Healthcare of the Future 2022
T. Bürkle et al. (Eds.)

doi:10.3233/SHTI220331

Towards a FHIR-Based Data Model for Coronary Angiography Observations

Florian HOLWEG[a,1], Stephan ACHENBACH[b], Noemi DEPPENWIESE[c], Luise GAEDE[b] and Hans-Ulrich PROKOSCH[a]

[a] *Chair of Medical Informatics, Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany*

[b] *Cardiology and Angiology Clinic, Universitätsklinikum Erlangen, Germany*

[c] *Center of Medical Information and Communication Technology, Universitätsklinikum Erlangen, Germany*

Abstract. Coronary heart disease is among the most frequent causes of death globally. Thus, our research project aims to develop prognostic models, to predict the risk of spontaneous myocardial infarctions based on a combination of clinical parameters and image data sets (invasive coronary angiograms). To train such models we use data from more than 30,000 coronary angiograms acquired at the cardiology department of Erlangen University Hospital. To linking such proprietary data with additional clinical parameters and to harmonize it for future cross-hospital federated machine learning approaches we defined a mapping for coronary angiography based on the symptom/ clinical phenotype HL7® FHIR® module of the German medical informatics initiative. In this paper we describe the final design of the coronary angiography information model and our mapping approach to ICD-10 and SNOMED CT. From the database we use a subset of 15 required values patient characteristics to create the HL7® FHIR® resource.

Keywords. HL7 FHIR, data model, coronary angiography, SNOMED CT

1. Introduction

Cardiovascular diseases, prominently including myocardial infarction, are the most frequent cause of death globally. Myocardial infarction often occurs in patients without prior symptoms since the most frequent mechanism is the sudden rupture of non-stenotic atherosclerotic coronary artery lesions [1-3]. Algorithms to identify at-risk individuals would therefore be of utmost clinical relevance.

A frequently performed test in patients with suspected coronary artery disease is invasive coronary angiography. In Germany alone, more than 900,000 coronary angiograms are acquired annually. In approximately 60% of these, coronary artery stenoses are identified and revascularization usually ensues. However, no guidance exists for further treatment in patients without stenosis. If, among these, individuals with increased risk for future myocardial infarction could be reliably identified, this would offer substantial opportunity for targeted, intensive risk factor modification with subsequent reduction of the risk for myocardial infarction and death. However, no such

[1] Corresponding Author, Florian Holweg, Friedrich-Alexander University Erlangen-Nürnberg, Chair of Medical Informatics, Am Wetterkreuz 15, D-91058 Erlangen, Germany, E-mail: florian.holweg@fau.de

algorithms exists today. Automated classification of angiograms is challenging due to their complex nature and large data volumes.

It is the long-term goal of our project ER-MIKKI (Determination of myocardial infarction risk from coronary angiography and clinical data with AI) to develop and validate risk prediction tools for future myocardial infarction based on coronary angiography image data and clinical parameters, making use of artificial intelligence-based models. The project is based on a cohort of 30,000 patients whose structured coronary angiography reports are stored in an MS Access® database (MSAd) using a proprietary format, without link to any further clinical data. In order to develop and validate the risk prediction algorithms, integration of the structured angiographic reports with clinical parameters such as laboratory data, procedures, medication and procedures performed, as well as data harmonization to a common data model was required.

This first subproject of ER-MIKKI therefore aims at developing a HL7® FHIR®-(HL7) based data model based on the core dataset definitions of the German medical informatics initiative (MII) [4-6], thus allowing to smoothly integrate the coronary angiography reports with clinical data within the Erlangen University Hospital data integration center (DIC) for subsequent processing.

2. Methods

Based on a comprehensive analysis of MSAd, the following parameters were identified as the basis for record linkage with further clinical data: First and last name, the date of birth and administrative gender. The date and start/end times of the clinical procedure serve as time stamps. Clinical parameters comprise the classified stenosis severity (% value) for the four major coronary arteries: Left Main Stem (LM), Left Anterior Descending (LAD), Ramus Circumflex (RCX) and Right Coronary Artery (RCA). In addition, a short free text describes the findings for each artery. The MSAd proprietory database contains one table with more than hundred attributes for saving the patient data from coronary angiography and 30 reference tables comprising value lists for categorical data items. The database contains more than 100 data points for each patient for a coronary angiography. Unfortunately, the structure of the MSAd is not based on a standard Entity Relationship Modell. One line in the output table corresponds to one observation.

