CCTIRS NON-INTERVENTIONNAL RESEARCH PROTOCOL

Study of the natural history of Alport Syndrome by establishment of an International database

RaDiCo-EURBIO-Alport

VERSION N°1.0 DU 09/November/2015

N° Gestionnaire	N° CPP	N° MESR	N°CNIL
C15-82	2016-janvier-14130 ND	DC-2015-2564	916204

CONFIDENTIAL

Organisme responsable : Inserm - ITMO Santé Publique - Pôle Recherche Clinique (PRC) Biopark, Bâtiment A, 8 rue de la Croix Jarry, 75013 Paris Carole Pierrart, carole.pierrart@inserm.fr Tél: 01 44 23.67.29 Fax: 01 44 23 6126 **Délégation Gestion** : **RaDiCo** Didier Beudin, Chef de projet Recherche Clinique Didier.beudin@radico.fr Tél : 01 71 73 87 94	Responsable de la recherche : **Dr Laurence Heidet** Co-Responsable de la recherche : **Pr Bertrand Knebelmann** Fonction : PH, PU/PH Unité Inserm d'affiliation : U1151 Adresse : Hôpital Necker, 149 rue de Sèvres, 75015 Paris Tel : 01 44 49 44 63 / 01 44 49 54 58 Fax : 1 44 49 44 60 / 01 44 49 54 50 Email : laurence.heidet@aphp.fr bertrand.knebelmann@aphp.fr

☐ **Etude monocentrique**
☐ **Etude multicentrique Français**
☐ **Etude multicentrique Européenne**
☒ **Etude multicentrique Internationale**

TABLE OF CONTENT

SIGNATURE PAGE:

Title : Study of the natural history of Alport Syndrome by establishment of an International database (RaDiCo-EURBIO-Alport)
Version N° 1.0 of 09/November/2015

The study will be conducted according to protocol and legal and regulatory requirements.

The research has received authorisation from CNIL on 13th March 2017 .

The research has received authorisation from CPP of Ile de France on 31 march 2016.

Scientific leader / Coordinating investigator Dr Laurence Heidet Service de Néphrologie Pédiatrique Hôpital Necker, 149 rue de Sèvres, 75015 Paris, FRANCE	Date : ……/……/…….. Signature :
Co-Scientific leader / Coordinating investigator Pr Bertrand Knebelmann Service de Néphrologie Hôpital Necker, 149 rue de Sèvres, 75015 Paris, FRANCE	Date : ……/……/…….. Signature :
Clinical Research Project Manager - RaDiCo-Inserm U933 Didier Beudin, Hôpital Armand Trousseau Bâtiment Lemariey – 3eme étage 26 Avenue du Dr Arnold Netter 75012 Paris, FRANCE	Date : ……/……/…….. Signature :
Investigator : Name : …….. Hospital / Site / Department : …… …….. Town : …….. Country : …….	Date : ……/……/…….. Signature :

GLOSSARY

AP-HP :	Assistance Publique-Hopitaux de Paris
AS:	Alport Syndrome
BNDMR :	Banque National de Données Maladies Rares / French National Bank for Rare Diseases Data
CCMR:	Centre de Compétences Maladies Rares / Rare Diseases Competence Centre
CCTIRS :	Comite Consultatif sur le Traitement de l'Information en matière de Recherche dans le domaine de la Santé
CEMARA :	Centre des maladies rares
CKD :	Chronic Kidney Disease
CNIL :	Commission Nationale de l'Informatique et des Liberté / French National Commission on Information Technology and Liberties
CPP :	Comité de Protection des Personnes / Ethics Committee
CRMR :	Centre de Référence Maladies Rares / Rare Diseases Reference Centres
CRF:	Case Report Form
eCRF:	Electronic Case Report Form
eGFR :	Estimated Glomerular Filtration Rate
ESRD:	End-Stage Renal Disease
F-CRIN:	French Clinical Research Infrastructure Network
GCP:	Good Clinical Practices
GEP:	Good Epidemiological Practices
HADS :	Hébergeurs Agréés de Données de Santé / Accredited Health Data Hosting Provider
HRCT:	High resolution Computed Tomography
ICH E6:	International Conference On Harmonisation for Good Clinical Practice
IEC:	Independent Ethics Committee
IRB:	Institutional/Independent Review Board
ORKiD:	Orphan Kidney Diseases
QoL:	Quality of Life
RAAS Blocker:	Renin–Angiotensin–Aldosterone System blocker
REDCap :	Research Electronic Data Capture
RRT:	Renal Replacement Therapy
SAP :	Statistical Analysis Plan
SFTP:	SSH File Transfer Protocol or Secure File Transfer Protocol

1. STATUS OF THE QUESTION AND STUDY OBJECTIVES

1.1. Status of the question

Alport Syndrome (AS) is an inherited disease characterized by the association of a glomerular nephropathy, a sensorineural deafness and retinal or corneal defect 1. Its frequency is about 1/5000. It is associated with mutations in one of the three genes encoding the alpha 3, 4, and 5 chains of type IV collagen, which form a distinct network in the glomerular basement membrane that is essential for the long-term stability of the glomerular filtration barrier. The disease can be inherited as a dominant X-linked, as an autosomal recessive or an autosomal dominant trait. Patients initially present with hematuria, and subsequently develop proteinuria followed by progressive renal failure. Median age at end stage renal failure is about 20, but there is a large inter- and intra-familial variability. The progression of the disease can be divided in 4 stages: isolated hematuria, microalbuminuria, macroproteinuria, progressive renal failure. Earing and ocular defects also have a progressive evolution.

