Journal of Biomedical Informatics 43 (2010) 1036–1043

Contents lists available at ScienceDirect

Journal of Biomedical Informatics

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

A quality improvement model for healthcare terminologies

Tae Youn Kim [a,*], Amy Coenen [a], Nicholas Hardiker [b]

[a] College of Nursing, University of Wisconsin-Milwaukee, 1921 E. Hartford Avenue, P.O. Box 413, Milwaukee, WI 53201, USA
[b] School of Nursing & Midwifery, University of Salford, Mary Seacole Building, Greater Manchester M6 6PU, UK

ARTICLE INFO

Article history:
Received 18 June 2010
Available online 17 August 2010

Keywords:
Controlled Healthcare Terminology
Terminology quality improvement
International Classification for Nursing Practice
Healthcare informatics

ABSTRACT

A number of controlled healthcare terminologies and classification systems have been developed for specific purposes, resulting in variations in content, structure, process management, and quality. A terminology quality improvement (TQI) model or framework would be useful for various stakeholders to guide terminology selection, to assess the quality of healthcare terminologies and to make improvements according to an agreed standard. A TQI model, thus, was formulated based on a review of the literature and existing international standards developed for healthcare terminologies. The TQI model, adapted from Donabedian's approach, encompasses structure, process, and outcome components in relation to a terminology life cycle – change request, editing, and publication. Multi-dimensional quality outcome measures also were identified in the areas of terminology content, modeling structure, mapping, and process management. A case study was developed to validate the TQI model using the International Classification for Nursing Practice (ICNP). The TQI model represented the complexity of activities involved in terminology quality management. The ICNP case study demonstrated both the applicability of the TQI model and the appropriateness of the criteria identified in the TQI model: openness and responsiveness, clarity and reproducibility, understandability, accessibility and usability, interoperability, and quality of documentation. The applicability of the TQI model was validated using ICNP. While ICNP exhibits many of the desirable characteristics of contemporary terminologies, the case study identified a need for further work on ICNP policy and on documentation.

© 2010 Elsevier Inc. All rights reserved.

1. Introduction

A number of classification systems and controlled terminologies have been developed and used in healthcare over the past several decades. A terminology is a structured organization of representative concepts and their relationships, mappings, descriptions, and translations [1]. The design and scale of terminologies vary as each terminology has its own purpose, target domain, and intended users. The most popular approach adopted to construct terminologies is to organize given concepts using a tree structure with subsumption relations according to its respective domain knowledge [2]. Recently there has been more interest within the biomedical informatics community in developing ontologies using formal languages that are interpretable by both humans and computers (e.g., Web Ontology Language or OWL). Although the definition of ontology differs according to a given context, here an ontology is considered a form of knowledge representation where the nature of a concept is formally defined through properties and relations with other concepts [3]. The strengths of an ontology is not only in organizing domain specific knowledge in a coherent manner but also in enhancing the reasoning power of a machine to support decision-making of various stakeholders.

Currently, more than one hundred terminologies and classification systems are available within the Unified Medical Language System (UMLS), a resource for integrating and distributing healthcare terminologies [4]. Also, more than 170 ontologies that have been developed in biomedical communities are accessible through BioPortal, which is maintained by the National Center for Biomedical Ontology [5]. Although there are overlapping terminologies or ontologies across the two terminology resources, the numbers indicate how rapid this field of practice and research is developing.

It is widely accepted that terminologies and ontologies can promote clear and consistent communication among healthcare providers and can extend utility of healthcare data beyond the clinical encounter. Further, the introduction of electronic health records has brought attention to the significance of adopting standardized controlled vocabularies in order to enhance interoperability among health care systems. Many existing terminologies, thus, strive to keep pace with developments. For example, approximately half of the terminologies integrated into the UMLS have been updated more than one time since their initial release [4]. An updated version of the Systematized Nomenclature of Medicine Clinical Terms (SNOMED CT) accompanies each 6-monthly release of the UMLS [4,6].

* Corresponding author. Fax: +1 414 229 6474.
E-mail address: kimty@uwm.edu (T.Y. Kim).

1532-0464/$ - see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.jbi.2010.08.006

Extensive work among terminology experts, researchers, and engineers has been conducted during the last two decades [7–12], resulting in some notion of best practice for terminology development and management. Guidance published through such collaborative work became a basis for establishing international standards developed under the auspices of the International Organization for Standardization (ISO). For example, the technical specifications of ISO TS 17117:2002, *Health Informatics – Controlled Health Terminology – Structure and High-Level Indicators*, describe the principles that are necessary and sufficient to evaluate a controlled health terminology at a high level [13]. Similarly, ISO TS 18104:2003, *Health Informatics – Integration of a Reference Terminology Model for Nursing*, identifies mandatory high-level categories and their semantic relations necessary to communicate nursing diagnoses and interventions across healthcare settings [14], based on terminologies and classifications that are used in nursing practice [15].

Although national and international efforts have established some quality criteria, these present only a limited view of healthcare terminology management. Given the fact that each terminology has unique maintenance policies and procedures, it should be noted that the quality of a terminology is affected by a terminology life cycle, consisting of multiple steps. This requires continuous inspection using robust criteria. While a range of quality assurance methods from fully automated error detection to manual reviews have been applied to healthcare terminologies, few frameworks or models to guide best practices for terminology development and maintenance have been discussed in relation to a terminology life cycle.

