

The International LAM Registry: A Component of an Innovative Web-Based Clinician, Researcher, and Patient-Driven Rare Disease Research Platform

Michael Nurok, M.D., Ph.D.,¹ Ian Eslick, M.S., M.Eng.,² Carlos R. R. Carvalho, M.D., Ph.D.,³ Ulrich Costabel, M.D., Ph.D.,⁴ Jeanine D'Armiento, M.D., Ph.D.,⁵ Allan R. Glanville, M.D., F.R.A.C.P.,⁶ Sergio Harari, M.D.,⁷ Elizabeth P. Henske, M.D.,⁸ Yoshikazu Inoue, M.D., Ph.D.,⁹ Simon R. Johnson, D.M., F.R.C.P.,¹⁰ Jacques Lacronique, M.D.,¹¹ Romain Lazor, M.D.,¹² Joel Moss, M.D.,¹³ Stephen J. Ruoss, M.D.,¹⁴ Jay H. Ryu, M.D.,¹⁵ Kuniaki Seyama, M.D., Ph.D.,¹⁶ Henrik Watz, M.D., Ph.D.,¹⁷ Kai-Feng Xu, M.D.,¹⁸ Elizabeth L. Hohmann, M.D.,¹⁹ and Frank Moss, Ph.D.²

Abstract

Background: A relative inability to capture a sufficiently large patient population in any one geographic location has traditionally limited research into rare diseases.

Methods and Results: Clinicians interested in the rare disease lymphangioleiomyomatosis (LAM) have worked with the LAM Treatment Alliance, the MIT Media Lab, and Clozure Associates to cooperate in the design of a state-of-the-art data coordination platform that can be used for clinical trials and other research focused on the global LAM patient population. This platform is a component of a set of web-based resources, including a patient self-report data portal, aimed at accelerating research in rare diseases in a rigorous fashion.

Conclusions: Collaboration between clinicians, researchers, advocacy groups, and patients can create essential community resource infrastructure to accelerate rare disease research. The International LAM Registry is an example of such an effort.

¹Divisions of Surgical Critical Care, Cardiac and Thoracic Anesthesia, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts.

²MIT Media Laboratory, Massachusetts Institute of Technology, Cambridge, Massachusetts.

³Pulmonary Division, Heart Institute (InCor), University of Sao Paulo, Sao Paulo, Brazil.

⁴Department of Pneumology/Allergy, Ruhrlandklinik-Universitätsklinikum, Essen, Germany.

⁵Division of Pulmonary and Critical Care, Department of Medicine, College of Physicians and Surgeons of Columbia University, New York, New York.

⁶Lung Transplantation Group, Division of Thoracic Medicine, St. Vincent's Hospital, Darlinghurst, NSW, Australia.

⁷Unità di Pneumologia e Terapia Semi-Intensiva Respiratoria, Ospedale San Giuseppe, Milan, Italy.

⁸Center for LAM Research and Clinical Care, Division of Pulmonary and Critical Care Medicine, Brigham and Women's Hospital, and Dana Farber Cancer Institute, Boston, Massachusetts.

⁹Department of Diffuse Lung Diseases and Respiratory Failure, Clinical Research Center, National Hospital Organization Kinki-Chuo Chest Medical Center, Osaka, Japan.

¹⁰Division of Therapeutics and Molecular Medicine and Nottingham Respiratory Biomedical Research Unit, University of Nottingham, Nottingham, UK.

¹¹Service de Pneumologie, Hôpital Cochin, Paris, France.

¹²Department of Respiratory Medicine, Clinic for Interstitial and Rare Lung Diseases, Centre Hospitalier Universitaire Vaudois, Lausanne, Switzerland.

¹³Translational Medicine Branch, National Heart, Lung and Blood Institute, National Institutes of Health, Bethesda, Maryland.

¹⁴Division of Pulmonary and Critical Care Medicine, Stanford University Medical Center, Stanford, California.

