Pulmonary arterial hypertension (PAH) is a fatal orphan disease. The global epidemiology of PAH is not well known and encourages combined national and international efforts to enhance understanding of the disease. A global database will help unify investigators and patients to foster collaboration and knowledge.

**Abbreviations:** FPAH = familial pulmonary arterial hypertension; HIPAA = Health Insurance Portability and Accountability Act; IPAH = idiopathic pulmonary arterial hypertension; NIH = National Institutes of Health; PAH = pulmonary arterial hypertension; PH = pulmonary hypertension; REVEAL = Registry to Evaluate Early and Long-Term PAH Disease Management; SMR = Scottish Morbidity Report; SPVR = Scottish Pulmonary Vascular Unit; WHO = World Health Organization

Despite this progress, although quality of life and exercise capacity are improved, these options are not curative. In addition, these therapies are costly and often not available worldwide. Furthermore, some causes of PAH, such as schistosomiasis, although very prevalent in developing countries, are typically ignored in research or the priorities of the pharmaceutical industry in the developed world. The pulmonary hypertension (PH) community needs to collectively collaborate to better understand the nature and global impact of PAH on patients and society. This will allow the community to design more representative and meaningful clinical trials and enhance the quality of PAH care in both developed and developing countries.

**Existing Databases**

The NIH registry, started in 1981 and concluded in 1987, was the first comprehensive evaluation of the epidemiology of “primary pulmonary hypertension” (now referred to as idiopathic PAH [IPAH], familial PAH [FPAH], and anorexin-associated PAH) in the United States (Table 1). The registry provided data from an era lacking specific PH therapies. Patients in the NIH registry were treated only with conventional therapy, which included diuretics, digoxin, supplemental nasal oxygen, and, in a minority of cases, anticoagulation.
incidence of 1.7 cases per million population (95% CI, 1.0-2.4), the first estimate of incidence for the disease.

An attempt to evaluate the trends in the United States by use of hospital discharge information and national mortality statistics by the Centers for Disease Control and Prevention helped describe a more general population. The Centers for Disease Control and Prevention surveillance study reported the mortality results from 1980 to 2002 from the National Vital Statistics database and the National Hospital Discharge Survey. Unfortunately, although the report provided valuable information on mortality, it was unable to report incidence and prevalence determinations in the United States. By using hospital discharge data, there may be underreporting or overreporting of the disease, because the data are not collected by specialists and are often not recorded accurately by physicians.

