

Patient Registries in Idiopathic Pulmonary Fibrosis

② Daniel A. Culver¹, Jürgen Behr^{2,3}, John A. Belperio⁴, Tamera J. Corte^{5,6}, Joao A. de Andrade⁷, Kevin R. Flaherty⁸, Mridu Gulati⁹, Tristan J. Huie¹⁰, Lisa H. Lancaster¹¹, Jesse Roman¹², Christopher J. Ryerson¹³, and Hyun J. Kim¹⁴

¹Cleveland Clinic, Cleveland, Ohio; ²Department of Internal Medicine V, Ludwig-Maximilians University of Munich, Munich, Germany; ³Asklepios Clinic Gauting, Member of the German Center for Lung Research, Gauting, Germany; ⁴David Geffen School of Medicine at UCLA, University of California, Los Angeles, Los Angeles, California; ^{5,6}Royal Prince Alfred Hospital, Camperdown, New South Wales, Australia; ⁷University of Sydney, Sydney, New South Wales, Australia; ⁸University of Alabama at Birmingham, Birmingham, Alabama; ⁹Division of Pulmonary and Critical Care Medicine, University of Michigan, Ann Arbor, Michigan; ¹⁰Yale University School of Medicine, New Haven, Connecticut; ¹¹National Jewish Health, Denver, Colorado; ¹²Vanderbilt University Medical Center, Nashville, Tennessee; ¹³Jane and Leonard Korman Respiratory Institute, Thomas Jefferson University, Philadelphia, Pennsylvania; ¹⁴University of British Columbia, Vancouver, British Columbia, Canada; and ¹⁵University of Minnesota, Minneapolis, Minnesota

Abstract

Over the past decade, several large registries of patients with idiopathic pulmonary fibrosis (IPF) have been established. These registries are collecting a wealth of longitudinal data on thousands of patients with this rare disease. The data collected in these registries will be complementary to data collected in clinical trials because the patient populations studied in registries have a broader spectrum of disease severity and comorbidities and can be followed for a longer period of time. Maintaining the quality and completeness of registry databases presents administrative and resourcing challenges, but it is important

to ensuring the robustness of the analyses. Data from patient registries have already helped improve understanding of the clinical characteristics of patients with IPF, the impact that the disease has on their quality of life and survival, and current practices in diagnosis and management. In the future, analyses of biospecimens linked to detailed patient profiles will provide the opportunity to identify biomarkers linked to disease progression, facilitating the development of precision medicine approaches for prognosis and therapy in patients with IPF.

Keywords: interstitial lung diseases; pulmonary fibrosis; observational study; biomarkers

A registry can be defined as a longitudinal, systematic collection of “real-world” data describing the health status and medical interventions in a defined population of individuals. Patient registries enable the collection of real-world data on the clinical course of diseases and their impact on patients and healthcare services. Registries are particularly useful in the case of rare diseases, for which only a small number of patients are seen at any individual center (1). Idiopathic pulmonary fibrosis (IPF) is a rare interstitial lung disease (ILD), with an estimated U.S. incidence of 3–10 per 100,000 person-years (2).

Before the creation of patient registries, most data on the natural history of IPF came from small, usually single-center, observational studies and from analyses of the placebo groups of clinical trials of investigational therapies (3). Although these studies provided valuable data on the clinical course of IPF, they were limited by their relatively small size and short duration of follow-up and the restricted patient populations studied. The establishment of patient registries has enabled the collection of data from cohorts of patients with IPF with a broader spectrum of disease severity and comorbidities managed

in clinical practice, thereby providing a better understanding of the real-world behavior and impact of IPF. In addition, many IPF registries collect a range of biological samples, with the aim of linking clinical features or outcomes to disease pathobiology. The vast number and heterogeneity of patients, the large number of outcome events, and the prolonged duration of follow-up in registries promise the possibility of leveraging biobank data to develop and validate individualized diagnostic and therapeutic approaches (i.e., precision medicine).

(Received in original form February 27, 2019; accepted in final form April 24, 2019)

②This article is open access and distributed under the terms of the Creative Commons Attribution Non-Commercial No Derivatives License 4.0 (<http://creativecommons.org/licenses/by-nc-nd/4.0/>). For commercial usage and reprints, please contact Diane Gern (dgern@thoracic.org).

Correspondence and request for reprints should be addressed to Daniel A. Culver, D.O., Cleveland Clinic, Main Campus, 9500 Euclid Avenue, Cleveland, OH 44195. E-mail: culverd@ccf.org.

CME will be available for this article at www.atsjournals.org.

Am J Respir Crit Care Med Vol 200, Iss 2, pp 160–167, Jul 15, 2019

Copyright © 2019 by the American Thoracic Society

Originally Published in Press as DOI: 10.1164/rccm.201902-0431CI on April 29, 2019

Internet address: www.atsjournals.org

Opportunities and Challenges Presented by IPF Registries

The insights gleaned from patient registries are complementary to those gained from clinical trials. Although clinical trials are designed to assess the effects of interventions in carefully defined patient populations, the populations studied in registries are larger and more heterogeneous (Table 1). Registry data provide a unique opportunity to answer the question of long-term effectiveness of antifibrotic medications, including in patients with greater disease severity than those studied in clinical trials and patients with specific comorbidities. In recent years, many IPF registries have been initiated across the world that will accumulate data on several thousand patients (Table 2). In aggregate, the number of patients enrolled in these registries is a tremendous advantage, providing the power to discover new clinical or biological variables and the capacity to validate or refute findings derived from smaller cohorts. These registries are collecting similar categories of data, and some have common data fields. However, caution should be used in comparing data collected in different registries, given the differences in the patient populations enrolled and the methodologies used to collect and categorize data; even “the same” parameter may not have been assessed in the same way in different registries. Although variability between

registries poses difficulties in comparison of their results, the differences between registries can theoretically also be exploited to glean novel insights about IPF.

