Mapping a Nursing Terminology Subset to openEHR Archetypes*

A Case Study of the International Classification for Nursing Practice

J. R. M. Nogueira1; T. W. Cook2; L. T. Cavalini3

1Department of Health Information Technology, Medical Sciences College, Rio de Janeiro State University, Rio de Janeiro, Brazil;
2National Institute of Science and Technology – Medicine Assisted by Scientific Computing, Petrópolis, RJ, Brazil;
3Department of Epidemiology and Biostatistics, Fluminense Federal University, Niterói, RJ, Brazil

Keywords
Information systems, electronic health records, nursing informatics, terminology

Summary
Background: Healthcare information technologies have the potential to transform nursing care. However, healthcare information systems based on conventional software architecture are not semantically interoperable and have high maintenance costs. Health informatics standards, such as controlled terminologies, have been proposed to improve healthcare information systems, but their implementation in conventional software has not been enough to overcome the current challenge. Such obstacles could be removed by adopting a multilevel model-driven approach, such as the openEHR specifications, in nursing information systems.

Objectives: To create an openEHR archetype model for the Functional Status concepts as published in Nursing Outcome Indicators Catalog of the International Classification for Nursing Practice (NOIC-ICNP).

Methods: Four methodological steps were followed: 1) extraction of terms from the NOIC-ICNP terminology; 2) identification of previously published openEHR archetypes; 3) assessment of the adequacy of those openEHR archetypes to represent the terms; and 4) development of new openEHR archetypes when required.

Results: The “Barthel Index” archetype was retrieved and mapped to the 68 NOIC-ICNP Functional Status terms. There were 19 exact matches between a term and the correspondent archetype node and 23 archetype nodes that matched to one or more NOIC-ICNP. No matches were found between the archetype and 14 of the NOIC-ICNP terms, and nine archetype nodes did not match any of the NOIC-ICNP terms.

Conclusions: The openEHR model was sufficient to represent the semantics of the Functional Status concept according to the NOIC-ICNP, but there were differences in data granularity between the terminology and the archetype, thus producing a significantly complex mapping, which could be difficult to implement in real healthcare information systems. However, despite the technological complexity, the present study demonstrated the feasibility of mapping nursing terminologies to openEHR archetypes, which emphasizes the importance of adopting the multilevel model-driven approach for the achievement of semantic interoperability between healthcare information systems.

1. Background

The Functional Status assessment is a common activity in nursing practice and other healthcare professions. Through the categorization of Activities of Daily Living (ADLs), healthcare providers can identify the maximum level of independence of subjects with different levels of disability [1]. There are applications being developed to specifically support nursing professionals in the assessment and care of disabled patients, including the concepts of Functional Status; however, such initiatives are incipient [2–4]. The data modelling approach commonly adopted by these applications and conventional Electronic Health Records turn them into “data silos” [5]. In other words, they are self-contained applications that are unable to communicate the full meaning of concepts with an Electronic Health Record of the patient, in a more general way. Thus, a challenge with conventional healthcare applications to this date is semantic interoperability, which can be defined as the ability to exchange and validate the semantics of data extracts among distributed, independently developed applications.

There are global efforts towards the development of healthcare terminologies in general, and nursing terminologies in particular. These vocabularies should allow data sharing among healthcare information systems [6]. Some of these initiatives include: the Clinical Care Classification (CCC) System, created by the American Nurses Association during the 1980s, consisting of two terminology sets for diagnosis and outcomes [7]; the North Ameri-
can Nursing Diagnosis Association (NANDA) terminology, created in 1982, based on nine response patterns to disease [8] and the International Classification for Nursing Practice (ICNP), developed by the initiative of the International Council of Nurses (ICN) since 1990 [9]. The knowledge representation of Functional Status, mostly used for the record of the Activities of Daily Living of disabled patients, is published as part of the ICNP Nursing Outcome Indicators Catalog (NOIC-ICNP) [10].