For providing such data in a harmonized data model, the core data set (MII cds) definitions from the German medical informatics initiative were evaluated. The MII cds comprises six basic modules (person, encounter, diagnosis, procedure, laboratory results, medication) and numerous extension modules. After an analysis of all such cds modules we decided to apply the module person to map the identification and demographic data of a patient and the module symptom/clinical phenotype for representing the clinical state or clinical observations/findings of a patient. Even though the implementation guide of the final HL7 profile of this MII module is not yet finalized and balloted, an ART-DÉCOR [7] based data model and UML class diagram were already available, mainly being built upon the HL7 resource Observation [8].

3. Results

By comparing the proprietary data structure of the original MSAd with the HL7 modelling options, two slightly different models would have been possible. Figure 1 shows the ultimately consented HL7 data model (in a reduced form). For transforming the demographic information of the patient, we use the core data set module PERSON. The patient identification field is used for transporting the unique key of the MSAd. The fields birth name, name prefix, name affix and family name are not available in the original system and are omitted in our model in figure 1. Linked with PERSON we apply the core data set module CLINICAL SYMPTOM/CLINICAL PHENOTYPE and reduce this to OBSERVATION only (illustrated in figure 1).

The observation is referencing to the HL7 resource Observation [8]. This resource has as a main component the code of any medical classification system e.g., ICD or SNOMED-CT (SCT). With these classification systems a comprehensive information model for coronary angiography findings can be established [9,10]. In our project, we create a main component with an ICD-10 Code to classify the type of cardiologic disease (e.g., I25.11 single-vessel-disease, I25.12 two-vessel-disease). For mapping the coronary angiography findings to the OBSERVATION resource, the data elements category, state, and methods were set to the fixed values PROCEDURE, FINAL and 3367005 as the SCT code for coronary angiography. To describe the detailed finding related to the respective coronary arteries a hierarchical observation tree is built by linking associated 14 SCT codes with each other, using the "Reference to" connection. In total 35 SCT codes and four ICD-10 Codes can be used for descripting the disease.

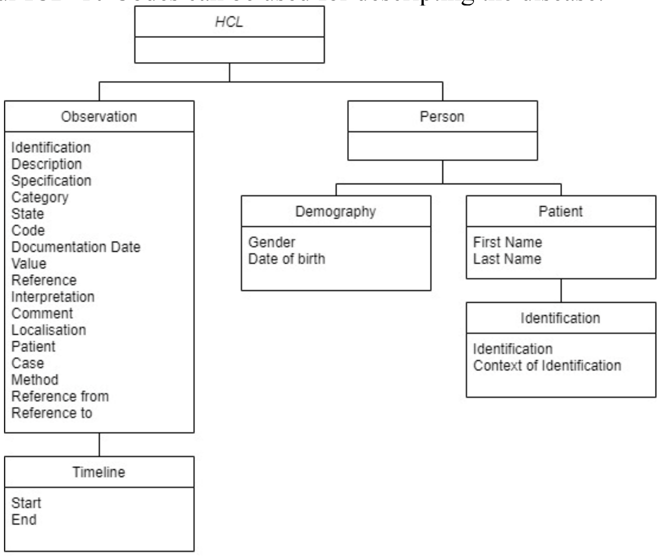

Figure 1. Art-Decor FHIR® Data Model for representing coronary angiography observations from the cardiology heart catheter laboratory (HCL)

Two hierarchically referenced OBSERVATION resources with SCT codes 3227004 and 3711007 for example describe that the left coronary artery main stem is smooth walled, whereas 3227004 referenced with 123641001 describe a left coronary artery occlusion. Documentation date, as well as start/end time are filled with the procedure date and times available in the MSAd. Data elements representing medical information not available within the original MSAd are left blank in the OBSERVATION.

