Journal of Biomedical Informatics 60 (2016) 1–13

Contents lists available at ScienceDirect

Journal of Biomedical Informatics

journal homepage: www.elsevier.com/locate/yjbin

# Formalizing the Austrian Procedure Catalogue: A 4-step methodological analysis approach

Sabrina Barbara Neururer [a,b,*], Nelia Lasierra [b,c], Karl Peter Peiffer [a,d], Dieter Fensel [b]

[a] Department of Medical Statistics, Informatics and Health Economics, Medical University of Innsbruck, Innsbruck, Austria
[b] Semantic Technology Institute Innsbruck, University of Innsbruck, Austria
[c] UMIT – University for Health Sciences, Medical Informatics and Technology, Hall in Tirol, Austria
[d] FH Joanneum University of Applied Sciences, FH JOANNEUM Gesellschaft mbH, Graz, Austria

ARTICLE INFO

*Article history:*
Received 29 June 2015
Revised 13 December 2015
Accepted 24 December 2015
Available online 30 December 2015

*Keywords:*
Ontologies
Austrian Procedure Catalogue
Definition analysis
Typological analysis
Health procedure classification modeling

ABSTRACT

*Objectives:* Due to the lack of an internationally accepted and adopted standard for coding health interventions, Austria has established its own country-specific procedure classification system – the Austrian Procedure Catalogue (APC). Even though the APC is an elaborate coding standard for medical procedures, it has shortcomings that limit its usability. In order to enhance usability and usefulness, especially for research purposes and e-health applications, we developed an ontologized version of the APC. In this paper we present a novel four-step approach for the ontology engineering process, which enables accurate extraction of relevant concepts for medical ontologies from written text.

*Methods:* The proposed approach for formalizing the APC consists of the following four steps: (1) comparative pre-analysis, (2) definition analysis, (3) typological analysis, and (4) ontology implementation. The first step contained a comparison of the APC to other well-established or elaborate health intervention coding systems in order to identify strengths and weaknesses of the APC. In the second step, a list of definitions of medical terminology used in the APC was obtained. This list of definitions was used as input for Step 3, in which we identified the most important concepts to describe medical procedures using the qualitative typological analysis approach. The definition analysis as well as the typological analysis are well-known and effective methods used in social sciences, but not commonly employed in the computer science or ontology engineering domain. Finally, this list of concepts was used in Step 4 to formalize the APC.

*Results:* The pre-analysis highlighted the major shortcomings of the APC, such as the lack of formal definition, leading to implicitly available, but not directly accessible information (hidden data), or the poor procedural type classification. After performing the definition and subsequent typological analyses, we were able to identify the following main characteristics of health interventions: (1) Procedural type, (2) Anatomical site, (3) Medical device, (4) Pathology, (5) Access, (6) Body system, (7) Population, (8) Aim, (9) Discipline, (10) Technique, and (11) Body Function. These main characteristics were taken as input of classes for the formalization of the APC. We were also able to identify relevant relations between classes.

*Conclusions:* The proposed four-step approach for formalizing the APC provides a novel, systematically developed, strong framework to semantically enrich procedure classifications. Although this methodology was designed to address the particularities of the APC, the included methods are based on generic analysis tasks, and therefore can be re-used to provide a systematic representation of other procedure catalogs or classification systems and hence contribute towards a universal alignment of such representations, if desired.



* Corresponding author at: Department of Medical Statistics, Informatics and Health Economics, Medical University of Innsbruck, Schoepfstrasse 41/1, 6020 Innsbruck, Austria. Tel.: +43 512 9003 70 912; fax: +43 512 9003 73 922.

*E-mail address:* sabrina.neururer@i-med.ac.at (S.B. Neururer).

http://dx.doi.org/10.1016/j.jbi.2015.12.018
1532-0464/

## 1. Introduction

The payment system of Austrian hospitals underwent a paradigm shift in 1997 [1]. Until then, hospitals had been reimbursed according to a so-called per-diem payment system [1]. This payment, which depended on the length of hospital stays of their patients, led to inefficient, long and expensive hospital stays [2]. To increase transparency and decrease the cost-explosion of health care, Austria implemented a performance-oriented hospital financing system based on modified diagnosis-related groups (DRG) [3]. Nowadays, in 2015, this system is still used in Austria. It allows for the billing of health services – including diagnoses as well as health interventions – to be rendered in hospitals. Such a payment model relies on accurate documentation and coding of diagnoses and health interventions. Diagnoses coding has been performed based on the International Classification of Diseases (ICD) [4] since 1989. While the ICD is an internationally accepted and commonly used standard, such a standard does not exist for procedure coding [5]. Therefore, Austria has developed its own country-specific classification system, which is called Austrian Procedure Catalogue (APC; German: Österreichischer Leistungskatalog) [6]. The Austrian Procedure Catalogue is the obligatory basis for procedure coding within the framework of the Austrian performance-oriented hospital financing system. In contrast to many mono-axial classification systems in the healthcare domain (e.g. ICD-10 [4]), the Austrian Procedure Catalogue (APC) offers a multi-axial architecture to classify health interventions [6]. APC codes are classified according to three independent axes: (1) Anatomical site, (2) Procedural Type, and (3) Access. The anatomical site axis describes the anatomical structure that is targeted by a specific medical procedure [6]. It is further subdivided into a general and a detailed anatomical site. While the general site provides a more general idea of the target region or body organ (e.g. "eye" [7]), the detailed site names the specific part of an organ targeted (e.g. "ocular muscles"). The procedural type explains the kind of medical procedures [6] (e.g. "therapy", "in vivo diagnostic investigation/in situ diagnostic investigation" [7]). The third axis provides information about how the targeted anatomical region of a medical procedure is accessed. For example, the medical procedure "suture of ocular muscles" is classified according to the three axes as follows: (1a) general anatomical site: "eye", (1b) detailed anatomical site: "ocular muscles", (2) procedural type: "therapy", and (3) access: "open access".

Each medical procedure is represented by a so-called procedure code. For example, the code "BJ010" represents the previously mentioned medical procedure "suture of ocular muscles". Additionally, a procedure has a procedure name and a textual description that provides more detailed information about the health intervention itself [7]. The APC contains approximately 1,500 procedure codes and is described in German. The APC was developed with the intention of it being used for billing and health policy making only, but it is also often employed for research purposes without paying any regard to the particularities of this field. Even though the Austrian Procedure Catalogue is an elaborate coding standard for procedure coding in general, it also has shortcomings that limit its usability, especially in the case of e-health applications or research purposes. A known shortcoming of the APC is that it offers a basic formal structure that allows a classification of terms according to three independent axes but does not support access to the full catalog in a machine-readable manner which is a first step towards semantic interoperability. Formal definition that allows the extraction of further information from the catalog is listed by Cimino [8] as an important desideratum for controlled vocabularies in the 21st century. Currently, the APC includes a list of procedure codes and descriptions in textual language that makes interoperability or the use for e-health applications difficult. Since formal definition and semantically enriched representation support semantic interoperability, it is a major requirement for useful e-health applications to share machine-readable knowledge. Ontologies provide a formal representation of concepts of a domain as well as the relations among them [9] and are becoming more and more important in terms of domain knowledge representation and sharing. Biomedical ontologies address most of the desiderata for controlled vocabularies [9], which makes ontologies a useful and powerful solution for the representation of medical coding schemes [10].

This lack of formal definition motivated our analysis and led us to design an ontology in order to extend and improve the Austrian Procedure Catalogue. The aim of this ontology is to explain and model medical procedures performed in Austrian hospitals in more detail than the APC does in its current state, and to offer a framework for the Austrian hospital financing system, as well as to provide benefits for research and e-health applications. Based on popular methodologies for guiding the ontology engineering process [11] and considering the particular features of a procedure catalog, we designed a four-step methodology which stresses a combination of systematic analysis tasks for deriving the main concepts of the ontology with the support of domain experts without involving them in long time-consuming tasks. As the APC plays an important role within the framework of the performance-oriented hospital financing system, we had to take important features of the existing system (e.g. procedure codes, axes codes, initial classes) into account.

This combination of analysis tasks enabled us to identify the vocabularies and relations, which had to be included in the ontology. We were able to abstract common characteristics of health interventions from textual descriptions that are contained in the APC. A domain expert's opinion was taken into account as the proposed analysis methods allowed us to receive feedback quickly and effectively and thus facilitated their involvement at the validation stage without excessive time demand. Hence, the aim of this paper is to present the four-step methodological approach used for the ontology engineering process that facilitates the identification of main concepts inside a procedure catalog system, which was successfully used to design an ontological model of the APC. These analysis tasks allow an accurate and systematic extraction of the required concepts and vocabularies and could also be applied to develop other ontologies in the medical field.

The remainder of this paper is organized as follows: Section 2 presents the background and state of the art, Section 3 includes a detailed description of the methods involved in the proposed methodological approach. In Section 4, the outcomes of each respective step are described. Section 5 discusses the obtained results and finally, conclusions are presented in Section 6.

## 2. State of the art

The following section presents related work and the contributions of our work to the state of the art are clarified.