The International Standards Organization’s (ISO) International Standard 18104 (ISO 18104:2014 – Health informatics – Categorical structures for representation of nursing diagnoses and nursing actions in terminological systems) is an important tool to guide the integration of nursing terminologies to computerized information systems. This reference was previously used to evaluate the mapping capabilities of three different nursing terminologies, and the results showed that the majority of their terms are interchangeable, which is promising to the harmonization efforts [11].

Despite the recent efforts on harmonizing nursing and general healthcare terminologies, such as the equivalence table between ICNP and the Systematized Nomenclature of Medicine – Clinical Terms (SNOMED-CT)[12], there remains many nursing terminologies being produced and maintained in an isolated manner. This brings many challenges to mapping [13] them to other controlled vocabularies in healthcare. In order to overcome this polysemny in healthcare informatics, the creation of a common language, or an universal model, in which all nursing terminologies could be articulated, has been suggested by some experts [6] and developed by some projects such as the Clinical Information Modeling Initiative (CIMI)[14].

However, there are remaining challenges to harmonize interface terminologies (that represent healthcare practice at the point of care) to controlled terminologies. These challenges occur out of the consensus approach used in defining controlled vocabularies and further to the harmonization process, since consensus in healthcare is difficult to achieve [15].

Given the scenario described above, there is a perceived need for the achievement of a semantically and syntactically unified representation of healthcare terminologies. This could allow independent development of systems and still maintain the semantic context for information exchange between healthcare applications. Thus, it is important to improve the analytical methods of nursing terminology mapping to structured, detailed clinical models that provide system interoperability at the semantic and syntactic level, such as the openEHR archetype model, an open specification based on a multilevel (or dual) model, in which the clinical knowledge is represented by constraints on a common and generic Reference Model [16]. The term binding capabilities of openEHR are well defined in the specifications but they have not been widely implemented. This study has the goal to add to the existing knowledge in nursing terminology mapping by proposing a new approach on the representation of term sets in structured models such as openEHR archetypes.

2. Objectives

The main objective of this study was to model the concepts related to the Functional Status as represented in the NOIC-ICNP according to the openEHR specifications.

3. Methods

The methodological approach adopted for this study comprised the following steps: extraction of relevant terms from the NOIC-ICNP terminology; identification of relevant openEHR archetypes already published; assessment of the adequacy of the selected openEHR archetypes for representing terms; development of new openEHR archetypes for terms not found in the existing ones.

3.1 Extraction of NOIC-ICNP Terms

The NOIC-ICNP terms related to the Functional Status were selected to guide the knowledge representation model in this study. This choice was determined based on the desire to create a well-defined healthcare concept for common use in nursing practice. Another reason to choose the NOIC-ICNP is because its term set is the most comprehensive for the knowledge representation of the concept of the functional status among the open and freely available nursing and general terminologies. For example, NOIC-ICNP term subset for the Functional status has 68 terms, whereas the second biggest term set analyzed, from the Logical Observation Identifiers, Names and Codes (LOINC) has 53 terms.

The correspondence between the NOIC-ICNP and the openEHR archetypes studied was performed by a certified nurse with formal training on the use of the openEHR specifications for knowledge modeling and a physician with expertise in health informatics and multilevel modeling technologies, under the supervision of a former member of the openEHR Architecture Review Board. The NOIC-ICNP Tables 1–4 (Section 1)[10] comprising the concepts of Functional Status related to acute care (assess for performance over full 24 hours), complex continuing care and long term care (assess for performance over all shifts during last seven days), and home care domain (assess for resident’s performance during last seven days), were the primary source of the term mapping.

3.2 Identification of Existing openEHR Archetypes

The openEHR specifications were chosen for being the original proposal for multilevel modeling which have been implemented in a verifiable manner in the scientific literature [17], and because the specifications and modeling tools are open and freely available for use. The openEHR archetypes are defined as the maximum data set for a particular healthcare concept [18], which means that all data elements that are part of a healthcare concept should be represented in the correspondent archetype [16, 19].