Important progresses have been made in the last two decades in the understanding of the molecular basis of the disease, and animal models have been established [2]. However, few advances have been made in the understanding of the mechanisms responsible for the progressive aggravation of the renal disease and in finding therapies, even if a few treatments have been shown to slow down the renal disease progression in animals (Renin Angiotensin Aldosteron system or RAAS blockers, TGFβ inhibitors, metaloproteinases, hematopoietic or amniotic stem cells, anti-micro-RNAs….) [3]. In humans, a European retrospective study has shown potential effect of RAAS blockers [4]. Most physicians now prescribe these drugs to AS patients having developed proteinuria [5]. However these results still need to be confirmed by controlled studies. The question of whether these treatments, if given before the occurrence of micro-albuminuria, may delay the progression of the disease, has never been addressed. Only a minority of patients has isolated hematuria (without proteinuria and/or renal failure). A study to address that question would thus require multicenter international collaboration. Also no study has evaluated the impact of the disease on school and professional life, the access of patients and families to genetic counselling, and the tolerance and compliance of the patients to RAAS blockers treatments.

Because the progression of the disease is very variable from an individual to another (even within a given family), and because a « window » for efficient treatment may occur before the occurrence of proteinuria, it would be very helpful to identify novel biomarkers able to predict the rate of progression of the disease. Given the complexity of the mechanisms involved in the progression of the disease, it is likely that a single molecule may not predict the evolution of the disease.

1.2. Objectives

1.2.1. Primary Objective

Main objective is to study the natural history of the Alport Syndrome.

1.2.2. Secondary Objectives

- Evaluation of tolerance and compliance to treatments with RAAS blockers (for French sites only)
- Evaluation of the impact of the disease on quality of life as well as social impact (such as school or professional eviction) (for French sites only) by comparison of calculated scores based on widely used survey completed by patients and parents (e.g. Short-Form Health Survey (SF-36) or equivalent adapted for children (SF-10)).

1.2.3. Exploratory objectives:

- To study the feasibility of therapeutic trial in term of number of patients that may be enrolled, at the different stages of the disease.
- To search for early biomarkers able to predict the progression of the disease and to evaluate the efficacy of treatments. This will be done after the establishment of a urine collection from patients affected with early stage AS

- Evaluation of the ease for patients and families of access to molecular diagnosis and genetic counselling (for French sites only)

1.2.4. Information Technology Objectives

- Develop and diffuse an electronic tool for data collection from various sources linked to a database integrating a system of management and follow-up of data-management. This electronic tool will allow the collection of data for Alport Syndrome in paediatric and adult patients over France, Germany, Spain, Italy, United Kingdom, Belgium and Hungary and potentially USA.
- Include data generated by patients and, where relevant, their parents and or carers (Quality of Life data for French sites only).

2. STUDY POPULATION

2.1. Description

The countries that are potentially interested in joining the project have identified in their local records about 100 (Germany), 200 (Spain), 100 (United Kingdom), 785 (Italy), 100 (Belgium) patients. Patients in Hungary are currently not recorded in a database. The database ASTOR in USA have included 800 patients. For France, the CEMARA database (CNIL authorisation number: 1187326) contains currently 680 Alport Syndrome patients while there are 950 identified patients in France. These patients are followed in nephrology and paediatric nephrology services belonging to the French network for rare renal diseases, called “Filière de santé maladies rares ORKiD” (4 Reference Centers and 16 Competence Centres, see Appendix 3). It is also expected to include adult patients followed by nephrologists out of the “Filière ORKiD”.

Prevalent cases will include cases already registered in the different existing databases in the different countries, after monitoring that the inclusion criteria are respected. It is estimated to include around 1500 prevalent patients in Europe and possibly about 2500 world-wild.

There is no known estimated incident rate, but according to the available information and the estimated prevalence rate (1/5,000), it is expected to enrol roughly between 100 and 200 incident patients/year in the study. With a recruiting period of 2 years the total number of new patients is expected to be between 200 and 400.

Therefore, it is possible to include a maximum of about 2000 prevalent and incident patients in Europe, and potentially up to a total of 3000 patients worldwide.

2.2. Inclusion Criteria

- Diagnosis of AS based on (i) electron microscopic examination of the renal biopsy and/or (ii) molecular studies and/or (iii) abnormal expression of type IV collagen chains on skin and/or glomerular basement membranes.
- Signed informed consent

2.3. Non-inclusion Criteria

NA

2.4. Recruitment Process

2.4.1. Sites selection

2.4.1.1. Method of selection

The clinical partners chosen to participate in this cohort have been selected based on their national status as centres of expertise or recognised with the expertise.

Clinical partners in France are members of the ORKID rare disease network (French "Filière de Santé Maladie Rare") for rare kidney diseases; therefore a lot of AS patients are seen in these centres.

French nephrologists who are not from part of the network of renal rare disease (ORKiD) will also be selected based on their expertise and the number of Alport Syndroms patients they follow. These sites mostly follow adult AS patients.

Although not as well structured as in France, comparable networks of AS care centres exist in the other countries. Non-French centres have been selected based on their experience in Alport syndrome and their involvement in the International Alport network.

2.4.1.2. Number of site

So far, a list of about 40 sites have been identified including: the ORKiD network (4 Reference Centers and 16 Competence Centres), and 6 additional sites in the following countries (1 in each country): Germany, Spain, Italy, United Kingdom, Hungary, Belgium and potentially USA.

See appendix 3 for detailed list.

2.4.2. Patients selection

In France, paediatric and adult patients will be mainly recruited through the network of reference, competence and recognised expert centres of rare kidney diseases. Investigators will inform patients meeting the inclusion criteria about the RaDiCo-EURBIO-Alport cohort and invite them to participate during regular care follow-up visit for prevalent patient and during their first regular care visit (post-diagnosis) for incident patient.
The international partners are centres seeing a large number of Alport patients; however their coverage is rather regional than national, therefore they will be encouraged to collaborate with other centres in their country to enrol more patients.

Informed consent form and patient information sheet will be provided and explained by the investigator.

Participation in another study is not an exclusion criterion as this is a follow-up of cohort type study. Patients will be given as much time as necessary to evaluate their participation to the study.