The purpose of this paper is to introduce a terminology quality improvement model formulated through a synthesis of the literature and validated using a case study with the International Classification for Nursing Practice (ICNP). Discussion about the applicability of the model and concluding remarks also are presented. It is anticipated that this model or framework will be useful for various stakeholders to guide terminology selection, to assess the quality of healthcare terminologies and to make improvements according to an agreed standard.

2. Framework for terminology quality improvement

The quality improvement (QI) of health care continues to be a national and international agenda across care settings, requiring a QI framework to guide various stakeholders to improve the quality of care provided in a systematic way. Donabedian's framework [16] consisting of structure, process, and outcome components is a representative model applied in numerous settings across countries. That is, healthcare outcomes should be understood in relation to structural conditions within which health care was provided during the process of delivering care [16]. Its comprehensive view on quality care has influenced the way various stakeholders evaluate healthcare outcomes and conduct health service related research.

In light of this, Donabedian's approach could be applicable to measure the quality of terminologies as their development and maintenance involve various structures, processes, and outcomes that affect their quality. The International Health Terminology Standards Development Organization (IHTSDO) is an example of an organization that has integrated a QI framework into their terminology management activities [17,18]. According to the IHTSDO framework, the human and technical resources necessary to develop and maintain SNOMED CT might be considered structure in Donabedian's approach, while a series of actions conforming to terminology requirements that satisfy the needs of internal and external stakeholders might be considered process. Quality assessment is conducted according to detailed characteristics; functionality, reliability, usability, efficiency, maintainability, and portability might be considered outcomes or outcome measures (these are specified by ISO/IEC 9126 as *software quality characteristics* [19]). Although the quality assurance framework of IHTSDO is still evolving, it broadens the scope of evaluation measures with respect to terminology quality as compared to traditional perspectives on terminology evaluation.

Similarly, the National Cancer Institute (NCI) established a set of review criteria in the areas of terminology content, structure, editorial process, and documentation for NCI Thesaurus (NCIt) [20,21]. These recent publications also expanded the view of existing terminology requirements or criteria from terminology developers to terminology users. In other words, understandability, reproducibility, usability, accessibility, community acceptance, and reporting requirements were considered terminology outcome measures in relation to terminology technical specifications such as concept orientation, polyhierarchy, and concept permanence [20,21].

Although the literature does not explicitly display a standard set of QI methods applicable during the course of development and maintenance across healthcare terminologies, recognition of a terminology or ontology as a fundamental resource of domain knowledge demands quality criteria. A mechanism established to ensure terminology quality, thus, needs to be transparent as a terminology or ontology evolves according to its developmental stage and end-users' needs. The lack of definitions of quality and of a quality improvement framework, however, provides a challenge not only to terminology developers in assuring terminology quality but to terminology users and researchers in evaluating a given terminology [22]. Also, variations in methods of auditing a terminology present potential for further research and efforts to promote best practices.

3. Model formulation and description

We formulated a terminology quality improvement (TQI) model based on a review of the literature and existing international standards for healthcare terminologies. As terminology quality is not guaranteed without proper structure and process in place, a TQI model, adapted from Donabedian's healthcare quality improvement framework, consists of *structure*, *process*, and *outcome* components. We define terminology quality improvement as a formal mechanism through which a terminology satisfies multi-dimensional requirements established through organizational structure and work processes. Quality controls should not be considered a stand-alone method; rather they should be understood within a terminology life cycle to identify areas for improvement in terminology management and ensure terminology quality.

3.1. Structure

In this TQI model, *structure* refers to organizational resources necessary to support all administrative and technical activities involved in terminology development, maintenance, and distribution. While the nature of supporting organizational resources may vary, they are a necessary feature of terminology development and maintenance [12,17,18,20,21]. Such organizational resources, for example, include funding for terminology management efforts, human resources for terminology review, editing, and dissemination, as well as technical resources for terminology storage, editing, and access. Within the TQI model, internal policy and procedural documents describing detailed work processes, decision rules and technical specifications are also considered part of the *structure* component.

3.2. Process

Within the TQI model, *process* is defined as a series of activities identified throughout a terminology life cycle that support the evolution of a terminology. A terminology life cycle denotes the recurring stages of the life of a terminology. The TQI model represents the life cycle as three phases: change request, editing, and publication (Fig. 1). While terminology development and maintenance might include other activities, these three phases appear to be common to terminology initiatives [6,17,20,21,23–25]. Each phase of the life cycle comprises a set of processes through which a terminology is characterized, including its quality.

Terminology Change Request is a formal mechanism to submit any changes with respect to a given terminology [6,20,23]. Change requests are submitted through an established mechanism documented in organizational policy and procedures of a given terminology. Changes include concept addition, revision, and deletion, along with sound rationale. Generally, change requests are made as pertinent sciences advance, there is room for further improvement for a domain of interest, or various types of errors are found in a given terminology as a result of terminology application and auditing. Such inputs from any internal and external users of a terminology are taken into consideration for further review, validation and subsequent documentation within a terminology life cycle, requiring collaboration of terminology developers and users as well as administrative supporting personnel.

Once requested changes are accepted, *Terminology Editing* begins to activate a new concept, revise an existing concept or inactivate a concept according to a formal or informal concept change and version management guideline. Concomitant with rapidly changing healthcare environments with massive amount of data accumulation, healthcare terminologies keep evolving and require a systematic method of maintaining information generated from the evolutionary work. This requires formalized processes to capture change so that it can be recognized and incorporated into healthcare information systems for data retrieval and aggregation [26]. Accordingly, change specifications are an essential component of the maintenance process, along with editing tools, to support the "execution" of core activities necessary for terminology management such as concept submission, validation, documentation, and versioning [24,27].