Table 1—Existing Databases

<table>
<thead>
<tr>
<th>Registry</th>
<th>Years</th>
<th>Strengths</th>
<th>Weaknesses</th>
</tr>
</thead>
<tbody>
<tr>
<td>NIH</td>
<td>1981-1987</td>
<td>Prospective evaluation</td>
<td>Not inclusive of all WHO category 1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>First national registry</td>
<td>Different standard of care</td>
</tr>
<tr>
<td></td>
<td></td>
<td>First characterization of the disease</td>
<td>Lack of advanced imaging modalities</td>
</tr>
<tr>
<td>IPPS</td>
<td>1992-1994</td>
<td>Prospective evaluation</td>
<td>Not inclusive of all WHO category 1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>220 Centers Europe: France, United Kingdom, Belgium, Netherlands</td>
<td>Different standard of care</td>
</tr>
<tr>
<td>French National Registry</td>
<td>2002-2003</td>
<td>17 expert centers</td>
<td>Only those with a catheter diagnosis</td>
</tr>
<tr>
<td></td>
<td></td>
<td>All category 1 PAH</td>
<td>Wide variation of prevalence</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Uniform diagnostic criteria</td>
<td>Often CHD or CTD not referred</td>
</tr>
<tr>
<td>Pulmonary Hypertension Connection</td>
<td>1982-2006</td>
<td>Large center over extended time period</td>
<td>Referral biases</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4 physicians caring for patients</td>
<td>Retrospective and prospective chart review</td>
</tr>
<tr>
<td></td>
<td></td>
<td>All category 1 PAH</td>
<td>Single center may not reflect national trends</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Incident and prevalent cases</td>
<td>Echocardiogram used for diagnosis in 10% of patients—may be misclassification</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Most referrals to these centers</td>
<td>Low HIV prevalence</td>
</tr>
<tr>
<td>Scotland</td>
<td>SMR: 1986-2001</td>
<td>Compares inexpert and expert centers diagnosis</td>
<td>Hospital discharge records: coding irrespective of guidelines for PAH</td>
</tr>
<tr>
<td></td>
<td>SPVU: 1997-2005</td>
<td>Most referrals to SPVU</td>
<td>Survey or inexpert diagnoses</td>
</tr>
<tr>
<td>China</td>
<td>1999-2004</td>
<td>Similar data as NIH registry, reflects global epidemiologic trends</td>
<td>Only idiopathic, patients with CHD and CTD</td>
</tr>
<tr>
<td>REVEAL</td>
<td>2004-present</td>
<td>Incident and prevalent cases</td>
<td>Small referral center</td>
</tr>
<tr>
<td></td>
<td></td>
<td>National registry &gt;50 centers</td>
<td>Therapies not available</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Standard entry criteria</td>
<td>Many did not have cardiac catheterization</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Prevalence may be underestimated because not all patients enrolled at every site</td>
<td>Industry sponsored</td>
</tr>
<tr>
<td>CMREF</td>
<td>2006-present</td>
<td>National registry</td>
<td>Only advanced disease</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Tissue, serum, and DNA</td>
<td>Serum not collected over time, limiting</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>biomarker creation</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Not inclusive</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>No translational projects all basic science</td>
</tr>
</tbody>
</table>

CHD = congenital heart disease; CMREF = Cardiovascular Medical Research and Education Fund; CTD = connective tissue disease; IPPS = International Primary Pulmonary Hypertension Study; NIH = National Institutes of Health; OCP = oral contraceptive pills; PAH = pulmonary arterial hypertension; REVEAL = Registry to Evaluate Early and Long-Term PAH Management; SMR = Scottish Morbidity Record; SPVU = Scottish Pulmonary Vascular Unit; WHO = World Health Organization.

with warfarin and/or vasodilators, such as calcium channel blockers and hydralazine. The registry found that more women than men had the disease, with a mean age at diagnosis of 35 years and a median survival of 2.8 years if untreated. Most of our knowledge of the natural history of PAH is derived from this landmark NIH registry, and thus the epidemiology of IPAH, FPAH, and anorexigen-associated PAH, in addition to the remaining World Health Organization (WHO) category 1 associations, in the current era is not well known.

This was followed by a collaborative European registry in 1996, the International Primary Pulmonary Hypertension Study, which included data from 220 centers in the United Kingdom, France, Belgium, and the Netherlands to explore the potential role of anorexic agents and other suspected risk factors for IPAH. The study also evaluated the incidence of IPAH in Belgium. The registry estimated an incidence of 1.7 cases per million population (95% CI, 1.0-2.4), the first estimate of incidence for the disease.

An attempt to evaluate the trends in the United States by use of hospital discharge information and national mortality statistics by the Centers for Disease Control and Prevention helped describe a more general population. The Centers for Disease Control and Prevention surveillance study reported the mortality results from 1980 to 2002 from the National Vital Statistics database and the National Hospital Discharge Survey. Unfortunately, although the report provided valuable information on mortality, it was unable to report incidence and prevalence determinations in the United States. By using hospital discharge data, there may be underreporting or overreporting of the disease, because the data are not collected by specialists and are often not recorded accurately by physicians.
More recently, three large academic observational registries from France, the United States (single center, Chicago: Pulmonary Hypertension Connection [PHC]), Scotland, and one small registry from China reported country- and institution-specific epidemiologic trends. The French network’s national prospective registry estimated the incidence to be 5 to 25 cases per million individuals per year of adult inhabitants for PAH. They calculated that the prevalence in France of PAH was 15.0 cases per 1 million adult inhabitants with 5.9 cases per 1 million for IPAH. A wide range of PAH regional prevalence was observed and most likely depended on proximity to diagnostic facilities. On an encouraging note, the 1-year survival of both the French and Chicago registries’ incident cohorts was significantly better than survival estimated by the NIH registry-derived equation (France 88%, Chicago 86%).