### 2.1. Ontologies and vocabularies for medical procedures

As of today, structured data about medical procedures is barely comparable on an international level [5]. Many different classification systems exist for health interventions like the German Operationen- und Prozedurenschlüssel (OPS) [12], the American Current Procedural Terminology (CPT) [13] or the French Classification Commune des Actes Médicaux (CCAM) [14]. The APC is a fairly small and simple catalog, mainly used for billing and health policy making, while the CPT is a complex catalog used for reporting to payors, developing medical review guidelines, medical research as well as education. The CPT consists of approximately 10,000 different codes and is therefore more detailed and complex than the APC. The German OPS is a large, mono-hierarchical classification of health care procedures with alphanumeric codes organized in six chapters [12]. It is used for procedure coding within the framework of the German DRG system (G-DRG) as well as for quality reports [12]. In contrast to Germany or the United States of America, where a specific person – a medical coder – takes care of the diagnoses and procedure coding, in Austria the medical personnel (usually physicians, sometimes nurses) conducts these tasks. This is a strong argument for an easy and less complex classification system for Austria, as coding can be a demanding and time consuming task. The French CCAM is a hierarchically structured procedure coding system, which served as a role-model for the conversion of the mono-axial Austrian Procedure Catalogue into its multi-axial successor in 2009. The CCAM as well as the Austrian Procedure Catalogue lack of formal semantics grounded in logic. Formal knowledge representation by means of ontologies provides a powerful solution to support semantic interoperability and knowledge sharing within the scope of e-health [15]. Besides national initiatives, the World Health Organization (WHO) started to develop the International Classification of Health Interventions (ICHI) [16], which was intended to become an international standard for procedure coding. The first (trial version, 2002) version of ICHI [17] was developed as a catalog of health interventions and provided a shallow hierarchy. The classes were coded as sequential numbers and they left no place for future extensions [5]. This version of ICHI (ICHI 2002) was withdrawn for revision by the WHO. A new alpha version is now available [18].

The power of formal definition is also underlined by a comprehensive multilingual terminology for electronic health records, the Systematized Nomenclature of Medicine – Clinical Terms (SNOMED CT) [19,20], a taxonomy of 311,000 concepts [21]. According to Schulz et al., SNOMED CT meets the needs for a global standardized terminology [21]. Nevertheless, this powerful instrument also contains several issues that are addressed in various different publications (e.g. [22–24]).

In recent years, many biomedical ontologies have been developed. An ontology is defined as a "formal, explicit specification of a shared conceptualization" [25]. In other terms, ontologies represent classes of entities of the real world and focus on the principled definition of concepts and relations between them [9]. They offer a good solution to addressing the challenge of machine-readable concepts in order to facilitate data sharing.

With a formalization of the Austrian Procedure Catalogue, which does not yet exist, we are contributing to the current state of the art. Furthermore, we provide a novel systematic and well-structured methodology that enables an accurate extraction from written text fragments of concepts relevant for medical ontologies. With the four-step approach for medical ontology engineering as well as with the formalized APC, we enable its re-usability and facilitate its further integration in one possible standard.

### *2.2. Methodologies and tasks for the ontology engineering process*

The Ontology engineering process refers to "the set of activities that concern the ontology development process, life cycle and the methodologies, tools and languages for building ontologies" [26]. Available literature within the context of the Semantic Web offers a wide variety of methodologies for guiding the ontology engineering process [27]. Some of the most popular and well known methodologies are the METHONTOLOGY [11], DILIGENT [28,29], UPON [30] or the "*Ontology Development 101*" presented by Noy and McGuinness [31]. These methodologies assist ontology literate developers (usually in collaboration with different stakeholders and domain experts) within the design of an ontology and, in general, contain tasks for: (1) identifying the scope and purpose of the ontology, (2) deriving required vocabularies and relations, (3) providing a conceptualization, formalization and implementation (4) and conducting an evaluation whilst planning its maintenance. As research in ontologies has been extended out of the Semantic Web community to other domains, innovations in the available methods considering the particular features of each application scenario becomes essential. Particularly, research in biomedical ontologies brings new challenges that demand deep involvement of domain experts and strong validation tasks assuring its integration to the clinical setting and usability [32]. Aside from literature reviews, these particular needs are usually covered by collaborative tasks involving clinical personnel and engineers, e.g. interviews and/or innovative participatory tasks for knowledge co-production [33,34]. New trends also propose to relay these tasks to domain experts and other users (inside and outside of the medical domain) through gamification and other crowdsourcing and motivational approaches, not only for the ontology process itself, but also to maintain quality Linked Data [35–37]. However, conducting these tasks is time-consuming, and time is of high value, particularly, in the medical domain. Hence, our innovative proposal attempts to define methods for the knowledge acquisition to be conducted by non-medical personnel who could provide rich and accurate vocabularies and relations to be further validated in a simple and quick manner by a medical domain expert. In particular, this includes a comparative analysis, definition analysis and typological analysis: well-known and effectively-proven methods used in social sciences (but not that commonly used in the domain of computer science and in particular, ontology engineering tasks) [38]. Innovations in our methodology refer to the analysis tasks within the process of deriving vocabularies and relations. Hence, domain experts have a principal role in the ontology process, while little effort is required from them. However, future evaluation tasks will be included in order to assess the final ontology version.

## 3. Material and methods

The aim of the Austrian Procedure Catalogue ontology is to explain and model medical procedures performed in Austrian hospitals in a more detailed manner than the APC does in order to answer a set of questions used as competency questions for the development of the ontology. These questions were identified when analyzing the APC in detail and after defining the possible applications of the APC ontology in the clinical setting. This includes questions such as the following: "Which medical procedures target a specific anatomical region?", "Which medical procedures target a specific body system?", "What is the aim of a specific medical procedure?", or "Is a specific medical procedure related to a specific population?" Once we had formulated the competency questions to be answered by the help of the APC ontology, and had defined its scope, we applied a set of analysis tasks to extract the knowledge to be encoded in the ontology, derived the correct vocabularies and provided it with a formalization of the conceptualized catalog. In fact, the APC ontology contains static information already available in the APC and extends it to additional characteristics of health interventions. Fig. 1 presents a graphic overview on the proposed four-step approach to systematically identify the main concepts of the Austrian Procedural Catalogue ontology.

The first step, named comparative pre-analysis, consisted of an analysis of strengths and weaknesses of the current version of the Austrian Procedure Catalogue. One way to highlight such strengths and weaknesses is to compare the APC with other well-established or elaborate and comparable health intervention coding systems.

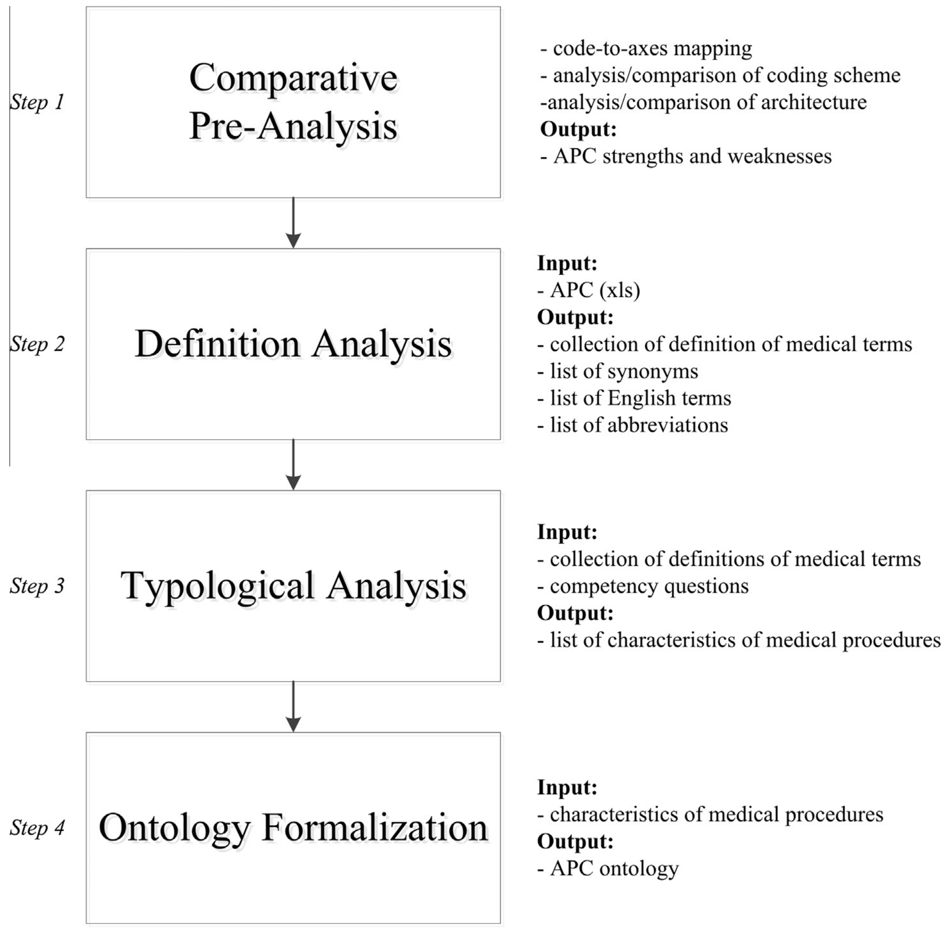


**Fig. 1.** Four-step approach for formalizing the Austrian Procedure Catalogue (APC).

In the second step – a definition analysis –, a list of definitions of medical terms used in the procedure classification was obtained. The list of definitions resulting thereof was used as an input for Step 3. Within Step 3, which consisted of a typological analysis, the most important concepts, which describe medical procedures, were identified. Therefore, the competency questions, which were posed at the beginning, were also taken into account, since the characteristics derived from this step had to answer these questions. Finally, this list of concepts was used in Step 4, which corresponds with the ontology formalization step as such.