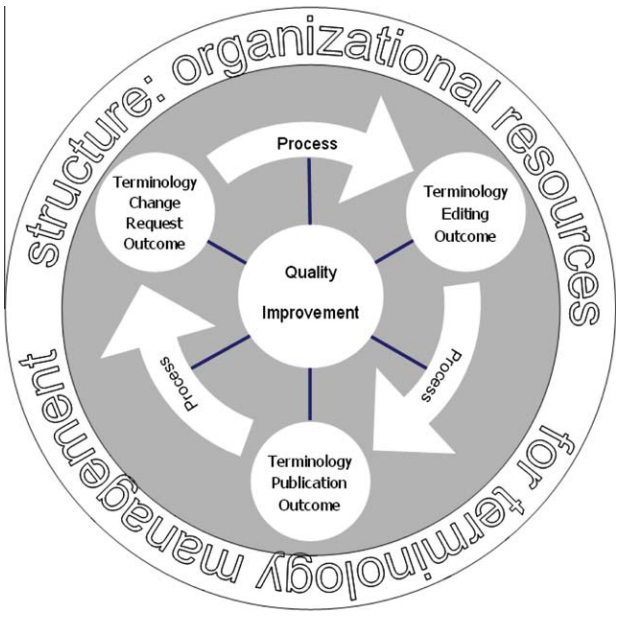

Fig. 1. Terminology quality improvement model applied to the terminology life cycle.

When editing is completed with subsequent documentation, a series of tasks are completed throughout the *Terminology Publication* phase to release a new version of terminology for public use. Some regular or ad hoc auditing necessary to maintain quality occurs during this phase of the terminology life cycle. In other words, a terminology goes through a number of inspections from fully automated auditing to human dependent validation [25]. Growing interest in logic-based terminology development in the biomedical informatics community has further established ontological principles for development and auditing of a terminology through semantic analyses. Such ontological principles address better ways of representing a concept that is semantically valid, complete, exclusive, univocal, and consistent within a hierarchical structure [28–30]. These principles have been applied in part to auditing various terminologies such as SNOMED CT [29,31], NCIt [32], Gene Ontology [33], and logic-based local terminologies [25,34].

In addition, semantic groups that were introduced to create clusters of concepts with similar characteristics by assigning categorical high-level concepts within a terminology prompted additional investigation of managing and auditing a terminology. Such methods called meta concept-based semantic analyses are especially useful for a large-scale terminology since they may reveal semantic inconsistencies among concepts with non-hierarchical relations [10,11]. The UMLS as a unified terminology resource in healthcare has been tested often to find inconsistent classifications of selected source concepts through semantic types and network [35–37].

Given the methodological principles for terminology auditing, some of the classical hierarchical structure of terminologies without formal definitions present challenges to employing explicit rules. Such terminologies can be audited using only implicit rules applied by human experts, which are more likely to be used on an ad hoc basis, hindering the consistent quality checking of terminologies. Once errors detected through terminology auditing are repaired as needed, a terminology goes through technical preparation of release to enhance the accessibility and usability in respective settings. Any documents related to a new release (e.g., technical specifications for implementation and user guidance) are also finalized during the publication phase [4,6].

3.3. Outcome

In the TQI model, *outcomes* denote any products (e.g., terminology, mappings, translations, subsets, etc.) generated through the terminology life cycle for public use, as well as a measure of terminology quality in terms of the degree of conformance to terminology requirements. Many researchers and terminology working groups since the early 1990s have addressed requirements or criteria necessary to maintain the quality of terminologies [9–12,38,39]. Most frequently mentioned requirements in the past decades are associated with terminology content, modeling structure, mapping, and process management. ISO/TS 17117:2002, *Health Informatics – Controlled Health Terminology – Structure and High-Level Indicators* lists a set of outcomes resulting from international efforts that encompasses technical specifications on various aspects of terminology management [13]. The ISO standard addresses a need to describe the purpose and scope of a terminology and the degree of domain coverage (comprehensiveness and completeness) of a terminology (for a given purpose and scope). In addition to criteria relating to terminology content, other criteria within the standard, such as concept orientation (non-redundancy, non-ambiguity, and non-vagueness), target structure and organization. Table 1 presents a synthesis of terminology requirements derived from the literature and descriptions of each requirement.

Table 1
Summary of Healthcare Terminology Requirements [9–13,20,21,24,26,38–40].

