THE EUROPEAN ATHEROSCLEROSIS SOCIETY

FAMILIAL HYPERCHOLESTEROLAEMIA STUDIES COLLABORATION

[EAS FHSC]

STUDY PROTOCOL

Version: June, 2016

THE EUROPEAN ATHEROSCLEROSIS SOCIETY

FAMILIAL HYPERCHOLESTEROLAEMIA STUDIES COLLABORATION

– EAS FHSC –

STUDY DESIGN

International registry of individuals with familial hypercholesterolaemia (FH). The EAS FHSC registry consists of a consortium of worldwide FH cohorts, registries, databases and data collections with access to information on individuals with a clinical and/or genetic diagnosis of homozygous (HoFH) and/or heterozygous (HeFH) FH.

Scope: multinational, multicentre; worldwide scope.

Data: cross-sectional and longitudinal (retrospective and prospective).

Selection of participants to be included in the registry:

- Inclusion criteria:
  - Individuals with a diagnosis of heterozygous or homozygous FH.
    Diagnosis: clinical and/or genetic.
    Also including positive clinical diagnosis with negative genetic test and vice versa.
    In case of clinical diagnosis:
    - Criteria used: Dutch Lipid Clinics Network, MedPed, Simon-Broome, Japanese criteria, or other to be specified.
    - Categories of possible, probable and confirmed FH.
  - Relatives of index cases without a diagnosis of FH where screening (cascade or other) is carried out.
- Exclusion criteria:
  - Secondary causes of dyslipidaemia (e.g. untreated hypothyroidism, cholestasis, nephrotic syndrome).
  - Where data collation does not conform to the local or country wide standards for anonymised data.

Duration: intended to be run for the long-term in order to increase the follow-up of patients already included, to include new individuals diagnosed with FH over time, to bring new cohorts to the registry, and to take into account contemporary changes in clinical practice, so that the findings remain relevant. The project is expected to evolve over time as will be further developed and expanded. The process is expected to be dynamic as we aim to continuously employ the best data sharing strategies and optimise our operational protocols based on the experience of the consortium, number of participants included, the results from the first wave of data collated, evolving aims, resources available and feedback from those utilising the data and the feedback from the wider medical community.

STUDY OBJECTIVES

The mission statement of the EAS FHSC aims to empower the medical and global community to seek changes in their respective countries or organizations regarding how FH is detected and managed and the clinical consequences thereof, with a view to promoting early diagnosis and more effective treatment of FH.

The general aims of the EAS FHSC are:

- To establish a global registry of FH with a view to gaining an in-depth understanding of the contemporary burden of both HoFH and HeFH: patient management, treatments, impediments, long-term risks, impact of patient-specific and societal factors, gene–drug interactions and the role of screening.
- To disseminate the information gained from the above-mentioned activities to an international audience including physicians and other healthcare professionals, as well as patient organizations, with a view to sensitising them to the contemporary burden of FH, encouraging open discussion on the management of FH patients, promoting a uniform, evidence-based standard of care, and encouraging them to contribute actively to research.

Objectives

1. To establish an international, standardised registry of patients with FH by bringing together regional, national and international cohorts and registries with access to patients with FH, in order to maximise the potential exploitation of the data and generate large scale, robust information to accurately and reliably investigate (i) the burden of both HoFH and HeFH, (ii) how it is detected and managed, (iii) the clinical consequences of current practice on delivery of care and outcomes, and (iv) the determinants influencing optimal management and LDL-C target attainment in FH. This will include but will not be limited to the following:

1.1. General and cross-sectional analyses:
- How FH patients are detected, comparison of different proposed diagnostic criteria, whether current screening strategies for FH are adequate and, if not, what could be done differently to maximise coverage.
- How patients are managed, treatments offered/advised, how their efficacy is monitored.
- What proportion of FH patients meet the targets (e.g. LDL-C goals, patients receiving therapy), the impediments in attaining LDL-C goals, the role played by societal factors (such as access to healthcare in different settings and the availability of specialist advice) in enabling treatment to achieve LDL-C goal, the influence of gene–drug interactions in attaining LDL-C goals.
- The impact of patient-specific factors, socio-economic factors and treatment-related factors on LDL-C goal attainment.
- Potential variations depending on different geographic settings as a result of factors such as population genetics, health care delivery systems and other patient-, socio-economic- or treatment-related factors.
- Where possible, evidence for economic evaluation of different screening strategies and of interventions will be addressed.

1.2. Risk stratification beyond LDL-C levels, e.g. by genotype (to establish the value of incorporating genetic data into the clinical diagnosis of FH) or lipoprotein(a) levels.

1.3. Analysis of risk/outcomes: Long-term risk of outcomes in patients with FH (including, where possible, estimates of the years lost due to FH), with a special focus on the following end-points:

- Primary end-point
  - The composite of Cardiovascular Disease Events (fatal and non-fatal).
  - Cardiovascular mortality.
  - All-cause mortality.
- Secondary end-points:
  - Each component of the Cardiovascular Disease Events end-point separately.
  - Aortic valve and supravalvular disease.
  - Statin intolerance (clinical, biochemical).
  - New onset of diabetes.
  - Cancer diseases.
  - Pregnancy outcome in female patients with FH.
- The impact of patient-specific factors, socio-economic factors and treatment-related factors on outcomes.

The following definitions shall apply: (i) Cardiovascular Disease Events: the composite of fatal and non-fatal coronary heart disease/acute coronary syndrome/myocardial infarction, cardiac sudden death, fatal and non-fatal stroke, transient ischemic attack, peripheral vascular disease, and revascularization (percutaneous or surgical); (ii) Aortic valve and supravalvular disease: confirmed by any imaging test; (iii) Statin intolerance (clinical and/or biochemical): defined according to the EAS consensus panel (Eur Heart J 2015;36:1012-22); (iv) New onset of diabetes mellitus: fasting glycaemia >126 mg/dL at least in two separately occasions, glycosylated haemoglobin ≥6.5%, and/or patient receiving a new prescription of an antidiabetic drug (Eur Heart J 2013;34:3035-87). (v) Pregnancy outcomes refer to maternal, obstetric and neonatal outcomes.