4. Discussion

The ER-MIKKI research team aims at the determination of myocardial infarction risk from coronary angiography and clinical data with AI technologies. Even though in the current project phase we plan to only use data from one cardiology department to train our AI model, future enhancements could build on datasets from multiple hospitals, as it is supported by the German medical informatics initiative. For this goal, data must be provided in an interoperable manner, implying the standardization of data, so it can be easily shared among various institutions, avoiding the creation of silos. Similar to González-Castro in the CASIDE project [11] we responded to this need with the design of the ER-MIKKI-MODEL, a standardized data model based on HL7 which allows structured representation of coronary angiography observations. As already experienced by other researchers [12] we found multiple ways to first create an ART DÉCOR model and then map the clinical findings to SCT. The flexibility that HL7 offers to adapt to specific use cases and the power of SCT to sometimes express clinical meaning in different manners also poses challenges, since there often are multiple ways to map the same concept using available resources and codes. As we also wanted to link our data with clinical data from the university hospital EHR system, which has already been provided according to the MII core data set basic modules within our DIC, we had decided to follow the MII HL7 profile implementation guides and also use the ART DÉCOR tool to describe our model. For defining the mappings of the ICD-10 and SCT Codes to the dataset in the MSAd we can apply the standard MIRACUM mapping service [13] developed in our DIC. For quality assurance a workflow was defined.

References

[1] Little WC. Angiographic assessment of the culprit coronary artery lesion before acute myocardial infarction. Am J Cardiol. 1990;66(16):44G-47G. doi:10.1016/0002-9149(90)90395-h

[2] Nittel D, Seltrecht A. Krankheit: Lernen im Ausnahmezustand?: Springer; 2013.

[3] Nittel D. Krankheit: Brustkrebs und Herzinfarkt aus interdisziplinärer Perspektive. Mit Online-Material für Fachleute. Dordrecht: Springer; 2012. 574 p. ger.

[4] Semler SC, Wissing F, Heyder R. German Medical Informatics Initiative. Methods Inf Med. 2018;57(S 01):e50-e56.

[5] Medizininformatik Initiative. Der Kerndatensatz der Medizininformatik Initiative [Internet] [cited 2021 Dec 21]. Available from: https://www.medizininformatik-initiative.de/de/der-kerndatensatz-der-medizininformatik-initiative

[6] Medizininformatik Initiative. Basismodule des Kerndatensatz der MII [Internet] [cited 2021 Dec 21]. Available from: https://www.medizininformatik-initiative.de/de/basismodule-des-kerndatensatzes-der-mii

[7] Art-Decor. Art-Decor [Internet]. 2021 [updated 2021 Oct 7; cited 2022 Jan 3]. Available from: https://art-decor.org/mediawiki/index.php?title=Main_Page

[8] HL7. Observation [Internet] [cited 2021 Dec 21]. Available from: https://www.hl7.org/fhir/observation.html

[9] Benson T, Grieve G. Principles of Health Interoperability: SNOMED CT, HL7 and FHIR. 3rd ed. Cham, s.l.: Springer International Publishing; 2016. 451 p. (Health Information Technology Standards). eng.

[10] Uysal S. ICD-10-CM Diagnosis Coding for Neuropsychological Assessment. Arch Clin Neuropsychol. 2019;34(5):721–30. doi:10.1093/arclin/acy084 Cited in: PubMed; PMID 30357265.

[11] González-Castro L, Cal-González VM, Del Fiol G, López-Nores M. CASIDE: A data model for interoperable cancer survivorship information based on FHIR. J Biomed Inform. 2021;124103953.

[12] Oeppert L, Wehner K, Schrader T. FHIR data model for data exchange between cancer registries and IQTIG. 2021.

[13] Mate S, Seuchter SA, Ehrenberg K, Deppenwiese N, Zierk J, Prokosch HU, Kraska D, Kapsner LA. A Multi-User Terminology Mapping Toolbox. Stud Health Technol Inform. 2021 May 24;278:217-223.