The real-world nature of patient registries poses challenges for data collection and analysis. For registries in IPF, even the application of diagnostic criteria is not straightforward. IPF patient registries set up before 2018 were largely designed to enroll patients who met the diagnostic criteria published in 2011 (4). However, some registries rely on local diagnosis by site investigators, whereas others have the diagnosis confirmed at the enrolling center or by central review. Substantial variability may exist between enrolling centers regarding the comprehensiveness of testing for connective tissue diseases, the assessment of exposures, and the interpretation of high-resolution computed tomographic scans. In some registries (e.g., CARE-PF [Canadian Registry for Pulmonary Fibrosis]), all the enrolling centers are specialized referral centers, whereas other registries (e.g., the Australian Idiopathic Pulmonary Fibrosis Registry) include a much broader range of pulmonary practices. These differences in methodology mean that the patient populations enrolled in different registries vary in the extent to which they would meet the strictest application of diagnostic guidelines. The publication of new diagnostic guidelines for IPF in September 2018 (5) will create additional

complications in analyzing data collected in registries over time.

The timing of diagnosis is another complicating factor. For many patients with IPF, there is a prolonged period between symptom onset and diagnosis due to delays in presentation and referral and the time needed to acquire diagnostic data (6–8). Furthermore, the enrolling center for a registry may not be the center at which the patient first received their diagnosis of IPF. This means that the patients enrolled in IPF registries are at different points in the course of the disease, and it is often not clear when their IPF first developed. Some registries split patients into those with “incident” (diagnosed in past 6 mo) and “prevalent” IPF (9), but this does not entirely resolve the problem, because many patients with newly diagnosed IPF will have had the (undiagnosed) disease for some time.

Registry data are maximally useful when the data are complete, with stringent quality control, but this poses administrative and resourcing challenges (10), which may vary across registries. Missing data may be substantial for some variables. A number of statistical techniques exist for handling missing data (11, 12), but none is without limitations. In particular, missing data complicate the analysis and interpretation of longitudinal analyses. Given that a particularly salient benefit of registries is the assessment of events over time (e.g., hospitalizations, continuation of antifibrotic therapy, change in health-related

Table 1. Key Differences between Clinical Trials and Patient Registries

	Clinical Trials	Registries
Diagnosis	Narrowly defined based on diagnostic guidelines	May not require strict adherence to diagnostic guidelines; may allow for evolution of disease definition if broad inclusion criteria are used and adequate descriptive data are collected
Severity of disease	Generally target middle ranges; several exclusions related to comorbidities	All strata
Follow-up	Usually more frequent and comprehensive than is typical of clinical practice	Typically reflects clinical practice
Outcomes studied	Focus on pulmonary function tests and health-related quality of life Generally too short to investigate mortality Few data on healthcare use and costs	Course and impact of disease over long term, including mortality Detailed data on healthcare use and costs
Duration	Months to a few years	May last several years
Data quality	Few missing data; stringent quality control	Substantial missing data; variable data quality

Table 2. Ongoing Multicenter National/International Registries Including Patients with Idiopathic Pulmonary Fibrosis

Registry Name (www.clinicaltrials.gov Identifier)	Country/Countries	Patient Population	Size
IPF-PRO/ILD-PRO Registry (NCT01915511)	United States	Patients with IPF or other progressive ILD that is newly diagnosed or newly confirmed at the enrolling center	1,000 patients with IPF (fully enrolled); 1,000 patients with progressive non-IPF ILDs to be enrolled
PFF-PR (NCT02758808)	United States	ILDs, including IPF	1,461 patients enrolled as of August 2018; target is 2,000 patients, ~60% of whom have IPF (19)
Pulmonary Fibrosis Foundation Contact Registry (NCT01935726)	United States	Pulmonary fibrosis of any cause (and caregivers of these patients)	Target enrollment is 50,000 patients
CARE-PF	Canada	ILDs, including IPF	>3,000 patients enrolled, >600 of whom have IPF
AIPFR	Australia	IPF	768 patients enrolled
Australasian ILD Registry	Australia and New Zealand	ILDs, including IPF	1,003 patients enrolled
EMPIRE	Croatia, Czech Republic, Hungary, Israel, Poland, Serbia, Slovakia, and Turkey	IPF	>2,048 patients enrolled (34)
eurIPFreg (NCT02951416)	Austria, Czech Republic, France, Germany, Hungary, Italy, Spain, and United Kingdom (open to all European countries)	ILDs, including IPF	525 patients with IPF enrolled as of October 2016 (20); target enrollment is 2,000 patients with ILDs (www.clinicaltrials.gov)
INSIGHTS-IPF registry (NCT01695408)	Germany	Incident (diagnosed within 6 mo) and prevalent IPF	Target is 1,000 patients with IPF (almost fully enrolled)
EXCITING registry (NCT02645968)	Germany	ILDs, including IPF	601 patients enrolled, including 151 patients with IPF (57); enrollment is complete
FinnishIPF registry	Finland	IPF	>700 patients enrolled (58)
PROOF and PROOF-NEXT (NCT03732859) registries	Belgium and Luxembourg	IPF	277 patients enrolled in PROOF (21); target enrollment in PROOF-NEXT is 600 patients
Swedish IPF-Registry	Sweden	IPF	>217 patients with IPF enrolled as of May 2018 (59)
FIBRONET registry (NCT02803580)	Italy	IPF	210 patients enrolled (completed) (www.clinicaltrials.gov)
INDULGE IPF (NCT03074149)	Greece	IPF	Target enrollment is 300 patients (60)
REGIS	Romania	ILDs, including IPF	>104 patients enrolled (61)
TURK-UIP (NCT02821039)	Turkey	ILDs with UIP on HRCT	Target enrollment is 2,000 patients (www.clinicaltrials.gov)
BTS	United Kingdom	IPF, sarcoidosis	2,000 patients with IPF and 400 patients with sarcoidosis enrolled
PORTRAY registry (NCT03666234)	China	Newly diagnosed IPF	Target enrollment is 800 patients
ILD-India registry	India	ILDs, including IPF	>1,084 patients enrolled (62)
JIPS registry (NCT03041623)	Japan	Newly diagnosed idiopathic ILDs, including IPF	>498 patients enrolled, >249 of whom have IPF (63); target enrollment is 600 patients