2. To develop and implement a novel bespoke electronic platform and FH data warehouse for data sharing, cleaning, harmonisation and analysis for patients with FH to support the FH registry. It will allow the collation of retrospective and prospective data from multiple sources and different formats, and will support the management of a large amount of data, critical to ensure sustainability.

3. To disseminate the information gained from the FH registry to an international audience including physicians, other healthcare professionals, policy-makers and patients organizations. To sensitise the different stakeholders involved in FH care to the contemporary burden of FH, encourage open discussion on the management of FH patients and their families, promote a uniform evidence-based standard of care, and encourage those involved in FH care to contribute actively to research.

4. To consolidate a network of investigators interested in FH through which collaborative research and networking on FH can be conducted on a large scale.

DATA SHARING

We will develop a system whereby the data sets will be transferred to and collected by the Coordinating Centre through a secure web platform. Data sharing plans involve instructions to the lead investigators for

data to be transferred and data to be collected, updated with feedback for the individual investigators providing the data. For each cohort/data set a two-stage approach will be followed: the first one comprises the collation of the cross-sectional and retrospective data readily available; the second stage will involve integration of longitudinal, prospective records (including follow-up data).

In order for these investigators' to provide reliable data for inclusion, a set of minimum predefined quality and operational criteria must be fulfilled:

- The data must be in electronic format;
- The amount of data shared should be considered enough to add relevant information to the registry;
- The data must have been collected rigorously following a standardised well-designed protocol; inclusion and exclusion criteria must be well defined; FH must be well defined according to accepted criteria and not rely only on self-reported history of FH; any measurements recorded must have been done with validated and standardised methods and devices; standardised definitions should have been used; the participants included should be representative of the target population;
- The parameters recorded must include at least the minimum set of data (core information) specified below;
- Local ethical and security aspects must be in place;
- The data sharing agreement must be signed prior to any data transfer.

A detailed data request form for providing anonymised data will be sent to the investigators contributing to the registry for them to provide data on individuals with FH and their relatives where available. A minimum set of data (core data) will be required to be included in the registry, comprising:

- How patients are being managed (e.g. cascade screening, availability of genetic testing);
- Characteristics of the clinics providing the data (e.g. number of physician and nurses, specialised clinic or general physicians, public or private clinic, etc.);
- General and demographic information;
- Familial and personal cardiovascular history; other cardiovascular risk factors;
- FH diagnostic criteria;
- Where genetic tests are available, these genetic data will form part of the core data requested.
- Lipid-lowering medication with doses;
- Laboratory lipid profile;
- Data on outcomes if follow-up exists (retrospective or prospective);
- In the case of HoFH, data about LDL apheresis will also be collected.

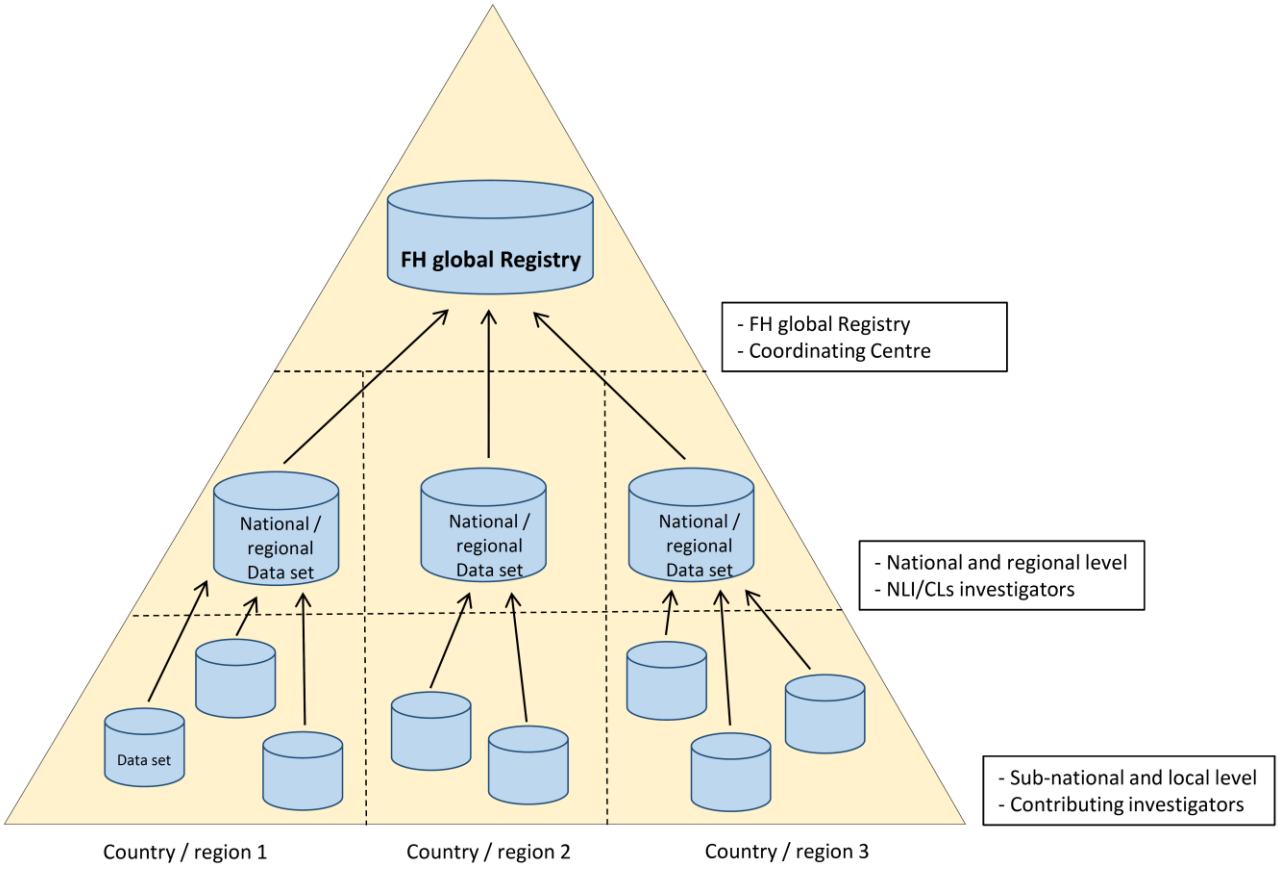
Additional data to be collected where available include a number of other parameters of interest such as more detailed information on comorbidities, other medications and laboratory results, tests such as electrocardiography, echocardiography, calcium coronary scores, angiography, etc. (annex 1).