PAH is still detected late in the course of the disease when patients already have severe functional and hemodynamic compromise. The French reported a delay of 27 months from onset of symptoms to diagnosis. Patients in these contemporary registries were older than those in the previous reports with a substantial number >70 years of age (French = 50 ± 15 years and 9.1%; PHC = 48 ± 14 years and 8.5%). Most patients enrolled had IPAH, with the French reporting a higher enrollment of subjects with HIV and anorexigen-associated PAH, whereas Chicago reported a higher number of patients with PAH associated with connective tissue disease.

Peacock et al examined the incidence and prevalence of PAH in Scotland from two perspectives: Scotland population-based hospitalization discharge records from 1986 to 2001 (Scottish Morbidity Report [SMR]) and a specialized tertiary center with PAH expertise, the Scottish Pulmonary Vascular Unit (SPVU), from 1997 to 2005. During this period 4,794 patients were discharged with a first-time diagnosis of primary PH or PAH, of which 374 cases met today’s accepted definition of PAH. Therefore, the annual incidence of PAH based on this report was 7.1 cases per million population, with a prevalence as of the end of the calendar year 2002 of 52 cases per million population. Data from the SPVU were available for 1997 to 2006. In 2005, the last year with a complete data set, the incidence was nearly identical to the SMR data at 7.6 cases per million population, but the corresponding prevalence was lower, at 26 cases per million population, a number more similar to the French registry results. Peacock et al also reported a higher annual incidence of IPAH than previously reported, with 2.5 and 4.0 cases per million population in men and women, respectively.

Registries from smaller centers in various countries are useful in determining practice patterns in addition to epidemiologic trends. In a study from China, Jing et al described 72 patients with IPAH and FPAH from 1999 to 2004. During this period, the Chinese did not have access to PAH-specific therapy (standard of care included diuretics and anticoagulation), and thus it is not surprising that the cohort is more similar to the NIH registry. The poor survival at 1, 2, and 5 years of 65%, 56.9%, and 20.8% is therefore not unexpected. The average age was 35.9 ± 12.2 years (range 9.7–73.8 years) with a 2.4:1 female-to-male patient ratio. The mean duration of symptoms to diagnosis was similar to the French group at 26.4 ± 1 27.6 months. Only 20 patients had a catheterization, the gold standard for diagnosis, emphasizing some of the limitations of international guidelines and how they can be difficult to follow in different regions of the world.

Industry-sponsored registries help amass large numbers of patients by providing sites with support to collect the data. The Registry to Evaluate Early and Long-Term PAH Disease Management (REVEAL) registry is a multicenter (54 US sites), observational, industry-sponsored study currently enrolling patients with PAH. It will include prospective data for 6 years (2006–2012), including but not limited to standard hemodynamics, pulmonary function tests, chest CT-scan, functional class, therapies, and mortality data. Results presented in abstract form are similar to the PHC and the French data. This large-scale registry has the ability to better phenotype patients, study survival, and identify prognostic factors among the standard diagnostic tests used in the management of PAH.

**GLOBAL DATABASE: IS IT NEEDED?**

The epidemiology uncovered by the older or current registries will help the community to better understand the disease. Each has its strengths and weaknesses (Table 1), but it is clear that important features of the disease are dynamically changing, limiting interpretation of this data. For example, following three international meetings, (most recent Dana Point, California, 2008) the classification of PH and PAH-related syndromes has changed each time. Although not the intention of this effort, the leaders classified several distinct syndromes under the common umbrella of PAH. This implies that several of these potentially unrelated syndromes have common treatments. Thus, the targeted or treated patients are now quite different compared with old and current registry patients, limiting the overall value of these past registries.