We were able to perform the analysis tasks of Step 1 to Step 3 without deep medical knowledge and background, but basic understanding of the medical domain as well as information modeling is needed. A medical domain expert was involved in the validation of our results. The validation was performed at two different stages of the analysis: (1) end of Step 1 and (2) end of Step 3 (validation of Step 2 and Step 3).

### 3.1. Comparative pre-analysis (Step 1)

During Step 1, we performed a comparative pre-analysis in order to identify strengths and weaknesses of the current Austrian Procedure Catalogue (APC). Particularly, we performed mappings of procedures of the APC to: (1) the French Classification Commune des Actes Médicaux (CCAM) and (2) to the International Classification of Health Interventions 2002 (ICHI 2002). A strength was defined as an APC attribute of particular worth that gives advantage over CCAM or ICHI 2002 (e.g. multi-axial architecture) and might be relevant for the ontology and therefore should be taken into account during the ontology development process. A weakness was defined an attribute of the APC that should be either revised or replaced following the example set by CCAM or ICHI 2002 during the ontology implementation process (e.g. poor classification).

We compared the APC architecture with special regard to the underlying coding system, hierarchy, content and granularity with the CCAM and the ICHI 2002. We followed the mapping approach described by Hanser et al. [5] and mapped APC procedures to classes of CCAM and ICHI 2002. First, we analyzed the code description of the APC procedures in order to conceive its meaning. For the APC-to-CCAM mapping, we then assigned each single procedure of the APC to one class of each axis of the CCAM: (1) Anatomical site, (2) Action, (3) Access. For example, the APC procedure "Reconstruction of the aortic valve" was mapped to the following CCAM classes: (1) Anatomical site = DB (Cardiac valve, endocardium), (2) Action = M (Repair, which includes "reconstruction" according to the CCAM coding system), (3) Access = A (Open access). ICHI 2002 offers a different architecture than the CCAM and arranges the health interventions according to a shallow hierarchy that includes the body system (chapter), the anatomical site (Section) and the procedural type (Subsection) [5]. Nevertheless, we tried to assign each procedure of the APC to an anatomical site as well as to one of the eleven listed procedure types of the ICHI 2002. For example, our APC procedure "Reconstruction of the aortic valve" was mapped to the following ICHI 2002 classes: (1) Anatomical site = Heart – aortic valve (which belongs to the chapter Cardiovascular system), (2) Procedure type = Reconstruction. For the mapping processes, we limited ourselves to

1:1-mappings. This means that we assigned only one Anatomical site, Action/Procedure type (or Access mode) of CCAM or ICHI 2002 for each medical procedure from the APC. Besides full 1:1-mappings, also partial 1:1-mappings were carried out, which means that it was not possible for a specific APC description to identify related classes within all three axes of the CCAM or within the Anatomical site as well as the Procedure type of ICHI 2002. For example, the APC procedure "Other procedure – endocrine glands" could not be mapped to exactly one anatomical site class of ICHI 2002, as ICHI 2002 offers more detailed classes and therefore, more than one corresponds to endocrine glands. This means that no 1:1-mapping for the anatomical site was possible. Nevertheless, we were able to identify a corresponding procedural type in ICHI 2002 that corresponds to "other procedures". So, in summary, not a complete 1:1-mapping (in terms of assigning an ICHI 2002 anatomical site as well as procedural type to a specific APC procedure) was possible, but a partial 1:1-mapping was successfully carried out. This limitation to strict 1:1-mappings helped us to identify differences in the granularity of the classification systems, if such a mapping could not be carried out. Performing 1:$n$, $m$:1 or even $m$:$n$-mappings would not have added additional information for our purposes. As Hanser et al. also performed 1:1 mappings of ICHI 2002 to CCAM, our results allowed for cross-comparison of the results.

All the mappings were performed manually in German language (as the APC procedures are described in German) by a knowledge engineer with background in medical informatics. One medical expert was involved for the validation of our results of Step 1. To conduct this task, we first provided basic information and explanation to the medical domain expert in order to make the mapping process understandable. Then, two Excel sheets (one for CCAM, one for ICHI 2002) containing all the mappings were given to the physician. These tables included a total of approximately 6250 mappings (mappings of one code to up to three/two axes of CCAM/ICHI 2002 are counted separately). This table consisted of the following columns: (1) APC code of each procedure, (2) the name of the medical procedure, (3) the mappings to the anatomical site classes of CCAM/ICHI 2002, (4) the mappings to the procedure type classes of CCAM/ICHI 2002 and (5) the mappings to the access classes of CCAM. Along with these Excel sheets, we provided a list of all anatomical site, procedural type and access classes (only for CCAM), that were offered by CCAM and ICHI 2002, to the physician. Thereafter, the physician assessed the correctness of the mappings performed by the author. The correctness of the mappings was decided based on the knowledge of the physician. For mappings classified as wrong, she was requested to indicate the appropriate one. The results of the validation process were discussed thereafter with the medical expert.

### 3.2. Definition analysis (Step 2)

For the definition analysis (Step 2), we concentrated on procedures referring to the "heart, large vessels and cardiovascular system". It was not our purpose to cover all procedures in detail (which would lead to the analysis of more than 1,500 procedures), but to come up with a methodology to systematically identify the main concepts of a medical procedure ontology. Therefore, we focused on a complex body structure with many different procedures, representative enough to conceptualize the APC in a generic manner. The definition analysis itself consisted of five different tasks as depicted in Fig. 2. These tasks needed to be performed for each medical procedure separately: (1) analysis of the code description, (2) extraction of search strings from the code description, (3) searching for strings in a clinical dictionary, (4) searching for search strings on the Internet and (5) document definitions as well as additional information.

The first task consisted of reading and analyzing the code description in order to provide an overview on the medical procedure itself and to conceive its meaning. This understanding was needed to perform the second task, in which the text of the code description of the medical procedure was split up into several different search strings. For example, the medical procedure "Reconstruction of the aortic valve" was split up into the search strings "reconstruction" and "aortic valve". We did not further subdivide it into "aortic" and "valve", as aortic valve was understood as a coherent medical term. The list of search strings was used as an input for Task 3 and Task 4. In Task 3, the search strings derived in the previous task were looked up in the Pschyrembel Clinical Dictionary (digital version) [39] in order to identify their definitions. For example, "reconstruction" is defined as the "reconditioning, regeneration (of the initial state)", while "resection" is defined as the "surgical removal of a sick part of an organ". If Task 3 could not be performed successfully, the definitions of the search strings were searched on the Internet using valid sources. As soon as an

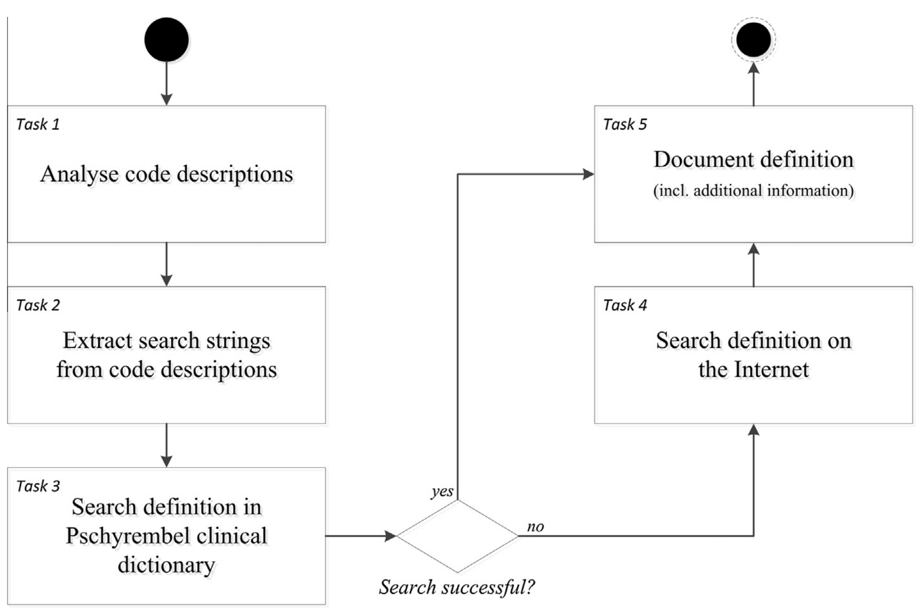


**Fig. 2.** Tasks to perform a definition analysis.

appropriate definition was discovered (either in the clinical dictionary or on the Internet), the definition as well as additional information (e.g. synonyms, abbreviations, Latin terms, English/German terms) were documented (Task 5). The additional information served to cover the medical terminology on a broader basis than it is covered by the APC in itself. The list of definitions was used as input to perform the typological analysis (Step 3).