Where possible, standardised definitions will be used and free text avoided. We will encourage that investigators provide updates of shared data sets within a reasonable timeframe.

We will develop data quality assurance strategy which will include, for example: measurement of data consistency and instituting a standard method for analysing the reasons for, and impact of, any data inconsistencies. We will continually improve the quality of the data with a data quality assurance system.

Individual investigators from each local site will contribute to the FH registry through the National Lead Investigators/Country-leads (NLI/CLs) for each region or country. These NLI/CLs will transfer and upload the data to the Coordinating Centre through a bespoke secure web portal. See next figure.

Where possible, depending on the funding and resources available at each time, the EAS FHSC will try to provide the lead investigators with a certain amount of funds to help in the management of data with the purpose of data sharing.



DATA MANAGEMENT

The preliminary data shared will be checked for consistency and accuracy by at least 2 independent investigators of the Coordinating Centre (scientific coordinator, data manager), and any discrepancies will be resolved by raising queries with the individual investigators. The Coordinating Centre, supported by a technical team, will develop cutting edge and artificially intelligent routines incorporating machine learning for automated data cleaning, harmonising and make the data uniform and compatible. Various suitable well-defined data models will be designed into which collected and cleaned data will be loaded. After ensuring satisfactory data quality from all registries (contributors), the information will be saved into a central database to be held securely on a server at the Coordinating Centre (Imperial College London, UK) in strict adherence with all data safety protocols and regulatory requirements i.e. Society for Clinical Data Management (SCDM), Clinical Data Interchange Standards Consortium (CDISC), Study Data Tabulation Model Implementation Guide for Human Clinical Trials (SDTMIG) and the Clinical Data Acquisition Standards Harmonization (CDASH) standards. A bespoke Data Warehouse hosted at the Coordinating Centre will be developed to consolidate heterogeneous data from multiple sources and formats through appropriate smart algorithms. A central, well-organised, robust, secure, compliant “data warehouse” with different well-defined data models hosted by the Coordinating Centre will be key for multiple reporting and analysis use-

cases, allowing researchers and stakeholders immediate access to processed, clinically relevant rich information.

During the course of the project various tools will be developed to support, improve, standardise and automate the Data Management. These tools are mainly:

- **Process Workflow:** The Data Management involves various steps and a uniquely focussed advanced workflow will be designed for automating the whole process. The workflow will streamline the entire process all the way from designing data models, data collection, data validation, data harmonization, data transformation and data analysis and reporting. The workflow itself will be a cloud based solution hosted as a software-as-a-service (SaaS) application to meet demanding requirements of the global project.
- **Data Monitoring Application:** The Data Management monitoring application will be designed and developed to complement the workflow to ensure homogenisation of the data to meet all regulatory requirements and will be updated regularly to ensure adherence to contemporary local and global regulatory standards. The monitoring application will help administrators to be involved at all stages of data management right from inception to completion. The monitoring of the various stages of data management will ensure the quality standards are maintained and any deviations from protocols are immediately reported and rectified. Due to the heterogeneous nature of data collected from multiple sources; it will be critical for various data management stages of the workflow to be assessed for quality at regular intervals during whole project including, database designing, data validation, discrepancy management, medical coding, data extraction, and database locking.
- **Process Logging System:** The data will go through various stages during the Data Management cycle. It is therefore important to have a system which logs changes made to the collected data during various stages i.e. data validation, data harmonisation, data standardisation and data transformation. The logging system will help administrators in reverting data to the original state if there are conflicts or errors. The logging system will be essential for auditing the whole process and to improve the Data Management according to the requirements and suggestions of stakeholders. In the absence of a proper logging system there is a potential to introduce un-detectable errors and merge sub-standard data with the master data for any global project. The process logging system will also bring transparency, assurance and instil confidence to national and international data contributors.

PLATFORM AND DATA WAREHOUSE

We will develop a central database that would consolidate subject-oriented, time-variant and non-volatile data from multiple sources into a robust large scale FH Data Warehouse (FHDW). The FHDW will provide through novel informatics the capacity for retrospective and ongoing prospective data pooling, harmonisation and analysis. The integration, validation, processing, and exploration of complex data is a technical challenge for clinical researchers and a major rate limiting step to clinical FH research projects. Through the bespoke data architecture and platform which we will develop we will support: secure data harvesting from multiple global repositories/databases/file systems; data cleaning, validation, quality control and merging into a Data Warehouse hosted by the Coordinating Centre at Imperial College London, UK, in strict adherence with all data safety protocols and regulatory requirements. The standardized central data will support intelligent data analysis. The FHDW is envisioned as the “master” repository for FH specific health care data that can integrate disparate sources of data on the same individuals for instance at later time points, in order to support the evolving information trail and analytical requirements of stakeholders involved in the management, evaluation and care of the FH patients. Once built, the FHDW will serve as a key resource

