The Q-Codes: Metadata, Research data, and Desiderata, Oh My! Improving Access to Grey Literature in Family Medicine

Melissa P. Resnick, University of Texas Health Science Center at Houston, Houston, TX, USA
Ashwin Ittoo, HEC Management School, University of Liège, Belgium
Marc Jamoulle and Marc Vanmeerbeek, Department of General Practice, University of Liège, Belgium
Frank S. Shamenek, Consultant, New York, USA
Chiehwen Ed Hsu, College of Management, National CK University, Tainan, Taiwan
Robert Vander Stichele, Heymans Institute of Pharmacology, University of Ghent, Belgium
Julien Grosjean and Stefan Darmoni, Department of Information and Medical Informatics (D2IM), University of Rouen, France
Elena Cardillo, Institute of Informatics and Telematics, National Research Council, Italy
Miguel Pizzanelli, Department of Family and Community Medicine, University of the Republic of Uruguay

***Abstract**:*
***Problem/Goal:** In GL19's "Indexing grey literature in General Practice: Family Medicine in the Era of Semantic Web," Jamoulle and colleagues (Jamoulle et al., 2018) propose the use of a relatively new terminology (3CGP) to allow for the indexing and retrieval of (GP/FM) knowledge which otherwise would be lost, or difficult to locate. Though designed to meet Cimino's (Cimino, 1998) twelve desiderata for the design of a controlled healthcare vocabulary, Jamoulle and colleagues (Jamoulle et al., 2018) acknowledge that a detailed requirement by requirement evaluation of 3CGP was not performed. The goal of this paper is to evaluate the Q-Codes component of the 3CGP terminology, in detail, with each of Cimino's twelve desiderata.*

***Research Method/Procedure:** In our work, we will focus on qualitative analysis, whereby our taxonomy, the Q-Codes, and in particular, its vocabulary satisfies a standard set of desiderata. Qualitative analysis provides a simple and yet effective way to assess the Q-Codes taxonomy's quality. We will briefly describe each of the desiderata and discuss how our taxonomy satisfies each one of them (or not).*

***Anticipated Results of the Research:** The qualitative evaluation is intended as an initial stage, which focuses on the Q-Codes taxonomy's contents, namely, its vocabulary (e.g. terms and definitions). Our aim with the qualitative evaluation is to investigate whether our proposed taxonomy, and in particular its vocabulary, satisfies a set of desiderata. This will enable us to determine whether the knowledge acquisition and (part of) the conceptualization steps of our ontology development process have been performed correctly. We consider that validating our vocabulary against a set of well-defined desiderata is paramount before evaluating other aspects of the taxonomy (such as the relations). As a set of desiderata, we chose that proposed by Cimino in his seminal study entitled "Desiderata for controlled medical vocabularies in the twenty-first century" (Cimino, 1998). These desiderata ensure that our taxonomy can be successfully deployed and exploited in actual GM/FM applications / activities, such as indexing grey literature. The desiderata define a set of (desired) characteristics that (ideally all) standard medical vocabularies should satisfy. Thus, these desiderata help in alleviating inter-operability issues, with the use of common standards ensuring the efficient integration of our taxonomy with other medical vocabularies and resources (taxonomies, ontologies). From the results of this study, improvements can be made to the Q-Codes component of, and thus, the 3CGP terminology. This, in turn, improves the ability to index the grey literature with the 3CGP terminology, providing greater access to needed information.*

***Indication of costs related to the project:** This project has not been funded.*

Introduction

In recent years, grey literature has become more important, especially in research areas, such as General Practice/Family Medicine (GP/FM) (Jamoulle, Grosjean, et al., 2017). Grey literature has different meanings to different people. Thus, there are many definitions for grey literature. Three of these definitions will be presented below.

Denda (Denda, 2002) notes that "grey literature is a body of information that is often not identified through standard acquisitions procedures or retrieved through research tools such as indexes, catalogs, or databases." Some state that grey literature is "material that is difficult to catalogue" (Mahood, Van Eerd, & Irvin, 2014; Tillett & Newbold, 2006). For others, grey literature is defined as: "that which is produced on all levels of government, academics, business and industry in print and electronic formats, but which is not controlled by commercial publishers" (Bellefontaine & Lee, 2014; New York Academy of Medicine, 2018; Paez, 2017; Pappas & Williams, 2011). This third and last definition will be used to define grey literature for this research.

Given this definition, grey literature can encompass many types of materials including, but not limited to, dissertations, conference proceedings, reports, book chapters, magazine articles, newsletters, blogs, wikis, conference abstracts, and preprints (Mahood et al., 2014; TextRelease, n.d.-b). Mahood and colleagues (Mahood et al., 2014) point out that GreyNet provides an extensive list of materials considered grey literature on their website (TextRelease, n.d.-a). However, most of these resources are difficult to find, as they lack bibliographic information, such as author/publisher or volume/issue/page numbers (Mahood et al., 2014) and are not even indexed (Denda, 2002; Mahood et al., 2014). The lack of bibliographic information or metadata, especially indexing terms, can often lead to loss of information. The use of a terminology, one type of metadata, can assist in improving this situation.