Requirements	Descriptions
1. Content	The content of the terminology will be appropriate for the stated purposes and domain of use
(1) Purpose and scope	Any terminology shall have its purpose and scope clearly stated in operational terms so that its fitness for particular purposes can be assessed and evaluated
(2) Coverage	The extent to which the terminology is representative shall be explicitly specified for each domain and purpose
a. Comprehensiveness	The extent to which the breadth of coverage is incomplete shall be explicitly specified for each domain and purpose
b. Completeness	The extent to which the depth of coverage is incomplete shall be explicitly specified for each domain and purpose
2. Modeling structure	The structure of the terminology will be appropriate for the stated purposes and domain of use
(1) Concept identifier	A unique identifier must be assigned to each concept
a. Context-free identifiers	Identifiers must not be tied to hierarchical position or other contexts
b. Persistence of identifiers	Identifiers shall not be re-used when a concept is obsolete or superseded
c. No duplicate identifiers	There shall not be more than one concept with the same identifier
(2) Concept orientation	The basic unit of a terminology must be a concept, which is the embodiment of some specific meaning, and not a code or character string
a. Non-redundancy	There shall not be more than one concept with the same meaning in the terminology
b. Non-ambiguity	There shall not be a concept with more than one meaning in the terminology
c. Non-vagueness	There shall not be concepts without explicit meaning in the terminology
(3) Concept composition	Composite concepts shall fit into a practical model that extends a terminology
(4) Concept representation	The compositional system shall contain a formal mechanism to represent concepts
a. Formal definitions Explicitness of attributes Explicitness of relations	The compositional system shall contain formal definitions for non-atomic concepts and formal rules for inferring subsumption from these definitions. The formal attributes of each concept and formal behavior of all relations among concepts in the terminology shall be explicitly defined
b. Normalization of semantics	The extent to which normalization of semantics can be performed formally by the terminology shall be clearly indicated
c. Multiple hierarchies with consistent view	Concepts shall be accessible through all reasonable hierarchical paths. The compositional system shall support multiple views of hierarchical paths in a consistent manner
d. Internal consistency	Relations between concepts (e.g., concept categorization and hierarchical relationships) should be uniform across parallel domains within the terminology
e. Lexical consistency	The terminology shall support accepted spelling rules, syntactic variants, etc.
(5) Version control by concept	Updates and modifications shall be referred to by consistent version identifiers
a. Date of changes made and source for changes	New and revised terms, concepts, and synonyms shall have their date of entry, along with pointers to their source and/or authority
b. Obsolete marking	Superseded entries shall be so marked, together with their preferred successor
(6) Language independence	The terminology shall incorporate multilingual support
3. Mapping	The terminology will incorporate mapping support for the stated purposes and domain of use
(1) Synonyms mapping	The extent to which terms with the same meaning are mapped to the concept in the terminology shall be explicitly stated
(2) Inter-terminology mapping	The degree to which the terminology is cross-mapped to other classifications shall be explicitly stated
4. Process management	The management process of the terminology will be explicitly stated
(1) Formal methods for – Change requests – Change and version management – Editing and modeling conventions – Auditing – Distribution of transformable release format – Concept composition, – Translations – Subset creation – Inter-terminologies mapping and harmonization	The terminology shall have policy and documents describing formal methods for change requests and review; concept change and version management; concept editing and modeling conventions; auditing (e.g., lexical rules, attribute inheritance); creating transformable release formats; concept composition (e.g., syntax and grammar); translations; subset creation; and permissible mapping and harmonization across terminologies

These requirements form the basis of the outcome measures in the TQI model. Specific criteria organized by structure, process, and outcome components in relation to the terminology life cycle are presented in Fig. 2.

4. Model validation through a case study

In order to validate the TQI model, we used the International Classification for Nursing Practice (ICNP) as a source terminology. The ICNP is a logic-based compositional nursing terminology registered with HL7 and recognized by the American Nurses Association [15,41]. The current ICNP Version 2 consists of 2009 primitive concepts and 833 pre-coordinated nursing assessment, diagnosis, intervention, and outcome concepts [23]. ICNP is recognized as a related member of the World Health Organization Family of International Classifications (WHO-FIC). Also, the International Council of Nurses (ICN) has an agreement with the IHTSDO to harmonize ICNP and SNOMED CT. The following describes the application of the TQI model in terms of structure, process and outcome.

4.1. Structure

Although the development of ICNP began in 1989, it was not until 2000 that ICN established a formal ICNP Programme. Since that time, many structures have been put into place to support the work of the Programme, including a governance structure with designated relationships among individual staff positions (i.e., ICNP operational team) and volunteer groups. For example, the Strategic Advisory Group, is a volunteer group of terminology and strategic experts brought together annually at the ICN headquarters, in Geneva. The terms of reference and membership of this group is aimed at advising the ICN Chief Executive Officer on strategic planning. In addition, the ICN, as a federation of 132 national nurses associations, provides a structure for networking and communicating with nurses worldwide, enhancing global participation