### 3.3. Typological analysis (Step 3)

The aim of the third step, the typological analysis, was to derive main characteristics of health interventions from the list of definitions created during Step 2. This step consisted of two tasks: (1) identification or construction of main characteristics of health interventions and (2) description of these main characteristics. The first task is an iterative process. For this purpose, the list of definitions was systematically analyzed using a content analysis method, where we tried to abstract the definitions step-wise and iteratively created more general types. They represent the main characteristics of health interventions and therefore the main concepts of the proposed ontology. For example, the abstraction of the "aortic valve" shows that it could be abstracted to the more general type "cardiac valve site" based on its definition (see Fig. 3). Therefore, we were able to say that the aortic valve, which is the most detailed representation of an anatomical site, is a cardiac valve site. This means, these two concepts are related by a so-called is-a (isA) relationship. Is-a relationships were represented by a class-subclass relationship in the ontology. The different cardiac valves were subsumed under "cardiac valve site" and abstracted to "heart site". The heart is part of to the "cardiovascular system" (isPartOf).

Fig. 4 shows another example of the definition and typological analysis process with regard to the procedural type for three different medical procedures: ("Endomyocardial biopsy", "correction of a congenital heart defect with life-support machine", "reconstruction of the aortic valve"). The abstraction from the definition to the characteristic (Procedural type) needed four abstraction levels (LVs). In order to guarantee a structured and guided approach, we were looking for best-practice examples for performing this abstraction. As the CCAM provides a mature system to classify such medical procedures, we decided to use this. Therefore, in Level 1 (LV 1), we were able to abstract our definitions to specific actions according to CCAM. In Level 2 (LV 2), we further abstracted them according to the action groups of CCAM. While "Biopsy" was abstracted to "Take", "Correction" as well as "Reconstruction" was abstracted to "Repair". Repair procedures belong to "Therapy" procedures according to the APC, while "Take" is an "In-vivo diagnostics" procedure. All these procedures describe specific types of medical procedures and are therefore abstracted to the characteristic "Procedural Type".

By performing the typological analysis following a definition analysis, we were also able to extract important relationships of the proposed ontology. For instance, we were able to find out that a medical procedure may include the implantation of a stent,

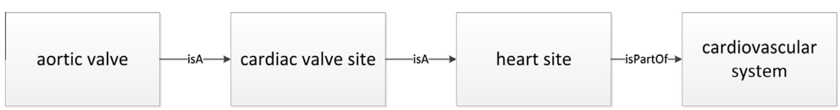


**Fig. 3.** Abstraction of aortic valve to cardiovascular system (including relations).

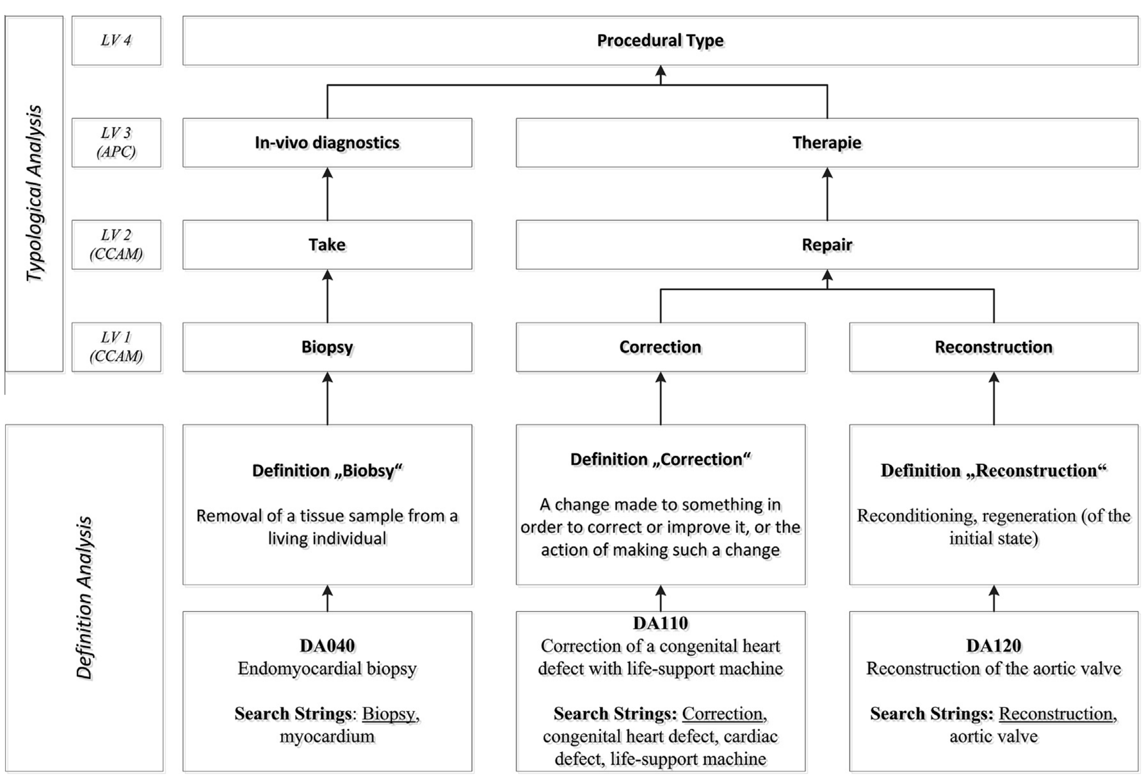


**Fig. 4.** Example of the definition and typological analysis approach.

which indicates a relationship of medical procedure to a specific medical device.

One medical expert, who was the same person as the one that was used for the validation of Step 1, was involved in the validation of our results of Step 2 and 3, which was performed at the end of Step 3. To conduct this task, we first provided basic information and explanation to the medical domain expert in order to make the definition and typological analysis understandable. Then, we gave an Excel list to the physician, which contained the following information: (1) procedure code, (2) the procedure description, (3) the search stings that were extracted from the procedure description, (4) a definition per search string, (5) additional information per search string (synonyms, abbreviations, English terms, etc.) and (6) the abstracted characteristics resulting from the typological analysis. The list contained a total of 501 German definitions. In the first instance, we asked the medical domain expert to validate the definitions that were derived by a knowledge engineer during Step 2. Therefore, she checked every single definition for every search string we have identified and documented. Based on her knowledge and medical experience as a general practitioner, she indicated the definition as "correct" or "incorrect". If definitions were marked as incorrect, we asked her to provide a correct definition based on available literature. The results of the first validation step were discussed thereafter by the medical expert and the author, in order to guarantee that no inaccuracies in terms of understanding the natural language occur. In a second step, the validity of the identification of the characteristics of health interventions was assessed by discussing the (stepwise) abstraction of the information gained from the definitions to more general types by performing a content analysis. It was necessary to assess whether our abstractions were plausible or not. In case an abstraction was inaccurate according to the physician's knowledge and in case that the physician was able to come up with a better generalization, this was documented and justified in the Excel list.

### *3.4. Ontology formalization (Step 4)*

The main characteristics of medical procedures derived in Step 3 were used for the ontology formalization itself. These characteristics represent the main concepts of the ontology. Once we had derived all the vocabularies, we used them to provide a conceptualization that answered our competency questions:

For the development of the ontology, we chose the OWL 2 Web Ontology language (informally OWL 2) [40]. Extending RDF [41] and RDF Schema [42], OWL 2 is one of the most popular languages for developing ontologies. It is an extension of RDF and RDF Schema that offers many powerful vocabularies to provide higher levels of expressiveness to the ontologies. Given the demands in terms of restrictions (e.g. qualified cardinality restrictions), this language was selected to develop APC. Particularly, the Protégé [43] software was used for the formalization of the ontology.

## 4. Results

The following section describes the main results of each step of our proposed four-step approach. Results from each step were used as input for the subsequent steps.

### *4.1. Comparative analysis (Step 1)*

#### *4.1.1. Mapping (CCAM)*

After a comparative analysis with the CCAM, we found out that the APC and the French CCAM are comparable regarding (1) their architecture, (2) the number and content of the majority of classes [44]. Nevertheless, important disadvantages of the ACP compared to the CCAM were determined and are thus aspects that need to be taken into account for the ontologized version of the APC. The procedural type axis of the APC and the Action axis of the CCAM are significantly different in number, content and granularity. While the APC offers only 5, the CCAM provides 41 different classes. This is supposed to be the major disadvantage of the APC compared to the CCAM. Therefore, either a revision of the APC procedural type axis or the introduction of additional classes or further sub-division of the APC axis into the CCAM axis structure is suggested. For the implementation of the APC ontology, we decided to introduce additional levels of detail for the procedural type according to the CCAM subdivision.