A concerted effort will be made to include most patients. Communication will be achieved through the physician societies (society of nephrology, society of paediatric nephrology, societies of transplantation), patient organisations, industry, dialysis and transplantation centres.

2.5. Patient Information and Patient Informed Consent Process

An informed consent form as well as an information notice adapted to the cohort study will be proposed for signature at the time of inclusion visit. Only investigators of the centres listed in appendix 3, will be entitled to include patients, either during their follow-up visit, or during their first post-diagnostic visit. An information notice will be read and explained to the patient or his/her parents or guardians by the investigator who will answer any question the patient or his/her guardians may have. The objective, methodology, duration of the study will be presented. The investigator will inform the patient whom consent is solicited about his/her right to refuse to participate to the study or to withdraw at any time

his/her consent from the study without incurring any prejudice nor impact on his/her routine medical care.

In case of minor/protected adult patient, the later will receive information according to its capacity of understanding, from investigators of experience with minors/protected adult, regarding the study, the risks (those related to the actual standard of care for these pathologies) and benefits (improvement of knowledge and standard of care). The legally authorized representative will also receive the information and help the minor/protected adult participant to understand. When a subject deemed legally incompetent, such as a minor child, is able to give assent to decisions about his / her participation in the study, the investigator must obtain that assent in addition to the consent of the legally authorized representative. The evaluation of whether or not a child/protected adult can give assent will not solely be based on chronological age, but will also depend on other factors such as developmental stage, intellectual capacities, life / disease experience, etc. This will be made after discussion of the parents / legal representative with the investigator, but the parents will normally know the child/protected adult best and hence will usually be in a position to decide on whether the child has understood the information as much as is possible.

The investigator will provide assistance in answering any question raised by the patient or his/her legally authorized representative(s). The explicit wish of a minor/protected adult who is capable of forming an opinion and assessing this information to refuse participation or to be withdrawn from the study at any time will be considered by the investigator.

In case of consent, the patient or his/her legally authorized representative on one hand and the investigator on the other hand will fill out and sign the consent forms. The patient and the investigator will each keep an original signed version of the documents. Inclusion date, date of informed consent withdrawal (if applicable) as well as modalities of collection of informed consent form will be mentioned in the patient file.

All patients included in the cohort before the age of 18 years old must, at the time of his(her) majority, sign a new informed consent form to participate in the study.

Participants (or guardians) will be asked to sign in confirmation of the following items:

- Consent for collection of personal medical data
- Consent that collected personal medical data can be shared as de-identified data with other research team for research purpose*
- Consent to document their quality of life through specific questionnaires (French only as a start)
- Consent for study researchers to view medical notes
- Consent for transfer of previous collected personal medical data into the RaDiCo-EURBIO-Alport database*
- Consent to sampling urine (French only)
- Consent to use samples for research purposes
- Consent to retain portions of samples for further research, as planned by the protocol, aiming to better characterize anomalies responsible for their disease*
- Consent to send portions of samples to collaborating centres for additional specialist analysis*
- Consent to be re-contacted to discuss further information (this may be requesting further information from the family, or to share new findings about the cause of disease, or to ask for participation to a new clinical study)*
- Consent to publish results of their medical information about the cause or progress of disease.
- Acknowledge that have had sufficient time to think about his/her participation to the study

**Records of consent will be kept in the RaDiCo-EURBIO-Alport database*

Understanding that the participant may at any time require the destruction of his/her clinical data, simply by informing his/her practitioner, with no impact on his/her routine medical care.

Patients whose personal medical data have been previously recorded in the CEMARA database, although they were informed and did not oppose to the collection of personal data in this database at the time, will go through the same informed and consent process and in particular will further be informed about and consent to the transfer of their personal data previously collected in CEMARA into the

RaDiCo-EURBIO-Alport database. Such participants (or guardians) will be asked to sign in confirmation of the following additional item:

- Consent for transfer of previous collected personal medical data into the RaDiCo-EURBIO-Alport database.

In any case, consent refusal or revocation by the patient or his/her legally authorized representative will not be ignored. Moreover, written authorization of parental authority holders is mandatory. However, this authorization might be given by the sole parental authority holder present, as risks and constraints of the study are those of the routine medical care and if the other parental authority holder cannot give his authorization in deadlines compatible with methodological requirements related to the research and its purposes.

Some patients whose data have been collected in the CEMARA database may be lost to follow-up. However, the medical data of these patients are crucial to answer the primary objective of the study (characterize the natural history of AS). These patients or their legal representatives did not oppose to the collection of personal medical data at the time they were informed about CEMARA, and therefore if approved by local authorities (i.e; CNIL in France), participating sites will look for the non-opposition of patients (or legal representative) for the transfer of medical data into the RaDiCo-EURBIO-Alport database. Therefore, unless these patients explicitly opposed to this data-transfer, their data will be included within the cohort database. For the same reason, the medical data of AS dead patients recorded in CEMARA are crucial to answer the primary objective of the study. As these patients or their legal representatives gave their consent and/or did not oppose to the collection of personal medical data at the time, a waiver to the information of the families of dead AS patients will be requested for ethical reason to each local authorities (i.e.; CNIL in France) before the transfer of their medical data into the RaDiCo-EURBIO-Alport database. The fact of contacting families about their dead relative and reminding them about this severe familial pathology, could upset them more than they are already.

The information and consent form is a document that has been approved prior to the implementation of the research by the Ethics committees, on the occasion of the protocol review.