Definition of abbreviations: AIPFR = Australian IPF Registry; BTS = British Thoracic Society; CARE-PF = Canadian Registry for Pulmonary Fibrosis; EMPIRE = European MultiPartner IPF Registry; eurIPFreg = European IPF Registry and Biobank; EXCITING = Exploring Clinical and Epidemiological Characteristics of Interstitial Lung Diseases; FIBRONET = IPF Italian observational study; HRCT = high-resolution computed tomography; ILD = interstitial lung disease; INDULGE IPF = Investigating Idiopathic Pulmonary Fibrosis in Greece; INSIGHTS-IPF = Investigating Significant Health Trends in Idiopathic Pulmonary Fibrosis; IPF = idiopathic pulmonary fibrosis; IPF-PRO/ILD-PRO = Idiopathic Pulmonary Fibrosis/Interstitial Lung Disease Prospective Outcomes; JIPS = Japanese Idiopathic Interstitial Pneumonias; PFF-PR = Pulmonary Fibrosis Foundation Patient Registry; PORTRAY = Idiopathic Pulmonary Fibrosis Registry China Study; PROOF = Prospective Observational Registry to Describe the Disease Course and Outcomes in Idiopathic Pulmonary Fibrosis Patients in a Real-World Clinical Setting; New and Extended Belgium-Luxembourg; REGIS = Romanian Registry for Interstitial Lung Diseases; TURK-UIP = Turkish Thoracic Society Usual Interstitial Pneumonia registry study; UIP = usual interstitial pneumonia.

quality of life [HRQoL]), biased loss of follow-up is a major hurdle that must be carefully considered in the design and implementation of registries. The fact that patients with worse disease are more likely to drop out of a registry (due to death and other reasons) means that data collected over time may show better patient outcomes than would have been observed if all patients had been retained. However, it is also possible that patients with milder or less progressive disease may be less motivated to return to tertiary centers, producing the opposite bias. If resources permit, missing data can be minimized using innovative approaches for communicating with patients, such as using call centers or telehealth technologies to follow up directly with patients who have not attended the enrolling center within a certain time window. If permission has been obtained, administrative data, such as from insurance claims or health records, can potentially be used to fill gaps in the registry database.

Key Insights from IPF Registries to Date

Registries have provided a wealth of information about the characteristics of patients with IPF. In most registries, the characteristics of patients at enrollment appear similar to those of patients enrolled in phase III trials such as the ASCEND (Assessment of Pirfenidone to Confirm Efficacy and Safety in Idiopathic Pulmonary Fibrosis) (13) and INPULSIS (14) trials, with patients being predominantly elderly, male, and ex-smokers with significant lung function impairment (Table 3) (15–21). The INSIGHTS-IPF (Investigating Significant Health Trends in Idiopathic Pulmonary Fibrosis) registry in Germany appears to be an exception in that the patients in this registry had more severe gas exchange impairment at baseline (mean DL_{CO} , 35.5% predicted) than patients in other registries or in most clinical trials (9). Patients in IPF registries generally have slightly worse scores on instruments assessing symptoms and HRQoL than patients in clinical trials (9, 22, 23). Data from several registries suggest that HRQoL is worse in patients who have more impaired lung function, have worse dyspnea or cough, or are using supplemental oxygen (22–25).

Data from patient registries have helped to illuminate the prevalence of comorbidities in patients with IPF. Among the first 525 patients enrolled in the

eurIPFreg registry (European IPF Registry and Biobank) (20) and the first 502 patients enrolled in the German INSIGHTS-IPF registry (9), arterial hypertension was present in 32% and 54% of patients, respectively. Among patients in the INSIGHTS-IPF, eurIPFreg, and IPF-PRO (Idiopathic Pulmonary Fibrosis Prospective Outcomes) registries, pulmonary hypertension was reported in 17%, 17%, and 8% and coronary heart disease/coronary artery disease was reported in 25%, 18%, and 29%, respectively (9, 16, 20). Sleep disorders such as sleep apnea were reported in approximately one-fourth of patients in the eurIPFreg and U.S. registries (16, 19, 20). Gastroesophageal reflux disease was reported in 28% and 29.5% of patients in the eurIPFreg and INSIGHTS-IPF registries, respectively (9, 20), and in 47%, 55%, and 62% of patients in the PROOF (Prospective Observational Registry to Describe the Disease Course and Outcomes of Idiopathic Pulmonary Fibrosis Patients in a Real-world Clinical Setting), IPF-PRO, and Pulmonary Fibrosis Foundation registries, respectively (16, 19, 21). Registry data will provide the opportunity to investigate the impact of specific comorbidities on morbidity, mortality, HRQoL, and healthcare resource use.