¹⁵Division of Pulmonary and Critical Care Medicine, Mayo Clinic, Rochester, Minnesota.

¹⁶Department of Respiratory Medicine, Juntendo University School of Medicine, Tokyo, Japan.

¹⁷Pneumologisches Forschungsinstitut GmbH am Krankenhaus Großhansdorf, Zentrum für Pneumologie und Thoraxchirurgie, Großhansdorf, Germany.

¹⁸Department of Respiratory Medicine, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences, Beijing, China.

¹⁹Partners Human Research Committees, Partners Healthcare, Boston, Massachusetts.

Introduction and Background

A RELATIVE INABILITY TO CAPTURE DATA concerning a sufficiently large patient population in any one geographic location has traditionally limited insight into rare diseases and complicated clinical trial design. To overcome this challenge, clinicians and researchers interested in lymphangioleiomyomatosis (LAM) from around the world worked with the LAM Treatment Alliance 501(c)(3), the MIT Media Lab, and Clozure Associates to cooperate on the design of a state-of-the-art encrypted web-based International LAM Registry that can be used for clinical trials and other research focused on the global LAM patient population. The International LAM Registry (ILR) is a component of a set of web-based resources, including a patient self-report data portal, aimed at accelerating research in rare diseases in a rigorous fashion. The ILR is due to launch with a discrete first project in 2010.

LAM is a rare multisystem disease affecting almost exclusively women, predominantly in their child-bearing years. The disease is characterized by the proliferation of abnormal smooth muscle-like cells that are associated with the formation of lung cysts and microscopic nodules, cystic structures in the axial lymphatics, and renal angiomyolipomas.^{1–6} LAM occurs in approximately one-third of women with the heritable genetic disorder Tuberous Sclerosis Complex.^{7–10} Sporadic LAM is much rarer, with an incidence of 2.6 cases per million women aged 20–69 years, although a significant proportion of women with LAM may be undiagnosed.⁶ While the disease is variable in its presentation and course, its most morbid impact is on the lungs, and progressive respiratory compromise may lead to death.^{2,3,6}

There are no effective treatments for LAM. Lung transplantation is the only known life-prolonging therapy.¹¹ LAM, however, can reoccur in the new graft, and genetic analysis of these cases suggests a metastatic basis for the disease.^{12,13}

Progress in cancer biology and research in both LAM and TSC yielded several potential molecular targets for this disease.¹⁴ The ongoing Multicenter International LAM Efficacy of Sirolimus Trial (trial registration identifier, NCT00414648, Frank McCormack PI) is currently investigating the effects of the mTOR inhibitor sirolimus in pulmonary LAM. In addition, a placebo-controlled trial investigating the effects of doxycycline is underway in the United Kingdom (EUDRA CT registration 2007-003745-32, Simon Johnson, PI).

Objectives

The International LAM Registry has the following objectives:

1. To accelerate research towards a treatment;
2. To provide scientists, clinicians, and patients with access to a state-of-the-art data coordination platform that can be used for clinical trials and other research focused on the global LAM patient population;
3. To enhance the understanding of the variability, progression, and natural history of LAM.

Towards these ends the registry will:

1. Create, reinforce, and expand a network of cooperating LAM clinician researchers and basic scientists;

2. Build clinical trials data and technology coordination functions into the online platform;
3. Increase access to harmonized clinical data by clinicians and researchers around the world;
4. Create a mechanism to address discrete scientific questions that would be difficult to answer without access to a global patient population;
5. Provide patients with access to:
 - a. A site for ascertaining their willingness to have their treating clinicians contacted for future trials;
 - b. Their own clinical information;
 - c. A centralized location for providing consent;
 - d. A portal to update their contact information to enable long-term tracking;
 - e. Web-based questionnaire surveys to improve clinician, patient, and researcher communication and simplify data collection.

