secondary complications in the United States (2).

Recent changes in the data collection platform, increased interest in registry-based research, and the launch of a postapproval research initiative highlight the need for high-quality observational data. The purpose of this article is to describe the data collection, patient population, and key limitations of the CFFPR for those interested in registries and for those specifically interested in using CFFPR data for research. Some of these results have previously been reported in the form of an abstract (10, 11).

Methods

Study Population

All individuals diagnosed with CF and associated disorders (CFTR-related metabolic syndrome and CFTR-related disorders) who are seen at CF Foundation–accredited care center programs and provide informed consent are eligible to participate in the CFFPR. The CF Foundation has developed and sustains a network of 121 accredited CF care centers (comprised of 121 pediatric care programs and 105 adult care programs) and 51 affiliate programs across the United States (*see* additional information available in the online supplement). All accredited care programs are required to participate in the CFFPR. A portion of care center funding is based on the number of patients enrolled in the CFFPR and the completeness of their records. Each program obtains institutional review board approval and written informed consent and assent, as appropriate, from participants and/or their legal guardians. The CF Foundation provides online user manuals, data entry guidelines, training sessions, and user support for care center staff who enter data into the CFFPR. The CF Foundation serves as the coordinating center for data collection and data analysis.

Data Collection

The CFFPR data are collected through a web-based portal, PortCF, which contains five electronic data capture forms: demographic, diagnosis, encounter, care episode, and annual review forms. All data are entered by staff at the care center programs from the data available in the medical record or in forms completed by patients or families. CFFPR questionnaires are available in the annual reports (2).

The demographic form captures date of birth, sex, race, Hispanic ethnicity, state of birth, and other information needed to verify that the record is for a unique patient. Information on the date and cause of death are also captured on this form. The diagnosis form is used to record the means by which a CF diagnosis was reached, describing symptoms, if any, at diagnosis, as well as the results of tests to confirm diagnosis, including pilocarpine iontophoresis sweat tests and *CFTR* genotype. The encounter form is used to

record data from each clinic visit to the care center and includes basic anthropometric measures, respiratory microbiology, CF medications, pulmonary function tests, CF-related complications, laboratory tests, and airway clearance therapies. Data from all hospitalizations and courses of home intravenous (IV) antibiotics are collected on the care episode form, including the reason for hospitalization and pulmonary function tests. Each year, the care centers are required to complete an annual review form for each patient seen at the care center. This form is used to collect information not included in the encounter forms, such as socioeconomic status, health insurance, and transplant information.

Data Security

The current web-based registry application is maintained on a secure registry platform that meets U.S. Food and Drug Administration standards for data integrity (i.e., compliant with 21 C.F.R. part 11). Only authorized care center staff have access to the password-protected online portal, which is protected by firewall and intrusion detection systems. All data entered into PortCF are transferred to the CF Foundation weekly using a secure SSH File Transfer Protocol. The data are then processed, stored, and encrypted to meet the standards of the Health Information Technology for Economic and Clinical Health Act (12).

Data Processing

The electronic data capture system contains basic data validations to reduce the entry of implausible values. For some data fields, previous values are highlighted to alert data entry personnel to potential errors. Key fields are required so that a minimum amount of information is collected for a given form. Users at care centers are able to edit only records that were entered by their care center. Data are reviewed to flag what appear to be erroneous values from measures such as height, weight, body mass index, and laboratory tests. The CF Foundation verifies death and transplant dates as well as suspicious dates of birth and other important variables as needed.

Estimating the Generalizability of Registry Data

The target population for the CFFPR includes all individuals with a diagnosis of CF in the United States. In practice, the CFFPR is restricted to patients who are seen at least once during the calendar year at a CF Foundation–accredited care center program and have consented to have their data included in the CFFPR. Currently available estimates of the number of individuals in the United States with CF are based on analyses from 20 years ago and therefore need to be updated (13).

To assess the extent to which a given year of the data in the CFFPR is representative of the entire U.S. population of people living with CF, we considered three aspects of the registry population. First, we estimated the total number of people with CF in the United States. Second, among those seen at an accredited care center, we estimated the proportion who consented to participate in the registry. Third, among those who have been included in the registry in recent years, we estimated loss to follow-up.