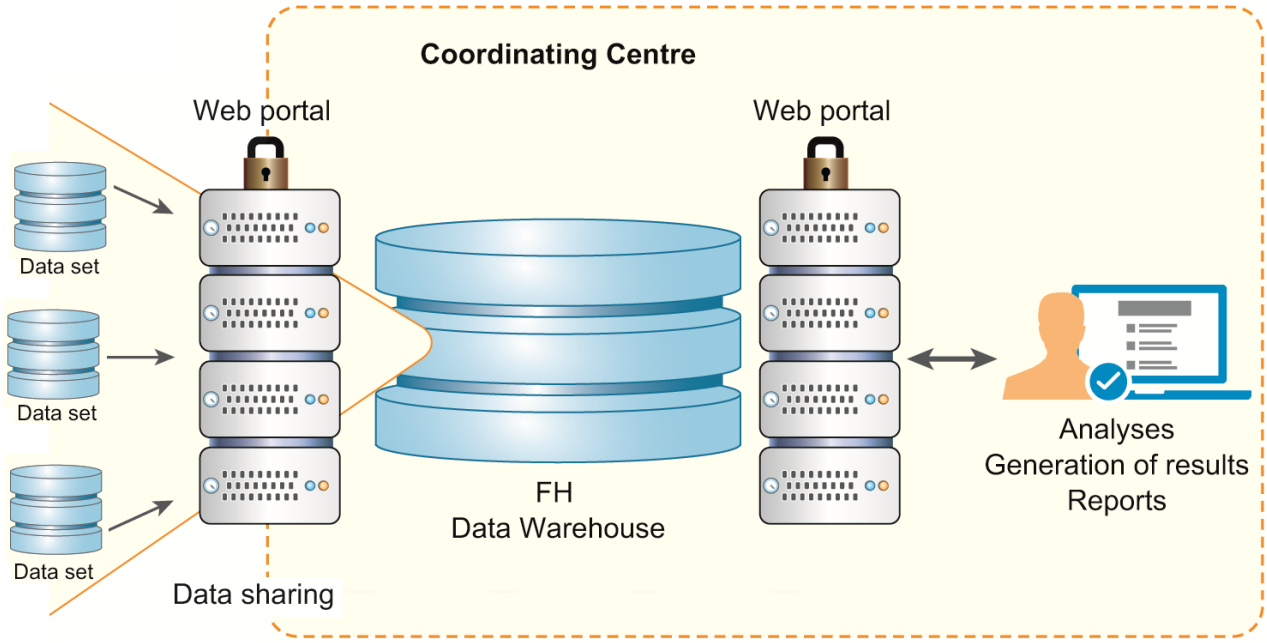
for implementing the various outputs e.g. quality measurement, quality improvement, analyse trends over time. Its goal is to make isolated data easily accessible, reliable, consistent, and secure to support informed planning, decision-making, and communication. The aggregated FH data will be the principal source for identifying valid, novel, potentially useful, and ultimately understandable patterns in large scale complex data across multiple countries, providing both power and precision, thus facilitating “Knowledge Discovery” to tackle the population burden of FH.

The FHDW will incorporate the data backup and recovery policies for minimal downtime of the platform. These policies will be flexible and adaptable to usage and demands of the platform. Main aspects of backup policies of the raw data (shared by each NLI/CL) and the master copy of the consolidated data are:

- Full backups of FH data will be performed weekly. Full backups will be retained for 3 months before being overwritten;
- Incremental backups of the FH data will be performed daily. Incremental backups are retained for 1 month;
- Backups will be run during minimal load time;
- Backups are stored in secure location/s different from master copy. A limited number of authorised personnel will have access to the backup data;
- Data will be available for restoring within a few minutes of a backup job completing;
- Data will be available during the retention policy of each backup job i.e. 3 months;
- The potential data loss during a working day will be avoided by utilizing incremental backup and process logging system.

The replication of master data into multiple copies will be stored on independent and remote servers.

The overall architecture of the FHDW and its corresponding platform is shown in the figure below. The FHDW will be a complete platform supporting various research activities and thus will be composed of the following components:



1. **Web Portal for Data Collection:** The secure web portal will be developed as a means for data coordination. National Lead Investigators/Country-leads will be asked to upload data via portal for transparency and auditing purposes.
2. **Data Validation:** Cutting edge open source technologies will be used to Extract, Transform and Load (ETL) data onto the Global Data Warehouse. The extracted data will be of the highest quality, structurally valid, and contextually rich.
3. **Data Collation:** The validated retrospective as well as prospective patient data (follow-up of patients) from different global sources will be merged within the central database.
4. **Data Analysis Tools:** Widely accepted data analysis, data mining and reporting tools will be employed on the merged data to support an arbitrary number of queries and research questions.
5. **Web Portal for Reporting:** A secure web portal will also be developed to generate various reports for stakeholders. The purpose of these reports is to highlight healthcare trends, limitations and improvement scope; only extractable from the merged central database.
6. **Data Storage:** Raw data will be stored in unstructured database (NoSQL) specific to contributors and merged data will be stored in the structured format in relational database with technical capacity to handle a large number of participants with lifetime events history.
7. **Privacy Tool:** To protect patient privacy, only anonymised data will be collected and aggregated, and a Global Unique Identifier (GUID) will be assigned to each patient's data. The GUID approach enables researchers to follow patients over time and across different events, studies and countries.
8. **Data Management Workflow:** An advanced workflow will be designed for automating the whole process of Data Management. The workflow will streamline the entire process all the way from designing data models, data collection, data validation, data harmonization, data transformation and data analysis and reporting.
9. **Workflow Monitoring:** The monitoring application will help administrators to be involved in all stages of data management right from inception to completion; thus ensuring the quality standards are maintained and any deviations from protocols are immediately reported. The system will make use of existing standards for wider acceptance and interoperability.
10. **Logging Tool:** The logging tool will log changes made to the collected data during various stages. The logging tool will help administrators to improve the Data Management process and in reverting data to the original state if there are conflicts or errors. The process logging system will also bring transparency for the confidence of national and international data contributors.