Concept, History, and Funding

The LAM Treatment Alliance (LTA), a nonprofit 501(c)(3) organization focused on aggressively driving and funding research towards a treatment for LAM, conceptualized and funded the ILR by entering into a strategic partnership with the MIT Media Laboratory. The MIT Media Lab's research seeks to envision and explore the impact of emerging technologies on society. Researchers are developing technologies to support new forms of collaboration among doctors, patients, clinicians, and researchers with a specific emphasis on the problems encountered in rare disease communities. The ILR is one component of a set of web-based tools, including a patient self-report data portal (LAMsight) that, with patient consent, can be linked to create a robust record of de-identified data for each subject. These tools are aimed at accelerating research in rare diseases in a rigorous fashion. They are part of a vision that engages patients as valuable participants in the research process.

Previous registry initiatives have failed to achieve high levels of patient and data accrual due to the time-consuming nature of data entry and poor perception of benefit to individual researchers. To address these issues, the ILR's organization is focusing on maximizing incentives to participation. Philosophically, the ILR seeks to ensure that raw patient data are available as a community resource which allows scientists to compete for insight into LAM while ensuring that de-identified patient data remains in the public domain.

There are multiple incentives for participation. The most important envisaged incentive is providing simple coordinated access to clinical trials. In addition, the ILR will provide access to data from the global LAM patient population estimated at between 1000–2000 subjects by report of interested clinicians, allow complete freedom to publish one's own data at any time, provide for clinician inclusion on ILR publications in an appendix to all articles, include the option for clinicians to become authors or refuse use of their own data if a study uses more than 20% of patients from one's Center, and permit clinician and researcher access to an easy, secure data management platform with technical support at no cost.

Conventional views of data gathering hold that a discrete question should be identified in a hypothesis testing fashion and that data should be gathered to solely address the specific question. While this approach is entirely valid, enquiry of

large data sets have also yielded novel insights into disease processes.¹⁵ Initially, the ILR will emphasize the hypothesis-testing approach to facilitate collaborations and to improve the data management platform. The long-term goals for the ILR are clinical trials coordination and a data repository used to test for novel and unpredicted associations and insights into LAM and other rare diseases. For this reason, the platform was conceptualized and designed to handle and harmonize large volumes of data. The ILR platform will have many of the features required for data management and administration that anticipate the challenges of undertaking clinical trials for a rare disease. A given trial often replicates effort that may include patient recruitment, data management facilities, multisite coordination protocols, and the trial design itself. A common platform for LAM clinical trials could significantly lower the activation energy required to initiate new trials, improve intra-trial coordination, and provide a clear point of contact for long-term patient engagement both through and independent of a given clinic.

Through the collaboration with the MIT Media Laboratory, novel capabilities for the registry will be designed that help to simplify patient recruitment, support harmonization of data across sites, automate event and other regulatory reporting, data safety monitoring, and analysis and visualization of the ongoing trial data (where approved by a trial protocol).

Companion portal: LAMsight

LAMsight (www.lamsight.org) is an online portal being developed contemporaneously to the ILR.¹⁶ The website is an interactive community that allows patients to enter self-reported data in a de-identified fashion. With appropriate regulatory approval, researchers can place questionnaires and surveys on LAMsight. In future iterations of the platform being designed for LAM, patients will be able to consent to have their de-identified LAMsight data linked to their clinician entered de-identified ILR data. As future portals are developed (tissue, or genomic portals), with appropriate regulatory approval, de-identified data from these will be linkable, creating a robust research tool. Data entered into LAMsight are always labeled as self-reported and clearly distinguished from clinician-entered ILR data.

International LAM Registry Governance, Rules, and Rationale

The ILR will be governed by a Scientific Committee modeled on a journal peer review board. Peer review of all applications that make use of the ILR will be undertaken by the Scientific Committee which is comprised of independent and accomplished non-LAM clinicians and scientists, LAM experts, a patient with sophisticated insight into LAM, and an expert in human subjects oversight and regulatory affairs. Rules for participation in the ILR appear in Appendix 1. The rules were designed with a view to being inclusive while providing participation incentives for those clinicians entering the most patients. Several mechanisms were created to ensure these aims. Although any clinician seeing LAM patients can participate in the ILR, to be considered a "Collaborator," an individual clinician is required to enter at least five patients. Collaborators are then named in the appendix to all scientific publications that use ILR data. This rule ensures a

minimum contribution before the rewards from participating are received.