The titles of the NOIC-ICNP Tables 1 to 4 were used as keywords, which were used for browsing the Clinical Knowledge Manager (CKM), the global and official repository of validated openEHR archetypes. The
term “ADL”, used in this context as the acronym for “Activities of Daily Living”, was entered on the CKM browser, and the Barthel Index archetype was retrieved and downloaded. The Barthel Index is an ordinal scale used to evaluate performance in activities of daily living. Each item is rated with a number of points, and the final index for a given patient is the sum of points scored for each item [20]. The archetype in its original form as retrieved from the CKM did not specify any terminology bindings to controlled vocabularies such as LOINC, SNOMED-CT or the ICNP.

3.3 Analysis of the Mapping between NOIC-ICNP Terms and the Archetype

The NOIC-ICNP terms were matched one by one to the corresponding archetype nodes of the Barthel Index archetype, trying to find the best similarity possible between them. Thus, when the archetype nodes matched the NOIC-ICNP terms in structure, granularity, data types and conceptual model, the match was regarded as “exact”. When there was no direct correspondence between the archetype nodes and the NOIC-ICNP terms, or when there was not complete semantic correspondence between them, the best match was chosen and the necessary adjustments were reported.

For instance, the archetype could be more specific than the NOIC-ICNP terms: as an example, for the term “Impaired ability to walk” (code 10001046), there were correlation with three archetype nodes, and one of them was chosen as the best match: the internal terminology “Walks with help” [at0020] of the DV_TEXT “Mobility” [at0017]. On the other hand, the NOIC-ICNP terms could be more specific than the archetype, such as the term “Positive ability to walk” (code 10028333), with the best match being the internal terminology “Independent” [at0039] of the DV_TEXT “Mobility” [at0017].

3.4 Modeling of New openEHR Archetypes

Where knowledge modeling required the production of a new archetype, if a certain term set could not be found in the CKM, it was performed with Ocean’ Archetype Editor’ version 2.2.779 Beta, which generates the archetype in Archetype Definition Language version 1.4 [19].

When a new archetype had to be modeled, the root class was set to OBSERVATION, since the information on the Functional Status is recorded from the patient observation performed by a healthcare provider during an encounter. This choice is in accordance to the openEHR Archetype Object Model specification.

The ‘data’ attribute of the OBSERVATION class has its value constrained by default to a HISTORY class instance, for which the ‘events’ attribute had the cardinality set from one to unbounded, being its value constrained to an EVENT class instance, defined as any type (Any event), allowing the record of periodic or occasional events. The ‘data’ attribute value of the EVENT class was constrained to an ITEM_TREE class instance, indicating that the IADL data elements were arranged as a hierarchical tree of items. For the ‘items’ attribute of the ITEM_TREE class, the cardinality was set from zero to unbounded, meaning that no data item of this archetype is mandatory, but any combination thereof, of any size, is allowed.

4. Results

The openEHR CKM search retrieved the “Barthel Index” archetype, composed of 11 ELEMENT, 43 archetype nodes and 30 values of internal terminologies. This archetype was then compared with 68 NOIC-ICNP terms related to the Functional Status, with 19 exact matches between a term and the correspondent archetype. A total of 23 archetype nodes were matched to the NOIC-ICNP terms. In all cases, the archetype nodes were mapped to more than one NOIC-ICNP term, since the terms were more specific than the archetype nodes, as shown on the Methods section and on Tables 1–3 in the supplementary online material. There were no matches in the archetype for 14 of the NOIC-ICNP terms, and nine archetype nodes did not match any of the NOIC-ICNP terms. The mapping between the NOIC-ICNP terms and the correspondent archetype nodes is presented on Tables 1–3 in the supplementary online material. It is important to note that some of the nodes of the “Barthel Index” archetype did not match to any NOIC-ICNP term (Table 4 in the supplementary online material).