Traditional data warehouses require metadata or a common vocabulary, which is already well defined and used in many other domains, such as banking or finance. The FHDW aims to define FH specific domain metadata in a similar fashion, based on common vocabularies and related standards in medicine. Furthermore, the elaborate structural meta-information will be used to actively support the user in tasks that usually require significant IT knowledge, such as defining complex search queries or data quality constraints, or applying advanced data visualization algorithms to the data. The proposed warehouse supports the domain expert through the whole process of knowledge discovery from data integration to exploration.

Data Extract, Transform and Load (ETL)

Once the data is safely transferred to the Coordinating Centre through the EAS FHSC web portal and after ensuring satisfactory data quality, the information will be harmonised into the central FHDW. The FHDW will consolidate heterogeneous data from multiple sites through bespoke smart algorithms. The heterogeneity of source data models, data navigation, integrity, unit compatibility and attributes to uniform data definition language (DDL) are resolved by Extract, Transform and Load (ETL). The machine learning and expert system

based declarative rules will ensure highest quality, structurally valid, and contextually rich data. The FHDW will manage two types of data: core datasets and additional datasets, accommodating future requirements. The raw source data (including additional datasets) will be stored in an unstructured database (NoSQL) and merged data (core dataset) will be stored in the relational database with technical capacity to handle Big Data with lifetime events history. Internal partitioning of data in data marts will support an arbitrary number of queries and research questions; conceptual cubes and dimensions will be employed for complicated multifaceted queries. The research subject based data marts and cubes will optimize analysis, reporting and visualization of large global data. Data acquisition and workflow merging will be managed and monitored by an interactive dashboard to maintain the status of all key steps.

Data servers physically hosted by ICL are encrypted to industry standards utilizing public–private key infrastructure. A Global Unique Identifier (GUID) will be assigned to each subject enabling investigators to follow their own subjects over time.

The resource makes use of existing standards for wider user acceptance, technical interoperability and scalability. Regular surveys of the users' (lead investigators') experience will capture their expectations and obtain key suggestions and feedback, allowing appropriate adaptation and improvements of the user/resource interface, enhancing end-user experience for ease-of-use and efficient data management, allowing the resource to properly grow and evolve.

Multi Stage Data Storage for enhanced Security

The nature of data and research requires robust architecture for the data storage. The data will be stored physically on servers hosted by ICL with different access controls. The main architecture for the data collection is shown in the figure below:

- **Data Extraction:** National Lead Investigators/Country-leads will extract data from their local data storage systems taking into consideration of the specific requirements/limitations locally. The FHDW technical team will provide reasonable support to lead Investigators to carry out this task.
- **Web Portal for Data Collection:** National Lead Investigators/Country-leads will use the secure Web Portal to upload his/her copy of extracted data. The Web Portal communication will be based on Transport Layer Security (TLS 1.2) and 128 bits Advanced Encryption Standard (AES 128).
- **Temporary Data Server for Data Storage:** The Web Portal will upload the data to the temporary server in the partitions specific to the corresponding National Lead Investigator/Country-lead. An event will be generated for the data to be moved to a more secure place once the data upload is complete. Once the data is copied to the next stage the data will be removed from Temporary Data Server. This is in line with standard practice.
- **Intermediary Data Server for Data Harmonisation:** The uploaded data is immediately moved to the secure data server which has no public access. This data server contains the data 'as is' and is used as a holding area to determine whether data meets quality and protocol requirements. Entry into this layer

should be through a number of 'data quality' gateways to ensure that erroneous or missing data is rejected or handled appropriately.

- **Secure Data Server for Data Warehouse:** The most secure data server will host the FHDW, with no external and public access. The communication with FHDW will only be via a secure messaging service and the Secure Data Server will be configured to read and write messages from/to the messaging service. This will contain a single table and a number of reference tables. Each related table may reference a number of additional tables. Data structured in this manner will be used to produce/populate Online Analytical Processing (OLAP) style cubes.
- **Application Server for Data Analysis and Reports:** The application server will have various executable tools and scripts with artificial intelligence. The application server will request from the FHDW a subset of data via secure messaging service for relevant data. Once the requested data is received, it will be analysed and only analysis results and reports will be saved on the application server; the requested data will be discarded and deleted after successful processing.

ANALYSIS PLAN

Where possible and after ensuring consistency and compatibility of the information and description of the variables, the data from different providers will be merged and analysed at an individual level as a composite dataset. Where this is not possible, analyses will be carried out at a country/provider level, by analysing each dataset individually and then pooling the results together.

The harmonised registry will allow us to conduct different types of analyses, such as verifying previous published findings with greater power and precision, conducting alternative analyses of the same data, novel new analyses, cross-cohorts comparisons, meta-analyses, analysing unpublished data (not previously conducted due to low sample size within each separate dataset), conducting exploratory analyses to generate new research hypothesis, etc.

A number of research questions will be addressed, such as those shown in the table below:

1	How FH patients are detected, comparison of different proposed diagnostic criteria, whether current screening strategies for FH are adequate and, if not, what could be done differently to maximise coverage.
2	How patients are managed, treatments offered/advised, how their efficacy is monitored.
3	What proportion of FH patients meet the targets (e.g. LDL-C goals, patients receiving therapy), the impediments in attaining LDL-C goals, the role played by societal factors (such as access to healthcare in different settings and the availability of specialist advice) in enabling treatment to achieve LDL-C goal, the influence of gene–drug interactions in attaining LDL-C goals.
4	Long-term risk of outcomes in patients with FH (including, where possible, estimates of the years lost due to FH), with a special focus on the following end-points: ▪ Primary end-point 1: the composite of Cardiovascular Disease Events (fatal and non-fatal). ▪ Primary end-point 2: Cardiovascular mortality. ▪ Primary end-point 3: All-cause mortality. ▪ Secondary end-points: - Each component of the Cardiovascular Disease Events end-point separately. - Aortic valve and supravalvular disease. - Statin intolerance (clinical, biochemical). - New onset of diabetes. - Cancer diseases. - Pregnancy outcome in female patients with FH.