An in-depth description of the methods used to estimate the number of individuals with CF are included in the online supplement. We employed two methods to estimate the number of individuals living with CF in the United States in 2012, using national statistics on births and deaths and published estimates of the incidence of CF from newborn screening programs (14–23). Briefly, with the first method, we estimated the cumulative number of people with CF over time using national birth data and estimates of CF incidence to estimate the expected number of births of infants with CF per year and subtracting the annual number of deaths due to CF. With the second method, we compared the number of CF deaths recorded in national death data with the number of deaths recorded in the CFFPR.

Assessing the Quality of Registry Data

The CF Foundation initiated an audit program to quantify the accuracy and completeness of critical data fields within the 2012 CFFPR as compared with the medical record. A stratified sample of 28 care center programs (approximately 10% of total programs) was selected for the audit. The goal was to obtain a diversity of programs with regard to pediatric versus adult programs, size, geography, and membership in the clinical trials research network (*see* online supplement for details of site selection). Key variables most often used for clinical care and research were selected for the audit. The records of 50–70 randomly selected patients were reviewed for each program selected.

The audit provided data on the completeness of the CFFPR with regard to clinic visits and hospitalizations, and accuracy among records where data existed in both the medical record and the CFFPR. For this evaluation, *accuracy* was defined as the proportion of cases where the values in the CFFPR and the medical record matched among cases where the data existed in both the medical record and the CFFPR. We presumed the medical record was the gold standard. Further details of the audit are provided in the online supplement.

Results

CFFPR Key Metrics

From 1986 through the end of 2014, the registry contained data on 48,463 unique patients representing 632,022 person-years of data, 2,497,178 clinic visits, and 241,984 hospitalizations and/or home IV treatment episodes. In 2014, there were 855 new diagnoses of CF, 119,716 clinic visits, and 10,081 courses of IV antibiotics in the hospital and/or home among 28,676 individuals with CF.

In 2014, 765 individuals were involved in entering patient data into PortCF. Timing of data entry in relation to patient clinical encounters differed by care center program. The mean data entry occurred within 4 weeks of the clinic visit among 33% of care center programs. Twenty-seven percent of centers entered data within an average of 5–8 weeks after the clinic visit, and 17% of centers entered data within 9–12 weeks of the clinic visit. The remaining 23% enter data, on average, more than 12 weeks after the clinic visit.

Patient Characteristics

In 2014, there were 28,676 people with CF included in the CFFPR, of whom 51.6% were male, 93.9% were white, 4.6% were African American, 3.1% were of another race, and 8.2% were Hispanic (of any race). An average of 4.5 clinic visits and 0.7 pulmonary exacerbations per patient were reported. A detailed summary of the 2014 registry population is available in the Patient Registry Annual Data Report (2).

Generalizability of Registry Data

Estimates of patients with CF in the United States. On the basis of the two methods of

estimating the number of persons with CF in the United States using national birth and death data, we derived estimates of 33,292 and 34,327 individuals with CF in the United States in 2012, respectively. In 2012, the CFFPR contained 27,804 individuals. Thus, approximately 81–84% of persons with CF were captured in the CFFPR in 2012, the most recent year for which national birth and mortality data were available.

Patients who do not consent to participate in the registry. As reported by the care centers during the annual reaccreditation process, there were 1,875 patients who were seen at accredited care centers in 2012 who did not provide consent to participate in the registry. This represents 6.3% of patients seen at CF Foundation–accredited care centers in 2012.