In the chapter of the Functional Status for acute care, complex continuing care, long term care and home care corresponded to the same archetype node. For instance, for the concept “Walking” ( categorized as: in general; in room or in corridor; Locomotion on unit or off unit; in home or outside of home), the terms “Positive ability to walk” (code 10028333) corresponded to the internal code “Independent” (archetype node = [at0039]) of the ELEMENT “Mobility” [at0017], while the term “Impaired ability to walk” (code 10001046) corresponded to three internal codes of the ELEMENT “Mobility” [at0017]: Immobile or < 50 meters [at0018], “Wheel chair independent” [at0019] and “Walks with help” [at0020]. Similar situations occurred for the NOIC-ICNP terms related to Transfer (in general or to the toilet), Toilet use, Bed mobility, Eating, and Dressing (in general; upper body or lower body). This situation demonstrates that there are differences in the granularity of the information between the archetype and the terminology. The only solution that preserves the archetype structure and still allows matching between the archetype nodes and the NOIC-ICNP terms, minimizing the semantic loss, is the definition of an openEHR template to further semantic modeling, associating the archetype node with the most equivalent meaning to the NOIC-ICNP term, as presented on Tables 1–3 in the supplementary online material for the examples above.

The NOIC-ICNP terms related to the Instrumental Activities of Daily Living (IADL) did not match any of the “Barthel Index” archetype nodes (Table 5 in the supplementary online material), since the Barthel Index does not evaluate them. Thus, it was necessary to model a new archetype, “Instrumental Activities of Daily Living”, for the direct mapping of the corresponding terms. The code in Archetype Definition Language of the resulting archa-
type is available in the supplementary online material.

The tree of items of the "IADL" archetype consists of the following constraint definitions of seven instances of the ELEMENT class: meal preparation [at0004]; ordinary housework [at0007]; managing finances [at0010]; managing medications [at0013]; phone use [at0016]; shopping [at0019]; and transportation [at0022]. The 'occurrences' attribute values of these ELEMENT class instances were set as 0..1, which means that the record of these data elements is optional. The 'value' attribute of these ELEMENT class instances was constrained to the DV_CODED_TEXT data type class whose values were associated to a local terminology, according to Table 5 (supplementary online material).

5. Discussion

This study presented the process of modeling the terms related to the concept of Functional Status as represented on the NOIC-ICNP as openEHR archetypes. The search for a compatible archetype resulted in the use of the "Barthel Index" archetype for the mapping, which needed to be supplemented by modeling a new archetype for IADL concepts.

It is important to note that the ideal modeling solution in this case would be linking each NOIC-ICNP term related from the IADL to the correspondent archetype node; however, the archetype editor adopted for this study did not include the ICNP in its terminology set. Although the openEHR specifications can handle the binding, since the binding process itself is covered in great detail in the specifications, the available archetype modeling tools do not cover all the possibilities allowed by the openEHR model. The alternative to this date is to use a plain text editor to directly edit the Archetype Definition Language file, which requires a deep knowledge of the openEHR specifications, which is difficult to achieve by healthcare providers, since there is no formal education available on the subject.

There have been a number of efforts to map between medical and nursing terminologies, usually producing matches between two or more reference terminologies. It is common to find high levels of matching between any given pair of terminologies. This suggests a superposition of concepts between them and a significant part of the discordances are related to lexical and semantic variability. Previous studies usually adopted a systematic mapping process, with two independent reviewers producing the matching results and the estimation of inter-reliability scores [22, 23]. This study, however, chose a consensus approach to perform the matching, following the qualitative methodologies adopted in similar studies [24, 25].

In order to better understand the technical challenges of the mapping proposed in this study, it is important to note that all terms defined in an archetype are pre-coordinated by definition. Thus, universal models such as archetypes can be a barrier to the adaptation of the knowledge representation of the dynamics of the nursing practice, at the point of care due to the top-down nature of archetypes. Although ICNP provides the capability of post-coordination, and taking into account that pre-coordination and post-coordination are complimentary processes, the constraint model is implemented to define a range or meaning for a node in the archetype. Whether the modeler uses pre- or post-coordinated terms from a vocabulary is a choice to be made based on the needs of the concept represented by the model. In this case study both the pre-coordination and post-coordination capabilities of the openEHR archetypes were exercised, although, according to the openEHR specifications, archetypes are not originally fitted for post-coordination, since they are representations of a maximal data set for any given healthcare concept, which give them a pre-coordinated nature by default.