If any manuscript includes data from more than 20% of patients from a single participating clinician, then this clinician must consent to the use of his/her patients in the study and be given the opportunity to co-author or refuse the inclusion of his/her patients. This rule is designed to ensure that significant contribution to the ILR is rewarded and not exploited.

The ILR recognizes that there is an inherent tension between the desire to ensure that data are maximally accessible and in the public domain and the need to protect contributors from having their data used without reward. To resolve these conflicting interests, anyone with appropriate skills can make a request to use the ILR data for research, priority however is given to contributors. The Scientific Committee of the ILR will be the arbiter of these decisions.

The ILR seeks to facilitate research while at the same time ensure appropriate human subject protections. To facilitate research, interested investigators will be able to request a feasibility query from the Scientific Committee to determine whether there are sufficient data to address one or more specific questions. The query can take place prior to gaining local human subjects review board (e.g., IRB or Ethics Committee) approval. If the feasibility query is favorable, then the investigator will be able to apply for human subjects approval and on documentation of this, raw data from the query are provided. Several ILR rules were created anticipating future ILR functions. These may include the potential to link a subject's ILR data to tissue and specimen data, to data that a patient can self-report, or to genomic data. By providing patients with a portal to self-report data that are linked to the ILR, clinicians and scientists will be able to perform survey studies of patients and have the responses linked to the ILR, subject to appropriate regulatory approval. The ILR platform is not able to perform these functions currently and if collaborators decide that these functions will be useful, patients will be permitted to create the linkage with appropriate informed consent procedures in place.

Quality Control and Validation

Data will be regularly controlled for quality and prior to entering data into the ILR, participants are required to provide information which permits assessment and analysis of data quality, integrity, and provenance for the data they provide. Given the known variability in Pulmonary Function Testing (PFT) and their current importance in measuring LAM progression and response to treatment, Centers will be required to report on the level of quality control that they employ. Data from each Center will be assigned a quality level based on a range of quality markers including but not limited to compliance with standardized criteria for calibration and performance of tests, the ATS/ ERS pulmonary function testing recommendations, and for individual tests, a review of at least three de-identified blows by a spirometry expert.

In addition to self reporting, the ILR will appoint a quality control officer, a non-LAM pulmonologist, who will visit a selection of participating Centers on an annual basis to observe testing and quality control procedures. During this visit, the quality control officer will request a random selection of patient medical records entered from that Center and a

portion of the data from those records will be re-entered into the ILR as a test of data accuracy and provenance. The officer will submit an annual report and recommendations to the Scientific Committee.

Data from Centers not found to meet the self-reported standard will have their quality label adjusted accordingly. ILR participants agree to facilitate access of the quality control officer to the Center's Pulmonary Function Test laboratory, and medical records. Centers will be provided with a copy of their report. This information will not be made available to other ILR participants.

Organization

The International LAM Registry provides a browser-based application that enables the entry, aggregation, query, reporting, and export of clinical patient data to and from a centralized database. The description here captures the intended final form of the site and while many of these features have been implemented or prototyped, the final version of the site may differ in some respects as development proceeds and user feedback is accommodated.

The ILR's web-based interface is divided into a collection of modules that serve different functions and can be enabled or disabled to support different classes of users.