Loss to follow-up. We analyzed loss to follow-up among the 24,115 patients who were in the CFFPR in 2009 and did not have a record of death before December 31, 2013. The registry contains all 5 years of data for 20,614 individuals (85.5%) in this 2009 cohort, 4 years of data for 1,586 (6.6%), 3 years of data on 796 (3.3%), 2 years of data on 576 (2.4%), and 1 year of data on 543 (2.2%). Among the 3,501 individuals in the 2009 CFFPR who did not have 5 years of data, 2,271 were defined as lost to follow-up as of December 31, 2013, while 1,230 had gaps in registry coverage between 2010 and 2012 but were included in 2013 CFFPR data. The overall retention rate for the 2009 cohort was 90.6%, which was calculated by determining the proportion of patients from the original cohort who were in the CFFPR in 2013. Table 1 compares those patients lost to follow-up (without reentry into the cohort) from 2009 to 2013 with those who had gaps but had returned to the CFFPR by 2013 and those reported to the registry every year between 2009 and 2013. Overall, those who were lost to follow-up were older and more likely to have undergone a lung transplant. Of those lost to follow-up, 11% are known to have had a lung transplant.

Data Quality

CFFPR records from 2012 were audited for 1,628 patients, including 1,453 hospitalizations and 8,427 clinic visits. Overall, the CFFPR contained 95% of clinic visits and 90% of the hospitalizations that were found in the medical record. Data entered into CFFPR matched the data entered into the medical record in 82.6–99.9% of records, depending on the variable examined (Table 2). Information on the medications prescribed at each visit was the least accurate: 1.3–17.4% of medications entered into the CFFPR did not match the medical record. This may be due in part to the way medications are prescribed and data are collected. Inhaled tobramycin and inhaled aztreonam lysine, for example, are often prescribed in 1 month on/1 month off cycles. Confusion about how to report the prescription of these medications when the clinic visit occurred during an "off" month may have resulted in the lower rates of visit-level accuracy for these variables (accuracy for tobramycin, 82.6%; accuracy for aztreonam, 82.6%). For annual reporting and research, the use of a particular medication is often aggregated at the annual level to represent any reported use during the year. If accuracy of these variables is considered at the annual level, accuracy is higher (tobramycin 87.2% and aztreonam 91.6%).

Discussion

The CFFPR has been recognized as an example of a leading disease registry in the United States (24). We estimate there were approximately 33,292–34,327 individuals with CF in the United States in 2012. Approximately 81–84% of this estimated population was seen at a CF Foundation–accredited care center in 2012 and consented to have data included in the CFFPR.

An additional 6% of the CFFPR population was seen in a CF Foundation–accredited care center but did not provide

Table 1. Characteristics of persons in the Cystic Fibrosis Foundation Patient Registry in 2009 by loss to follow-up status

	Persons in CFFPR All 5 Years, 2009–2013	Persons with Gaps*	Persons Lost to Follow-Up†	All Reported to CFFPR in 2009
n (%)	20,614 (85%)	1,230 (5%)	2,271 (9%)	24,115
Age in 2009, mean (SD)	17.5 (12.7)	22.7 (13.1)	23.0 (12.9)	18.3 (12.9)
Age ⩾18 yr in 2009, n (%)	8,386 (41%)	783 (64%)	1,464 (64%)	10,633 (44%)
Age at diagnosis, mo, mean (SD)	3.0 (7.4)	4.7 (9.4)	6.4 (11.3)	3.4 (8.0)
Genotype delF508/delF508, n (%)	9,946 (49%)	471 (41%)	720 (35%)	11,137 (48%)
FEV_1, % predicted, age <18 yr, mean‡ (SD)	89.0 (18.3)	87.0 (20.0)	86.4 (21.6)	88.8 (18.6)
FEV_1, % predicted, age ⩾18 yr, mean‡ (SD)	67.0 (22.3)	71.7 (22.6)	70.1 (25.1)	67.7 (22.6)
BMI, mean‡ (SD)	22.8 (3.8)	23.3 (4.2)	23.0 (4.5)	22.9 (3.9)
BMI percentile, mean‡ (SD)	51.0 (26.1)	50.9 (27.9)	48.3 (29.2)	50.8 (26.4)
Persons with CF-related diabetes, n (%)	2,963 (14%)	163 (13%)	313 (14%)	3,439 (14%)
Transplant recipient, n (%)	992 (5%)	82 (7%)	258 (11%)	1,332 (6%)
Positive *Pseudomonas aeruginosa* culture, n (%)	9,685 (47%)	474 (38%)	709 (31%)	10,868 (45%)
Federally or state-funded insurance, n (%)	9,133 (44%)	393 (32%)	748 (33%)	10,274 (42.6%)
Number of PEX in 2009, mean§ (SD)	0.6 (1.1)	0.3 (0.8)	0.5 (1.2)	0.6 (1.1)