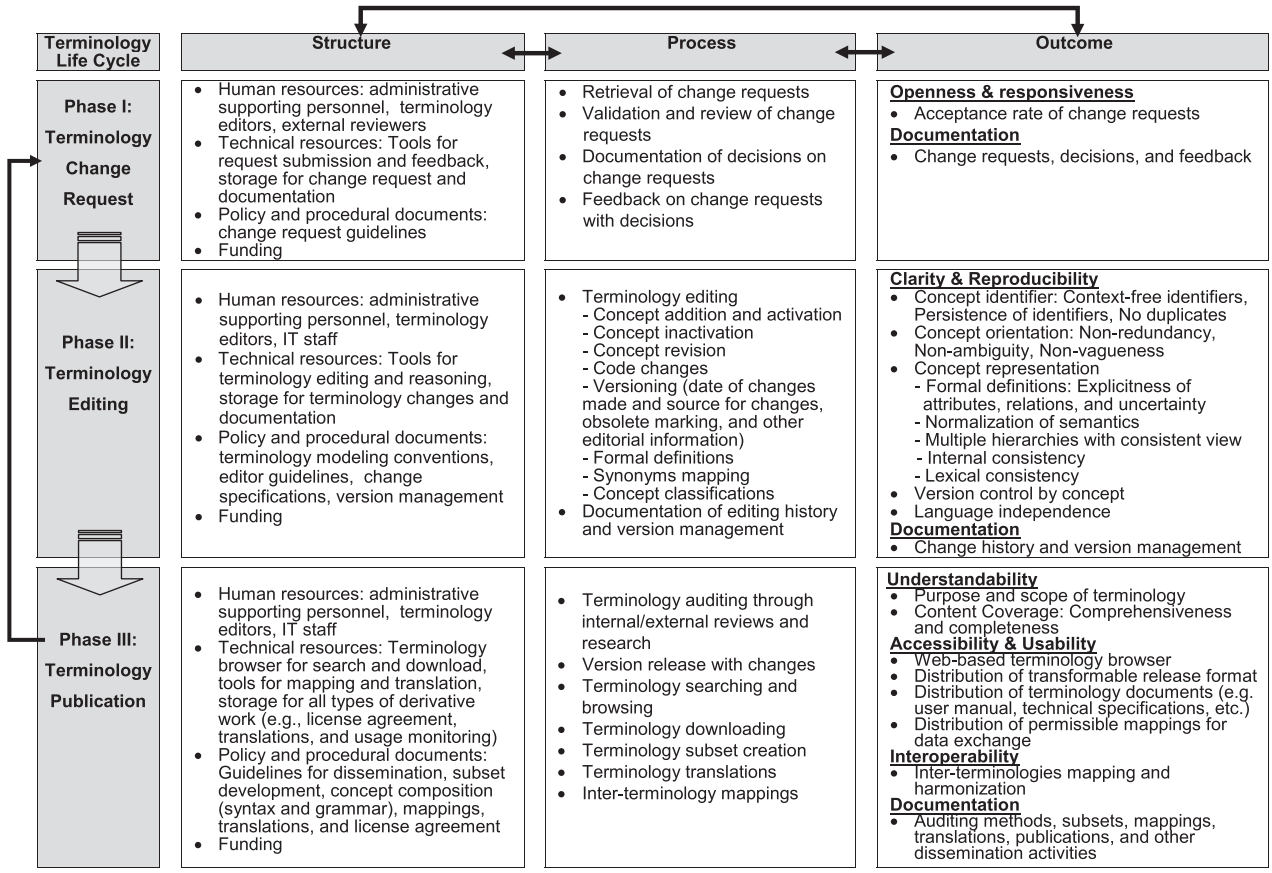

Fig. 2. The terminology quality improvement model.

of stakeholders. The mission of ICN, to represent nursing worldwide, advancing the profession and influencing health policy, guides the direction and policies of the ICNP Programme.

Another major structural initiative to decentralize ICNP research and development (R&D) activities is the establishment of ICNP Centres that are accredited and renewed by ICN according to policy. There are currently six accredited Centres with defined organizational structure and 3-year plans. Each Centre identifies the specific aspects of their work (e.g., translation or testing of ICNP). Centres may be organized at a number of levels, including, but not limited to country, inter-country, regional, and global. The Centres are all represented at the ICNP Consortium, which is held every 2 years in conjunction with the ICN Congress, in revolving regions of the world. Guidelines for application and a description of each Centre are available at the ICN website [42].

In addition to ICN personnel and groups to support the ICNP Programme, a number of tools have been implemented or developed to support ICNP processes and outcomes. For example, ICNP is designed using open source tools for authoring (Protégé) [43] and reasoning (Fact++) [44]. ICNP Collaborative Space (C-Space) is a Web-based tool developed to support collaborative work for mapping clinically-relevant subsets to ICNP and for distribution of the most recent version via a browser with download capabilities [45]. A browser and translation tool (BaT Tool) has also been developed by the ICNP R&D Centre for German-speaking Countries [46]. In line with current good terminology practice [13,14], ICN has written policies and procedures that cover concept submission and review, modeling, translation, version management, subset creation, and release formats.

4.2. Process

4.2.1. Phase I: Terminology Change Request

ICNP change requests are submitted by experts in nursing, other healthcare terminologies, system developers, and others worldwide. These can be individuals or groups, including the ICNP technical advisory group, the operational team, R&D Centres, and vendors. Once change requests are submitted, the ICNP operational team initiates a review of use case and application, contacting submitters for clarification. The review process facilitates the collection and review of any recommendations for additions or changes to a current ICNP version. More than 200 expert reviewers, representing 53 countries, participate in the review process, bringing clinical knowledge and experience. All activities involved in this phase are guided by the ICNP *Concept Submission and Review Guidelines* [47] and the process is supported by ICNP C-Space [45].

With the release of the ICNP Version 1.0, the ICNP Programme emphasized the need for research in the implementation of ICNP. Mechanisms to gather feedback from those implementing ICNP, especially in use in health information systems, facilitate continuing research on ICNP utility. Accordingly, the Terminology Change Request component of the life cycle facilitates on-going application and evaluation of ICNP as well as collaboration with other terminology developers, which has become a means of terminology improvement over the past years.

4.2.2. Phase II: Terminology Editing

The process of Terminology Editing is guided by the ICNP *Style Guide for Modeling* and *Policy and Procedures for Version*

Management. These procedural documents derived from current terminology standards and research [13,14,40,41] describe a systematic mechanism for concept modeling, editing changes, and version control.

The ICNP is designed using Web Ontology Language (OWL) which is underpinned by description logic; this helps in determining subsumption relations among classes [48]. The current ICNP ontology comprises OWL classes, properties, conditions, and annotations. While a class, as the basic unit of ICNP, represents a nursing entity, conditions associated with classes represent formal definitions for entities. Tools such as Protégé and Fact++ facilitate concept modeling and reasoning to assist in examining the human component of modeling (e.g., inconsistency) and performing formal normalization of semantics by automated description logic-based reasoning [43,44]. Further, ICNP uses a concept annotation function in Protégé to maintain (a) unique identifiers (arbitrarily assigned sequential numeric codes), (b) editorial information (e.g., preferred terms, concept descriptions, versions, dates, sources), (c) synonyms, and (d) subsets. Information about inactivated concepts, replacement codes and translations is managed in separate tables.