For the APC as well as for the CCAM, the combinations of the axis codes build multi-axial triples. Several different procedures can be assigned to one multi-axial triple. For example the APC procedure "suture of ocular muscles" is described by the following axes: (1) general anatomical site: eye (axis code: B); detailed anatomical site: ocular muscles (axis code BJ), (2) procedural type: therapy (axis code: T), (3) access: open access (axis code: A). The combination of these axis codes is "BJTA". Several other medical procedures can be assigned to the same combination of axis codes. The difference between the APC and the CCAM regarding the basic coding system is the design of the procedure code itself. While the CCAM procedure codes are built by the whole multi-axial triple followed by a three digit number (7-digit code), only the Anatomical site axis is reflected in the procedure code (5-digit code) within the APC. For the example of the "suture of ocular muscles" the procedure code is BJ010. The Procedural type and Access codes are dropped within the APC procedure codes. Therefore the APC codes neither reflect the multi-axial triple nor the multi-axial architecture anymore. For example, the medical procedure "Reconstruction of the aortic valve" is represented by the procedure code "DB030", which only reflects the general and detailed anatomical site axes of the APC (D = Heart and large vessels, cardiovascular system; DB = Cardiac Valves, endocardium). The axes codes of the procedural type of this medical procedure (T = Therapy) and the access (A = Open access) cannot be assessed by the procedure code itself.

Both classification systems lack a formal definition as required by Ciminio [8]. We also found out that the lack of formal definition, which was a known shortcoming of the APC, lead to a further issue, which concerns the quantity of information that is not accessible by the multi-axial architecture of the APC but implicitly available in textual code descriptions [45]. For instance, while the APC allows to search for medical procedures related to the heart very easily by filtering for the heart-related classes via the anatomical site axis, procedures using a life-support machine have to be searched manually or by using a full-text search on the code descriptions [7].

#### *4.1.2. Mapping (ICHI 2002)*

The advantages of the APC over the International Classification of Health Interventions (ICHI 2002) can be found in the multi-axial architecture and the characteristics of the basic coding system. While the APC offers a multi-axial architecture, ICHI 2002 is a classic list of health interventions. Therefore, the APC has an advantage over ICHI 2002 when it comes to extending the content. The APC can be extended easily with additional medical procedures without any effect on the architecture and coding system. ICHI 2002 classes are numbered consecutively and therefore an extension of the content would lead to inconsistencies in the architecture.

Even though ICHI 2002 does not provide a multi-axial architecture, it sorts its codes according to: (1) an organ system, (2) an anatomical site and (3) a procedural type. As this structure is not reflected in the procedure codes, ICHI 2002 cannot benefit from this classification. A major shortcoming of the Austrian Procedure Catalogue is the procedural type axis. When comparing the APC procedural type axis with the procedural type structure of ICHI 2002, it is shown that the subdivision provided by ICHI 2002 is far more mature than the one provided by the APC [46]. For example, while ICHI 2002 sorts its codes according to procedure types such as Examination, Incision, Destruction, Excision, Reduction, Repair, or Reconstruction, the APC only provides procedure type classes such as Medical Treatment, Therapy, Special organization effort, or in vitro examination. These classes are even overlapping. For example, a medical treatment is a specific kind of therapy and therefore rather a subclass than an independent, distinct class.

The major disadvantage of the APC compared to the International Classification of Health Interventions (ICHI 2002) is the immature procedural type axis.

Regarding the anatomical sites, both procedure catalogs use combinations (e.g. “Pelvis, Hip” or “Neck and Thorax” in ICHI 2002, “Head and Neck” in the APC). This leads to (1) problems in terms of varying granularities (2) limited comparability with other medical procedure classification systems, (3) limited support for proper statistical analysis [5] of the performed health intervention. The ICHI 2002 also lacks formal definition.

#### *4.1.3. Physician validation (Step 1)*

After performing the mappings (APC to CCAM and APC to ICHI 2002) by a non-medical expert, we provided a total of 6250 mappings (84% of full mappings, 16% of partial mappings) to a medical expert who validated or corrected our results. The validation took the physician two working days, while the mapping process itself required several weeks. The majority of mappings were indicated as correct (93%) and only 460 mappings (7%) were incorrect. Many of the incorrect mappings were caused by reproduction errors (approx. 300 of the mappings). For example: If we were not able to correctly map a procedure targeting the jaw joint once, it is very likely that other procedures targeting the jaw joint will be mapped incorrectly as well. Nevertheless, a percentage of wrong mappings under 10% can be considered as acceptable, especially when taking into account the fact that this approach vastly decreased the work load for the physician.

### *4.2. Definition analysis (Step 2)*

For the definition analysis, 121 procedures of the Austrian Procedure Catalogue relating to the heart, large vessels and the cardiovascular system were identified. Out of the short text and the code descriptions, we were able to extract 501 search strings. For example, the “Reconstruction of the tricuspid valve” (Code DB050) was split up into the search strings “tricuspid valve” and “reconstruction”. For these two search strings as well as for all other search strings, the definitions were documented including additional terms.

For example: Search string: ventricle (Germ. Ventrikel)
Definition: small gaster; chamber
English term(s): gaster; ventriculus; ventricle
Synonym: Ventrikulus
Abbreviation: (none)

In 73% ($n = 336$) of all cases, we were able to identify the definitions of search strings in the clinical dictionary. For approximately one quarter of the search strings, we used Internet sources (Internet based clinical dictionaries) in order to identify their definitions.

Along with the definitions, we were able to identify English terms, synonyms as well as abbreviations. In 80% ($n = 403$) of all cases, for the German search strings, the English translation was documented in order to provide a broader scope of medical terminology. For one third of the search strings, alternative terms (synonyms) were identified. Abbreviations were provided by the dictionaries or found on the Internet in 15% of all cases ($n = 74$). English terms, synonyms and abbreviations were also used for the formalization of the ontology.

### *4.3. Typological analysis (Step 3)*

Using the list of definitions derived from the definition analysis, we were able to create the types in Table 1, which represent the

**Table 1**
Abstracted main characteristics of health interventions (typological analysis results).

| Type | Description | Example |
|---|---|---|
| Procedural type | The procedural type describes the kind of action that was performed by a particular medical procedure. For example, it allows classifying, whether some device was implanted or not or whether an anatomical structure was repaired or not (e.g. suture of ocular muscles)<br>(The procedural type was iteratively abstracted and therefore offers different levels of granularity; subclass-structure) | Implant repair remove |
| Anatomical site | An Anatomical site is the anatomical target region that is targeted by a medical procedure<br>(The anatomical site was iteratively abstracted and therefore offers different levels of granularity; subclass-structure) | Heart |
| Device | A medical device describes an instrument, apparatus or machine that was used, implanted or removed during a medical procedure | Life-support machine stent |
| Pathology | Pathology describes the disease that makes a specific medical procedure necessary | Sarcoma |
| Access | Access describes how an affected Anatomical site (target region) was approached during a specific medical procedure | Open transdermal |
| Body System | A body system is a group of Anatomical sites (usually organs) that are working together | Cardiovascular system |
| Population | Population is a specific subgroup of the populace that is targeted by a certain medical procedure (e.g. a hysterectomy usually concerns women) | Children, women, men |
| Aim | The Aim of a medical procedure describes its desired outcome and the goal, why a specific procedure is performed | Cardiac resynchronization |
| Discipline | The discipline describes the medical field within which a medical procedure is usually performed | Cardiology |
| Technique | Technique indicates that a specific medical procedure is performed by a certain approach that is often named after its inventor | Ross procedure |
| Body Function | The body function describes a process that takes place in the body and is affected by a medical procedure | Sleep |

main characteristics of health interventions and thus the higher abstract concepts.

As a result of the typological analysis, we were able to abstract more than 90% of all definitions to abstract characteristics of health interventions. For only 8% of all definitions, no adequate type was identified (e.g. intracavitary, intracranial).

The first three characteristics listed in Table 1 (Procedural Type, Anatomical Site, Medical Device) were abstracted from subtypes in an iterative process. This abstraction procedure helped to define the taxonomy of the ontology. For example, the search string "tricuspid valve" was abstracted to the more general site "cardiac valve". The cardiac valve was then abstracted in a further iteration to the type "heart site", which again belongs to the type "anatomical site". Using this approach, we were able to identify characteristics (types) as well as sub-types. The original value sets of the APC's three axes (anatomical site, procedure type, access) including the codes and additional information needed to be restored in order to meet the requirements of the performance-oriented hospital financing system. For further subdivisions and for extending the original value sets of the APC, several standardized vocabularies were used. For the further subdivision of the procedural type axes, we used CCAM. For example, the procedural subtypes "correction" and "reconstruction" were summarized to "repair" procedures according to the CCAM classification. In the following abstraction iteration, "repair" procedures were abstracted to the general type "Procedural Type". The "Anatomical Site" is further subdivided in more detailed sited according to MeSH (Medical Subject Headings) [47]. For the medical devices, we used the classification scheme of the Austrian Registry for Medical Devices [48], which offers the classes (1) active implantable products, (2) anesthesia and respiratory devices, (3) dental devices, (4) electromedical and electro-mechanical devices, (5) hospital inventory, (6) non-active implantable products, (7) ophthalmic and optic devices, (8) re-usable products, (9) single-use products, (10) technical support for handicapped people, (11) X-ray or other imaging products, (12) products for complementary medicine, (12) products of biological origin and (14) technical assistance devices. The Access axis' value set as well as the additional information (available in the APC) value sets (chargeable unit, side) are restored from the original APC. MeSH is also used for the value sets of the new classes such as pathology and aim.

#### 4.3.1. Physician's validation (Step 2 and 3)

After a non-medical expert performed the definition and typological analysis, we provided a total of 501 definitions to a medical expert who validated or corrected our results. The validation took the physician four hours, while the analysis process itself required several weeks. All definitions were indicated as correct by the physician. Nevertheless, the physician was able to provide nine additional synonyms, which were not identified during the initial analysis.