Definition of abbreviations: BMI = body mass index; CF = cystic fibrosis; PEX = pulmonary exacerbations.
Of individuals in the Cystic Fibrosis Foundation Patient Registry (CFFPR) in 2009, 85% were included in the CFFPR for all 5 years between 2009 and 2013. Individuals lost to follow-up or with gaps in care are generally older and are more likely to have received a lung transplant.
*Individuals with one or more missing years of data, but who are in the CFFPR in 2013.
†Not in the CFFPR in 2013.
‡Average of the highest measurement per quarter for up to four quarters. FEV_1 measurements are reported for individuals 6 years of age and older. BMI percentile measurements are reported for individuals ages 2–19 years. BMI is reported for individuals ages 20 years and older.
§Number of PEX, defined as treatment with intravenous antibiotics.

Table 2. Results of the Cystic Fibrosis Foundation Patient Registry data quality audit for key variables

Variable	Records* Evaluated for Missingness (n)	Missing† (n [%])	Records Evaluated for Accuracy (n)	Inaccurate‡ (n [%])
Demographic variables				
Date of birth	1,627	0 (0.0%)	1,627	19 (1.2%)
Sex	1,627	0 (0.0%)	1,627	2 (0.1%)
Race	1,536	0 (0.0%)	1,536	31 (2.0%)
Hispanic origin	1,568	0 (0.0%)	1,568	28 (1.8%)
CF mutation	1,509	88 (5.8%)	1,421	66 (4.6%)
Anthropometric measures				
FEV_1	6,380	85 (1.3%)	6,295	70 (1.1%)
Weight	3,965	48 (1.2%)	3,917	76 (1.9%)
Height	4,046	112 (2.8%)	3,934	187 (4.8%)
Respiratory infections				
Respiratory cultures	5,287	84 (1.6%)	—	—
MRSA	—	—	5,203	67 (1.3%)
MSSA	—	—	5,203	97 (1.9%)
Haemophilus influenzae	—	—	5,203	49 (0.9%)
Pseudomonas aeruginosa	—	—	5,203	30 (0.6%)
Stenotrophomonas *maltophilia*	—	—	5,203	10 (0.2%)
Medications				
Pancreatic enzymes§	—	—	8,427	108 (1.3%)
Azithromycin	—	—	3,845	389 (10.1%)
Aztreonam	—	—	1,678	292 (17.4%)
Dornase alfa	—	—	5,705	344 (6.0%)
Hypertonic saline	—	—	4,161	450 (10.8%)
Tobramycin	—	—	3,047	531 (17.4%)
Hospitalizations				
Hospitalizations	1,453	150 (10.3%)	—	—
Reason for hospitalization	—	—	1,303	10 (0.8%)
Duration of hospitalization	—	—	1,302	47 (3.6%)

Definition of abbreviations: CF = cystic fibrosis; MRSA = methicillin-resistant *Staphylococcus aureus*; MSSA = methicillin-sensitive *Staphylococcus aureus*. An audit of Cystic Fibrosis Foundation Patient Registry (CFFPR) data suggests high accuracy and low missingness of most variables compared with the medical record.

*For demographic variables, there is one record per patient. For anthropometric and medication variables, each clinic visit is evaluated, with patients each contributing multiple clinic visits. For respiratory infection data, all respiratory cultures are included, and multiple cultures per patient are possible. Each patient may also contribute more than one record of hospitalization per year.

†*Missing* is defined as data fields with no value entered in the CFFPR when there was a value entered in the medical record. It is not possible to evaluate "missing" microbacterial species or medication data because of the way this information is recorded in the CFFPR.

