

An international registry of patients with plasminogen deficiency (HISTORY)

Amy D. Shapiro,¹ Marzia Menegatti,² Roberta Palla,³ Marco Boscarino,² Christopher Roberson,¹ Paolo Lanzi,⁴ Joel Bowen,⁵ Charles Nakar,¹ Isaac A. Janson¹ and Flora Peyvandi^{2,3}

¹Indiana Hemophilia & Thrombosis Center, Indianapolis, IN, USA; ²Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico, Angelo Bianchi Bonomi Hemophilia and Thrombosis Center and Fondazione Luigi Villa, Milan, Italy; ³Università degli Studi di Milano, Department of Pathophysiology and Transplantation, Milan, Italy; ⁴Misto s.r.l., Milan, Italy and ⁵Rho, Inc., Durham, NC, USA

©2020 Ferrata Storti Foundation. This is an open-access paper. doi:10.3324/haematol.2019.241158

Received: October 22, 2019.

Accepted: January 20, 2020.

Pre-published: January 30, 2020.

Correspondence: *AMY D. SHAPIRO* - ashapiro@IHTC.org

Database design

The study will include a minimum of seven (7) data entry points per subject (baseline and every 6 months for 3 years); data may be collected on other non-scheduled visits. Retrospective baseline data collection (demographic and clinical patient history for one year prior to enrollment unless <1 year of age) will be performed at each study site, with follow-up prospective data acquisition occurring by telephone by experienced clinical study physicians, nurses or study staff. In-person evaluation will occur in cases of suspected clinical manifestations indicative of pseudomembranes or other intermittent medical events. Laboratory evaluation will be performed at baseline and at end of study; if pregnancy occurs in affected females, PLG levels will be drawn at the time of pregnancy identification, during the third trimester and at delivery.

The design of the database will allow the collection of general information about the state of health of each subject with particular attention applied to: phenotype and genotype analysis; original diagnosis (age and reason for screening); type, site and number of clinical manifestations indicative of pseudomembrane formation either spontaneous or triggered; detailed information on type, intensity, and duration of any treatment administered for prophylaxis with reference to the type of product used, frequency and doses of treatment; treatment administered to treat acute or chronic pseudomembranes with reference to the type of product used, frequency and doses of treatment; laboratory parameters; use of any concomitant therapy; detailed information of management of surgical procedures; and complications associated with treatment. The database is housed in Italy at UNIMI.

Database security

All data for a given individual will be entered into the database under an assigned, unique alphanumeric code.

Data privacy within the project is the responsibility of the participating institutions (i.e. sites) and their operators (users entering data), as well as the database manager at UNIMI. Each institution or site can have multiple operators/users.

Each database operator/user is responsible for entering and/or updating data for that given site's study subjects. An operator's level of access will be defined prior to receiving database access; access will be limited to the data needed for the task at hand and controlled by the database manager. Each operator will be responsible for taking all necessary precautions to ensure data confidentiality and integrity.

Database program management and maintenance (backup and storage, etc.) and corresponding IT systems are the responsibility of the study database manager. The database manager will consult with and communicate to third parties only with regard to scientific and statistical results of the study data.

Once a year, a list of all individuals with database access and their tasks permitted will be produced. The risks associated with the data and their impact on security will be evaluated along with measures and procedures to mitigate those risks. The data protection systems will be evaluated and updated every six months to correct any vulnerabilities.

Every practical precaution will be taken to ensure that the study labelled biological samples and clinical, demographic and epidemiological information are maintained in a highly secure database system that employs state-of-the art firewalls, and password-protected technologies, and

enforces the procedures and policies necessary to ensure security of participants' data. Specifically:

1. Physical security is provided at the location of participant data storage. These measures include actions which limit and monitor physical access, including mandated use of computer controlled, identity-specific access devices, and a double locked environment
2. Network security is provided through various methods including the use of dedicated network lines, virtual private networks, firewalls, access-controlled lists, and a minimum of 128 bit encryption for information transfers involving unsecured networks
3. Databases are password protected, regularly backed up and supportive of the actions needed to determine role-based access. Audit logs are created tracking changes made to the database with a date/time stamp as well as the user making the changes
4. Passwords and other access control methods are changed on a regular basis and remain guided by specific policies and software to ensure that they are appropriate
5. Personnel are properly trained in the appropriate use of protected health information as supported by local policies and procedures
6. Access to protected information is on a need to know basis and only to authorized employees of the two principle investigative sites or their designees. All individuals receiving access will have been properly trained on the procedures and policies for conducting human subject research with such training documented

Data review

Data review will occur on a regular basis: baseline data within two weeks of entry into the database; 6-month follow-up visits, monthly as they accrue; and close out visits within two weeks of a subject's termination from the study.

