DEVELOPING CARDIAC ELECTROPHYSIOLOGY ONTOLOGY: MOVING TOWARDS DATA HARMONIZATION AND INTEGRATION

Hadi Kazemi-Arpanahi1, Mostafa Shanbehzadeh2*, Saeed Jelvay3, Hassan Bostan3

1Assistant Professor, Department of Health Information Technology, Abadan School of Medical Sciences, Abadan, Iran.
2Assistant Professor, Department of Health Information Technology, School of Allied Medical Sciences, Ilam University of Medical Sciences, Ilam, Iran.
3MSc in Health Information Technology, Department of Health Information Technology, Abadan School of Medical Sciences, Abadan, Iran.

ABSTRACT

Introduction:
Cardiac electrophysiology (EP) studies the electrical heart conduction system which is used for diagnosis and treatment of cardiac arrhythmias. In this context, a huge amount of data is generated, requiring efficient and effective access, interpretation, and data analysis from multiple sources in a unified view. To resolve this challenge, this paper presents an ontology to reconcile data heterogeneity problems in this domain.

Material and Methods:
The cardiac EP ontology was constructed according to the life cycle of ontology building. Structural, functional, and expert evaluation was performed to ensure its quality and usability.

Results:
Cardiac EP ontology was developed using protégé environment and implemented in OWL editing tool. It presented a detailed hierarchical structure of the cardiac EP domain with around 324 instances describing cardiac EP-related concepts.

Conclusion:
Cardiac EP ontology provides an explicit formal description of the concepts, relationships, and properties associated with cardiac electrophysiology making seamless data integration between multiple heterogeneous databases. It also is a useful framework for knowledge representation in knowledge-based systems, as well as for explicit communication between experts in the EP domain.

Keywords:
Ontology, Heart Electrophysiology, Semantic Data Integration, Interoperability.

INTRODUCTION

Nowadays, health information systems (HISs) have been transformed from local to large complex systems [1]. With the advent of novel technologies such as Internet of Things (IOT), cloud computing, semantic web, wearable and portable medical devices, and increasing adoption of health information technologies, a huge amount of health data are generated [2-6]. In this big data era, there is an emerging need to facilitate interoperability and reusability among a variety of data sources. Nevertheless, the technical solutions to integrating data are still a key challenge. Data integration is a daunting process as heterogeneous data are collected in different data sources [7-10].

To reconcile this challenge, ontology has been emerging as an interesting and commonly used solution to represent concepts, properties, and semantic relationships in a specific domain. Ontologies are the conceptual models that give context and meaning to the data [11-14].

Accordingly, consideration of ontology is one of the most important prerequisites for achieving a
comprehensive and standardized approach in cardiac electrophysiology \[15\].

According to world health organization reports, by 2020, about 25% of healthy life years would be missed as cardiovascular disease complications, especially in developing countries. The results of studies have also demonstrated that cardiovascular disease is the leading cause of death in Iran \[16\]. Cardiac arrhythmias are an important contributor to many diseases, and have considerable effects on patient morbidity and mortality. Cardiac electrophysiology (EP) studies electrical heart activities, which provides diagnostic and therapeutic strategies including drug therapy, catheter ablation, device implantation, and upgrade/repair/replacement/explant devices for the treatment of heart-rhythm disorders \[17-19\]. In this field, huge amount of data are generated from different HISs and cardiovascular implantable electronic devices (CIEDs) \[20, 21\]. Each CIED and HIS has its own nomenclature, technical standards, and communication protocols. In such situation, interoperability will face some difficulties. Hence, of the absence of a unified conceptual framework for heart interventions has led to data redundancy and rework \[22, 23\]. With such issues in mind, the objective of this study is to develop an ontology called cardiac EP ontology in order to address the semantic integration and interoperability between independent HISs used in EP, and also between them and CIEDs applicable in heart electrophysiology.

**MATERIAL AND METHODS**

Currently, there is no valid and coherent ontology design guide for designing or assessing medical ontologies. Indeed, the nature of biomedical ontologies required application of a customized approach to identifying and presenting unique concepts and their relationships in a certain domain \[24\]. The ontology of heart electrophysiology has been developed based on METHONTLOGY and according to the ontology life cycle \[25\]. This is an accepted method in the construction of many developed ontologies \[26-28\]. In this study, initially, EP ontology domain and scope were clearly defined using needs assessment techniques in some focus group meetings with cardiologists and Health Information Management (HIM) experts. Prior to ontology construction, we used a list of sources recommended upon cardiologists including medical text books, clinical (GOLD) guidelines, and other online references to extract the key concepts. Then, cardiac electrophysiologists were asked to enumerate all known variables associated with cardiac EP. Following the knowledge extraction, a hierarchical conceptual model was developed to aid in the organization and design of the ontology classes and properties prior to implementing the ontology in a formal representation. The conceptual model was designed in tabular and graphical models, and revised throughout the development process to ensure precision, consistency, and extendibility, as well as to decrease redundancy and to support the functional specifications.