4.2.3. Phase III: Terminology Publication

This phase includes (a) terminology auditing, (b) distribution of ICNP in various formats, (c) distribution of ICNP-related publications and materials, (d) translation, (e) execution of non-commercial user agreements and commercial distribution agreements, (f) distribution of subsets, and (g) education. Terminology auditing focuses on consistency on concept definitions, hierarchical classifications, lexical rules, and concept identifiers. The format of ICNP files is determined by users' needs and is outlined in *Policy and Procedures for Version Management*. The current release includes OWL, CSV, TXT, and Excel files. Combined with human inspection, all the machine-aided processes for auditing and dissemination are performed using a database management system (MySQL).

The ICNP operational team also publishes and distributes a bi-annual newsletter, ICNP Bulletin, subsets, and version changes which are available in print and online. Processes to support ICNP education include generating and facilitating presentations and publications, and updating web content. Together with monitoring and reporting ICNP education by nurse experts worldwide, ICNP Programme members submit abstracts or respond to invitations to make oral and poster presentations, develop journal and book manuscripts, and prepare general educational materials. A final example of ICNP processes regards translations and mapping with other terminologies, which are guided by formal procedures associated with ICNP C-Space and the BaT tool [45].

4.3. Outcome

With the continuous support from ICN, ICNP Version 2 was released in June 2009 and is available through ICNP C-Space [45], BioPortal [5], and UMLS [4] in order to meet a practical need of nurses in describing their practice with respect to clients, families and communities [23]. The degree of conformance of ICNP Version 2 to the terminology quality requirements summarized in Table 1 is described as follows.

The ICN provided a clear statement of purpose and scope for ICNP: "The ICNP is a unified nursing language system. It is a compositional terminology for nursing practice that facilitates the development of and the cross-mapping among local terms and existing terminologies [23]." Accordingly, ICNP Version 2 is comprised of nursing phenomena (diagnoses), actions, and outcomes. Establishment of ICNP R&D Centres as well as terminology subsets development resulted in an expansion of ICNP [49]. It would be difficult, however, to demonstrate comprehensiveness of this terminology without specifying context of use. Comprehensiveness, thus, is assessed and assured on a case-by-case basis (e.g., subset development). Similarly it would be difficult to demonstrate completeness without specifying acceptable depth of coverage for each entity represented within the terminology. Hence completeness is assessed and assured on a case-by-case basis (e.g., mapping).

ICNP Version 2 (designed using OWL) conforms to current best practice for formal ontologies [41]. ICNP codes are never duplicated or re-used and editorial information on concept changes is maintained in a consistent manner using annotations to OWL classes. Entities, represented as OWL classes within Version 2, are described by fully-specified knowledge names. In other words, context is included in the names of classes so that they can be interpreted independently of their hierarchical placement, often to the detriment of readability (e.g., ActualNegativeAbilityToWalk).

Since the use of OWL within Protégé prohibits the use of duplicate knowledge names, implied relations, and multiple parents in the asserted hierarchy [43], the possibility of ambiguity is substantially limited (non-ambiguity). However, OWL does permit multiple parents for individual classes after description logic-based reasoning so that the inferred ICNP Version 2 is multi-hierarchical [43]. Internal consistency would be difficult to demonstrate in practice without performing pair-wise comparisons across the entire terminology. The high degree of compositionality within Version 2 and consistent patterns of pre-coordination (through the use of the *Style Guide for Modeling*), however, do serve to promote consistency across parallel domains. For example, GrandfatherRole is a child of FamilyMemberRole so that, through composition, Grandfather is necessarily a child of FamilyMember. Lexical consistency is assured through the use of British English spelling for knowledge names and accepted spelling conventions for translations. It has been possible to translate ICNP into multiple languages as preferred terms, synonyms, and concept descriptions.

Previous work has demonstrated that it is possible to map other terminologies to ICNP through processes that reveal resolvable deficiencies in source and target terminologies, and in the mapping process itself [50–52]. However, there is no explicit statement that indicates the degree to which mapping is possible. This would be difficult to demonstrate in practice without specifying degree of ability to map to all known classifications.

ICNP policy and procedures developed cover concept submission and review, modeling, translation, version management, subset creation, and release formats. Areas of improvements include posting submissions as well as tracking reviews, perhaps with an open report on the ICN website. Updating documentation continues to be a challenge and requires ongoing attention to represent current reality. For example, documentation regarding auditing of ICNP and mapping between ICNP Version 2 and other terminologies as well as local or national subsets is in production. While further work is required on policy, and particularly on documentation, compliance with the criteria around content, modeling structure, mapping and process management serves to ensure that ICNP is characterized by (a) openness and responsiveness; (b) clarity and reproducibility; (c) understandability; (d) accessibility and usability; (e) interoperability; and (f) quality of documentation pertaining to all activities involved in terminology management [12,17,20,21,53].

5. Discussion and conclusion

Wide dissemination of various terminologies and information systems has spurred terminology adoption in practice. However, challenges remain in assessing and improving the quality of terminologies because of variations in development, maintenance, auditing, and governance. Further, quality control has been often

viewed as a separate activity in terminology management rather than an integrated activity across all phases of the terminology life cycle. Having a generalizable terminology quality improvement model may be beneficial to both terminology developers and users.

The purpose of this paper was to introduce the TQI model formulated through a synthesis of literature and existing standards and to validate the model using a case study (ICNP). We developed the TQI model in relation to all phases of the terminology life cycle and have demonstrated through the application of the TQI model that a well-defined work process, along with organizational structure, is essential to produce best outcomes.