- **Viewer:** the data viewer provides an overview of a subset of patients in the database. Individual patient forms and data can be viewed and links are provided to the data entry module.
- **Entry:** the entry module facilitates the entry of data from clinical sites and registries worldwide. It provides a compact view of one or more forms. Special support is provided to handle the unique issues of provenance and harmonization across different sites and countries as described below.
- **Query and Analysis:** this module allows a user to submit queries to the database to retrieve field data for a subset of patient records. The query system further supports query refinement based on data provenance and privilege. It can provide direct or aggregated views of the data. Basic statistical tools and data visualization features are provided to characterize aggregated query results.
- **Dataset review:** the system records the results of approved queries and makes them available to the requester, if the requester has an account on the system. A record of query results and exported datasets are maintained for future download and review.
- **The system supports multiple user account types** to customize the site interface for the diverse rights and needs of clinicians, data entry personnel, researchers, individual patients, site curators, and site administrators. Each account type will have access to a specific subset of the system's modules and data (Appendix 2).

The ILR's database supports a wide variety of records and can be easily extended to accommodate new data types, such as blood tests, or content types such as CT scans. The primary data storage is a highly flexible schema for defining forms, each of which contains a sequence of, potentially complex, data fields. There are a large number of field types to support different data formats and layouts (such as numbers, mea-

surements, dates, and fixed rating scales). Each data entry associates a data value for a given field with a unique patient ID in the database. The interface can provide different views of the forms, fields, and data for different users and purposes.

Technical Challenges and Solutions

Creating infrastructure to support multiple sites in multiple countries presents a variety of specific design challenges for the ILR. Specific features and technologies are introduced to address each of the major challenges.

The site employs an object-oriented database architecture.¹⁷ This approach provides a highly flexible, and powerful query system, reducing the cost of making changes to the system over time. The design of the records supports the rapid creation and evolution of forms and fields to support different sites or different trials. A consequence of this facility is that data for a given patient may be recorded on different forms, while representing a common parameter (e.g., PFTs), that may be valuable in the course of an ILR query. A given field can be marked as being able to be harmonized with fields in other forms.

These and other relationships among data lead to potentially very complex provenance. Data may have been subject to different collection conditions or there may be multiple interviewers recording patient history. Data may be edited or updated over time. To enable query and analysis in the presence of these complex relationships, the database maintains a complete history and provenance of all database entries including date, individual who entered the data, originating site, and any additional necessary annotations.

Data harmonization is a significant challenge that arises when aggregating data across multiple sites. To adequately capture subtle differences between question phrasing or measurement procedure, the database provides support for per-answer annotations that record characteristics of the entered data. For ease of entry, pre-defined annotations can be selected from a drop-down control during data entry. Specific sites can set defaults to simplify data entry overhead. As with data provenance, additional annotations can be associated with a data field as required, such as whether patients have the ability to access particular forms or fields in their records.

Electronic security is ensured through several means. The HTTPS browser protocol used to secure most web-based electronic transactions is employed to encrypt data flowing to and from the ILR. Access to the ILR is enabled by a unique username and a password assigned to each user. Each user is assigned to a specific role category on the site by the Scientific Committee. A password policy is instituted that ensures the strength of the password and encourages regular updating. All data that are backed-up off-site are encrypted by a public-key encryption protocol that will ultimately be under control of the Scientific Committee.

One of the goals of the ILR is to enable long-term tracking of LAM patients by motivating patient involvement and self-report. Scenarios where this would be useful include, but are not limited to, information about patients who have moved to new clinics, de-duplication of patient records and tracking patients that have stopped seeing their treating clinicians. Updates over time and long-term tracking may be supported in two ways:

1. Patients can be provided their unique ID to log into the ILR to review and print their own clinical records. Patients can provide their electronic mail (e-mail) address in order to be contacted in the future for trials or to check on their current status.
2. Patients can participate in the companion, patient-oriented site LAMsight and their unique account there can be related to their ILR unique ID.

The availability of records and access to information about LAM provide incentive for individual patients to access these resources and to provide their e-mail address. The ILR and LAMsight both make the e-mail address invisible to all users. However, a clinician, a researcher, or the site curator can send an e-mail to users by using their unique ID. The system will map from the unique ID to the e-mail address anonymously, e-mail the user via an encrypted mail server interface, and record this interaction. The site can be enabled to allow patients to review their record and to initiate a query to both their treating clinician and site curator regarding data entered under their identifier that they believe to be inaccurate. Patients will have the ability to validate their own record in a similar fashion.