‡*Inaccuracy* is defined as the proportion of cases where the values in the CFFPR and the medical record were not the same among cases where there were data entered for that field in both the medical record and the CFFPR.

§Pancreatic enzyme replacement therapy.

consent to participate in the CFFPR. Although this is a dynamic cohort, loss to follow-up is minimal: over 90% of individuals in the CFFPR in 2009 remained in the CFFPR 5 years later. Audit data from the CFFPR demonstrate that, compared with the medical record, missingness of data in the CFFPR is low and accuracy is high for a number of key variables, including demographics, lung function, nutritional status, and hospitalization. Data quality, however, is not uniform, with increased missingness seen in key variables related to treatments. Overall, the CFFPR has become an invaluable tool for clinical management and patient care, national surveillance, quality improvement, and research.

Countries with national CF registries reporting data to the European Cystic Fibrosis Society Patient Registry estimate inclusion of 70–100% of persons with CF in these countries (25). Some of these registries likely cover nearly the entire CF population in their respective countries but have more limited data collection, doing so only on an annual basis. Currently, efforts are underway to develop methods to harmonize data across registries. In addition to data collected, methods used to calculate key metrics also need to be standardized, as highlighted by the work by Sykes and colleagues on factors influencing survival statistics (26).

Although our work demonstrated that the majority of clinical visits and hospitalizations are recorded in the CFFPR and that a number of key clinical and demographic variables had high accuracy, a finding that is supported by a recent evaluation of missingness in the CFFPR (27), data missingness remains an important limitation to the CFFPR. Care centers included in this data quality audit were selected to achieve balance on size, clinical trial participation, and timeliness of data entry. While the sampling strategy is consistent with the standard for such audit studies, we cannot be certain that the results reflect the care center network as a whole. Linking CFFPR data to other sources of data may be one avenue to improve data quality and extend the utility of data collected in the CFFPR; for example, the CFFPR data have previously been linked to United Network for Organ Sharing transplant data (28). In addition to improving data quality, prior work has

linked the CFFPR with other data sets that collect data elements not typically found in registries (e.g., data from the U.S. Environmental Protection Agency's Aerometric Information Retrieval System contains air pollution and temperature and humidity data) (29, 30).

Several limitations are highlighted by the analyses reported here. Data are entered into the CFFPR for individuals with CF seen at accredited care center programs in the United States. There is no centralized reporting for new cases of CF, which makes it challenging to enumerate persons with CF in the United States. Our estimates of the number of persons with CF in the United States are limited by the quality of available data. Published data from newborn screening programs are limited by state-by-state differences in newborn screening panels and racial/ethnic composition. As a result, there are discrepancies between estimates, especially for nonwhites.

In addition, older estimates of incidence may no longer be accurate, owing to CF diagnostic capabilities evolving over time, leading to early diagnoses of milder cases of CF. Alternatively, changes in prenatal counseling and screening for *CFTR* mutations may be leading to fewer persons being born with CF (31). Furthermore, previous comparisons of the CFFPR and national mortality data have suggested that the percentage of deaths missed by the CFFPR is highest among infants who die in the first year of life and those who survive beyond middle age (32). Therefore, we may underestimate the number of individuals in the United States who were born before 1968, when birth data became available, and are still alive. In addition to uncertainty about the number of persons with CF who are not represented in the CFFPR, clinical and demographic characteristics of persons with CF who are not represented in CFFPR are unknown and may differ in important ways from individuals included in the CFFPR. Therefore, the estimates presented here are preliminary, and further work is necessary to provide a more accurate estimate of the CF population in the United States and to compare those individuals in the CFFPR who are represented in the CFFPR with those who are not in the CFFPR. This may be one area that could benefit from linkages to other sources of data.

Research is an important use of the CFFPR. In regard to using the CFFPR to analyze clinical associations and outcomes in CF, many of the standard limitations hold true, as outlined in the Strengthening the Reporting of Observational Studies in Epidemiology guidelines (33). Specifically, for CF, given that some individuals with CF have milder manifestations of the disease and are more likely to survive well beyond the median survival in CF, there is potential for survivorship bias in the analysis and interpretation of cross-sectional data.