The multi-dimensional TQI model reflects the complexity of activities involved in terminology quality management. It should be noted, however, that the TQI model does not intend to include all activities involved in the terminology life cycle; just the major components of ongoing development and maintenance across different terminologies or ontologies. The case study with ICNP demonstrated the applicability of the TQI model for terminology quality measurement and improvement. This also revealed that certain quality indicators identified in the literature (e.g., comprehensiveness, completeness, and mapping) are not easily quantifiable without specifying context of use and measurement criteria.

The TQI model may help informatics practitioners and researchers as well as end-users conduct on-going terminology evaluation in a consistent manner. Along with recommendations made through international organizations such as ISO and IHTSDO, best practices for terminology quality improvement should be constantly sought and advanced. The continuing development of tools and techniques for terminology management will also be necessary [2].

Acknowledgments

This study was supported by the U.S. National Science Foundation grant award (PFI-0650323; T.K., A.C.) and the International Council of Nurses (ICN; T.K., A.C., N.H.). The ICN holds the copyright of International Classification for Nursing Practice.

References

[1] International Organization for Standardization. Health informatics – vocabulary for terminological systems (ISO/TC 17115:2007). Geneva: ISO; 2007.
[2] Rogers JE. Quality assurance of medical ontologies. Methods Inf Med 2006;45(3):267–74.
[3] Gruber T. Toward principles for the design of ontologies used for knowledge sharing. Int J Hum Comput Syst 1995;43(5–6):907–28.
[4] National Library of Medicine. UMLS Metathesaurus. 2009. Available from: http://www.nlm.nih.gov/pubs/factsheets/umlsmeta.html.
[5] National Center for Biomedical Ontology. NCBO BioPortal. 2009. Available from: http://bioportal.bioontology.org/.
[6] The International Health Terminology Standards Development Organization. Release of SNOMED CT. 2009. Available from: http://www.ihtsdo.org/snomed-ct/release-of-snomed-ct/.
[7] Cimino JJ, Clayton PD. Coping with changing controlled vocabularies. Proc Annu Symp Comput Appl Med Care 1994:135–9.
[8] Tuttle MS, Olson NE, Campbell KE, Sherertz DD, Nelson SJ, Cole WG. Formal properties of the Metathesaurus. Proc Annu Symp Comput Appl Med Care 1994:145–9.
[9] Cimino JJ, Clayton PD, Hripcsak G, Johnson SB. Knowledge-based approaches to the maintenance of a large controlled medical terminology. J Am Med Inform Assoc 1994;1(1):35–50.
[10] Cimino JJ. Desiderata for controlled medical vocabularies in the twenty-first century. Methods Inf Med 1998;37(2):135–9.
[11] Campbell JR, Carpenter P, Sneiderman C, Cohn S, Chute CG, Warren J. Phase II evaluation of clinical coding schemes: completeness, taxonomy, mapping, definitions, and clarity. CPRI Work Group on Codes and Structures. J Am Med Inform Assoc 1997;4(3):238–51.
[12] Chute CG, Cohn SP, Campbell JR. A framework for comprehensive health terminology systems in the United States: development guidelines, criteria for selection, and public policy implications. ANSI Healthcare Informatics Standards Board Vocabulary Working Group and the Computer-Based Patient Records Institute Working Group on Codes and Structures. J Am Med Inform Assoc 1998;5(6):503–10.
[13] International Organization for Standardization. Health Informatics – Controlled health terminology – Structure and high-level indicators (ISO/TX 17117:2002). Geneva: ISO; 2002.
[14] International Organization for Standardization. Health Informatics – Integration of a reference terminology model for nursing (ISO/TX 18104:2003). Geneva: ISO; 2003.
[15] American Nurses Association. ANA recognized terminologies and data element sets. 2006. Available from: http://www.nursingworld.org/npii/terminologies.htm.
[16] Donabedian A. The definition of quality and approaches to its assessment. Ann Arbor: Health Administration Press; 1980.
[17] International Health Terminology Standards Development Organization. IHTSDO Quality Assurance Framework: discussion document for quality assurance framework. Copenhagen (Denmark): IHTSDO; 2008.
[18] International Health Terminology Standards Development Organization. IHTSDO Quality Framework Toolkit: framework description and example/prototype application. Copenhagen (Denmark): IHTSDO; 2008.
[19] International Organization for Standardization. Software engineering – product quality – Part 1: quality model (ISO/IEC 9126-1:2001). Geneva: ISO; 2001.
[20] de Coronado S, Wright LW, Fragoso G, et al. The NCI Thesaurus quality assurance life cycle. J Biomed Inform 2009;42(3):530–9.
[21] Cimino JJ, Hayamizu TF, Bodenreider O, Davis B, Stafford GA. The caBIG terminology review process. J Biomed Inform 2009;42(3):571–80.
[22] Elkin PL, Brown SH, Carter J, et al. Guideline and quality indicators for development, purchase and use of controlled health vocabularies. Int J Med Inform 2002;68(1–3):175–86.
[23] International Council of Nurses. ICNP® Version 2. Geneva (Switzerland): ICN; 2009.
[24] Bakhshi-Raiez F, Cornet R, de Keizer NF. Development and application of a framework for maintenance of medical terminological systems. J Am Med Inform Assoc 2008;15(5):687–700.
[25] Baorto D, Li L, Cimino JJ. Practical experience with the maintenance and auditing of a large medical ontology. J Biomed Inform 2009;42(3):494–503.
[26] Oliver DE, Shahar Y. Change management of shared and local versions of health-care terminologies. Methods Inf Med 2000;39(4–5):278–90.
[27] Raiez F, Arts D, Cornet R. Terminological system maintenance. A procedures framework and an exploration of current practice. In: Engelbrecht R et al., editors. Connecting medical informatics and bio-informatics. Amsterdam: IOS Press; 2005. p. 701–6.
[28] Zhang S, Bodenreider O. Law and order: assessing and enforcing compliance with ontological modeling principles in the Foundational Model of Anatomy. Comput Biol Med 2006;36(7–8):674–93.
[29] Bodenreider O, Smith B, Kumar A, Burgun A. Investigating subsumption in SNOMED CT: an exploration into large description logic-based biomedical terminologies. Artif Intell Med 2007;39(3):183–95.
[30] Verspoor K, Dvorkin D, Cohen KB, Hunter L. Ontology quality assurance through analysis of term transformations. Bioinformatics 2009;25(12):i77–84.
[31] Jiang G, Chute CG. Auditing the semantic completeness of SNOMED CT using formal concept analysis. J Am Med Inform Assoc 2009;16(1):89–102.
[32] Min H, Perl Y, Chen Y, Halper M, Geller J, Wang Y. Auditing as part of the terminology design life cycle. J Am Med Inform Assoc 2006;13(6):676–90.
[33] Yeh I, Karp PD, Noy NF, Altman RB. Knowledge acquisition, consistency checking and concurrency control for Gene Ontology (GO). Bioinformatics 2003;19(2):241–8.
[34] Cornet R, Abu-Hanna A. Description logic-based methods for auditing frame-based medical terminological systems. Artif Intell Med 2005;34(3):201–17.
[35] Cimino JJ, Min H, Perl Y. Consistency across the hierarchies of the UMLS Semantic Network and Metathesaurus. J Biomed Inform 2003;36(6):450–61.
[36] Gu H, Perl Y, Elhanan G, Min H, Zhang L, Peng Y. Auditing concept categorizations in the UMLS. Artif Intell Med 2004;31(1):29–44.
[37] Vizenor LT, Bodenreider O, McCray AT. Auditing associative relations across two knowledge sources. J Biomed Inform 2009;42(3):426–39.
[38] Arts DG, Cornet R, de Jonge E, de Keizer NF. Methods for evaluation of medical terminological systems–a literature review and a case study. Methods Inf Med 2005;44(5):616–25.
[39] Zhu X, Fan JW, Baorto DM, Weng C, Cimino JJ. A review of auditing methods applied to the content of controlled biomedical terminologies. J Biomed Inform 2009;42(3):413–25.
[40] Hardiker NR, Coenen A. Standards to support the ongoing development and maintenance of nursing terminologies. In: Oud N, Sheerin F, Ehnfors M, Sermeus W, editors. ACENDIO 2007 6th European Conference of Acendio. Amsterdam: Oud Consultancy; 2007. p. 39–43.
[41] Hardiker NR, Coenen A. Interpretation of an international terminology standard in the development of a logic-based compositional terminology. Int J Med Inform 2007;76(Suppl. 2):S274–280.
[42] International Council of Nurses. ICN Accredited Centres for ICNP® Research & Development: Guidelines for Application and Re-accreditation. Geneva (Switzerland): ICN; 2008. Available from: http://www.icn.ch/pillarsprograms/icn-accredited-centres-for-icnpr-research-a-development/.
[43] Knublauch H, Dameron O, Musen MA. Weaving the biomedical semantic web with the Protégé OWL plugin. KR-MED Proc 2004:39–47.
[44] Tsarkov D, Horrocks I. FaCT++ description logic reasoner: system description. Lecture Notes Comput Sci 2006;4130:292–7.
[45] International Council of Nurses. ICNP® Collaborative Space (C-Space). 2009. Available from: http://icnp.clinicaltemplates.org/.