Registries are perhaps the most important source of data on current practices in the diagnosis of IPF and ultimately will allow evaluation of temporal trends in diagnostic procedures. For example, surgical lung biopsies have been performed in 13–35% of patients enrolled in registries (9, 16, 19–21, 26). In eurIPFreg, 20–30% of patients diagnosed with IPF in 2010–2011 had open or thoracoscopic lung biopsy, but these numbers declined in the following years (20). Rates of bronchoscopy and analysis of BAL fluid (BALF) vary widely across registries, with analysis of BALF conducted in 85% of patients in eurIPFreg (20) and 62% of patients in INSIGHTS-IPF (9), whereas bronchoscopy was performed in only about 20% of patients in the Australian IPF registry (15) and an even smaller proportion of patients in the IPF-PRO Registry (26). Data from the INSIGHTS-IPF, Pulmonary Fibrosis Foundation, and IPF-PRO registries suggest that multidisciplinary discussion (MDD) was performed at the enrolling center in 22%, 40%, and 42% of patients, respectively (9, 19, 26). It is unclear to what extent diagnoses made without MDD reflect clinicians' belief that MDD was not required to make a diagnosis of IPF in a particular patient, rather than a lack of access to a multidisciplinary team. In the Australian IPF Registry, a central

multidisciplinary review was implemented to overcome the paucity of local access to MDD. It should also be noted that there is no standard definition of MDD in the diagnosis of IPF, and different sites/registries may have interpreted "MDD" in very different ways. Interestingly, recent data from the Australian IPF Registry showed that patients who had received a clinical diagnosis of IPF but were judged not to meet 2011 international diagnostic guidelines in a central multidisciplinary review exhibited disease behavior and mortality similar to those of patients who met those diagnostic guidelines (27). These findings highlight an important function of patient registries: their ability to confirm or challenge the validity of diagnostic categorizations.

Registry data support the findings of patient surveys which indicate that patients often experience a delay in diagnosis. Among the first 525 patients participating in eurIPFreg, the average time between onset of symptoms and diagnosis of IPF was 21.8 months (20). Other registries have reported similar findings (9, 17). However, determination of the time from onset of symptoms to diagnosis of IPF in patient registries is confounded by the fact that most patients with IPF are diagnosed and managed by local physicians; thus, the date that a patient is referred to an enrolling center is not the date of diagnosis, and the time taken for a patient to be referred to an enrolling center may be highly variable between sites.

One of the greatest benefits provided by patient registries is the opportunity to study the clinical course of diseases in the real world. Data from IPF registries demonstrate the progressive nature of IPF (17, 22, 28–30). Among 514 patients participating in the EMPIRE registry (European MultiPartner IPF Registry), 23.5% of patients had a decline in FVC greater than or equal to 10% predicted over the course of 12 months (17). In the Australian IPF registry, over a median follow-up period of about 2 years, FVC declined by approximately 5% predicted per year (28). Registry data also illustrate the very high mortality associated with IPF. In the INSIGHTS-IPF registry, 36.5% of patients died in the first 2 years after enrollment (29). The factors shown to be associated with higher mortality in patients with IPF in registry studies are largely consistent with the findings of clinical trials and smaller observational studies. Advanced age and worse lung function (FVC and DL_{CO})

Table 3. Demographic and Clinical Characteristics of Patients in Idiopathic Pulmonary Fibrosis Registries and Clinical Trials

Characteristics	IPF Registries						Clinical Trials		
	INSIGHTS-IPF Registry (n = 502) (9)	Australian IPF Registry (n = 647) (15)	IPF-PRO Registry (n = 662) (16)	Czech Patients in EMPIRE Registry (n = 514) (17)	eurIPFreg (n = 525) (20)	PFF Registry (n = 835) (19)	CARE-PF Registry (n = 317) (18)	ASCEND Trial (n = 277)* (13)	INPULSIS Trials (n = 423)* (14)
Age, yr	68.7 ± 9.4	70.9 ± 8.5	69.7 ± 7.6	67 (50-82)	68.1 ± 11.1	71 ± 8	70.9 ± 8.5	67.8 ± 7.3	67.0 ± 7.9
Male sex	78	68	75	70	74	74	72	77	79
Current or former smoker	61	72	68	53	69	63 [†]	—	61 [†]	71
FVC, %	72.2 ± 20.6	81.0 ± 21.7	69.7 (60.1-79.8)	80.0 (48.7-116.3)	68.4 ± 22.6	68 ± 17	72.8 ± 19.5	68.6 ± 10.9	79.3 ± 18.2
D-co, % predicted [‡]	35.5 ± 15.5	48.4 ± 16.7	40.6 (31.3-49.3)	45.6 (21.3-72.3)	42.1 ± 17.8	41 ± 18	49.9 ± 16.7	44.2 ± 12.5	47.0 ± 13.4

Definition of abbreviations: ASCEND = Assessment of Pirfenidone to Confirm Efficacy and Safety in Idiopathic Pulmonary Fibrosis; CARE-PF = Canadian Registry for Pulmonary Fibrosis; EMPIRE = European MultiPartner IPF Registry; eurIPFreg = European IPF Registry and Biobank; INSIGHTS-IPF = Investigating Significant Health Trends in Idiopathic Pulmonary Fibrosis; IPF = idiopathic pulmonary fibrosis; IPF-PRO = Idiopathic Pulmonary Fibrosis Prospective Outcomes; PFF = Pulmonary Fibrosis Foundation.

*Data are mean ± SD, median (25th-75th percentile), or percentage.

[†]Former smokers only.

[‡]It is not known whether the same equations were used to calculate percent predicted values across the studies.

have consistently been shown to be predictors of mortality in IPF registries (15, 17, 30). A recent analysis of data from 662 patients in the IPF-PRO Registry demonstrated that use of supplemental oxygen at rest was the strongest predictor of mortality over a follow-up period of 30 months (30). In addition, analyses of data from the IPF-PRO Registry and the Australian IPF Registry (15) suggest that worse scores on patient-reported outcomes such as the St. George's Respiratory Questionnaire provide important prognostic information beyond that provided by demographic/physiological factors.