3. ADOPTED METHODS OF OBSERVATION

This is an International multi-centre, longitudinal, non-interventional study that uses observational study methods to collect descriptive, retrospective and prospective data, allowing the collection of consistent and comparable clinical, biological, molecular, histological and quality of life data to study the natural history of the Alport Syndrome.
Regulatory requirements will first be initiated in France by Inserm, and once all French authorisations are obtained, regulatory submissions in the other countries will be managed by the local sites in each countries.
The study will be associated to a urine collection. All corresponding regulatory requirements for the bio-collection will be performed at start of the study. However, in a first step only the site of Necker Hospital, APHP Paris will collect urine samples. The other participating sites will start to collect urine samples, only when funding will have been secured to cover the samples shipments and storage.
After obtaining the consent of patient and / or legal guardians for minors or those under guardianship, data will be collected during regular visits of routine care. Data collection arrangements are detailed in the paragraph: "Duration and organization of the cohort." No additional examination to patient current care is necessary. However, patients followed by French sites, will have to complete surveys at every visit to assess the repercussions of their pathology on their quality of life (SF-36 for adults and SF-10 for children) and their compliance to treatment. They will be asked to complete a survey about access to genetic counselling/diagnostic at inclusion visit. It is also recognized that a cohort study should improve standards of care by formalizing questionings and clinical tests that must be performed at each follow-up visit.
Stratification according to gender and mode of heritability will be performed prior analyses.

3.1. Main objective evaluation criteria

Description of the symptoms and variations in the diseases' course:

- Renal function: eGFR, age at ESRD, requirement of Renal Replacement Therapy (RRT) and type of RRT
- Urine bio-analysis results: Presence or not and quantification of hematuria, microalbuminuria and proteinuria
- Presence or not of hypertension
- Level of Hearing loss
- Ocular symptoms (presence or not of lenticonus, cataract, retina and cornea impairment)

3.2. Secondary objectives evaluation criteria

- Records of adverse events for the long-term safety of RAAS blockers treatment.
- For French sites only: Impact of disease on QOL will be evaluated through scores of QOL questionnaires SF36 for Adult et SF10 for paediatric patients
- For French sites only: Compliance will be evaluated using X. Girerd Compliance Questionnaire

3.3. Exploratory objectives evaluation criteria

- Stratification of patients according to their disease stage; patients' distribution analysis among countries
- Association assessment between the urinal concentration of the five molecules recently described by Terzi's lab as predicting progression of CKD (or other putative biomarkers) with the rate of decline of the GFR (according of the estimated GFR) on a 3 year- period.
- Result from survey about patients and families access to molecular diagnosis and genetic counselling (for French sites only).

4. ORIGIN AND NATURE OF NOMINATIVE COLLECTED DATA AND JUSTIFICATION FOR USE

4.1. Origin and nature of collected data

Data collected from French sites in CEMARA (CNIL authorisation number: 1187326) will be transferred to the RaDiCo-EURBIO-Alport database. From that moment, French sites will enter no additional data in the CEMARA database that is replaced by the RaDiCo- EURBIO-Alport database.

In Germany, data on prevalent AS patients have been collected in an European Alport Registry. In Italy, prevalent data on AS patients have been collected in the Italian Alport National Registry. In United Kingdom, prevalent data on AS patients have been collected in the RADAR Registry. In USA, data on AS patients are collected in the ASTOR database.

Regulatory requirements in France will be the responsibility of Inserm and regulatory requirements in Belgium, Italy, Spain, The Netherlands, Hungary, United Kingdom, Germany and USA will be the responsibility of the local sites.

Time of importation of retrospective data and start of inclusion of new patients will differ by country as this will rely on timelines to complete regulatory requirements and initiate the study in each country.

Once all local regulatory requirements are completed for a site/country, relevant AS data for the cohort from the local existing databases, will be transferred to the RaDiCo-EURBIO-Alport database in RedCap™.

From that point, all new data (either from follow-up visits of prevalent patients or from new patients to be included) will be directly entered into the RaDiCo- EURBIO-Alport database by the investigators in charge of the patients in each participating country. If non-French sites are willing to continue collecting data in their local database, regular exports to RaDiCo-EURBIO-Alport database will be performed according to the process described in the following section: "5 METHODS OF DATAFLOW".

Data from patient's auto-questionnaires will be recorded via a specific designed secured webpage / REDCap™ form directly into the RaDiCo- EURBIO-Alport database.

Nature of the data	***Collected data for the study***
Identification by full name	No
Identification by number order	Yes
Other way of identification (define): IdMR	Yes
Health (i.e.: personnel and family medical history, treatments …)	Yes
Family status (i.e.: single, married…)	Yes
Military Situation	No
Training and Education (i.e. : school level, professional training)	No
Professional Situation (i.e. : working, unemployed, occupational categories …)	No
Economic and financial Situation (i.e.: income level, social benefits…)	No
Good and services consumptions (i.e.: home care, transportation…)	No
Lifestyle, behaviour (i.e.: alcohol, tobacco consumptions, practises of sports ...) *	Yes
Social security number or RNIPP (National Heath Number)*	No
Ethnical origins or political, philosophical, religious opinions, trade union membership, or mores*	No
Identifying Biological samples	No
Crimes, sentences or security measures*	No

* Some information on lifestyle behaviours will be collected as part the Quality of Life questionnaire.

4.2. Collection of human biological samples

4.2.1. Description of the collection

Urine samples will be collected only for patients not at ESRD stage (not under RRT) and followed by French sites. Two tubes of 15 ml will be collected: one with and one without protease inhibitors. The first tube will be used to measure the 5 biomarkers previously shown by Terzi's group to be able to predict the evolution of the CKD better than albuminuria. This bio-collection will be limited to France.
In case this candidate approach will not allow to predict CKD progression in this cohort of patients, an unbiased analysis (mass spectrometry) will perform on the second sample of urine collected without anti-proteases.

4.2.2. Logistics of the collection

Urines will be collected four times, at T0, and one, two and three years later. They will be labelled with the RaDiCo code, patient initials, sex of patient and date of sample. According to resources on site and location, tubes will be processed as follow:

- Tubes are stored at +4°C up to 24 hours maximum and shipped strictly at + 4°C to the storage laboratory.
- Tubes are centrifuged at 1000 g 15 minutes at +4°C and decanted and stored supernatant with or without protease inhibitors at -20°C.Shipment in dry ice in the week.
- Tubes are centrifuged at 1000 g 15 minutes at +4°C and decanted and stored supernatant with or without protease inhibitors at -80°C. Grouped shipment in dry ice.