Regulatory and Ethical Concerns

Recognizing the diverse requirements for performing human subjects research throughout the world, the ILR rules require that researchers using the ILR demonstrate the receipt of local regulatory approval to participate in data entry and undertake research. In addition, the ILR has received approval for use at the corresponding author's institution by the Partners Human Research Committee and is subject to continuous review from this body. Sample institutional review board, ethics, and individual patient consent forms are made available to participants in order to minimize the necessary effort in acquiring approval.

When adding a new patient to the ILR, clinicians are encouraged but not required to enter the patient's name and date of birth. These data are not retained, and are used to create a unique one-way hash,¹⁸ or identifier, for the patient which prevents duplicate records from being created if the same patient is seen at multiple Centers. This information is kept in a secure part of the site to which users do not have access and which cannot be used to directly identify patients. A separate ID is issued to the clinician for that patient for all subsequent data entry. These procedures minimize the storage of personal identifiers in the actual dataset as specified by US standards on healthcare privacy as currently defined by the Health Insurance Portability and Accountability Act of 1996 (HIPAA).¹⁹

Where there is a compelling reason to store personal identifiers such as examination and diagnosis dates at a resolution more accurate than a year, or treating clinic and physician, these fields are marked as identifiers that can be automatically excluded from queries, resulting in fully de-identified datasets. A mapping between patient names and unique IDs may be maintained on physical media by the Scientific Committee for purposes of data recovery in the event a clinic loses the mapping or, for example, if a patient is requesting access to the records directly.

The ILR institutes procedural controls that complete these technical capabilities. All users will be validated and ap-

proved by the Scientific Committee. Only the site curator and administrator have access to the entire database and use of this interface must be approved by the Scientific Committee.

Finally, for technical support and other uses, the ILR maintains a record of user account access to the site and associated information such as the IP address of the requests. The site maintains these identifiers in the system log for the duration of a week and then expunges them from the system.

Future Functionalities

The ILR is currently focused on the essential issues of data capture, data quality, security, privacy, and governance. The ILR is being designed to be an integrated data management portal aimed at coordinating clinical trials. The ILR is a component of a set of web-based tools that, with patient consent, will be linked to create a robust record of de-identified data for each subject. Specific functions that may be added in the future include:

- Direct patient-researcher interaction modules designed to support questionnaires and other studies using instruments such as home peak flow meters that can record information electronically and upload to the site.
- A comprehensive tissue program module to further integrate global patient data and specimen resources by creating a centralized location for consent and allowing linkage between clinical, research, and tissue data.
- Genetic information including germline and somatic mutations in candidate causative genes, polymorphisms in candidate susceptibility genes, and genome-wide DNA, RNA, microRNA, and proteomic analyses.

Acknowledgment

Source of Support: The LAM Treatment Alliance, 501(c)(3)

References

1. McCormack FX. Lymphangioleiomyomatosis: A clinical update. *Chest* 2008;133:507-516.
2. Hayashida M, Seyama K, Inoue Y, Fujimoto K, Kubo K, Respiratory Failure Research Group of the Japanese Ministry of Health, Labor, and Welfare. The epidemiology of lymphangioleiomyomatosis in Japan: A nationwide cross-sectional study of presenting features and prognostic factors. *Respirol* 2007;12:523-530.
3. Johnson SR. Lymphangioleiomyomatosis. *Eur Respir J* 2006; 27:1056-1065.
4. Ryu JH, Moss J, Beck GJ, et al. for the NHLBI LAM Registry Group. The NHLBI lymphangioleiomyomatosis registry: Characteristics of 230 patients at enrollment. *Am J Respir Crit Care Med* 2006;173:105-111.
5. Taveira-DaSilva AM, Steagall WK, Moss J. Lymphangioleiomyomatosis. *Cancer Control* 2006;13:276-85.
6. Urban T, Lazor R, Lacronique J, et al. Pulmonary lymphangioleiomyomatosis: A study of 69 patients. *Groupe d'Etudes et de Recherche sur les Maladies "Orphelines" Pulmonaires (GERM"O" P)*. *Medicine (Baltimore)* 1999;78:321-337.
7. Franz DN, Brody A, Meyer C, et al. Mutational and radiographic analysis of pulmonary disease consistent with lymphangioleiomyomatosis and micronodular pneumocyte hyperplasia in women with tuberous sclerosis. *Am J Respir Crit Care Med* 2001;164:661-668.