As the field determines intrapatient and interpatient genetic and environmental susceptibilities and differences between PAH causes, it will be important...
to gather information from a multitude of centers throughout the world. This will allow researchers to better understand genetic variability and environmental influences, such as health-care access, nutrition, exposures, and so forth. None of our current databases captures information about the disease from a global perspective and none is able to capture enough data to validate potential novel biomarkers to be used in clinical practice and clinical research. Thus, the need for a large global database is critical for the future of our understanding and management of this still-deadly disease.

New Information a Global Database Should Include

PAH Incidence and Prevalence: The true epidemiology remains unknown for both IPAH and its related syndromes. This is mainly due to earlier diagnosis, the availability of treatment options, the emergence of organized PAH programs that provide systematic care and follow-up, and the changes in the definition of PAH and its subtypes. There is no information about the impact of diagnoses that are prevalent in nondeveloped countries, such as schistosome-associated PAH, hemolytic syndromes, or HIV-related PAH, among others. For example, in South America or sub-Saharan Africa these conditions are the most common of all the PAH syndromes and have a significant overall impact in public health. Tertiary care programs now exist to capture the majority of cases in a region or a country. This allows for relatively accurate estimation of PAH incidence or prevalence as long as these strategically important centers participate fully in a registry.

Basic Demographics: The existing dogma (based on the NIH registry data) is that the typical patient with PAH is a young woman in her 30s. Although the reason for the female predisposition, an intriguing phenomenon, remains unknown, recent anecdotal evidence suggests that this is likely true, and in fact the relative frequency of women with PAH may be higher than previously reported. This might be due to a more accurate capturing of PAH cases but also because of the inclusion of PAH syndromes that affect women more than men, such as scleroderma. In addition, there is evidence that the age of the typical patient is significantly higher than before. There is no information at all about the demographics of PAH outside North America and Europe.

Natural History and Survival: At this point, neither is known in the developed or developing world. There is some evidence that the recently approved therapies may improve symptoms and quality of life, but there is no evidence that they modify the natural history of the disease. Capturing WHO functional class status, exercise capacity, therapy regimens, and mortality data will provide invaluable global information, particularly in the long term (ie, >5 years). For example, the REVEAL registry captures data only within a relatively narrow window of time (6 years), which is likely shorter than the expected survival of many patients with PAH after the time of diagnosis. It is therefore unlikely that accurate information on natural history and survival will be provided, and this will definitely not reflect the world outside the United States.

Another important global issue is the skepticism of the need for placebo-controlled clinical trials at this point, particularly in newly diagnosed patients. The true impact of the currently used therapies may never be truly assessed because of the design of past clinical trials. The blinded-randomized phase of these studies typically extended ≤4 months. There is no question that there will be several parts of the world that will not have access to several of the modern PAH therapies for a while, because of cost and access. It will be very difficult for a center in India or Africa to offer an endothelin receptor antagonist or a parenteral prostacyclin, because the price of these drugs often exceeds $50,000 or even $100,000 per year. Many developed countries are now questioning the feasibility of providing such expensive therapies without clear and definitive impact on the patient lives. Only occasionally, patent laws in some areas of the world allow for earlier-than-normal availability of generic drugs; one such example is sildenafil in India. Thus, the ability to compare the natural history and survival of matched patients with and without a certain therapy might be an important alternative.

Global Database: Is It Feasible?