Tubes are sent to: Laboratoire Inserm U1151, 6ème étage, Bâtiment Lavoisier, Hôpital Necker, 75015 Paris. On arrival, tubes are added a number on their label corresponding to their order of arrival in the laboratory and are stored in secured location at -80°C.

Analysis for biomarkers will be performed in a second step when all samples will have been sent back.

Budget for providing tubes and shipment needs to be consolidated. Therefore the collection will be imitated at the Necker Hospital and extended to other sites in a second step.

5. Methods of Data Flow

RaDiCo-EURBIO-Alport will not require export of data from the national SNIIRAM database.

5.1. Purposes and environment of the RaDiCo Information System

Software applications, client or server processing computers, infrastructures, networks and physical equipment used for the RaDiCo-EURBIO-Alport (Rare Diseases Cohort) cohort study are all destined to provide services for running tasks for each clinical study and for no other purpose whatsoever.
Computer equipment made available by Inserm-UMR S 933-RaDiCo, is used to assist the teams conducting clinical studies of the selected cohort.

5.2. Data sources and inputting databases

5.2.1. Prospective and retrospective data

The RaDiCo-EURBIO-Alport cohort study collaboration is based on an existing network of healthcare centres (Rare Diseases Reference Centres or Rare Diseases Competence Centres (CRMR and CCMR,)) expert in the rare disease(s) under focus, physically located in public hospitals. Each hospital participating in the study provides a de facto guarantee that the operating environment of its care data is compatible with legislation. Data are either collected from existing source information databases or are manually entered from specific forms:
Case of prevalent patients occasionally followed for years and for whom it is very important to be able to use data previously collected by centres. In these cases and from a logistical and regulatory standpoint, Inserm-UMR S 933-RaDiCo will organize the recovery of useful retrospective data for the cohort. Staff responsible for exporting data will be qualified individuals in the CRMR or CCMR of the hospital concerned.
Case of new patients, called incident patients, diagnosed during the study and eligible. Their medical data will be entered directly in the ad hoc forms of electronic questionnaires (e-CRF; REDCap software suite, see below). If some centres require a double input (paper plus computer), Inserm-UMR S 933-RaDiCo will organise the latter for technical reasons.
In some cases, a part of prospective data will continue to be collected by Inserm-UMR S 933-RaDiCo from local systems at regular intervals. This is the case when clinicians use input applications and associated database for other studies and trials, or for local registers.
In these three cases, data processing is managed Inserm-UMR S 933-RaDiCo but data input is the responsibility of the CCMR and/or CRMR.

5.2.2. Data transfer mechanisms

In the case of data sources from healthcare centres, Inserm-UMR S 933-RaDiCo prefers exports in text format with UTF-8 character codes, delimited by comma separators (,) and with names in page header fields. These files are compressed and deeply encrypted before transmission.
Public/private key infrastructure (PKI) (either 1024 bits RSA or, under study, 160 bit elliptical curve cryptography). The negotiation and exchange protocol will be addressed below.
At subsequent steps of the processing chain, exports will be imported for processing and homogenisation in the latest stable versions of MySQL and MS SQL Server databases. This processing system is a data warehouse with access restricted to persons authorised to carry out data management and deduplication operations.

If patients are no longer seen for consultations and no change is recorded in the database, the databases and as well files are frozen and so only one export, called migration, is planned to feed the database of the cohort study.
In both cases, after ad hoc importing and processing of the data recovered, the databases underlying the e-CRF data collection application will be populated by data from prevalent patients.
An asynchronous transfer system, similar to Post Office Box principle, is proposed. It uses a transit zone (data stored temporarily without being divulged) between the healthcare centre and Inserm.
This transit zone separates information producer sites from consumer sites. Transfer technology will be based on individual sessions from secure servers with an SFTP, along with individual certificates for each server.
SFTP: SSH File Transfer Protocol or Secure File Transfer Protocol. FTP transfer is encapsulated in a protected SSH tunnel with a public/private key combination and with AES 256 encryption.
After data are recovered, they are expunged from the transit zone. Operations are logged.

5.3. Tools for clinical studies

The Inserm-UMR S 933-RaDiCo platform has selected REDCap as its clinical research software suite.
REDCap is a Web software suite operating in client/server mode. It ensures input, processing and management of clinical research data. REDCap is recognised internationally and has been adopted by almost 1000 clinical research organisations throughout the world, including for instance the HEGP (Georges Pompidou European Hospital) in France and the Harvard School of Public Health in the USA.
REDCap enables researchers to easily implement research and survey databases on an "industrial" scale, at the same time as complying with Good Clinical and Good Epidemiological Practices (GCP and GEP). The system includes measurements of quality control and safety, and enables research partners to work at levels of data isolation and confinement compliant with legislation.

5.4. Architecture of the information system infrastructure

5.4.1. Infrastructures and environment of operations

All systems described below run in a virtual, redundant and supervised architecture in a certified "health data" hosting and facilities management centre (HADS in French) in the framework of a service contract with the Inserm. The equipment is dedicated and isolated.
The required operating system is LINUX long term support, Apache, the MySQL database engine and the PhP application scripting language, i.e. the LAMP suite.
Technologies of the infrastructure system:
- Debian Linux, version 8 "Jessie", is the distribution adopted at the time this document was published. Inserm-UMR S 933-RaDiCo will ensure upgrading of technologies and versions of its software.
- Ancillary servers such as e-mail (Exim), directories (AD), resolution of DNS domains, certificate servers and encryption are all in a secure infrastructure.
- A second infrastructure contains servers for collection of files in transit and the data warehouse used for deduplication and data management. This infrastructure uses MS Windows 2012 R2 Data Center Edition servers (and upgrades to the next stable version), orchestration servers of the MS System Center suite, MS SQL SERVER 2012 R2 Enterprise database servers (and upgrades to the next stable version) and MS SharePoint Web servers for reports and statuses.