Last, universal newborn screening for CF has been in place in all 50 states in the United States since 2010, although many states introduced it earlier. Therefore, the clinical characteristics of very young patients included in the CFFPR in recent years are likely to be different from those of young patients included in the CFFPR in previous decades. Now, asymptomatic and potentially healthier infants are being diagnosed with CF and are included in the CFFPR at much younger ages than they previously would have been. When examining cohorts of individuals with CF over time, we need to be aware of the changing case mix in the CFFPR as the proportion of individuals diagnosed through newborn screening increases. There is also the potential for infants with CFTR-related metabolic disorder, who have positive CF newborn screening results but unclear CF diagnostic test results, to be incorrectly classified as having CF in the CFFPR (34).

Conclusions

The CFFPR contains data on almost 50,000 unique patients and has been used for research reported in over 120 peer-reviewed manuscripts, in addition to numerous quality improvement and benchmarking initiatives. The CFFPR captures a substantial portion of the U.S. patient population with CF and has robust and high-quality data in key variables of interest, such as lung function, nutritional status and hospitalizations. Data in the registry have been used for many years to compare center-level variation in care and outcomes. As CF registries are implemented in other countries, the CFFPR has also been used to compare treatment and outcomes between the United States and other countries (35). International comparisons leverage variation in availability and physician preference of therapies and can be facilitated with standardization of data collection across registries.

Registries such as the CFFPR are important tools for collection of real-world data and are thus an ideal tool for conducting comparative effectiveness research (CER). Compared with randomized controlled trials, registry studies offer large study populations, long observational periods, and are less expensive, especially in a rare disease setting (36, 37). The CFFPR is beginning to be used for both CER research and pragmatic clinical trials (38). Increased interest in CER, linkages to other data sources, and international comparisons will continue to drive the need for high-quality observational data such as that included in the CFFPR. ■

Author disclosures are available with the text of this article at www.atsjournals.org.

Acknowledgment: The authors thank the patients, care providers, and clinic coordinators at CF centers throughout the United States for their contributions to the Cystic Fibrosis Foundation Patient Registry.

References

1 Farrell PM, Rosenstein BJ, White TB, Accurso FJ, Castellani C, Cutting GR, Durie PR, Legrys VA, Massie J, Parad RB, *et al*.; Cystic Fibrosis Foundation. Guidelines for diagnosis of cystic fibrosis in newborns through older adults: Cystic Fibrosis Foundation consensus report. *J Pediatr* 2008;153:S4–S14.

2 Cystic Fibrosis Foundation. Cystic Fibrosis Foundation Patient Registry: 2014 annual data report. Bethesda, MD: Cystic Fibrosis Foundation; 2015 [accessed 2016 Apr 27]. Available from: https://www.cff.org/2014-Annual-Data-Report/

3 Doershuk CF. Cystic fibrosis in the 20th century: people, events, and progress. Cleveland, OH: AM Publishing; 2001.

4 Gliklich RE, Dreyer NA, Leavy MB, editors. Case Example 29. Challenges in transitions and changes in data collection. In: Registries for evaluating patient outcomes: a user's guide. AHRQ publication 13(14)-EHC111. 3rd ed.Rockville, MD: Agency for Healthcare Research and Quality; 2014. pp. 336–337.

5 Marshall BC, Penland CM, Hazle L, Ashlock M, Wetmore D, Campbell PW III, Beall RJ. Cystic fibrosis foundation: achieving the mission. *Respir Care* 2009;54:788–795, discussion 795.