[46] Schrader U, Tackenberg P, Widmer R, Portenier L, Konig P. The ICNP-BaT – a multilingual web-based tool to support the collaborative translation of the International Classification for Nursing Practice (ICNP). MedInfo 2007;129(Pt 1):751–4.
[47] International Council of Nurses. ICNP® Concept Submission and Review Guidelines. Geneva (Switzerland): ICN; 2010. Available from: http://www.icn.ch/pillarsprograms/icnpr-concept-submission-and-review-guideline/.
[48] W3C. OWL Web Ontology Language Overview. 2004. Available from: http://www.w3.org/TR/owl-features/.
[49] Coenen A, Kim TY. Development of terminology subsets using ICNP. Int J Med Inform 2010;79(7):530–8.
[50] Hardiker NR, Casey A, Coenen A, Konicek D. Mutual enhancement of diverse terminologies. AMIA Annu Symp Proc 2006:319–23.
[51] Matney SA, DaDamio R, Couderc C, et al. Translation and integration of CCC nursing diagnoses into ICNP. J Am Med Inform Assoc 2008;15(6):791–3.
[52] Dykes PC, Kim HE, Goldsmith DM, Choi J, Esumi K, Goldberg HS. The adequacy of ICNP version 1.0 as a representational model for electronic nursing assessment documentation. J Am Med Inform Assoc 2009;16(2):238–46.
[53] Bodenreider O, Burgun A, Botti G, Fieschi M, Le Beux P, Kohler F. Evaluation of the Unified Medical Language System as a medical knowledge source. J Am Med Inform Assoc 1998;5(1):76–87.