An important finding of this mapping is related to the difference between the granularity of the NOIC-ICNP terms and the "Barthel Index" archetype nodes, and the need to adopt the same ELEMENT for the representation of various NOIC-ICNP terms. When implementing an application for real use, in order to accommodate this difference of granularity and perform the differentiation between these terms, the solution proposed in the openEHR specifications is the modeling of an openEHR template [26]. In order to be semantically validated, a data instance generated through this hypothetical application would have to be represented not only by the semantics contained in the archetype, but also the semantic differentiation made at the template level. Thus, much of the semantics of the terminology would be restricted to the template level, which creates an additional layer of complexity in the implementation of information systems based on the openEHR specifications.

The results of this study suggest that, for this specific case in nursing informatics, modeling a healthcare concept in full compliance to the openEHR specifications can become a very complex task, which is corroborated by other authors [27, 28]. However, the mapping goal of this study was achieved, since the semantic mapping according to the openEHR specifications was sufficient to represent the semantics of the healthcare concepts represented in NOIC-ICNP, both in terms of structure and data types, allowing connections between ELEMENT instances and this terminology subset, as proposed by previous studies [29].

Furthermore, some authors have suggested that the level of success in mapping healthcare terminologies to openEHR archetypes is specific to each terminology, since there are reports of modeling deficiencies between archetypes and the SNOMED-CT terminology. On one hand, the lack of consensus on the structure and organization of data during design leads to overlapping constructs and inconsistencies in the data representation. On the other hand, the clinical modeling constructs made separately from the SNOMED-CT leads to differences between the archetypes and that terminology, such as lexical differences in terms of names, semantic variability or as seen above, a disparity in the granularity of documentation [30].

Despite the technical difficulties, there is an increasing consensus on the superiority of multilevel model-driven approaches such as the openEHR specifications in playing the role of the desired "common language" or the "universal model" that unifies nursing (in particular) and healthcare (in general) terminologies. Thus, it should be put in the supplementary online material.
possible to develop purpose-specific but semantically interoperable, healthcare applications for the record of information and the computerized aid to the continuing care of the elderly, in which the Functional State is a core concept. The present study adds evidence to support these efforts, since the proposed mapping was well-succeeded, although technically difficult to implement.

Future developments of the present study include the development and application of the principles of Evidence-Based Nursing (especially the systematic review and meta-analysis) to produce a set of validated data collection instruments to guide the modeling of openEHR archetypes and their mapping to reference terminologies. Other knowledge representation models, such as the Health Level 7 (HL7) standards and other controlled vocabularies, such as the Logical Observations, Names and Codes (LOINC) terminology, as well as different functional status assessment, such as the Lawton Personal Self-Maintenance Scale, can be implemented on prototype applications, with specific tests of data exchange, in order to produce further scientific evidence of the capabilities of health-care informatics standards to provide semantic interoperability for healthcare information systems.

6. Conclusion
The openEHR specifications were regarded as sufficient for modeling concepts of the NOIC-ICNP mapped to the archetype Barthel Index. However, the mapping process proved to be extremely complex and it is possibly difficult to implement in a real situation of software development for deployment in healthcare settings. Because an archetype should represent the “maximum data model” defined for a given clinical concept and in order to reach the required representation of the Functional Status, an existing archetype was used, and a new archetype was modeled only for the data that had no correspondence with any archetype on the CKM.

The potential contribution of multilevel modeling to improve the quality of healthcare is significant, and it is expected that real implementations will be realized in the near future, since the technology required is already available today. Thus, knowledge representation models of a more stable and well-defined structure could be implemented in prototype applications. Design of specific tests for semantic interoperability is required in order to produce scientific documentation that multi-level modeling provides a foundation for automated clinical decision support in healthcare.

Acknowledgments
This paper was partially supported by the National Institute of Science and Technology – Medicine Assisted by Scientific Computing (grant no. E-26/170.030/2008).

References