6 Schechter MS, Fink AK, Homa K, Goss CH. The Cystic Fibrosis Foundation Patient Registry as a tool for use in quality improvement. *BMJ Qual Saf* 2014;23:i9–i14.
7 Cystic Fibrosis Foundation. Patient registry data requests [accessed 2016 May 24]. Available from: https://www.cff.org/Our-Research/For-Researchers/CF-Research-Tools-and-Resources/Patient-Registry-Data-Requests/
8 Mall MA, Elborn JS. Cystic fibrosis. Sheffield, UK: European Respiratory Society; 2014. p. 332.
9 Mogayzel PJ Jr, Dunitz J, Marrow LC, Hazle LA. Improving chronic care delivery and outcomes: the impact of the cystic fibrosis Care Center Network. *BMJ Qual Saf* 2014;23:i3–i8.
10 Knapp EA, Fink A. Estimating the number of individuals with cystic fibrosis in the United States [abstract 502]. *Pediatr Pulmonol* 2015;50 (S41 Special Issue: The 29th Annual North American Cystic Fibrosis Conference October 8–10, 2015):S381–S382.
11 Dowd CG, Ostrenga J, Elbert A, Petren K, Fink A, Marshall BC. The Cystic Fibrosis Foundation Patient Registry (CFFPR) data quality audit program: an analysis of data completeness and data accuracy [abstract 538]. *Pediatr Pulmonol* 2015;50(S41 Special Issue: The 29th Annual North American Cystic Fibrosis Conference October 8–10, 2015):S396.
12 Health Information Technology for Economic and Clinical Health (HITECH) Act. Title XIII, Division A, and Title IV, Division B, American Recovery and Reinvestment Act of 2009 (ARRA). Pub. L. 111-5, 42 U.S.C. §§ 201 et seq., 123 Stat. 226.
13 FitzSimmons SC. The changing epidemiology of cystic fibrosis. *Curr Probl Pediatr* 1994;24:171–179.
14 Centers for Disease Control and Prevention. CDC WONDER: compressed mortality file [updated 2015; accessed 2016 Apr 27]. Available from: http://wonder.cdc.gov/wonder/help/cmf.html
15 Martin JA, Hamilton BE, Osterman MJK, Curtin SC, Matthews TJ. Births: final data for 2012. *Natl Vital Stat Rep* 2013;62(9):1–68.
16 American Lung Association. State of lung disease in diverse communities 2010. Washington, DC: American Lung Association; 2010 [accessed 2016 Apr 27]. http://action.lung.org/site/DocServer/state-of-lung-disease-in-diverse-communities-2010.pdf?docID=8744
17 Kaye CI; Committee on Genetics. Newborn screening fact sheets. *Pediatrics* 2006;118:e934–e963.
18 Comeau AM, Parad RB, Dorkin HL, Dovey M, Gerstle R, Haver K, Lapey A, O'Sullivan BP, Waltz DA, Zwerdling RG, *et al*. Population-based newborn screening for genetic disorders when multiple mutation DNA testing is incorporated: a cystic fibrosis newborn screening model demonstrating increased sensitivity but more carrier detections. *Pediatrics* 2004;113:1573–1581.
19 Sontag MK, Hammond KB, Zielenski J, Wagener JS, Accurso FJ. Two-tiered immunoreactive trypsinogen-based newborn screening for cystic fibrosis in Colorado: screening efficacy and diagnostic outcomes. *J Pediatr* 2005;147(3, Suppl):S83–S88.
20 Rock MJ, Hoffman G, Laessig RH, Kopish GJ, Litsheim TJ, Farrell PM. Newborn screening for cystic fibrosis in Wisconsin: nine-year experience with routine trypsinogen/DNA testing. *J Pediatr* 2005;147 (3, Suppl):S73–S77.
21 Prach L, Koepke R, Kharrazi M, Keiles S, Salinas DB, Reyes MC, Pian M, Opsimos H, Otsuka KN, Hardy KA, *et al*.; California Cystic Fibrosis Newborn Screening Consortium. Novel *CFTR* variants identified during the first 3 years of cystic fibrosis newborn screening in California. *J Mol Diagn* 2013;15:710–722.
22 Feuchtbaum L, Carter J, Dowray S, Currier RJ, Lorey F. Birth prevalence of disorders detectable through newborn screening by race/ethnicity. *Genet Med* 2012;14:937–945.