At first, the prospect of a global database might appear to be a colossal daunting endeavor. However, because PAH is a “specialist’s” disease, it is managed in specialized clinics, both in the developed world and, if they exist, in developing countries, typically in tertiary care or academic centers. This allows clustering of the cases within relatively few centers within each country and in centers that would have definite access to the internet and basic computer expertise. This is a much easier task compared with diseases that are seen typically in the community, such as hypertension or hyperlipidemia, which require national programs and national policies to capture the majority of the cases. These global programs do exist and serve as models of the global PAH database. For example, the WHO has launched the Surveillance of Chronic Disease...
Risk Factors projects, which aim to record the incidence of common risk factors (hypertension, hyperlipidemia, obesity, smoking, diabetes, and others) globally (https://apps.who.int/infobase/report.aspx). The project uses a variety of sources from which it obtains data, but it is governed by three basic principles: transparency, accessibility, and traceability.

Transparency refers to the ability to easily access and work with the stored data. Accessibility refers to the ability of all interested parties to easily contribute, access, and act on the information contained in the database. Traceability refers to the availability of an “audit trail” for all entered data; that is, any user can trace data points back to their original source and identify the methods used to assess, collect, and record the data. These three principles that govern the massive undertaking of the WHO could govern the much easier task of a global PAH database.

Transparency and Accessibility

Transparency and accessibility are most important at a time in medicine when conflicts of interest and relationships with industry require intense scrutiny. Recently the Institute of Medicine has made quite aggressive recommendations about how to define conflict of interest and the suitability of many clinical researchers with significant conflicts of interest to even participate in clinical research.13 Industry-sponsored databases will always suffer from biases (referral bias, for example, because they usually include data from patients in areas where the industry’s products are used or the industry has presence) and from potential conflicts, because it will be quite difficult (if not impossible) for data and conclusions that would hurt the interests of the sponsor to be recorded or disseminated. A global database with a potential huge impact in the field should be free of true or perceived conflicts of interest related to for-profit industries.

To ensure easy access, the data should be stored in a Web-based secure server with universal access and ideally simple navigation where data can be entered and accessed easily for all participants or registry users. In the United States, the database needs to comply with Health Insurance Portability and Accountability Act of 1996 (HIPAA) standards. HIPAA is a federal law passed in 1996 that sets basic requirements that health insurance plans must meet, including keeping a person’s medical information private.14 HIPAA responded to the health-care industry’s desire to reduce administrative cost by encouraging adoption of electronic transactions standards ensuring privacy and security of the patient details for public concerns.13 A very strict set of standards to ensure the protection of the identity and the security of the stored data has been set.

Although all registered users need to have equal and unrestricted access, a steering committee will need to supervise this process and set the standards, including membership of institutions and physicians to the registry. The analysis of the global data would be coordinated by the steering committee, whose members should not have true or perceived conflicts of interest, and the results would be available to the public in a free and unrestricted manner.

Traceability

Accuracy of the recorded data is of paramount importance, particularly because this would be the first attempt to capture data globally and because it is currently impossible to estimate what the data would actually look like globally (ie, there would be nothing to compare with). The diagnosis of all the diseases under the umbrella of the PAH diagnosis as well as their differentiation from the many secondary causes of PH is not only complex but also requires many and often expensive invasive (ie, cardiac catheterization) diagnostic tests. Unfortunately, not all practices, even in the developed countries, consider catheterization a mandatory test for the diagnosis of PAH. In addition, many developing countries might not be able to afford such testing. Obviously, inappropriate diagnosis will completely destroy the validity of this effort. There need to be very clear directions on how the disease is diagnosed, and cardiac catheterization needs to be mandatory. Appropriate catheterization will not only confirm the diagnosis but will also eliminate misdiagnosis of secondary PH conditions (ie, left ventricular diastolic dysfunction). CT scans and pulmonary function tests will be required to exclude other causes. The clinicians entering the data need to be experienced and comply with accepted criteria for entry. At the same time, the accuracy needs to be testable at any point, to strengthen the reliability of the project. Thus, although the patients’ personal data will not be entered in the database, the recording of a patient code, a physician code, and a center code will allow an auditor to trace the data to their origin and the records kept, including cardiac catheterization results. Thus, careful record keeping would also be a part of the database in each center, linked directly to the patient, the physician, and the center code numbers.