Data are backed up to prevent losses from failures, according to the HADS hosting restrictions and archives are also backed up for restitution when needed in the entire post-study period.

5.4.2. Control of the Information system

In compliance with the consortium agreement and the charter signed by the institutional partners of the RaDiCo-EURBIO-Alport cohort study, the obligations for interactions and good information technology practices are represented and programmed via "agents".
An agent is an automatic process with a computer login that acts on behalf of the person or company it represents and that executes actions in the name that entity.

The system that orders the correct sequence of actions by agents, that manages waiting loops, and that re-prompts agents in case of an operation error, is called an orchestrator. Orchestration supervises the correct operation of the entire system. Such orchestrators are required in industrial class process management systems.
The system chosen is Microsoft Orchestrator of the MS System Center suite. This system is used to orchestrate sensitive processes, in particular in data centres (Azure).
Users interacting with REDCap and other IT infrastructure administration systems involving the REDCap environment log on securely with disposable certificates (tokens). The entire system is centrally administered in a directory by staff certified by both the host and the Inserm-UMR S 933-RaDiCo. Logs of all operations are systematic and are archived for subsequent tracing.

5.5. Security, protection and confidentiality of data

The "Référentiel Général de Sécurité/General Security Database published by the ANSSI (National agency for security of information systems), version 2.0 and associated appendices were used to establish rules and recommendations.
With respect to interoperability with healthcare systems, Inserm-UMR S 933-RaDiCo employs the latest version of the national plan for the security of health data (PGSSI-S in French).

5.5.1. Confidentiality of patient identity

Every patient included in the RaDiCo cohort is assigned a unique ID in compliance with directives of the DGOS (Directorate general for care provision) implemented by the BNDMR (National data bank of rare diseases) (http://www.bndmr.fr/) managed by the AP-HP (Paris public hospitals authority) responsible for generating unique IDs for patients with a rare disease. It maintains the list of patients with a rare disease with the IdMR (national rare disease ID). The BNDMR ensures encryption of the IdMR as a 20-digit code created with the hashing technique using last and first names, birthdates and sex. The BNDMR edits and maintains the algorithm used and ensures correspondence between the IdMR and patient identity in case of need. Inserm-UMR S 933-RaDiCo cannot ensure correspondence.
A unique inclusion number is also planned, called “RaDiCo code”, to order patients according to the chronology of their inclusion in the e-CRF, to the site of inclusion and to the study. This code is composed of three parts of fixed length:

- First part coding for the RaDiCo cohort study name: eight alpha characters
- Second part coding for the inclusion site: 6 alphanumeric characters
- Third part coding for the patient’s inclusion number: 5 numeric characters.

The RaDiCo code will be automatically generated by the managing system of the REDCap eCRF.
Investigators will keep in a locked and safe place with limited access corresponding identity tables with patient’s name, RaDiCo’s code and the IdMR number. In no matter, these tables will not be communicated to RaDiCo. If the tables are lost, investigator will be able to request to BNDMR a re-identification of the patient.
A de-identification operation of patients to include is carried out in the processing of exported source data. This operation is conducted after data recuperation from the transit zone and after deduplication. This operation requires a step prior to de-identification: deduplication of medical data source files.

5.5.2. Security of connections and data protection

Under coordination by the principal investigator of the RaDiCo-EURBIO-Alport cohort study, hospital inclusion sites participating in the study prepare an explicit and detailed matrix that defines the roles, profiles and groups of persons involved and Inserm-UMR S 933-RaDiCo is then managing access rights to information system resources:

- Databases containing health data;
- Databases of administration accounts and delegation lists (directories);
- Automatic execution agent acting "in lieu of";
- Organisation of processes scheduled in the e-CRF (schedule, visits, data management etc.);
- Right to sign and lock forms;
- System for re-identification in exceptional cases according to established rules.

Connections are rendered secure by asymmetrical key exchange protocols.

- Protocol TLS 1.2 and 1.3 if published in a stable versions;
- The key exchange algorithm is Kerberos and RSA (1024 bits, even 2048 bits (under study));
- Servers are authenticated by digital certificates (X509);
- Sessions are encrypted (AES 256 bits, or ECC 160(under study));
- The integrity of data exchanged is controlled (HMAC – under study to choose between MD5 or SHA1).

Users authenticate themselves securely with a disposable code (token) and accesses are logged. Accounts are centrally managed in central directory servers (MS Active Directory) by a single accredited entity for account administration.
Sensitive data such as health data, but also account data, are all transmitted and stored in encrypted form. These data are saved using daily, weekly and monthly backup plans.
For long term follow-up cohorts, the information system will be updated according to technologic and regulatory evolutions. Preliminary and complete security audits are conducted, followed by audits at regular intervals, to guarantee the security levels.
There will be an access portal used to approve and refuse user access. Before beginning the cohort study, it is indispensable to define different access levels with respect to the roles of users.

5.5.3. Security of the user's environment

Concerning users, whether they are involved in care, research, study administration or the technique itself, and in light of the heterogeneity of work stations (desktop computers, portables, tablets, Apple, Mac OS, MS Windows or Linux) and of the diversity of organisations with this equipment in the RaDiCo-EURBIO-Alport cohort study, the correct use of this hardware and operating systems is the direct responsibility of the participating sites using them (local information technology charters).
Concerning utilisations and during tasks associated with the clinical study, Inserm-UMR S 933-RaDiCo formulates recommendations for good practices with the goal of reducing risks to data security.
These good practices are found in the appendices of the consortium agreement.
Use of client navigators and Java environments is covered by the consortium agreement for the study.
If the needs of mobility require the use of nomad terminals and a local network connected wirelessly, e.g. Wi-Fi, the operations will comply with instructions of the ANSSI.
All relevant data for patients with Alport syndrome entered in CEMARA (CNIL authorisation number: 1187326) will be de-identified and transferred to the RaDiCo-EURBIO-Alport database (REDCap™ software).