IPF is known to be associated with high use of healthcare resources and high costs, particularly related to hospitalization (31, 32). This is supported by real-world data from patient registries. In the INSIGHTS-IPF registry, 38.8% of patients were hospitalized at least once over a 2-year period (28). In the IPF-PRO Registry, the probability of hospitalization was 30.2% over 12 months and was higher among those with more severe lung function impairment at enrollment (33). Mean length of hospital stay in patients who did not undergo lung transplant was 5 days (33).

Finally, IPF registries are starting to provide data on the use of approved drug therapies (nintedanib and pirfenidone). Data from eurIPFreg illustrate that approved antifibrotic therapies now dominate the drugs used to treat IPF (20). However, many patients with IPF still do not receive an antifibrotic therapy. Recent data from IPF registries in the United States showed that 55–62% of patients with IPF were receiving nintedanib or pirfenidone at enrollment (16, 19). Although registries have already provided some data suggesting that the use of antifibrotic therapies may extend life expectancy in patients with IPF (15, 20, 34, 35), the magnitude and durability of this benefit will become clearer with the availability of longer-term data.

Biobanking

“Biobanking” refers to the collection, processing, storage, and distribution of biological specimens, such as tissue; serum; plasma; whole blood; or other bodily fluids, such as urine, sputum, or BALF. Although biobanks may be created with specific

hypotheses in mind, many biobanks also have the forward-thinking purpose to store samples that might be used in the future to address new hypotheses (36). However, given the substantial cost of biobanking, registries may decide to prioritize certain types of samples. The evolution of biobanks into large interconnected operations has led to the development of best practice guidelines covering their complex administrative and operational requirements (37–39). In addition to planning related to governance, funding, equipment, storage, and personnel, there are many ethical and legal aspects to consider, including policies for obtaining and withdrawing consent, patient confidentiality, return of incidental findings, and the transfer of biospecimens and data (37, 38, 40). The success of a biobank depends on adequate sample quality. A stringent quality management system is critical to minimizing the effects of preanalytic and analytic factors and ensuring uniformity in specimen handling (36–38, 41).

A myriad of techniques are available to analyze biospecimens, including transcriptomics, proteomics, metabolomics, epigenomics, RNA sequencing, whole-exosome sequencing, genome-wide association studies, immunofluorescence staining, and analysis of neopeptides (42–45). Modern high-throughput techniques can generate a massive quantity of data. Systems biology, which uses predictive computational modeling to analyze complex biological systems and their interactions, is a field that has evolved largely from the development of these technologies (46). Advances in artificial intelligence applications have enabled very large and complex datasets to be analyzed through machine learning. These methods can be used to identify distinct phenotype profiles in a dataset, either by classifying samples on the basis of a known set of features or by unveiling groupings by clustering samples on the basis of their similarities (47–49). Coupled with clinical data, this biological information is likely to radically alter the diagnosis and management of IPF in future years (50).

Biobanks have particular value when the biological samples are matched to patients who have been well characterized with respect to their characteristics and outcomes (51, 52). In

the case of IPF, these characteristics include demographics, diagnosis-related data (e.g., serologies, radiographic patterns, pathologic patterns), measures of disease severity (e.g., pulmonary function tests, patient-reported outcomes), treatments, and comorbidities (53). Longitudinal data can be particularly useful because they enable changes in biomarkers to be linked to changes in health status. Biobanks are already beginning to generate data, including on blood biomarkers as predictors of disease severity and disease progression (54, 55). Importantly, the existence of multiple biobanks provides the opportunity to validate and compare findings in different patient populations to improve the application of precision medicine approaches for prognosis and therapy.

Needs and Future Directions in IPF Registries

Registries will help fill the gaps in the data collected in clinical trials because they cover a broader patient population, can collect information over patients' lifetimes, and provide insights into the delivery of care in the real world. Future registry efforts should consider less conventional/more broadly defined disease definitions, recruitment in understudied populations, novel phenotyping using biomarkers, and creative analytic methodologies, perhaps focused on specific goals or hypotheses, rather than reprising the ballooning number of current efforts. Such innovative registries are likely to provide new observations that propel the field forward. It is hoped that current registries will provide insights into particular endotypes and phenotypes of patients with IPF (perhaps based on physiology, imaging, serum biomarkers, or rates/patterns of progression) that are relevant to prognosis and treatment. Registries will also help illuminate the impact of therapies on patient outcomes, including life expectancy and healthcare use, in the real world. These findings may align with the Food and Drug Administration's Real World Evidence initiative, which aims to incorporate real-world data into decisions regarding drug labeling and indications and comparative effectiveness or safety analyses (56).

Several registries are collecting data on other ILDs as well as IPF; these will help improve understanding of the natural history and outcomes of these ILDs, which have not been studied as extensively as IPF and for which there are no approved therapies or established treatment algorithms. Enrolling broader populations will allow close examination of the usefulness of transitory diagnostic constructs and allow analyses of biological questions that may cross over various ILDs. Sharing data across registries will increase the power to look at particularly rare diseases, genotypes, phenotypes, and events. Collaboration, transparency, and effective application of technologies will be critical to maximizing the value of

registries to researchers and ultimately to patients.