Following review of baseline visit forms, any data for a subject that is discrepant with study inclusion/exclusion criteria (eg, PwPLGD with PLG activity >50%), that cannot be resolved through data query follow-up, will be brought to the attention of the study team to determine if the subject should remain in the study. A team member with an appropriate clinical background will review concomitant medication records to determine that medication, reason for medication, and dose are clinically relevant.

Database maintenance and data quality checks will be performed routinely. Data discrepancies will be flagged for follow up with the study sites. Discrepancies may include missing data, abnormal lab values, inconsistent dates (eg, study dates occurring prior to a subject's birthdate), or clinically implausible data (eg, male subject with pregnancy data). Data discrepancies and missing data identified during data review will be documented for each site; the investigators will follow up with the appropriate site to address any identified items. After a site has corrected the database, the corresponding study team member will confirm that data entry problems have been satisfactorily resolved.

Population and data analysis

The database will store variables measured once (such as baseline data) and repeated events including and concerning laboratory variables, clinical symptoms, treatment and side effects collected at follow-up visits. Variables will be reported as median values, interquartile ranges,

frequencies and percentage and will be compared by the Mann-Whitney U-Test and the Chi-square or Fisher's exact test. Descriptive statistics will be used for variables where appropriate. Frequency of the outcomes will be evaluated on all patients included. Because follow-up is continuing for years and inclusion is ongoing, not all patients will have the same follow-up in the database; to account for differences in follow-up, survival analyses or multivariable Cox regression analysis will be used to evaluate the occurrence of the outcome of interest. The analysis will include a comparison of major confounding variables.

The association between PLG levels and clinical phenotype parameters (when the latter are binary outcomes) will be evaluated by univariate and multivariate logistic regression. In the logistic model, restricted cubic splines with three knots will be used to search for possible non-linear relationships between PLG level and the outcome, expressed as log odds. Following the relationship between log odds and probability ($\text{probability} = \frac{\exp(\text{log odds})}{1 + \exp(\text{log odds})}$), a probability curve will be obtained, which represents the predicted probability of the outcome according to different levels of PLG. The receiver operating characteristic (ROC) curve will be used to assess what PLG level better discriminated between patients with or without the outcome. Sensitivity and specificity for different PLG levels will be calculated. The area under the ROC curve (AUC) with 95% CI will be used as an estimate of the predictive capability of the logistic model. All p-values will be two-sided and correspond to a significance level of < 0.05 . All analyses will be performed with the statistical software R (release 3.5.0 or higher; R Foundation for Statistical Computing).

Biorepository

A sample biobank will be created and housed at the IHTC and UNIMI. Samples will be stored for advanced testing and analysis to potentially include global hemostatic assays to investigate phenotypic correlation and whole genome sequencing to investigate further genetic modifiers of disease expression. Samples will only be stored with the consent of the participating subject.

The bio-repository will contain DNA, plasma and serum. Samples will be stored for a period of up to 15 years. Both the IHTC and UNIMI will be responsible for secure and safe storage of all collected samples and study participant information. Access to participant data and samples will be limited to study staff or their designees. All individuals receiving access will have been properly trained on the procedures and policies for conducting human subject research with such training documented.

The two main investigative sites will only release de-identified blood samples and information to collaborating researchers. Researchers that will be eligible to receive samples and information from the biorepository are expected to be a collaborative project; have a written protocol that documents scientific merit and is expected to lead to improvements in patient outcomes and public health; be approved by the researcher's local IRB; and have an investigator, recognized by peers as a person of integrity and merit in research.

At no time will personal or medical information in the study database be released to an insurance company, government agent or agency (excluding government public health agencies) or to the study participant's employer, potential employer, family member or friends. However, the information may be disclosed if required by law or to federal regulatory agencies and to the Institutional Review Board and their designees. Any published results from research using study samples will not identify study participants.

The biorepository will store and document access to samples and donor information. The samples will be maintained reliably with minimal deterioration over time, and they will be protected from physical damage, both accidental and intentional. The registration of each sample entering and exiting the system will be stored centrally on a computer that will be backed up frequently. Archival systems will de-identify samples to respect the privacy of donors and allow blinding of researchers to analysis.

To control temperature of storage, remote monitoring systems will be used to survey vital storage equipment in real-time throughout the facility. These systems will monitor equipment parameters 24/7 and can also send out instant, customisable notifications of power or mechanical failures. Some remote monitoring systems can also assist in determining when equipment is aged and needs maintenance or replacement. Biobank facilities will have back-up equipment, such as an alternative power source that is automatically activated when needed, and back-up storage, and have procedures in place to respond to equipment failure, weather emergencies and other critical situations.

Urine analysis

Local urine analysis will be requested for each patient to include dipstick, microscopic evaluation and testing for microalbuminuria. Local laboratory methodology for these tests from participating centers will be used and will not be mandated by the study but will be recorded.

Collection of blood samples

Blood will be collected by atraumatic venipuncture into plastic tubes containing 1/10th volume 0.129 mol/L buffered trisodium citrate. Plasma will be obtained by centrifugation at 2500 X g for 15 min at room temperature, transferred into plastic tubes, and stored along with leukocytes at -80°C until use.