8. Moss J, Avila NA, Barnes PM, et al. Prevalence and clinical characteristics of lymphangioleiomyomatosis (LAM) in patients with tuberous sclerosis complex. *Am J Respir Crit Care Med* 2001;164:669–671.
9. Costello LC, Hartman TE, Ryu JH. High frequency of pulmonary lymphangioleiomyomatosis in women with tuberous sclerosis complex. *Mayo Clin Proc* 2000;75:591–594.
10. Crino PB, Nathanson KL, Henske EP. The tuberous sclerosis complex. *N Engl J Med* 2006;355:1345–1356.
11. Glanville AR. Current status of lung transplantation. *Methods Mol Biol* 2006;333:105–130.
12. Bittmann I, Rolf B, Amann G, Löhns U. Recurrence of lymphangioleiomyomatosis after single lung transplantation: New insights into pathogenesis. *Hum Pathol* 2003;34:95–98.
13. Karbowniczek M, Astrinidis A, Balsara BR, et al. Recurrent lymphangioleiomyomatosis after transplantation: Genetic analyses reveal a metastatic mechanism. *Am J Respir Crit Care Med* 2003;16:976–982.
14. Juvet SC, McCormack FX, Kwiatkowski DJ, Downey GP. Molecular pathogenesis of lymphangioleiomyomatosis: Lessons learned from orphans. *Am J Respir Cell Mol Biol* 2007;36:398–408.
15. Warren JL, Klabunde CN, Schrag D, Bach PB, Riley GF. Overview of the SEER-Medicare data: Content, research applications, and generalizability to the United States elderly population. *Med Care* 2002;40:IV-3–18.
16. LAMsight website. www.lamsight.org; accessed 8 September, 2009.
17. Kim W. *Introduction to Object-Oriented Databases*. Cambridge: The MIT Press; 1990.
18. <http://csrc.nist.gov/publications/fips/fips180-2/fips180-2.pdf>; accessed 8 September 2009.
19. <http://www.hhs.gov/ocr/privacy/hipaa/administrative/privacyrule/index.html>; accessed 8 September 2009.

Address correspondence to:

Michael Nurok, M.D., Ph.D.