Conclusions

Patient registries provide data that are complementary to those provided by clinical trials and administrative databases. The heterogeneity of the patients in registries, as compared with clinical trials, should be regarded as one of registries' strengths, although the different patient populations enrolled and methodologies used across registries complicates comparison of their findings. Data from patient registries have already improved knowledge of the clinical course and impact of IPF and of diagnostic

and treatment practices. The biobanks associated with registries will be a valuable resource to examine questions related to pathogenetic mechanisms and prognostic biomarkers and to facilitate the introduction of personalized management strategies for IPF. ■

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

Acknowledgment: Writing support was provided by Elizabeth Ng, B.Sc., and Wendy Morris, M.Sc., of FleishmanHillard Fishburn, London, United Kingdom, which was contracted and funded by Boehringer Ingelheim Pharmaceuticals, Inc. Boehringer Ingelheim was given the opportunity to review the manuscript for medical and scientific accuracy as well as intellectual property considerations.

References

- Council of the European Union. Council recommendation of 8 June 2009 on an action in the field of rare diseases (2009/C 151/02). *Official Journal of the European Union* 2009;C151:7–10.
- Ley B, Urbania T, Husson G, Vittinghoff E, Brush DR, Eisner MD, et al. Code-based diagnostic algorithms for idiopathic pulmonary fibrosis: case validation and improvement. *Ann Am Thorac Soc* 2017;14:880–887.
- Raghu G. Idiopathic pulmonary fibrosis: lessons from clinical trials over the past 25 years. *Eur Respir J* 2017;50:1701209.
- Raghu G, Collard HR, Egan JJ, Martinez FJ, Behr J, Brown KK, et al.; ATS/ERS/JRS/ALAT Committee on Idiopathic Pulmonary Fibrosis. An official ATS/ERS/JRS/ALAT statement: idiopathic pulmonary fibrosis. Evidence-based guidelines for diagnosis and management. *Am J Respir Crit Care Med* 2011;183:788–824.
- Raghu G, Remy-Jardin M, Myers JL, Richeldi L, Ryerson CJ, Lederer DJ, et al.; American Thoracic Society, European Respiratory Society, Japanese Respiratory Society, and Latin American Thoracic Society. Diagnosis of idiopathic pulmonary fibrosis: an official ATS/ERS/JRS/ALAT clinical practice guideline. *Am J Respir Crit Care Med* 2018;198:e44–e68.
- Collard HR, Tino G, Noble PW, Shreve MA, Michaels M, Carlson B, et al. Patient experiences with pulmonary fibrosis. *Respir Med* 2007;101:1350–1354.
- Schoenheit G, Becattelli I, Cohen AH. Living with idiopathic pulmonary fibrosis: an in-depth qualitative survey of European patients. *Chron Respir Dis* 2011;8:225–231.
- Cottin V. Current approaches to the diagnosis and treatment of idiopathic pulmonary fibrosis in Europe: the AIR survey. *Eur Respir Rev* 2014;23:225–230.
- Behr J, Kreuter M, Hoepfer MM, Wirtz H, Klotsche J, Koschel D, et al. Management of patients with idiopathic pulmonary fibrosis in clinical practice: the INSIGHTS-IPF registry. *Eur Respir J* 2015;46:186–196.
- Kodra Y, Weinbach J, Posada-de-la-Paz M, Coi A, Lemonnier SL, van Enckevort D, et al. Recommendations for improving the quality of rare disease registries. *Int J Environ Res Public Health* 2018;15:1644.
- Graham JW. Missing data analysis: making it work in the real world. *Annu Rev Psychol* 2009;60:549–576.
- Mack C, Su Z, Westreich D. Managing missing data in patient registries: addendum to registries for evaluating patient outcomes: a user's guide, 3rd ed. AHRQ Publication No. 17(18)-EHC015-EF. Rockville, MD: Agency for Healthcare Research and Quality; Feb 2018 [accessed 2019 Jun 10]. Available from: www.effectivehealthcare.ahrq.gov.
- King TE Jr, Bradford WZ, Castro-Bernardini S, Fagan EA, Glaspole I, Glassberg MK, et al.; ASCEND Study Group. A phase 3 trial of pirfenidone in patients with idiopathic pulmonary fibrosis. *N Engl J Med* 2014;370:2083–2092.
- Richeldi L, du Bois RM, Raghu G, Azuma A, Brown KK, Costabel U, et al.; INPULSIS Trial Investigators. Efficacy and safety of nintedanib in idiopathic pulmonary fibrosis. *N Engl J Med* 2014;370:2071–2082.
- Jo HE, Glaspole I, Grainge C, Goh N, Hopkins PM, Moodley Y, et al. Baseline characteristics of idiopathic pulmonary fibrosis: analysis from the Australian Idiopathic Pulmonary Fibrosis Registry. *Eur Respir J* 2017;49:1601592.
- Culver D, Yow E, Neely ML, Belperio JA, Bender S, de Andrade JA, et al. Characteristics of patients with idiopathic pulmonary fibrosis (IPF) in the US: data from the IPF-PRO Registry [abstract]. *Chest* 2018;154(4 Suppl):397A–398A.
- Doubková M, Švancara J, Svoboda M, Šterclová M, Bartoš V, Plačková M, et al. EMPIRE registry, Czech part: impact of demographics, pulmonary function and HRCT on survival and clinical course in idiopathic pulmonary fibrosis. *Clin Respir J* 2018;12:1526–1535.
- Fisher J, Shapera S, Algami M, Morisset J, Johansson KA, Fell CD, et al. Baseline characteristics and comorbidities in the Canadian Registry for Pulmonary Fibrosis [abstract]. *Am J Respir Crit Care Med* 2018;197:A7478.
- Flaherty K, De Andrade J, Lancaster L, Limb S, Lindell K, Nathan S, et al. Baseline characteristics of 1461 participants in the Pulmonary Fibrosis Foundation Patient Registry [abstract]. *Eur Respir J* 2018;52(Suppl 62):PA2199.
- Guenther A, Krauss E, Tello S, Wagner J, Paul B, Kuhn S, et al. The European IPF registry (eurIPFreg): baseline characteristics and survival of patients with idiopathic pulmonary fibrosis. *Respir Res* 2018;19:141.
- Wuyts WA, Dahlqvist C, Slabbynck H, Schlessner M, Gusbin N, Compere C, et al. Baseline clinical characteristics, comorbidities and prescribed medication in a real-world population of patients with idiopathic pulmonary fibrosis: the PROOF registry. *BMJ Open Respir Res* 2018;5:e000331.
- Glaspole IN, Chapman SA, Cooper WA, Ellis SJ, Goh NS, Hopkins PM, et al. Health-related quality of life in idiopathic pulmonary fibrosis: data from the Australian IPF Registry. *Respirology* 2017;22:950–956.
- Kreuter M, Swigris J, Pittrow D, Geier S, Klotsche J, Prasse A, et al. Health related quality of life in patients with idiopathic pulmonary fibrosis in clinical practice: INSIGHTS-IPF registry. *Respir Res* 2017;18:139.
- de Andrade JA, Whelan T, Luckhardt T, Lancaster L, Gamerman V, Kulkarni T, et al. Clinical characteristics of patients with advanced idiopathic pulmonary fibrosis [abstract]. *Am J Respir Crit Care Med* 2017;195:A3457.
- Kreuter M, Swigris J, Pittrow D, Geier S, Klotsche J, Prasse A, et al. The clinical course of idiopathic pulmonary fibrosis and its association to quality of life over time: longitudinal data from the INSIGHTS-IPF registry. *Respir Res* 2019;20:59.