Collection of serum samples

Blood will be collected into plastic serum tubes containing an activator and inverted at least 4 times (glass tubes without an activator will be clotted for 2 hours at room temperature). The collected blood sample will be kept at room temperature prior to being centrifuged; serum will be obtained by centrifugation (1500 X g for 15 min at room temperature) and transferred into cryovials. Samples will be stored at -80°C within 30 min of being centrifuged.

DNA isolation and in vitro amplification using PCR

Genomic DNA will be isolated from peripheral blood leukocytes by salting-out method. The coding regions, intron/exon boundaries, and 5', 3' untranslated region (UTR) of the gene of interest will be amplified by polymerase chain reaction (PCR).

Sequencing

PCR fragments will be purified and directly sequenced. Any mutation identified by direct sequencing will be confirmed by repeat sequencing and restriction enzyme digestion if associated with the creation or loss of a specific restriction enzyme. Each mutation will be verified in at least 200 alleles of control populations to discriminate between mutations and polymorphisms.

Specifically identified SNPs previously discussed including; 1. rs4252129 in PLG gene associated with a lowered PLG level; 2. rs1084651 in an intron of LPA gene associated with a

lowered PLG level; 3. rs10412972 5' upstream of the SIGLEC 14 gene associated with an increased PLG level; and 4. PAI-1 4G also will be analyzed.

Next generation sequencing

Next generation sequencing (NGS) may be required for patient samples in which a genetic defect using traditional sequencing approach cannot be found. NGS is a high-throughput sequencing method to sequence DNA and RNA quickly and cheaply. There are a number of different NGS platforms using different technologies, all perform sequencing of millions of small fragments of DNA in parallel. NGS can be used to sequence entire genomes or constrained to specific areas of interest, including all 22,000 coding genes (a whole exome) or small numbers of individual genes. The application of NGS could assist with:

1. The identification of mutations, such as large genomic deletions of exons or whole genes, inversions and translocations that could not be discovered with the traditional sequencing
2. The identification of possible other causative genes (including familial studies) and to identify novel pathways in coagulation disorders
3. Personalized medicine-based approach using targeted sequencing of genes involved in homeostasis to identify novel mutations and single nucleotide variants (SNVs)
4. Prenatal diagnosis as requested for plasminogen deficiency
5. Pharmacogenomics.

PLG antibody assay

An ELISA for qualitative determination of anti-human plasminogen immunoglobulin G (IgG) is proposed. Clinical serum samples will be incubated in a 96-well plate coated with human plasminogen with an anti-human plasminogen antibody as a positive control. Secondary labeled anti-IgG antibodies and a colorimetric substrate will be added; the endpoint measurement will be recorded as an optical density. Results will be reported as negative, positive, or equivocal.

Functional PLG activity-coagulation assay

Reference activity levels for PLGD will be determined using a standard frozen plasma pooled from healthy controls (usually 20 men and 20 women not taking oral contraceptive, except where noted). This reference is assigned an arbitrary value of 100%.

PLG in a plasma sample will be activated by the reaction of excess streptokinase in presence of fibrinogen. The plasminogen-streptokinase complex will be determined by the rate of hydrolysis of a chromogenic substrate. Plasminogen activity in subject plasma will be measured by an automated chromogenic assay on IL Coagulation System in two stages:

1. Incubation of the plasma with streptokinase reagent in the presence of fibrinogen;
2. Quantification of the plasminogen-streptokinase complex with a chromogenic substrate.

The released paranitroaniline will be kinetically monitored at 405 nm. Plasminogen results will be reported in % activity.

Compliance

In compliance with North American, European, and local country regulations and laws, participant information sheets and consent/assent forms will be obtained prior to the collection of clinical information and biological samples. Central and/or local IRB approval will be required prior to enrolling subjects; this will include translations of participant information sheets and consent/assent forms into local languages. Only study staff will consent patients; consent/assent will be documented by the signature of the subject and/or their legal guardian. A legal guardian/authorized representative will provide consent for a subject under the age of majority (18 in the US) or for a subject who is not able to provide consent due to diminished capacity. Diminished capacity will be determined either by clinical documentation from a qualified professional or through attestation of a legally authorized representative.

Privacy: HIPAA and GDPR concerns

Patients will be provided plain language patient information sheets and consent forms to collect clinical information and biological samples in accordance with the Health Insurance Portability and Accountability Act of 1996 (HIPAA) or General Data Protection Regulation (GDPR) in the US and Europe respectively. Specific measures will be implemented to safeguard data and privacy; personal identification data will not be entered into the web-based database. In all geographical locations, Europe, the US, or elsewhere, patients will be able to opt-in and provide prior written authorization prior to participation in HISTORY. Subjects will be notified of their ability to withdraw consent at any time, consistent with research data requirements of HIPAA and GDPR.