Our approach to building cardiac EP ontology is top down where we have first outlined the most general nodes and subsequently added the descendant nodes. A set of ontology design principles and domain expert review have been applied to measuring EP ontology content and structural validity. Finally, The Protégé Web Ontology Language (OWL) (https://protege.stanford.edu/) has been used as a tool for building the EP ontology in OWL format \[29\].

**RESULTS**

The EP ontology concepts were extracted from cardiovascular reference books, disease (GOLD) guidelines, and then discussed with heart electrophysiologists and ontology engineers. The ontology presents a detailed taxonomic overview of the cardiac EP domain with around 324 instances describing cardiac EP-related concepts. Examples are "Cardiac hypertrophy", "Blood pressure signs" or "Heart murmurs". These concepts are interconnected with super-class and sub-class properties into a hierarchical tree-like structure. At the basic level, there are three super-classes: "pre-operative", "intra-operative", and "post-operative". In the field of ontology, instances are members of the classes and typically represent a list of concrete concepts relevant to the class. For example, the "Cardiac hypertrophy" class has the following six instances: "Cardiomegaly", "Combined ventricular hypertrophy", "Left atrial hypertrophy", "Left ventricular hypertrophy", "Right atrial hypertrophy" and "Right ventricular hypertrophy". In total, the EP ontology includes more than 320 instances (Fig. 1).

---

**Fig1. Statistical characteristics of cardiac EP ontology**

The first class in each ontology is the "OWL thing" class, which is the parent class for all real-world
entities such as “heart electrophysiology” which are considered in the current project. The ontology hierarchy structure is partitioned based on three major stages: preoperative, intraoperative, and postoperative. They are regarded as the three major sub-categories under the “heart electrophysiology” root class.

The preoperative superclass contains administrative, medication, and CAT-LAB Visit classes. The intraoperative category is the parent of the “operation” class, which covers catheter ablation and implantation electronic devices. The postoperative super class is the parent class of the health condition and discharge information concepts. EP ontology classes and subclasses are explained as follows:

“Patient-administrative” class contains patient’s demographical characteristics as well as information related to the current episode of care. “Cat Lab visit” class consists of concepts about which information is collected during a cardiac electrophysiology visit session, including “risk factors”, “medical history”, “pervious care surgery”, “heart conduction system”, “physical examination”, “laboratory testing” and “prior diagnostic study”. “Medication history” captures “cardiovascular drugs” and “non-cardiovascular drugs” organized into medication groups. Also, medications are organized into medication groups. "Operation" class consists of medical procedures used in the treatment process, including medications, devices, invasive and non-invasive procedures, and recommendations regarding cardiac arrhythmia. Invasive procedures are broken down into catheter ablation, permanent pacemaker implantation, implantable cardioverter-defibrillator, and upgrade/repair/change/explant device.

Intraoperative domain involves “general procedure” and “Implantation Electronic Devices” classes describing the events occurring or performed during the course of surgical EP ablation. “Procedure information” class consists of information about date of procedure, duration of procedure, sedation type, ablation type, and indication of catheter ablation. In order to describe the implantable electronic devices class, catheterization (ablation), pacemaker implantation, and device replacement or upgrade subclasses were used. The concepts in this domain are implantable cardioverter-defibrillator (ICD), Cardiac Resynchronization Therapy (CRT) and pacemaker implant, lead information, upgrade/repair, /replacement/explant procedure.

Postoperative sub domain contains "post procedure complication", “discharge information” and “discharge medication”. Post-procedure complication is segregated into “minor complication” and “major complication” subclasses. Discharge information includes “discharge date”, “discharge location”, “discharge status”, “cause of death”, and “date of follow up”. Finally, discharge medications are categorized into medication groups.

In Fig 2-7, all triple EP domains have been structured in the protégé environment. It provides an explicit description of the concepts, relationships, and properties associated with heart electrophysiology.
Fig 4. Graphical overview of "CAT lab visit" subclasses in protégé environment

Fig 5. Graphical overview of "invasive EP procedures" subclasses in protégé environment

Fig 6. Graphical overview of "post procedures complication" subclasses in protégé environment
DISCUSSION

Cardiac EP ontology is the first standardized human- and computable format that reflects explicit formal specifications of heart electrophysiology concepts and their relations. Cardiac EP ontology is intended to be used for facilitating the understanding, reusability, and sharing knowledge, and ultimately encouraging collaborative work among domain experts. This ontology was derived through an exhaustive, iterative, and collaborative construction process based on the expertise of multiple disciplinary teams including cardioelectrophysiologists, health informatics, and HIM experts.

The primary application of the cardiac EP ontology is knowledge management function in heart electrophysiology domain. In this regard, management and integration of large datasets provide enormous opportunities for elicitation of new knowledge [30]. Furthermore, most scientific knowledge is still kept in the format of natural language, which is mostly narrative, ambiguous, and subjective. With the incremental complexity of biomedical knowledge, a method is required for standardized and well-defined knowledge representation (KR). In other words, the knowledge representation is fundamental in the field of medical informatics. Ontology is the most popular way for knowledge representation in today’s web-based scenario where information is constantly being shared among different applications [31]. Thus, the EP ontology can be used as a KR technique to represent data and knowledge in cardiac electrophysiology.