### 4.4. Ontology formalization (Step 4)

The characteristics of health interventions as well as the static information from the Austrian Procedure Catalogue are the main concepts of the Austrian Procedure Catalogue ontology (see Fig. 5). The concepts representing static information from the APC itself that has to be preserved within the APC ontology are highlighted with dotted lines and gray background color in Fig. 5. This content is extended with the additional characteristics from the typological analysis. The core of the ontology is the Medical procedure concept, which is directly related to most of the other concepts. It represents the Medical procedure itself according to the APC. An example of a medical procedure is "Reconstruction of the aortic valve".

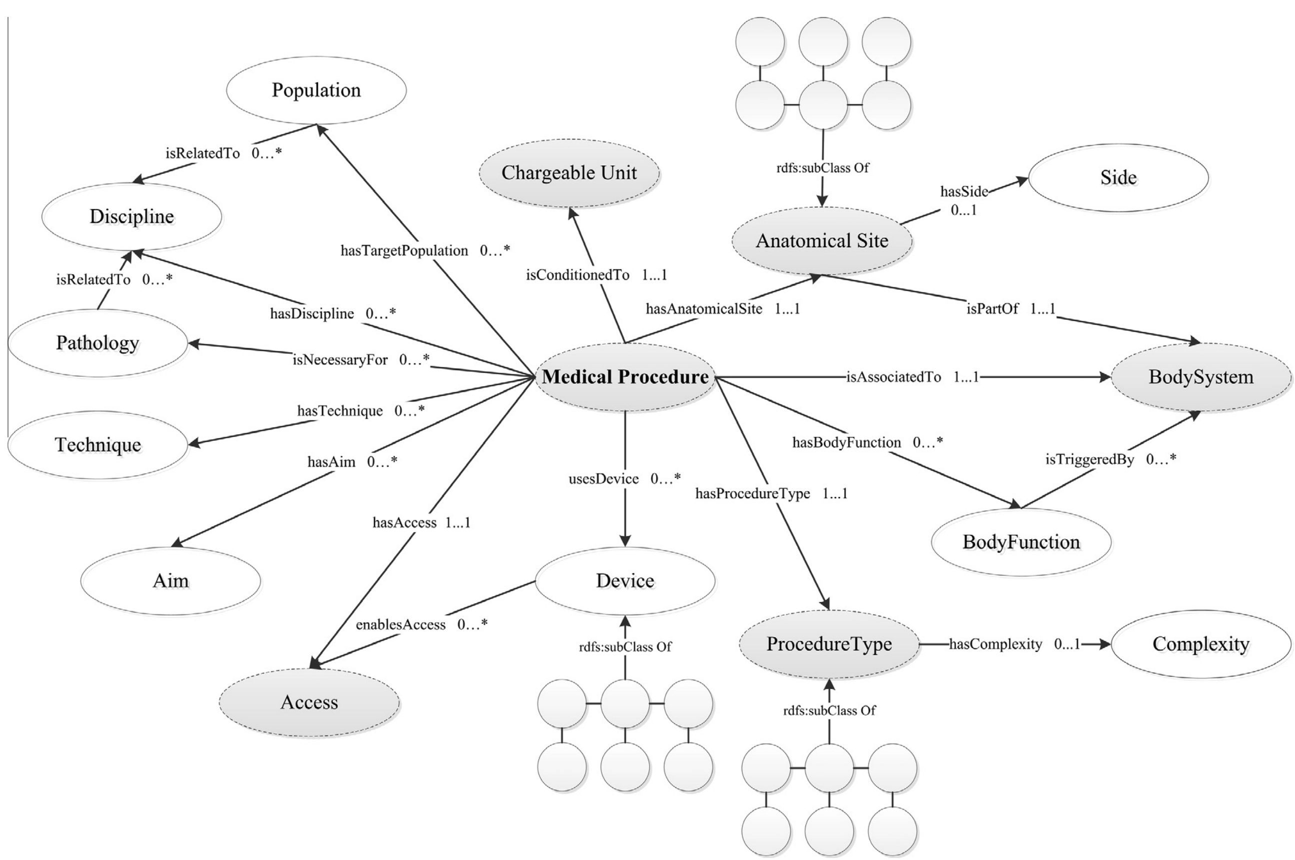


**Fig. 5.** Main concepts and object properties of the APC ontology.

Further main concepts of the APC ontology directly related to the Medical Procedure concept are listed and explained in more detail in the following: (1) Access, (2) Aim, (3) Anatomical Site, (4) Chargeable Unit, (5) Complexity, (6) Body Function, (7) Body System, (8) Device, (9) Discipline, (10) Pathology, (11) Population, (12) Procedure Type, (13) Side and (14) Technique.

*The Access* class contains types of access related to a specific medical procedure. It explains how the target region of the medical procedure was reached. Access is static information already available in the APC. An example for concept *Access* is "open access", which describes that it was necessary to open the patient's body in order to accomplish a specific health intervention. "Transdermal access", for example, shows that the target region of a medical procedure was reached through the skin.

*The Aim* class describes and models the aim, which underlies a specific procedure. It provides response to the question what the best possible outcome of a medical procedure is. An example for class *Aim* would be "cardiac resynchronization". It describes that a specific heart-related procedure was performed in order to resynchronize the cardiac rhythm.

The *AnatomicalSite* class models a generic summary of anatomic regions of the human body. This class has a hierarchy of subclasses in order to provide more detailed information on the anatomical site of a medical procedure. While the two most generic subclass levels represent the classes of the Anatomical site axes (general and detailed) of the APC (static information), further subclasses are introduced to provide more detailed information.

The *ChargeableUnit* class provides information for the coding of health interventions and explains how often a medical procedure can be coded (e.g. per session, per side, per stay). The ChargeableUnit class offers static information available from the APC.

The *Complexity* class offers additional information on the complexity of a medical procedure (e.g. partial, total).

*The Body function* class models functions of the human body. Body function is static information available in the APC. An example for a human body function is "sleep". In the original Austrian Procedure Catalogue, body functions are often treated like anatomical sites.

The *Body system* class describes the system of a human body that is affected by a specific medical procedure. The body system contains static information from the Austrian Procedure Catalogue. An example for a *Body system* is "cardiovascular system".

The *Device* class models the medical devices used while performing a specific medical procedure. It offers a subclass structure according to the classes of medical devices of the Austrian Registry of Medical Devices [48]. An example for a medical device is a "pace maker", which is an "active implantable device".

If a medical procedure is related to a certain medical discipline, it is modeled in the class *Discipline*. An example for a discipline would be "cardiology" or "pediatrics".

The *Pathology* class gathers instances of diseases or malfunctions for which a medical procedure is necessary. For example, the Resection of an intracavitary sarcoma instance has been linked to the instance sarcoma of type Pathology. This class can be used as some sort of connection points, if the ontology needs to be extended with diseases or mapped to a disease or diagnosis coding system or ontology (e.g. the International Classification of Diseases – ICD).

If specific populations are affected by a medical procedure, it is modeled in the class *Population*. Some health interventions are related to a certain group of people. An example for a population would be "children", "elderly", "female", or "male".

The *ProcedureType* class models the type of procedure that was used to perform a specific medical procedure in several different levels of detail. The most generic level represents the classes of procedural type according to the APC (static information). These classes were further subdivided into subclasses following the taxonomy of the CCAM Actions (see Fig. 4).

The *Side* class describes if a medical procedure targets a specific side of the AnatomicalSite. For example, some lung-related procedures may differ in the affected side of a patient (left, right), since the lung has two parts.

The *Technique* class models the exact approach of a medical procedure. They are often named after their inventor. An example for the class *Technique* is the "Ross procedure". This procedure is also often referred to as a "pulmonary autograft". It describes a cardiac surgery where a malfunctioning aortic valve is replaced with the patient's own pulmonary valve. The pulmonary valve, on the other hand, is replaced by an allograft. The procedure is named after Donald Ross, who was the first to perform this procedure.

Apart from the static information from the original APC modeled as classes and highlighted in Fig. 5, there is other important original information that needs to be restored within the APC ontology for financial and administrative reporting, e.g. the codes of the medical procedures but also the original axis codes for Anatomical sites (general and detailed), Procedure type and Access have to be preserved. For the ontology formalization process, we solved this by including the code associated with each medical procedure as well as the axes codes using data properties. Further data properties are e.g. synonyms, abbreviations, English terms and annotations for coding instruction.

The relations (object properties) between the concepts of the ontology as illustrated in Fig. 4 are described in more detail in Table 2.