26. Holtze CH, Salisbury M, Sheth JS, Belloli E, White ES, Neely ML, *et al.* Description of diagnostic processes undertaken for patients with idiopathic pulmonary fibrosis enrolled in the IPF-PRO Registry [abstract]. *Am J Respir Crit Care Med* 2017;195:A1552.
27. Jo HE, Glaspole I, Goh N, Hopkins PMA, Moodley Y, Reynolds PN, *et al.* Implications of the diagnostic criteria of idiopathic pulmonary fibrosis in clinical practice: analysis from the Australian Idiopathic Pulmonary Fibrosis Registry. *Respirology* 2019;24:361–368.
28. Jo HE, Glaspole I, Moodley Y, Chapman S, Ellis S, Goh N, *et al.* Disease progression in idiopathic pulmonary fibrosis with mild physiological impairment: analysis from the Australian IPF registry. *BMC Pulm Med* 2018;18:19.
29. Behr J, Kreuter M, Prasse A, Wirtz HR, Pittrow D, Klotsche J, *et al.* Exacerbations, hospitalisations and mortality in patients with idiopathic pulmonary fibrosis: 5-year follow-up of the INSIGHTS-IPF registry [abstract]. *Am J Respir Crit Care Med* 2018;197:A4525.
30. Snyder L, Neely ML, Hellkamp AS, O'Brien E, de Andrade J, Conoscenti CS, *et al.* Predictors of death or lung transplant after a diagnosis of idiopathic pulmonary fibrosis: insights from the IPF-PRO registry. *Respir Res* 2019;20:105.
31. Collard HR, Chen SY, Yeh WS, Li Q, Lee YC, Wang A, *et al.* Health care utilization and costs of idiopathic pulmonary fibrosis in U.S. Medicare beneficiaries aged 65 years and older. *Ann Am Thorac Soc* 2015;12:981–987.
32. Yu YF, Wu N, Chuang CC, Wang R, Pan X, Benjamin NN, *et al.* Patterns and economic burden of hospitalizations and exacerbations among patients diagnosed with idiopathic pulmonary fibrosis. *J Manag Care Spec Pharm* 2016;22:414–423.
33. Fan Y, Bender S, Conoscenti C, Knight JD, Davidson-Ray L, Cowper PA, *et al.* Hospital-based resource use and costs among patients with idiopathic pulmonary fibrosis enrolled in the IPF-PRO registry [abstract]. *J Manag Care Spec Pharm* 2018;24 (10-a):S71.
34. Vašáková M, Šterclová M, Mogulkoc N, Kus J, Müller V, Hajkova M, *et al.* Real world idiopathic pulmonary fibrosis in the EMPIRE registry [abstract]. *Eur Respir J* 2018;52(Suppl 62):PA2203.
35. Zúrková M, Kriegová E, Kolek V, Lostáková V, Šterclová M, Bartos V, *et al.*; ILD section; IPF registry. Effect of pirfenidone on lung function decline and survival: 5-yr experience from a real-life IPF cohort from the Czech EMPIRE registry. *Respir Res* 2019;20:16.
36. Vaught J. Biobanking comes of age: the transition to biospecimen science. *Annu Rev Pharmacol Toxicol* 2016;56:211–228.
37. International Society for Biological and Environmental Repositories (ISBER). Best practices: recommendations for repositories. 4th ed. 2018 [accessed 2019 Jun 10]. Available from: www.isber.org/page/BPR.
38. National Cancer Institute (NCI). Best practices for biospecimen resources. 2016 [accessed 2019 Jun 10]. Available from: <https://biospecimens.cancer.gov/bestpractices/>.
39. Organization for Economic Co-operation and Development (OECD). Guidelines on human biobanks and genetic research databases. 2009 [accessed 2019 Jun 10]. Available from: <http://www.oecd.org/health/emerging-tech/guidelines-for-human-biobanks-and-genetic-research-databases.htm>.
40. Widdows H, Cordell S. The ethics of biobanking: key issues and controversies. *Health Care Anal* 2011;19:207–219.
41. Grizzle WE, Bell WC, Sexton KC. Issues in collecting, processing and storing human tissues and associated information to support biomedical research. *Cancer Biomark* 2010;9:531–549.
42. Karsdal MA, Krarup H, Sand JM, Christensen PB, Gerstoft J, Leeming DJ, *et al.* Review article: the efficacy of biomarkers in chronic fibroproliferative diseases. Early diagnosis and prognosis, with liver fibrosis as an exemplar. *Aliment Pharmacol Ther* 2014;40:233–249.
43. Warr A, Robert C, Hume D, Archibald A, Deeb N, Watson M. Exome sequencing: current and future perspectives. *G3 (Bethesda)* 2015;5:1543–1550.