5.6. Data Management:

A Clinical Data Management Plan, specific to the study, will be prepared prior to the first inclusion of data.

The management of data such as clinical data requests a number of specific elements:

5.6.1. Case Report Forms

The study will be conducted using the REDCap electronic data capture system. The e-CRF developed by RaDiCo with REDCap will be compliant with 21CRF11 and FDA 1678 Guidance, GAMP and GCP guidelines. Data entry will be performed by or on the principal investigators or sub-/co-investigators behalf in the electronic case report form (e-CRF) provided by Inserm-UMR S 933-RaDiCo, and saved in the associated database.
Only authorized persons (principal investigator and sub-/co-investigators) can access the e-CRF at the study sites by means of personalised username and password. The principal investigators at the study sites who have the right to enter data and sign study documents should be listed on a “Staff list/authorisation Log”.
Data transmission to the web-server is performed by mean of an Internet connection and no specific software has to be installed at the study sites. REDCap system requirements are the following:

- Hardware and software: Hardware and software requirements are modest and the system runs in Windows/IIS and Linux/Apache web server environments. Internet connection with a 56kbit/sec modem (or higher), ISDN, ADSL, LAN (i.e. Ethernet), Fastweb.

- Browser: Microsoft Internet Explorer 5.0 or higher

Subjects will be identified on the eCRF by a unique identifier compliant with the NIH/NCATS Global Rare Diseases Patient Registry Data Repository/GRDRSM program recommendations. The investigator must make a separate, confidential record of any details (Enrolment Log) to permit identification of any subject enrolled in the trial in case follow-up is required.

Electronic e-CRFs are used to record clinical trial data as defined in other sections of the protocol, and are an integral part of the trial and subsequent reports. It is the responsibility of the Investigator to maintain adequate and accurate e-CRFs, by keeping them current to reflect status at each phase during the course of the trial. Data may be corrected until the visit is closed by the investigator. Subsequent data corrections can only be done by the data management centre on the basis of eQueries posted to the investigators or sub-/co-investigators. A track of any correction is maintained in an electronic audit trail.

A declaration ensuring accuracy of data recorded in the e-CRFs must be acknowledged by the principal investigator or, if permitted by local regulations, by his authorized sub-/co-investigator when the e-CRF is complete. A print out of the e-CRF will be possible at any stage of the study. The records of e-CRF should be archived at the study site.

The principal investigator will be responsible for retaining/archiving all records pertaining to the study in accordance with local regulatory guidelines.

5.6.2. Source documents

Source documents provide evidence for the existence of the subject and substantiate the integrity of the data collected. Source documents are filed at the investigator's site (patient medical record). Data reported on the Case Report Forms must be consistent with the source documents. Worksheets may be used to facilitate the data collection process. Data from these worksheets will be transferred to the official e-CRF; such worksheets will be considered source data and should be stored and archived at the site.

Direct access to source data / documents:

The investigator / institution will permit trial-related monitoring, audits, IEC / IRB review and regulatory inspection, providing direct access to all related source data / documents.

Signed informed consents and all source documents, including progress notes and copies of laboratory and medical test results, must be available at all times for review by the Sponsor's clinical trial monitor and/or inspection by health authorities. Electronic case report forms (e-CRFs) will be accessible and printable at any moment by the Sponsor's clinical trial monitor.

The accuracy of the data will be verified by performing site visits / on selected sites to review study procedures, accuracy of data and review of the clinical documents.

5.6.3. Processing of data

Study data will be entered into the electronic case report forms (e-CRFs) by the authorized principal investigators or sub-/co-investigators. The data-management of the study database maintenance will be performed by Inserm-UMR S 933-RaDiCo according to its own Standard Operating Procedures. The REDCap internal eQuery tool or among other software, R/S plus-generated PROC TEMPLATE, will be used by to clarify data inconsistencies. The investigators or sub-/co-investigators can also post additional information or data correction to be implemented on an already locked e-CRF visits. All eQueries and data corrections will be electronically tracked.

6. DURATION AND ORGANISATION OF THE COHORT

The visits will take place at the patients' local centre. The protocol visits are synchronized to regular outpatient or inpatient visits to avoid additional travel. Patients are seen according to standard clinical practice; no additional tests are required.

Once having signed the informed consent form and being included in the RaDiCo-EURBIO-Alport cohort, patients will be followed for 3 years and dedicated staff at participating sites (see Appendix 3)

will fill once a year the data in the e-CRF needed for cohort follow-up and collected as part of routine care of patients.
To be considered complete, the patient should have a completed inclusion visit form and a minimum of 1 follow-up visit form.
As part of routine care, patients may be seen more than once a year. In such cases, study dedicated staff at inclusion sites will complete a follow-up visit form for each visit conducted between the yearly evaluations as much as possible, (see "Study visits schedule" table below)..
Patients' inclusion and follow-up may continue over the initial period. Study dedicated staff at inclusion sites will then complete a follow-up visit form for each additional visit and patient.
These data collected during additional visits will be collected as exploratory data.