## 5. Discussion and outlook

This study describes a four-step approach to identifying main concepts of an ontology that models main concepts of medical procedures performed in Austrian hospitals in a systematic manner. We were able to extract implicit information contained in the APC and make it explicit. It allowed us to perform the analysis without the direct involvement of the physicians until we reached the validation steps. It should be pointed out that, although no strong background in medicine is required to conduct these tasks, some medical background/medical informatics background (such as in our case) is desirable in order to accomplish the steps in a precise and structured manner. For more complex classification system, the proposed approach could benefit from the inclusion of a medical professional in the team that is performed the analyses and modeling tasks. The first step – the comparative pre-analysis – gave an idea of the strengths and weaknesses of the Austrian Procedure Catalogue, but also provided best-practice examples regarding the weaknesses that might be useful to be taken into consideration in the APC ontology. The mappings also provided a quantification of the content and the granularity of the APC compared to CCAM and ICHI 2002. Its success depends on the classification systems that are chosen for comparison. We decided to use the French Classification Commune des Actes Médicaux, which served as a role-model for the transformation of the APC from a mono-axial to a multi-axial classification system. As the comparison (in terms of mappings) was part of the APC ontology development process which enabled us to identify strengths and weaknesses of the APC, we aimed at a comparison of classification systems with different architectures. Therefore, the initial version of ICHI (ICHI 2002) was selected for comparison with the APC because, in contradiction to the APC, the ICHI 2002 provided a shallow hierarchy, while the APC offered a three-axes approach. By choosing the ICHI 2002, we wanted to find out how the APC performs compared to a classification system with a simpler architecture. Thanks to this analysis, we were able to identify advantages

**Table 2**
Object properties of the Austrian Procedure Catalogue ontology.

| Name | Domain | Range | Description and Cardinalities |
|---|---|---|---|
| hasAnatomicalSite | MedicalProcedure | AnatomicalSite | An instance of Medical Procedure is related to **exactly one** instance of Anatomical Site |
| hasProcedureType | MedicalProcedure | ProcedureType | An instance of Medical Procedure is related to **exactly one** instance of Procedure Type |
| isAssociatedTo | MedicalProcedure | BodySystem | An instance of Medical Procedure is related to **exactly one** instance of Body System |
| hasAccess | MedicalProcedure | Access | An instance of Medical Procedure is related to **exactly one** instance of Access |
| isConditionedTo | MedicalProcedure | ChargeableUnit | An instance of Medical Procedure is related to **exactly one** instance of Chargeable Unit |
| hasDiscipline | MedicalProcedure | Discipline | An instance of Medical Procedure is related to **none, one or more** instances of Discipline |
| isNecessaryFor | MedicalProcedure | Pathology | An instance of Medical Procedure is related to **none, one or more** instances of Pathology |
| hasTechnique | MedicalProcedure | Technique | An instance of Medical Procedure is related to **none, one or more** instances of Technique |
| hasAim | MedicalProcedure | Aim | An instance of Medical Procedure is related to **none, one or more** instances of Aim |
| usesDevice | MedicalProcedure | Device | An instance of Medical Procedure is related to **none, one or more** instances of Device |
| hasBodyFunction | MedicalProcedure | BodyFunction | An instance of Medical Procedure is related to **none, one or more** instances of Body function |
| enablesAccess | Device | Access | An instance of Device is related to **none, one or more** instances of Access |
| isRelatedTo | Pathology **or** Population | Discipline | An instance of Pathology **or** Population is related to **none, one or more** instances of Discipline |
| isPartOf | AnatomicalSite | BodySystem | An instance of Anatomical Site is related to **exactly one** instance of Body system |
| isTriggeredBy | BodyFunction | BodySystem | An instance of Body Function is related to **none, one or more** instances of Body System |
| hasComplexity | ProcedureType | Complexity | An instance of Procedure Type is related to **none or one** instances of Complexity |
| hasSide | Anatomical Site | Side | An instance of Anatomical Site is related to **none or one** instances of Side |
| hasTargetPopulation | MedicalProcedure | Population | An instance of Medical Procedure is related to **none, one or more** instances of Population |

of the ICHI 2002 that helped us to determine suggestions for improvement of the APC ontology. Note that the ICHI Alpha [18] was not available when the initial comparison was conducted. No doubt that this could be a good option for the comparison for future improvements of the APC ontology. The first step requires a precise and highly structured way of working. Therefore, the outcome quality of the first step of the 4-step approach presented by us relies on several different factors: (1) the selection of appropriate classification systems for comparison in terms of quality and comparability and (2) the precision of how the systematic mappings are carried out. The second step, the definition analysis, provided a basic understanding of the procedures contained in the APC. An important prerequisite of this step is a textual description of the health interventions. The second step enabled us to identify information that was hidden in the code descriptions and not explicitly represented by one of the APC axes. It also provided additional information (e.g. synonyms, English and Latin terms, abbreviations), which helped to enrich the ontology with further information and to cover more terms of the medical terminology used by physicians. The success of the second step depends on the following aspects: (1) an accurate search string extraction, (2) the quality of the used vocabularies and therefore the identification of high quality definitions and (3) the validation of the definitions by a medical professional. To extract adequate search strings as well as to identify high quality definitions, we relied on clinical dictionaries. If we needed to search the Internet, we looked for trustworthy sources (online dictionaries) instead of relying on internet sources of unknown authorship. The definitions needed to be validated by a physician in order to guarantee their correctness. The third step, named typological analysis, allowed us to identify the main concepts in order to formalize the Austrian Procedure Catalogue. Therefore, a content analysis was performed on the list of definitions derived in the prior step. The success and the quality of the outcome of the third step rely on the following requirements: (1) this approach needs a strong focus on a systematic practice. Therefore, experience with qualitative content analysis methods is required and (2) accurate validation by a medical expert, guidance for the construction of the main characteristics as well as discussion of controversial results is needed. If performed accurately and based on a systematic practice, the third step results in a description of the taxonomy needed for the APC ontology. Thanks to the typological analysis, the fourth step resulted in a formalized Austrian Procedure Catalogue ontology containing all the main concepts identified. The proposed four-step approach can be re-used to develop and formalize other medical procedure coding classifications. The usage for other types of ontologies needs to be further analyzed.

The first step, a comparative analysis with other procedure classification systems, showed the Austrian Procedure Catalogue as an elaborate coding scheme. Nevertheless, various shortcomings were identified either. The main weakness of the APC that motivated this analysis is the lack of formal definition. Therefore, the usability of the APC is limited. This weakness also causes another shortcoming of the APC: the information that was included in the code description but not reflected in the multi-axial architecture of the APC. Thanks to the conducted definition and typological analysis, we were able to identify this information. The hidden data can be modeled explicitly by using formal definition. Providing a structured and formalized version of this information would enhance the usability and usefulness of the catalog for different daily and research applications. Another weakness of the APC was the ambiguous procedural type axis, which provided few classes with overlapping meanings. For example, the APC offered the procedural type classes "Therapy" and "Medical Treatment". This means, a medical procedure was either a therapeutic procedure or a medical treatment. Nevertheless, in common sense, these classes were overlapping, since a medical treatment is a sort of a therapy. With the presented approach, also a clear categorization of procedure types was identified.

For steps 2 and 3, we focused on an excerpt of the Austrian Procedure Catalogue – medical procedures that refer to the heart, large vessels and the cardiovascular system. Since it was not our intention to cover all codes of the APC but to come up with a systematic approach to formalize classification systems, we focused on codes related to a complex system such as the cardiovascular system. Nevertheless, given the nature of health care procedures contained in the APC, no changes are expected in the abstracted concepts as regards the coverage of further code descriptions. The participation of the medical personnel throughout the analysis tasks allowed us to evaluate the coverage of concepts, vocabulary as well as the taxonomy in an iterative manner.

Since the Austrian Procedure Catalogue is an essential part of the performance-oriented hospital financing system in Austria, it was crucial to include all the information including codes and coding guidelines provided by the APC in the proposed ontology. However, the need to preserve original information (e.g. procedure codes, axis codes), could limit the implementation of a proper semantic structure for codes within the developed ontology, as the APC procedure code does not contain all axes information (anatomical site, procedural type, access). The main axes of the

APC (anatomical site, procedural type, access) were confirmed by the typological analysis. We were able to extend the APC core of the ontology with important information, which can be used for different purposes, e.g. to enhance semantic interoperability for e-health applications or research purposes. Particularly the procedural type was further subdivided to provide more detailed and unambiguous information. Therefore, we adopted the procedural type structure of the French Classification Commune des Actes Médicaux. Other classifications, for example, the one provided by the SNOMED CT, could have been useful but also too detailed for a small classification system such as the APC. The CCAM turned out to be a very elaborate standard in the course of our comparative analysis.

Currently, we are conducting an assessment to determine the usefulness of the catalog with the collaboration of clinical personnel of the APC ontology. For this, we have planned two different assessment approaches: (1) semi-automated competency evaluation and (2) a task-based user evaluation. Additionally, a comparison of the APC ontology with SNOMED CT (structure of procedure hierarchy) is planned to be carried out in order to evaluate the proposed ontology. The assessment will lead to a refinement and curation of the ontology, constituting the fifth step of the proposed methodology. For future improvement, we proposed to extend the ontology with further codes referring to body structures other than the heart, large vessels and cardiovascular system. The ontology also provides a framework for extending the content of the Austrian Procedure Catalogue in order to cover more procedures.

## 6. Conclusion

We conclude that the proposed four-step approach for formalizing the Austrian Procedure Catalogue provides a systematically developed and strong framework to semantically enrich procedure coding classifications. Although this methodology was designed to address the particularities of the APC, the included methods are based on generic analysis tasks. Therefore our methodology can serve as a reference for the formalization of other classification systems. Additional steps or iterations can be incorporated if the complexity of the classification system requires it and hence contribute towards a universal alignment of such representation if desired.