A “How To” Example

A new or old patient is seen and receives a code number, which is entered in the system along with the codes of the physician and the center (Fig 1). Date of birth and sex are then entered. Subtype of
PAH diagnosis is selected from a pull-down menu. The date of diagnosis is listed; this should be the date of the ultimate diagnostic test (ie, cardiac catheterization). The WHO class is then entered as well as the 6-min walk performance (provided that each center uses a protocol regarding the standardization and the consideration of the training effect for the 6-min walk). The treatment is then recorded from another pull-down menu and concludes the entry. This process is simple, easy, and should not take more than 3 min per patient to enter. The next entry for that patient should occur within a minimum of 1 year and ideally within 6 months.

**Figure 1.** A “how-to” example. Ca = calcium; ERA = endothelin receptor antagonist; IPAH = idiopathic pulmonary arterial hypertension; PDE5i = phosphodiesterase t 5 inhibitor; PG = prostaglandin.

**A Three-Step Process**

This basic database of the bare essentials could form the skeleton around which more sophisticated databases can be built in centers that have the ability or the interest to participate in the future (Fig 2). For example, the same code that was used for a specific patient would be used to capture additional data, whether genetic, molecular/biomarker, or imaging. In such sub-databases, a specific mutation or a right ventricular function index or a circulating biomarker would be able to be analyzed along with the natural history and survival data, in step II of the process. Such databases would need more sophisticated and organized teams and would require funding. Step III could be taken in the future to maximize the impact of such a global project, and an attempt could be made to merge this database with existing databases wherever possible.

**A Global Network**

Similar to all bold and novel efforts, a spearheading global network is required for this database to be successful. Investigators across nations will need to find common ground gathering basic statistics, and conceivably based on their findings, they could then pursue more focused research projects.

**Challenges**

**Legal/Ethical:** Although in some parts of the world, as discussed above, there are very strict and detailed
rules about how personal data could be stored and accessed, in other parts of the world there is complete lack of rules, guidelines, or consensus in definition of a registry. Thus the requirements of each center’s human ethics committees will likely be extremely diverse. Consent would be required for each patient to participate, particularly as more material (including genetic data and biomarkers) may be added in the future. It will be quite challenging to reach the ideal solution (ie, a common consent form). Global meetings would provide a starting point for such discussions. The intellectual rights for the global database would be owned by the investigators as a group. The possibility of merging the database with existing ones in the future (particularly those that are industry owned) will likely bring additional challenges but can be managed with guidance from the primary investigators.

Funding: A common “brake” for many novel, complex projects is the perception that nothing can be achieved without ample funding. The perception of the need for funding may stop many projects from even starting, even though actual funding is not required. The lack of any relationship with industry will be a significant strength but will obviously limit the funding sources. Such a project would of course be open to non-industry funding from either federal sources (NIH or European Union) or philanthropies with a global interest (Gates Foundation). More important, however, the need for funding for step one of the global database (Fig 2) is minimal, if any is required at all. The purpose of the how-to example above is to demonstrate the ease with which a clinician can enter data in a Web-based database from the office, spending no more than 3 min per patient. Dedicated data-entry personnel are not needed. There is also promise for the donation of the Web space for the database by a global data storage company. The ability of each center to enter patient data to the global database and to review their own center-specific data will allow several centers to create their own database, a significant advantage for many centers in the developing world that do not have these resources. To enter steps II and III of the project will likely require a significant amount of funding, but just completing this initial effort will be a great achievement with significant impact in this field.

Acknowledgments

Financial/nonfinancial disclosures: The authors have reported to CHEST the following conflicts of interest: Dr Gomberg-Maitland receives research support from Actelion, Gilead, Lilly ICOS, Novartis, Pfizer, and United Therapeutics. She is a consultant for Biomarin, Gilead, Medtronic, Millenium, and Pfizer. Over the past 3 years, Dr Michelakis has received a $1,500 honorarium as a lecture fee from Encevisy Pharmaceuticals.

References