5	To establish the value of incorporating genetic data and other factors (e.g. lipoprotein[a]) into the clinical diagnosis of FH and estimation of the risk of outcomes.
6	The impact of patient-specific factors, socio-economic factors and treatment-related factors on LDL-C goal attainment and cardiovascular risk.
7	Potential variations depending on different geographic settings as a result of factors such as population genetics, health care delivery systems and other patient-, socio-economic- or treatment-related factors.
8	Where possible, evidence for economic evaluation of different screening strategies and of interventions will be addressed.

Standard validated statistical procedures and models for observational studies and weighted meta-analyses (individual participant data meta-analysis where possible; meta-analysis of aggregated data alternatively) will be applied. Statistical methods described in studies such as the Emerging Risk Factors Collaboration [Thompson S et al, Int J Epidemiol 2010;39(5):1345-59; Pennells L et al, Am J Epidemiol 2014;179(5):621-32] or the Non-Communicable Diseases Risk Factor Collaboration [NCD-RisC, Lancet 2016;387:1513-30] will be used where appropriate. After exploratory analyses with description of the variables of interest, cross-sectional analyses and correlates will be performed. Comparisons of continuous parameters will be done with parametric and non-parametric tests as appropriate and categorical variables using Chi-squared test. We will estimate exposure–outcome associations (unadjusted and adjusted) and epidemiological interactions using standard regression models. Risk prediction models using measures of discrimination (concordance index [C-index], discrimination measure [D-measure]) and reclassification (net reclassification index [NRI]) are intended to be performed. Cox proportional hazards regression models stratified by certain variables of interest (e.g. by gender, country, etc.) and Kaplan–Meier estimates of survival will be generated where time-to-event data is available. Measures of heterogeneity and risk of bias (e.g. regression dilution over time) will be applied.

We will account for different variables of interest, carry out appropriate adjustments systematically, consider key subgroups and correct for differences in definitions. Where feasible, reliable, and data available we will also try to roughly approximate the prevalence of FH by extrapolating the data to the general population. The different sources of incoming data will also help assess potential variations depending on different geographic settings, including factors such as population genetics, healthcare delivery systems and other patient-, socio-economic- or treatment-related factors.

Particular emphasis will be placed on exploring differences between key subgroups including but not limited to the following: HoFH and HeFH; different genetic subtypes of FH (e.g. broadly whether due to mutations in the LDL-receptor, ApoB, PCSK9 or LDL-receptor adaptor protein pathways); diagnosis based on different diagnostic criteria; clinical vs. genetic diagnosis; FH treated by a generalist vs. specialist; FH with pre-existing CVD vs. those without; FH with a family history of premature CVD vs. those without; FH individuals who attain LDL-C targets on standard treatments vs. those who do not; geographical region; gender and ethnicity; age at baseline; age at first treatment initiation.

The different sources of incoming data will help us assess potential variations depending on different geographic settings, including factors such as population genetics, health care delivery systems and other patient-, socio-economic- or treatment-related factors.

Following a “database-query” model, investigators involved in the registry will be able to submit research questions they would like to evaluate through the Coordinating Centre; if the request is deemed to be

scientifically relevant by the Executive Committee and considered feasible by the Coordinating Centre, the latter will construct and run the analyses and return the results to the requester.

STUDY OPERATIONAL/GOVERNANCE STRUCTURE

Due to the complexity and scale of the project, to ensure the smooth operation within the wide EAS FHSC global network, the EAS FHSC is supported by an Executive and a Steering Committees, a number of National Lead Investigators/Country-leads and a Coordinating Centre.

- **Executive Committee (EC):** the EC consists of a panel of experts in the field of FH led by Prof. Kausik K. Ray (Imperial College London, UK) and it represents the academic leadership of the EAS FHSC. It is constituted by the following members:
    - Prof. Kausik K. Ray (UK)
    - Dr. Handrean Soran (UK)
    - Prof. John J. Kastelein (The Netherlands)
    - Prof. G. Kees Hovingh (The Netherlands)
    - Prof. Pedro Mata (Spain)
    - Prof. Gerald F. Watts (Australia)
    - Prof. Frederick Raal (South Africa)
    - Prof. Raul Santos (Brazil)
    - Prof. Alberico L. Catapano (Italy)
    - Dr. Tomas Freiberger (Czech Republic)

  The EC represents the core group for overarching management decisions, driving the collaboration, obtaining support and funding to ensure sustainability. The EC will evaluate all proposals for research, data mining, and expansion of the EAS FHSC. The EC are expected to play an active role in raising awareness nationally and internationally for the EAS FHSC and disseminating data that arises from this collaboration. The EC may delegate day-to-day tasks to the Coordinating Centre such as the day-to-day running and decisions related to data collation, harmonisation and pooling. The EC will typically meet once per year; these meetings will be organised by the Coordinating Centre. More frequent interactions will be by teleconference, as necessary.

- **Steering Committee (SC):** the SC represents the advisory committee to provide guidance on the EAS FHSC project development. It consists of all those investigators/stakeholders/experts, usually appointed as EAS FHSC National Lead Investigators/Country-leads for their respective countries/regions.

- **National Lead Investigators/Country-leads:** lead investigators act as a national leader for the EAS FHSC within their respective countries or regions; as such they coordinate and are responsible for the identification and collaboration from the individual sites, promote participation of other national physicians and researchers in the collaboration, liaise with relevant societies and patient organizations, obtain appropriate local approvals and permissions as required, and gather the data from the respective region/country to be transferred to the Coordinating Centre.

- **Coordinating Centre:** The Coordinating Centre for the FH registry is based at the School of Public Health, Imperial College London, UK. It will act as a nodal point for data collation and management, standardisation, consolidation, analyses and queries, and will communicate, support and coordinate the activity of the investigators involved. The Coordinating Centre will ensure the availability and updating

of various computational tools needed to collate, manage and share the data and will make use of analytical partners (e.g. informatics, biostatistics) to ensure the accurate generation of results.