Divisions of Surgical Critical Care, Cardiac and

Thoracic Anesthesia

Brigham and Women's Hospital

Harvard Medical School

75 Francis Street

Boston, MA 02115

E-mail: mnurok@partners.org

APPENDIX 1. INTERNATIONAL LAM REGISTRY (ILR) RULES

1. The LAM Treatment Alliance (LTA) will fund the ILR as a service to the global LAM community on a renewable basis provided that the ILR adheres to its stated objectives, rules, and regulations. A minimum requirement is that the ILR remains transparent and accountable in its process and that the interests of the global LAM community, patients in particular, are put first.
2. Peer review and governance will be conducted by a Scientific Committee comprised of independent and accomplished non-LAM clinicians and scientists, LAM experts, a patient with sophisticated insight into LAM, and an expert in human subjects oversight and regulatory affairs. The Committee will be created by The LAM Treatment Alliance and a representative of LTA will sit on the Committee. The Committee will review and approve all applications to make use of the ILR data within one month of application. Decisions will be based upon scientific merit and the pre-stated goals of the ILR. Other considerations will include potential duplication of proposed projects, resource utilization, and legal, regulatory, and ethical concerns. The Committee will have the right to send proposals to domain-specific experts for review.
3. Participants are required to have enrolled a minimum of 5 patients to gain access as an ILR “collaborator”.
4. Clinicians being provided with or collecting data from colleagues in smaller Centers are encouraged to publish these data independent of the ILR and with their colleagues from small Centers as co-authors. This provision is intended as an academic incentive to smaller Centers to provide their data to the ILR.
5. Upon patient request, clinicians will provide an individual patient their unique identifier for the ILR portal. This will allow patients to see, but not modify their own clinical data on the ILR portal. Patients will not have access to their research data. In addition, patients will be allowed to link their ILR data with self reported data entered through a different portal on the ILR or LAM site. Self-reported data will always be clearly coded as self-reported.
6. Any health care provider or scientific researcher may generate proposals to publish aggregate data.
7. The publication policy that follows does not apply to site-specific data and, as such, clinicians are free to publish their own data at any time.
8. ILR collaborators, health care providers, or researchers interested in publishing ILR data must submit their proposal for publication and request for data analyses on the standardized registry form.
9. ILR collaborators, health care providers, or researchers can submit a request to the Scientific Committee for a feasibility analysis on a particular question in which they are interested. The standardized form must be used. A feasibility response will be returned within one month of application. Full data on the question will only be provided after the Scientific Committee has approved the request and proof of IRB approval to perform the study is provided.
10. Abstracts and manuscripts generated from pooled ILR data must be reviewed and approved by the Scientific Committee prior to submission for use in any public format. Review will take place within one month of submission.
11. If the data set used for the analysis is comprised in-total of more than 20% of patients from a single contributing physician, that contributing physician should consent to the publication and be invited to co-author provided his or her contribution is consistent with authorship criteria as described by the International Committee of Medical Journal Editors. The principal author will discuss the project and authorship directly with the contributing physician. The contributing physician will have the option to withdraw his or her data from this publication.
12. The International LAM Registry must be cited as an author on any publication; ILR collaborators will be listed in an appendix to the manuscript.
13. Manuscripts generated from registry data are required to comply with the International Committee of Medical Journal Editor’s requirements.
14. The International LAM Registry must be appropriately acknowledged as the source of all data. Funding sources for research and the ILR must be acknowledged in order of Funding significance. In addition, proper acknowledgement must also be given to those non-authors who contributed to the data analyzed.
15. This policy does not apply to requests for ILR data for the purposes of clinical interest or individual patient care provided that such data will not be presented in a public forum. Such requests will be addressed expeditiously and ideally within two weeks of receipt.
16. ILR data will be regularly validated.
17. Standards of data quality will be established by the Scientific Committee; ILR data will be labeled in accordance with its known provenance and quality.
18. In order to ensure that patients who are seen at multiple Centers are not entered under different identifiers, the ILR will encourage, but does not require, that each patient’s name and date of birth are entered. This information will only be available to generate a unique identifier for each patient.

APPENDIX 2. INTERNATIONAL LAM REGISTRY ACCESS LEVELS

- Clinician: Clinicians will have access to the full records of their own patients as well as the results of any approved queries run on their behalf and the resulting anonymized or aggregated datasets. Clinicians will be able to add new patients and edit the records of their existing patients. In addition, clinicians will have a basic query interface to search the records of their own patients.
- Data entry personnel: Data entry personnel will have access to de-identified patient records in order to add, edit, or update patient data.
- Researcher: A researcher account provides access to approved query results and aggregated data sets that can be downloaded from the website, but provides no access to a specific patient’s clinical records.
- Patient: The system can be modified so that a patient can be given access to an account that provides a read-only view of their own clinical data via the data viewer modules. Not all fields are necessarily visible as information collected and stored as part of a trial may not be appropriate to expose to the individual patient.
- Site curator: The site curator is responsible for maintaining the content of the site, adding and maintaining user accounts, and managing the clinical data fields and record definitions for clinical data. The curator runs data queries and analyses for all requests approved by the committee. The curator will also then make data available to specific clinician or researcher accounts.

