Context and Objective

Fibromuscular dysplasia (FMD) is a heterogeneous group of idiopathic, non-atherosclerotic, relatively rare vascular diseases, leading to the narrowing of medium-sized arteries, mostly the renal and internal carotid arteries [1]. Identifying patients with FMD at the international level will have a considerable value for understanding the epidemiology, clinical manifestations and susceptible genes in this arterial disease but also for including eligible patients in clinical trials or cohorts. In France, data are collected in heterogeneous databases at different points of care. In USA, the FMD Society of America decided to begin a registry in 2007 [2]. The large diversity of cultures, laws, regulations and operational implementation regarding personal health data processes (access, gathering, sharing, etc.) across countries induces variation in the nature of the collected data.

In this context, frameworks accessing and comparing distributed and heterogeneous data with consistent semantics are needed to enable cooperative research and progress in the comprehension of the disease. Our main challenge is to develop such framework that includes an e-registry collecting standardized FMD data across Europe and that gives access and re-use of existing local FMD registries in other regions like USA [3].

Material and Methods

A French network of nephrologists, neurologists, radiologists and specialists in hypertension is responsible to document phenotypic and genetic traits of the disease and the progression of FMD lesions in patients with renal and/or cervical artery FMD. Information regarding >500 patients with FMD is currently scattered in several non-standardized, redundant databases. A template has been defined to collect 318 items. Clinical data include elements such as date of diagnosis, types of tests conducted and results of these tests, past medical history, family history, subsequent clinical events and any clinical outcomes [1]. In parallel, the FMDSA began a registry to better understand the disease and its treatment [2]. Standards knowledge organization systems (KOS) and tools are available for the project. The coding of the medical domain values is adapted to a country’s existing practices. For example, the rare disease diagnosis can be coded in Orphanet (http://www.orpha.net/consor/cgi-bin/index.php) OMIM, or SNOMED CT depending on the context and the required granularity. In France, diagnostic coding relies on Orphanet codes and is done by clinicians. Multilingual academic portals such as UMLS (http://www.nlm.nih.gov/research/umls/) or HeTOP (http://www.hetop.eu/) provide alignments between reference KOS.

A three steps methodology will be followed:

1) A bottom-up approach to build a multidimensional model: A data set includes data elements (DE) corresponding to specifications that stipulate the sequence of inclusion of the DEs, whether they are mandatory,
what verification rules should be employed, and the scope of the collection. A core data element (CDE) is a DE that is used in various data sets and recognized by the experts as standard information. A specific DE (SDE) is a DE specifically defined for a given purpose (e.g., the “race” is a SDE used as a public health indicator in US). The first step is to build CDE and SDE data sets for FMD. The target model is a fixed set of CDEs.

2) Using reference terminologies to align SDEs: One limit of using CDE come from the fact that data are often collected at a local level for specific studies and may include specific data, semantically related to FMD but not explicitly expressed in the multidimensional model. For instance, the item “Summary_Smoking” is collected in [1] to answer a given query, it is semantically related to the item “number-of-cigarettes_daily” of the multidimensional domain but some alignment is needed. In this case, semantic interoperability can be achieved through data integration guided by an ontology. Reference KOS developed and maintained by Standard Development Organisations (e.g., IHTSDO) may be used to explicit the semantics of the data. Such approach will be used to define mappings between reference KOS and local interface terminologies used in local registries for SDEs [4]. KOS updates will be taken into account to update the mappings.

3) Development of an e-registry: To develop a platform that implements the multidimensional model, the semantic alignments and some web services to query data across several heterogeneous databases to answer research questions or to target specific profiles.


PREREQUIS DEMANDE, FORMATION : MASTER IN BIOMEDICAL INFORMATICS AND KNOWLEDGE ENGINEERING IN E-HEALTH

---

**EXISTE-T-IL UN CANDIDAT POUR CETTE THESE**

- OUI
- NON

**NOM DU CANDIDAT :**

**CURSUS :**

- OUI
- NON

**DEMANDE DE CONTRAT DOCTORAL DE L’ED**

- OUI
- NON

**LE CANDIDAT SE PRESENTE-T-IL AU CONCOURS DE L’EHESP DE LA FRM**

- OUI
- NON

**FINANCEMENT DU SALAIRE DU DOCTORANT PAR UNE AUTRE RESSOURCE**

- OUI
- NON

**SI OUI, LAQUELLE :**

**CONTACT POUR CE SUJET :**

**EMAIL :**

**TELEPHONE :**

**SPECIALITE DE LA THESE**

- EPIDEMIOLOGIE (AVEC EVENTUELLEMENT MENTION CLINIQUE OU SOCIALE OU GENETIQUE)
- BIOSTATISTIQUE/BIOBIOMATHEMATIQUE
- INFORMATIQUE BIOMEDICALE (RECOUVREANT AUSSI L’IMAGERIE BIOMEDICALE ET LA BIOINFORMATIQUE)
- RECHERCHE INFIRMIERE
- PERFORMANCE DU SYSTEME DE SOINS
- AIDE A LA DECISION (DONT COUT-EFFICACITE)
- AUTRES (PRECISER) : ...

**VISA DU DIRECTEUR DE L’UNITE**

- OUI

**SIGNATURE**