Our approach also facilitates data integration between different databases. To integrate data across different HISs, it is necessary to have a formal description of the mental concepts individuals have about different entities. Considering data integration, ontology is used to identify correspondence between entities of local information sources which are semantically related [32]. Indeed, ontology-based data integration (OBDI) denotes the use of ontologies that capture implicit knowledge across independent databases to achieve semantic interoperability between them [33].

Using ontologies regarding data integration, first a standard definition of data elements is provided so that it can be comprehended by humans and computers. Next, it draws the semantic relations among data elements and their allowed values. Thereafter, it models the constraints of data elements...
as well as automating data validation and quality assurance. Subsequently, it encodes different data integration scenarios explicitly using metadata [34]. Indeed, ontology provides a semantic interface which is independent of the database pattern. It supports synchronized management and identifies incompatible data. In addition, ontology provides a mechanism for defining concept-based queries and presents results in structured and uniform ways [35-37]. In this study, cardiac EP ontology can provide a semantic layer and conceptual interface for solving the data heterogeneity problems.

Hotchkiss et al. described the Hearing Impairment Ontology (HIO) which would allow standardized HI-related knowledge in a single location and promote HI data integration, interoperability, and sharing [38].

Mate et al. presented an ontology-driven method which employed abstraction as an alternative to defining Extract, Transforms and Loading (ETL) processes at the database levels; they used ontologies to represent and define the medical concepts of both the source and target system [39].

Lucas and their colleagues developed the cancer cell ontology according to the open biological and biomedical ontology (OBO) foundry principles. This ontology facilitates integration of different data sources by providing a structured semantic representation and explicit definitions that are human and computer readable [40].

Min et al. dealt with application of an ontology to support integration and querying of heterogeneous information across a prostate cancer database and tumor registry [37]. Yamada et al. constructed an ontology for integrating databases of different natures within the mental health domain [6].

Cardiac EP ontology is expected to provide clear, consistence, and formal information descriptions to standardized clinical and administrative processes through encouraging common understanding in the field. In the present study, cardiac EP ontology has been a stable conceptual interface for database systems as it offers a rich and predefined vocabulary for consistency and standard definition of data elements.

To the best of our knowledge, EP ontology has not previously been developed; however, several ontologies were available for the cardiovascular domain. For example, CIEDs ontology have been developed for the annotation of implantable devices and medication, automatic detection of therapeutic changes, and even analyzing data from other device registries. [41]. Cardiac-centered Frailty Ontology has been developed as a machine-interoperable description to allow making decisions on patient treatment. The full hierarchy has three top classes (clinical history finding, instrument finding, and physical examination finding), 12 qualifiers, 156 concepts, with 246 terms [42].

Heart Rate Turbulence (HRT) ontology was developed by Ruiz et al. based on SNOMED-CT codes for semantic interoperability's in her [43]. Romero et al. designed an OWL-based system for monitoring the vital signs of patients with acute cardiac disorders which processed a huge volume of complex data to accelerate vital decisions [44]. In this work, an ontology was developed for intelligent supervision and standardized data collection within and across cardiac EP departments.

It is hoped that EP ontology will enable standardized data collection from narrative and unstructured documents such as history and operative reports in EP department. This ontology can be applied for elicitation of valuable knowledge from EP huge data. It also facilitates data integration and data reusability across multiple heterogeneous EP databases. Nevertheless, this work had a limitation; it was extracted based on opinions of cardio electrophysiologist in Tehran Heart Center (THC). Nonetheless, the working group made these required data elements based on the best currently available appropriate evidence and a vast collective wealth of experience.

CONCLUSION

In computer and information sciences, ontology is the most comprehensive and standardized human- and machine-interpretable resource which formally represents knowledge as a set of concepts in a given domain. Accordingly, in this research, cardiac EP was conceptualized to serve as a power semantic framework for data harmonization in EP heterogeneous data environments. This paper could (i) concisely name EPS concepts, (ii) offer common understanding between clinicians and researchers, and (iii) promote more efficient data integration and interoperability with other existing systems. Given that biomedical ontology design processes are usually collaborative, repetitive, flexible, extensible and continuous [11]. So, cardiac EP ontology will be modify based on corrective, continuous enrichment, and development collaboration [45] by specialists in the field of medical informatics and cardiac electrophysiology. Additional future work is suggested to undertake enrichment of the relational ontologies and resolve any redundancies that may occur during the importation process of the different domain. Next, the remaining domain specific concepts and relations to cardiac EP will be added to finalize the extension. The final step will be evaluating the ontology itself.

ACKNOWLEDGEMENT

Authors would like to thank all cardiologists who participated in this study and played a role in the validation of ontology classes.
REFERENCES


Developing Cardiac Electrophysiology Ontology: Moving Towards Data Harmonization
Hadi Kazemi-Arpanahi et al.