ETHICS, SECURITY AND REGULATORY CONSIDERATIONS

- Major principles guiding the present project development include the added scientific and social value, rigorous adherence to local ethical standards, protection of patients' data privacy and confidentiality, responsible sharing and use of the data, transparency in data management, and creating a secure workspace. This project and related research will be conducted in accordance with the principals of the Declaration of Helsinki (World Medical Association), EU Good Clinical Practice (GCP) and the Committee on Bioethics of the Council of Europe (DH-BIO).
- If required, approval from the corresponding ethical and/or research committees will be requested by the investigators to collect, transfer and share the data. Where needed, informed consent from participants for his/her data to be shared and included in the registry will be requested at the local level by the corresponding investigator. These actions will be coordinated by and will be the responsibility of the local sites/investigators. Requirements in this respect will depend upon the respective institutional/country policies and regulations.
- Data sharing agreements will be executed between the lead investigator sharing the data and the Coordinating Centre as the recipient of the data, specifying, amongst others, how the data are shared; under what conditions; data and results ownerships; recognition and acknowledgment of the work and intellectual contributions from the different parts according to their involvement; and restrictions on using the data for purposes other than those intended for the EAS FHSC project development (annex).
- The analyses will be performed according to rigorous plans, making use of scientific, standardized and validated analytical methods, and carried out by investigators having a skill level appropriate for the tasks assigned to them.
- Disclosures of conflicts of interests will be requested from all investigators involved in any report or publication.
- The investigators and the Coordinating Centre will use all reasonable safeguards in connection with any transfer, communication or remote access connectivity involving the data. A secure web portal will be developed as a means of data entry/coordination and controlled environment. Lead investigators will upload data via the portal for transparency and auditing purposes.
- Only de-identified, pseudo-anonymised data will be collected, and removal of evident identifiers such as names, health numbers, addresses, etc., replacement of dates with time intervals where possible, and categorization of characteristics such as ethnicity, education and others, will be carried out. Certain data such as age, gender or geographical location (e.g. country/region, rural/urban) will be requested as they are needed to maintain scientific value and utility of the present registry; appropriate judgement to protect participants' privacy and prevent identification will be done in these cases.
- The data shared will not be made publicly available, but limited to the outputs from the present project centralized at the Coordinating Centre. The combined data at the Coordinating Centre will not be checked against other external databases, thus preventing re-identification of participants. Information will not be available nor downloadable for any other part apart from specific scientific proposals agreed with steering group and conducted through the Coordinating Centre; it will not be disclosed/transferred to third parties; nor used for commercial purposes, although the outputs of the work will undoubtedly have implications for translational research and use of novel therapies for the pharmaceutical industry.

Clinical data will be handled in compliance with various ethical standards. Collated patient details will be collected in multi-stage infrastructure as discussed earlier. The project partners will upload the data securely on the intermediary storage server through a proper SSL configuration. Access to the data sharing platform is encrypted using Transport Layer Security (TLS 1.2) and 128 bits Advanced Encryption Standard (AES 128). Uploaded data will be transferred to another secure server and will be stored in a password-protected, secure database for instance with no public access. The database will be housed in an access-controlled server room, hosted by the Imperial College Information and Communications Technology department. The database will be encrypted using the standard encryption methods, and will be physically stored across hundreds of disks in a storage area network – this reduces the possibility for needing to reconstruct the database should a drive be lost. The database will have no public access and will be configured to be accessed from approved terminals/computers via firewall configurations.

The stored data will always incorporate the following components, security policies and procedures: authorization, authentication, availability, confidentiality, data integrity and non-repudiation. The methods available for authorization or access controls include single sign-on databases or lists assigning rights and privileges of users to access certain resources, automatic account logoff after a specified period of inactivity to prevent access by invalid users, and physical access controls.

Access to the database will be restricted to named individuals in the Imperial College London project team, with read/write access to data elements defined as required for their role. Access will be further restricted to specific computers on the Imperial College network by firewall configuration. The fool-proof storage mechanism will result in secured storage, manipulation, analysis and reporting of the data. The anonymised data securely held will only be shared according to the terms of consent under the data sharing agreement. The data storage and handling will be compliant with European Data Protection Act and local Data Protection Officers will be consulted in case of any doubt and ambiguity.

ANNEX 1 – POTENTIAL VARIABLES OF INTEREST FOR THE EAS FHSC REGISTRY

Shown as a guidance. See corresponding section in the text for explanation.

General	Demographics	Familial Hypercholesterolaemia	Clinical history
- Patient ID (code) - Country - Date of birth: month/year - Date study entry: month/year	- Age - Gender - Ethnicity - Occupation - Education level - Geographical area: rural/urban	- Date of diagnosis - Age at diagnosis - FH type: homozygous/heterozygous - Clinical criteria used for diagnosis: DLCN/MedPed/Simon-Broome/Japanese FH criteria/other (specify) - Clinical criteria score - Clinical criteria diagnosis: no diagnosis/possible/probable/confirmed - Consanguinity: yes/no - Index case: yes/no - Availability of family tree: yes/no - Family tree code	- Hypertension: yes/no - Diabetes: type 1/type 2/other/no - Hypertriglyceridemia: yes/no - Smoker: current/former/never - Pack-years smoking: number of cigarettes smoked per day divided by 20 and then multiplied by the number of years smoked - Alcohol - Physical activity - CAD: yes/no - Premature CAD: yes/no [premature: male <55y, woman <60y] - MI: yes/no - Coronary revascularization: yes/no - Cerebral vascular disease: yes/no - Peripheral artery disease: yes/no - Aortic valve disease: yes/no - Supravalvular disease: yes/no - Premature non-coronary vascular disease: yes/no [premature: male <55y, woman <60y] - CKD: yes/no - CKD stage: KDOQI - Hepatic steatosis (suggested by US): yes/no - Achilles tendon lesions: tendinitis/injury-surgery/no
Family history	**Examination**	**Genetic study**	
- Family history of FH: yes/no - Family history of hypercholesterolaemia: yes/no - Family history of CAD: yes/no - Family history of other CVD (stroke, PAD): yes/no - 1st degree relative premature CAD: yes/no [premature: male <55y, woman <60y]	- Weight: kg - Height: cm - BMI: kg/m^2 - Waist circumference: cm - Corneal arcus: yes/no - Xanthomas: yes/no - Xanthelasma: yes/no - Systolic BP: mmHg - Diastolic BP: mmHg - Heart rate: bpm	- Genetic study: positive/negative/not done - FH defect: heterozygous/compound heterozygous/homozygous/ARH - Mutation gene: LDLR/ApoB/PCSK9/LDLRAP1/Unknown - Mutation type: stop/insertion/deletion/ etc. - Mutation result: defective allele/null allele - Residual LDLR activity: %/unknown - Mutation: nomenclature, name	