Table1: Study visits schedule:

Observations/Investigations	**Inclusion Visit (prevalent or incident patient)**	**Follow-up Visit (1/year, during 3 years) and exploratory visit**			**Performed by**
	V1	**V2**	**… → …**	**… Vn**	
Record demographic data, including genealogy of the patient	X				Study staff at investigational site
Record diagnostic data: Mode of discovery of the disease (hematuria, proteinuria, renal failure, deafness, familial testing…)	X				Study staff at investigational site
Record molecular data	X				
Record of medical history, including self-relevant and family history	X				Study staff at investigational site
Clinical evaluation: Including deafness (audiogram), eye examination, renal function	X	X	X	X	Investigator
Record of biological results (blood and urine samples): creatininemia, estimated glomerular filtration rate, albuminemia, proteinuria, micro-albuminuria	X	X	X	X	Study staff at investigational site
Histology (optic, electron microscopy, immunofluorescence analysis of type IV collagen expression)	X				
Record of treatments	X	X	X	X	Study staff at investigational site
Record of Renal Replacement Therapy	X	X	X	X	Study staff at investigational site
Completion of Quality of Life questionnaire	X	X	X	X	Patient

Observations/Investigations	Inclusion Visit (prevalent or incident patient)	Follow-up Visit (1/year, during 3 years) and exploratory visit			Performed by
	V1	V2	… → …	… Vn	
Treatment compliance	X	X	X	X	Study staff at investigational site
Access to genetic counselling and diagnosis questionnaire	X				Patient

The detail of the data to be collected at each visit are described in Appendix 4 (Mask of the Case Report Form).

There will be a 24 months inclusion period to include prevalent patients (1500 expected patients over Europe and up to 2500 worldwide) and incident patients (incidence rate: about 100-200 / year for the whole participating countries) in the RaDiCo-EURBIO-Alport cohort to reach a total of about 2000 included patient in Europe and up to 3000 worldwide. Thus by adding the time for inclusions and time for patient monitoring and analyses, the total duration of the study will be of about 6 years.

Particular attention will be given to the management of adolescent and adult patients, and to the transition from adolescent to adult care.

Clinical and biological data will be updated annually (eGFR, age at ESRD, RRT, follow up, transplantation, follow up). Creatininemia will be measured by enzymatic dosage, GFR will be estimated by MDRD for adult patients and by the Schwartz formula for paediatric patients.
For prevalent patients, available data on evaluation performed at different time prior to inclusion in the study will also be collected.

The evaluation of the impact of the disease on quality of life, school and professional life will be collected through SF36 for adult patients and SF10 for paediatric patients. Self-administered questionnaires will be filled by patients via website (accessible through internet browser on any device), every 12 months (once a year).
For record of compliance, X. Girerd Compliance Questionnaire will be completed for French patients once a year at a care follow-up visit.
For access to molecular diagnosis and genetic counselling, French patients will complete a survey at inclusion visit.

7. Data Analysis

Although this is a non-interventional study based on a model of longitudinal observational and descriptive investigation, a statistical analyses plan (SAP), compliant to the ICH E6 guidelines, will be written by Inserm-UMR S 933-RaDiCo's statistical team in collaboration with the investigators and other relevant experts (notably a collaboration with Partner Inserm F-CRIN).
Stratification according to gender and mode of heritability will be performed prior analyses.
Considering the context of rare disease and the low number of patients per sub-groups, all available patients willing to participate will be included. To date, there are no retrospective data that would allow relevant hypotheses leading to sample size calculations;

All analyses will be performed using R statistical analyses software or equivalent software.

7.1. For the primary objective

First of all a descriptive statistical analysis will be performed taking into account all patients included.
Adequate graphic method (Bar chart, box plot, forest plot…) will be selected in order to underline

relevant statistical information. Regarding qualitative item, percentage will be calculated with their corresponding confidence interval and means will be provided for quantitative variables with their standard deviation. Result will be display for all the study duration and also per visit (or another appropriate time unit) as a first approach of time distribution of the criteria selected for the primary objective. Some calculated items (derivation calculation, new classification) maybe create in order to get a best statistical information illustration.

Then, descriptive multivariate analysis method (PCA, MCA, etc.) will be used to reveal relevant specific patient profiles. Once defined, appropriate comparison test will be performed (chi2 test or exact fisher test for percentage or student test or other non-parametric test for means, ANOVA method for repeated values etc…)

In addition survival curve will be computed using the Kaplan Meier method for the relevant event (disease and or treatment outcome, complication) in order to assess the cumulative probability density. If some groups are find out during the statistical analysis, appropriate comparison test will be performed (Log Rank test for survival curves)

7.2. For the secondary objectives

Analyses will be comparable to that described for the primary objective.

7.3. For the exploratory objectives

- To search for early urinal biomarkers able to predict the progression of the disease and to evaluate the efficacy of treatments.

  Identification of biomarkers
  After a descriptive analysis in order to give global and detailed progression disease in the cohort, an explicative multivariate analysis will be performed in order to identify potential biomarkers associated with various stages of the pathology or to therapeutic success. The correlation will be calculated for each potential biomarker in order to quantify the strength of the association.

  Longitudinal association will then be tested with survival analysis
  Survival curve based on the Kaplan Meier method will be computed for the relevant event (e.g.time for RRT in the presence or absence of the biomarker). If appropriate, comparison between these survival curves will be performed with the Log rank test. Otherwise survival median will be compared using t-test.

  Identification of biomarkers weights
  A logistic regression model will assess the contribution of the various biomarkers for each variable (progression of disease or efficacy of treatment).
  If relevant, a cox model will be performed to test among time the association of the biomarkers and each variable

- Evaluation of the possibility that was offered to patients and families to access molecular diagnosis and genetic counselling (for French patients only)
  For this specific objective, the descriptive analyses will be comparable to that of the primary objective.

8. Size of study population

Prevalent patients : From 1500 among Germany, Spain, United Kingdom, Italy, Belgium, and France and up to 2500 with USA.

Estimated incident patient: 400: 200 per year (global incident rate for participating countries: Germany, Spain, United Kingdom, Italy, Belgium, Hungary, France and possibly USA) during 2 years (inclusion time).

Therefore, it is possible to enrol a maximum of about 2000 patients or with the possible addition of the ASTOR database (USA), up to about 3000 patients. However, only patients willing to participate will be

included. To date, there are no retrospective data that would allow relevant hypotheses leading to sample size calculations.

However even with a high attrition rate, the number of 1000 included patients seems reachable and will allow to answer study objectives.