23 Korzeniewski SJ, Young WI, Hawkins HC, Cavanagh K, Nasr SZ, Langbo C, Teneyck KR, Grosse SD, Kleyn M, Grigorescu V. Variation in immunoreactive trypsinogen concentrations among Michigan newborns and implications for cystic fibrosis newborn screening. *Pediatr Pulmonol* 2011;46:125–130.
24 Larsson S, Lawyer P, Garellick G, Lindahl B, Lundström M. Use of 13 disease registries in 5 countries demonstrates the potential to use outcome data to improve health care's value. *Health Aff (Millwood)* 2012;31:220–227.
25 Zolin A, McKnone EF, van Rens J, Fox A, Iansa P, Preftitsi A, Pypops U, Gulmans V, Jung A, Mehta A, *et al*. ECFSPR annual report: 2010 data. Karup, Denmark: European Cystic Fibrosis Society; 2014 [accessed 2016 Apr 27]. Available from: https://www.ecfs.eu/files/webfm/webfiles/File/ecfs_registry/ECFSPR_Report10_v12014_final_020617.pdf
26 Sykes J, Stanojevic S, Goss CH, Quon BS, Marshall BC, Petren K, Ostrenga J, Fink A, Elbert A, Stephenson AL. A standardized approach to estimating survival statistics for population-based cystic fibrosis registry cohorts. *J Clin Epidemiol* 2016;70:206–213.
27 Mendelsohn AB, Dreyer NA, Mattox PW, Su Z, Swenson A, Li R, Turner JR, Velentgas P. Characterization of missing data in clinical registry studies. *Ther Innov Regul Sci* 2015;49:146–154.
28 Liou TG, Adler FR, Cox DR, Cahill BC. Lung transplantation and survival in children with cystic fibrosis. *N Engl J Med* 2007;357:2143–2152.
29 Goss CH, Newsom SA, Schildcrout JS, Sheppard L, Kaufman JD. Effect of ambient air pollution on pulmonary exacerbations and lung function in cystic fibrosis. *Am J Respir Crit Care Med* 2004;169: 816–821.
30 Psoter KJ, De Roos AJ, Mayer JD, Kaufman JD, Wakefield J, Rosenfeld M. Fine particulate matter exposure and initial Pseudomonas aeruginosa acquisition in cystic fibrosis. *Ann Am Thorac Soc* 2015; 12:385–391.
31 Hale JE, Parad RB, Comeau AM. Newborn screening showing decreasing incidence of cystic fibrosis. *N Engl J Med* 2008;358: 973–974.
32 Nick JA, Chacon CS, Brayshaw SJ, Jones MC, Barboa CM, St Clair CG, Young RL, Nichols DP, Janssen JS, Huitt GA, *et al*. Effects of gender and age at diagnosis on disease progression in long-term survivors of cystic fibrosis. *Am J Respir Crit Care Med* 2010;182: 614–626.
33 von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP; STROBE Initiative. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: guidelines for reporting observational studies. *Int J Surg* 2014;12:1495–1499.
34 Ren CL, Fink AK, Petren K, Borowitz DS, McColley SA, Sanders DB, Rosenfeld M, Marshall BC. Outcomes of infants with indeterminate diagnosis detected by cystic fibrosis newborn screening. *Pediatrics* 2015;135:e1386–e1392.
35 Goss CH, MacNeill SJ, Quinton HB, Marshall BC, Elbert A, Knapp EA, Petren K, Gunn E, Osmond J, Bilton D. Children and young adults with CF in the USA have better lung function compared with the UK. *Thorax* 2015;70:229–236.
36 Motheral B, Brooks J, Clark MA, Crown WH, Davey P, Hutchins D, Martin BC, Stang P. A checklist for retrospective database studies–report of the ISPOR Task Force on Retrospective Databases. *Value Health* 2003;6:90–97.
37 Psoter KJ, Rosenfeld M. Opportunities and pitfalls of registry data for clinical research. *Paediatr Respir Rev* 2013;14:141–145.
38 Heltshe SL, Goss CH, Thompson V, Sagel SD, Sanders DB, Marshall BC, Flume PA. Short-term and long-term response to pulmonary exacerbation treatment in cystic fibrosis. *Thorax* 2016;71:223–229.