Lipid-lowering therapy	Lipoprotein apheresis (LA)	Lab profile (fasting)	Additional tests

- Lipid-lowering therapy: no/monotherapy/combination therapy
- Statin: yes/no
- Date starting statin: month/year
- Type of statin: simvastatin/pravastatin/lovastatin/ fluvastatin/atorvastatin/rosuvastatin/ pitavastatin
- Dose of statin: mg/d
- Statin intolerance: yes/no
- Ezetimibe: yes/no
- Fibrate: yes/no
- Bile acid sequestrants: yes/no
- N-3 fatty acids: yes/no
- Niacin: yes/no
- Sterols/stanols: yes/no
- Lipoprotein apheresis: yes/no
- PCSK9 inhibitor: yes/no
- Lomitapide: yes/no
- Mipomersen: yes/no
- CEPT inhibitor: yes/no
- Other: e.g. whether treatment has been continuous or intermittent, data/measurements of adherence to therapy, others specify
	- LA: yes/no
	- Indication: homozygous FH/severe heterozygous FH/other (specify)
	- Date starting: month/year
	- Pre-LA lipid profile
	- Pre-LA lipid-lowering drug
	- Post-LA lipid profile
	- Method of apheresis
	- System used
	- Calculated volume
	- Anticoagulation during apheresis
	- Place apheresis performed: lipoprotein apheresis unit/renal department/haematology department/other (specify)
	- Via: native veins/AV fistula/AV shunt
	- Lines: central lines/other lines
	- Duration episode of apheresis
	- Frequency of apheresis
	- Complications: no/yes (specify)
		- Total Cholesterol
		- LDL-C
		- HDL-C
		- Non-HDL-C
		- Triglycerides
		- Apolipoprotein A1
		- Apolipoprotein B
		- Lipoprotein (a)
		- Creatin kinase (CK)
		- Aspartate aminotransferase (AST)
		- Alanine aminotransferase (ALT)
		- C-reactive protein (CRP)
		- Glucose
		- Glycated haemoglobin (HbA1c)
		- Estimated glomerular filtration rate (eGFR): MDRD
		- Microlabuminuria
			- ECG:
			  ▪ Rhythm: sinus rhythm/atrial fib/flutter/other
			  ▪ LVH: (Sokolow-Lyon)
			  ▪ Other disorder: specify
			- Echocardiography:
			  ▪ LVEF: %
			  ▪ LV hyperthrophy: yes/no
			  ▪ Septum: mm
			  ▪ Posterior wall: mm
			  ▪ LV mass index: g/m^2
			  ▪ Abnormal wall motion: yes/no
			- Carotid Ultrasonography:
			  ▪ Carotid plaque: yes/no
			  ▪ Carotid (CCA) IMT: yes/no
			  ▪ Carotid atherosclerotic stenosis: %/no
			- Liver Ultrasonography: Liver steatosis: yes/no
			- Renal arteries Ultrasonography: stenosis: yes/no
			- Lower extremity Ultrasonography: Stenosis: iliac/femoral/distal/multiple/no
			- Ankle-brachial index: Result
			- Coronary calcium score:
			  ▪ Method:
			  ▪ Score:
			- Coronary angiography/revascularization:
			  ▪ Coronary angiography: no lesions/diffuse disease/focal stenosis
			  ▪ If stenosis: <50%/≥50%/>70%
			  ▪ Coronary arteries affected: none/LCA/LAD/LCx/RCA/several
			  ▪ PCI: yes/no
			  ▪ CAGB: yes/no
			- Other test available: Describe. E.g. MRI, etc.

Other therapy	**Follow-up**		
- Antiplatelet: yes/no - ACEI: yes/no - ARB: yes/no - B-blocker: yes/no - CCB: yes/no - Other BP-lowering drug: yes/no - Oral glucose-lowering drug: yes/no - Insulin: yes/no	- Date FH diagnosis - Date of enrolment - Date of first data collection - Date of subsequent follow-up visits - Date of event - Number of visits - Changes in medication		

Outcomes / Events			

- Date of event
- Age at event
- Cardiovascular event
  • Acute coronary syndrome: fatal/non-fatal
  • Myocardial infarction: fatal/non-fatal
  • Sudden cardiac death
  • Stroke: fatal/non-fatal
  • Transient ischemic attack
  • Peripheral vascular disease
  • Revascularization: percutaneous/surgical
  • Heart failure
  • Aortic valve/supravalvular disease
- Death: Cause of death:
  ▪ CAD
  ▪ Sudden cardiac death
  ▪ Cerebrovascular disease: ischemic/haemorrhagic
  ▪ Heart failure
  ▪ Acute aortic syndrome
  ▪ Cancer
  ▪ Other (specify)
- Statin intolerance: clinical/biochemical/both
- New onset of diabetes
- Cancer disease: type
- Pregnancy outcomes:
  ▪ Maternal
  ▪ Obstetric
  ▪ Neonatal			