44. Spitzer MH, Nolan GP. Mass cytometry: single cells, many features. *Cell* 2016;165:780–791.
45. Yu G, Ibarra GH, Kaminski N. Fibrosis: lessons from OMICS analyses of the human lung. *Matrix Biol* 2018;68–69:422–434.
46. Wang RS, Maron BA, Loscalzo J. Systems medicine: evolution of systems biology from bench to bedside. *Wiley Interdiscip Rev Syst Biol Med* 2015;7:141–161.
47. Abraham Y, Zhang X, Parker CN. Multiparametric analysis of screening data: growing beyond the single dimension to infinity and beyond. *J Biomol Screen* 2014;19:628–639.
48. Grys BT, Lo DS, Sahin N, Kraus OZ, Morris Q, Boone C, *et al.* Machine learning and computer vision approaches for phenotypic profiling. *J Cell Biol* 2017;216:65–71.
49. Tarca AL, Carey VJ, Chen XW, Romero R, Drăghici S. Machine learning and its applications to biology. *PLoS Comput Biol* 2007;3:e116.
50. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. *Nat Med* 2019;25:44–56.
51. Hayek SS, Sever S, Ko YA, Trachtman H, Awad M, Wadhvani S, *et al.* Soluble urokinase receptor and chronic kidney disease. *N Engl J Med* 2015;373:1916–1925.
52. Woo TE, Lim R, Surette MG, Waddell B, Bowron JC, Somayaji R, *et al.* Epidemiology and natural history of *Pseudomonas aeruginosa* airway infections in non-cystic fibrosis bronchiectasis. *ERJ Open Res* 2018;4:00162–2017.
53. White ES, Brown KK, Collard HR, Conoscenti CS, Cosgrove GP, Flaherty KR, *et al.* Open-access biorepository for idiopathic pulmonary fibrosis: the way forward. *Ann Am Thorac Soc* 2014;11:1171–1175.
54. Jo HE, Corte TJ, Glaspole I, Grainge C, Goh NS, Hopkins P, *et al.* Biomarkers can predict disease progression in idiopathic pulmonary fibrosis: analysis from the Australian IPF Registry [abstract]. *Am J Respir Crit Care Med* 2017;195:A7007.
55. Todd JL, Vinisko R, Neely ML, Overton R, Flaherty KR, Noth I, *et al.*; on behalf of the IPF-PRO Registry investigators. Peripheral blood matrix metalloproteinase profiling in the multicenter IPF-PRO Registry cohort. Presented at the International Colloquium on Lung and Airway Fibrosis. October 2018 [accessed 2019 Jun 10]. Available from: <http://uspubs-posters.com/iclaf2018/todd>.
56. U.S. Food and Drug Administration (FDA). Framework for FDA's real world evidence program. 2018 [accessed 2019 Jun 10]. Available from: www.fda.gov.
57. Kreuter M, Kabitz HJ, Hagemeyer L, Hammerl P, Esselmann A, Wiederhold C, *et al.* Outcome differences between idiopathic pulmonary fibrosis (IPF) and other interstitial lung diseases (ILD) - data from the EXCITING registry [abstract]. *Eur Respir J* 2018;52(Suppl 62):PA2197.
58. Rajala K, Lehto JT, Sutinen E, Kautiainen H, Myllärniemi M, Saarto T. Marked deterioration in the quality of life of patients with idiopathic pulmonary fibrosis during the last two years of life. *BMC Pulm Med* 2018;18:172.
59. Pesonen I, Gao J, Kalafatis D, Carlson L, Myllärniemi M, Ferrara G, *et al.* Forced vital capacity, total lung capacity and diffusion capacity in idiopathic pulmonary fibrosis [abstract]. *Eur Respir J* 2018;52(Suppl 62):PA2200.
60. Bours D, Daniil Z, Papakosta D, Antoniou KM, Markopoulou K, Kolilekas L, *et al.*; INDULGE IPF Study Group. Design, rationale, methodology, and aims of a Greek prospective idiopathic pulmonary fibrosis registry: investigating idiopathic pulmonary fibrosis in Greece (INDULGE IPF). *Respiration* 2018;96:41–47.
61. Strambu I, Salmen T, Traila D, Traila D, Croitoru A. Romanian Registry for Interstitial Lung Diseases (REGIS): inclusion of patients in 3 years [abstract]. *Eur Respir J* 2017;50(Suppl 61):PA868.
62. Singh S, Collins BF, Sharma BB, Joshi JM, Talwar D, Katiyar S, *et al.* Interstitial lung disease in India: results of a prospective registry. *Am J Respir Crit Care Med* 2017;195:801–813.
63. Fujisawa T, Suda T, Mikamo M, Mikamo M, Mori K, Enomoto Y, *et al.* Nationwide cloud-based integrated database with clinical, radiological and pathological data of idiopathic interstitial pneumonias for multidisciplinary discussions [abstract]. *Am J Respir Crit Care Med* 2017;195:A1551.