## Acknowledgment

This work was accompanied by and established with the help of Dr. Wegene Borena, Ph.D. who provided us with her medical expertise.

## References


[1] Austrian Federal Ministry of Health, The Austrian DRG system <http://www.bmg.gv.at/cms/home/attachments/9/4/3/CH1015/CMS1292515593745/austrian_drg_bmg_2010_end_eng_revised.pdf> (accessed: 11-Mar-2015).

[2] F. Breyer, P. Zweifel, M. Kifmann, Gesundheitsökonomik, Springer, Auflage: 5. Berlin u.a., 2004.

[3] M.M. Hofmarcher, H.M. Rack, Gesundheitssysteme im Wandel Österreich Gesundheitssysteme im Wandel, Polit. Sci. (2006).

[4] World Health Organization, International Classification of Diseases (ICD) <http://www.who.int/classifications/icd/en/> (accessed: 11-Mar-2015).

[5] S. Hanser, A. Zaiss, S. Schulz, Health care procedures, Meth. Inf. Med. 48 (6) (2009) 540–545.

[6] Austrian Federal Ministry of Health, Leistungskatalog BMG 2014 <http://www.bmg.gv.at/cms/home/attachments/8/6/4/CH1166/CMS1128332460003/leistungskatalog_bmg_2014.pdf> (accessed: 16-Mar-2015).

[7] Austrian Federal Ministry of Health, Leistungskatalog BMG 2014 – Tabelle (gesamt) <http://www.bmg.gv.at/cms/home/attachments/8/6/4/CH1166/CMS1128332460003/leistungskatalog_bmg_2014ges.xls> (accessed: 16-Mar-2015).

[8] J.J. Cimino, Desiderata for controlled medical vocabularies in the twenty-first century, Meth. Inf. Med. 37 (4–5) (1998) 394–403.

[9] A. Burgun, Desiderata for domain reference ontologies in biomedicine, J. Biomed. Inform. 39 (3) (2006) 307–313.

[10] T. Tudorache, S. Falconer, N.F. Noy, C. Nyulas, T.B. Üstün, M.A. Storey, M.a. Musen, Ontology development for the masses: creating ICD-11 in WebProtégé, in: Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), LNAI, vol. 6317, 2010, pp. 74–89, 2010.

[11] M. Fernández-López, a. Gómez-Pérez, N. Juristo, Methontology: from ontological art towards ontological engineering, Assessment SS-97-06 (1997) 33–40.

[12] DIMDI, OPS Operationen- und Prozedurenschlüssel <https://www.dimdi.de/static/de/klassi/ops/index.htm> (accessed: 11-Mar-2015).

[13] American Medical Association, CPT – Current Procedural Terminology <http://www.ama-assn.org/ama/pub/physician-resources/solutions-managing-your-practice/coding-billing-insurance/cpt.page?> (accessed: 29-Aug-2015).

[14] ATIH, CCAM – Classification des Actes Médicaux <http://www.ccam.sante.fr/> (accessed: 11-Mar-2015).

[15] N. Lasierra, A. Alesanco, S. Guillén, J. García, A three stage ontology-driven solution to provide personalized care to chronic patients at home, J. Biomed. Inform. 46 (3) (2013) 516–529.

[16] World Health Organization, International Classification of Health Interventions (ICHI) <http://www.who.int/classifications/ichi/en/> (accessed: 11-Mar-2015).

[17] Word Health Organization, International Classification of Health Interventions (ICHI), Alpha Version (Trial Version), 2002.

[18] World Health Organization, International Classification of Health Interventions (ICHI).

[19] IHTSDO, SNOMED CT – The Global Language of Healthcare <http://www.ihtsdo.org/snomed-ct> (accessed: 11-Mar-2015).

[20] J. Ingenerf, Die Referenzterminologie SNOMED CT: von theoretischen Betrachtungen bis zur praktischen Implementierung, MMI, Neu-Isenburg, 2007.

[21] S. Schulz, B. Suntisrivaraporn, F. Baader, M. Boeker, SNOMED reaching its adolescence. Ontologists' and logicians' health check, Int. J. Med. Inform. 78 (April) (2009) S86–S94.

[22] J.R. Campbell, J. Xu, K.W. Fung, Can SNOMED CT fulfill the vision of a compositional terminology? Analyzing the use case for problem list, in: AMIA Annu. Symp. Proc., vol. 2011, 2011, pp. 181–188.

[23] A. Agrawal, Y. Perl, Y. Chen, G. Elhanan, M. Liu, Identifying inconsistencies in SNOMED CT problem lists using structural indicators, in: AMIA Annu. Symp. Proc., vol. 2013, January 2013, pp. 17–26.

[24] K. Dentler, R. Cornet, Intra-axiom redundancies in SNOMED CT, Artif. Intell. Med. (November) (2014).

[25] R. Studer, V.R. Benjamins, D. Fensel, Knowledge engineering: principles and methods, Data Knowl. Eng. 25 (1998) 161–197.

[26] P. Gómez, Ontological Engineering – With Examples From the Areas of Knowledge | Springer, first ed., Springer-Verlag, London, 2004.

[27] E. Simperl, M. Mochol, T. Bürger, I. O. Popov, Achieving maturity: the state of practice in ontology engineering in 2009, in: Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), LNCS, vol. 5871(1), 2009, pp. 983–991.

[28] H.S. Pinto, S. Staab, C. Tempich, DILIGENT: towards a fine-grained methodology for DIstributed, Loosely-Controlled and evolvInG Engineering of oNTologies, in: 16Th Eur. Conf. Artif. Intell. – Ecai, 2004, pp. 393–397.

[29] S. Pinto, S. Staab, Y. Sure, C. Tempich, OntoEdit empowering SWAP: a case study in supporting DIstributed, Loosely-Controlled and evolvInG Engineering of oNTologies (DILIGENT), in: C.J. Bussler, J. Davies, D. Fensel, R. Studer (Eds.), The Semantic Web: Research and Applications, vol. 3053, Springer, Berlin Heidelberg, 2004, pp. 16–30.

[30] A. De Nicola, M. Missikoff, R. Navigli, A software engineering approach to ontology building, Inf. Syst. 34 (2) (2009) 258–275.

[31] N. Noy, D. McGuinness, Ontology development 101: a guide to creating your first ontology, Development 32 (2001) 1–25.

[32] A.C. Yu, Methods in biomedical ontology, J. Biomed. Inform. 39 (3) (2006) 252–266.

[33] N. Lasierra, A. Kushniruk, A. Alesanco, E. Borycki, J. García, A methodological approach for designing a usable ontology-based GUI in healthcare, Stud. Health Technol. Inform. 192 (2013) 1040.

[34] C.E. Kuziemsky, F. Lau, A four stage approach for ontology-based health information system design, Artif. Intell. Med. 50 (3) (2010) 133–148.

[35] K. Siorpaes, M. Hepp, OntoGame: towards overcoming the incentive bottleneck in ontology building, Engineering 4806 (2007) 1222–1232.

[36] K. Siorpaes, OntoGame: Games with a purpose for the Semantic Web, in: CEUR Workshop Proc., vol. 358(3), 2008, pp. 66–70.

[37] K. Siorpaes, M. Hepp, OntoGame: weaving the semantic web by online games, in: Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), LNCS, vol. 5021, 2008, pp. 751–766.

[38] P. Mayring, Einführung in die qualitative Sozialforschung, 5. Aufl. Weinheim: Beltz, 2002.

[39] W. Pschyrembel, Pschyrembel Klinisches Wörterbuch, De Gruyter, München, 2013. 265., neu.

[40] W3C, OWL 2 Web Ontology Language Document Overview, second ed. <http://www.w3.org/TR/owl2-overview/> (accessed: 16-Mar-2015).

[41] W3C, RDF – Semantic Web Standards <http://www.w3.org/RDF/> (accessed: 16-Mar-2015).

[42] W3C, RDF Schema 1.1 <http://www.w3.org/TR/rdf-schema/> (accessed: 16-Mar-2015).

[43] Stanford University, protégé – A free, open-source ontology editor and framework for building intelligent systems <http://protege.stanford.edu/> (accessed: 16-Mar-2015).
[44] S.B. Neururer, K.-P. Pfeiffer, A systematic comparison of the Austrian Procedure Catalogue and the classification commune des Actes Médicaux, in: 23rd International Conference of the European Federation of Medical Informatics (MIE 2011), 2011.
[45] S.B. Neururer, K.-P. Pfeiffer, Characteristics of health interventions: a systematic analysis of the Austrian Procedure Catalogue, Stud. Health Technol. Inform. 180 (Jan. 2012) 1090–1092.
[46] S. Neururer, W. Borena, K. Pfeiffer, Vergleich des österreichischen Leistungskatalogs mit der International Classification of Health Interventions (ICHI), in: GMDS-Jahrestagung, 2010.
[47] C.E. Lipscomb, Medical subject headings (MeSH), Bull. Med. Libr. Assoc. 88 (3) (2000) 265–266.
[48] Österreichisches Register für Medizinprodukte, Österreichisches Register für Medizinprodukte, 2002 <www.medizinprodukteregister.at/> (accessed: 16-Mar-2015).