During the GL19 conference, a relatively new terminology, Core Content Classification in General Practice Family Medicine (3CGP), was proposed to allow for the indexing and retrieval of GP/FM knowledge, which otherwise would be lost or difficult to locate (Jamoulle et al., 2018). This terminology is composed of two components: (i) the International Classification of Primary Care (ICPC-2), used for clinical concepts, and (ii) the Q-Codes taxonomy, used for contextual concepts (Jamoulle et al., 2018). The remainder of this paper is concerned with the Q-Codes taxonomy component of 3CGP.

The Q-Codes taxonomy is comprised of eight top-level categories, as shown in Table 1 (Jamoulle et al., 2018, 2017). Each of the eight top-level categories has a simple hierarchy of up to three levels. These eight simple hierarchies together form the taxonomy (Jamoulle, Grosjean, et al., 2017). In building this taxonomy, Jamoulle and colleagues (Jamoulle et al., 2018) have attempted to meet the twelve desiderata (guidelines) proposed by Cimino in 1998 (Cimino, 1998).

Table 1. Q-Codes top-level categories overview

Q-Code top-level category	Label	Examples of covered topics
QC	Patient's category	age, gender issues, abuse
QD	Family doctor's issue	communication, clinical prevention, medico legal issues
QE	Medical ethics	bioethics, professional ethics, info ethics
QH	Planetary health	environmental health, biological hazards, nuclear hazards
QP	Patient issue	patient safety, patient centeredness, quality of health care
QR	Research	research methods, research tools, epidemiology of primary care
QS	Structure of practice	primary care setting, primary care provider, practice relationship
QT	Knowledge management	teaching, training, knowledge dissemination

Over the past decades, terminologies have been constructed for several reasons including, but not limited to: capturing clinical findings, natural language processing, indexing medical records, indexing medical literature, and representing medical knowledge (Cimino, 1998). Terminology users tried to use various standard terminologies, but found this difficult, as no one standard terminology was appropriate for all of their needs (Cimino, 1998). Thus, users began to express a need for requirements for terminologies.

By the early 1990s, terminology researchers began writing about various requirements believed useful in building terminologies (Cimino, 1998). As stated by Cimino (Cimino, 1998), researchers have gone past discussing only the definitions in a vocabulary and started discussing the "deeper representational aspects" or other components of a vocabulary. In his seminal paper titled "Desiderata for Controlled Medical Vocabularies in the Twenty-First Century", Cimino (Cimino, 1998) presents these requirements or desiderata, which include: Vocabulary Content, Concept Orientation, Concept Permanence, Non-Semantic Concept Identifiers, Polyhierarchy, Formal Definitions, Rejection of "Not Elsewhere Classified" Terms, Multiple Granularities, Multiple Consistent Views, Context Representation, Graceful Evolution, and Recognized Redundancy.

Our motivations for selecting this set of desiderata are as follows. First, the desiderata have been articulated and formulated based on actual requirements of medical informatics application developers and end-users. Thus, these desiderata ensure that the Q-Codes taxonomy can be successfully deployed and exploited in actual GM/FM applications/activities, such as indexing of the grey literature. Second, the desiderata define a set of desired characteristics that ideally all standard medical vocabularies should satisfy. Thus, these desiderata help in alleviating inter-operability issues with the use of common standards, ensuring the efficient integration of this taxonomy with other medical vocabularies and resources. These desiderata will be described in more detail below.

Goal

Though designed to meet Cimino's (Cimino, 1998) twelve desiderata for the design of a controlled healthcare vocabulary, Jamoulle and colleagues (Jamoulle et al., 2018) acknowledge that a detailed requirement by requirement evaluation of 3CGP was not performed. The goal of this paper is to evaluate the Q-Codes component of the 3CGP terminology, in detail, with each of the twelve desiderata proposed by Cimino (Cimino, 1998).

Methods

In our work, we focused on qualitative analysis, whereby our taxonomy, the Q-Codes, and in particular, its vocabulary, satisfies a standard set of desiderata. Qualitative analysis provides a simple, and yet, effective way to assess the Q-Codes taxonomy's quality.

A copy of version 2.5 of the Q-Codes taxonomy was downloaded from: www.3cgp.docpatient.net/ and used for analysis. The twelve desiderata were obtained from Cimino's 1998 paper titled "Desiderata for controlled medical vocabularies in the twenty-first century" (Cimino, 1998).

The Q-Codes taxonomy was evaluated desideratum by desideratum. This analysis was performed as follows: (1) each desideratum was read, and (2) the taxonomy was examined for the presence or absence of the desideratum. In the next section, we will briefly describe each of the desiderata and discuss how the taxonomy satisfies each one of them (or not).

Results and Discussion

Desideratum 1: Content

The main issue to address concerning the Content Desideratum is the need for a systematic, explicit, and reproducible method for expanding content (Cimino, 1998).

Cimino (Cimino, 1998) suggests two main approaches for increasing content. In the first approach, all atomic units (e.g. single-word terms) of a domain terminology, for example GM/FM, are enumerated. Users are then allowed to combine them in order to compose

more complex multi-word terms (Côté & Robboy, 1980). The main benefit of this approach is that by allowing compositional extensibility, it facilitates domain-coverage (D. A. Evans, Rothwell, Monarch, Lefferts, & Cote, 1991). However, this approach suffers from several limitations. First, identifying all domain-specific atomic units is a nontrivial endeavor. Second, the composition of individual atomic units should preserve the semantics. In other words, the meaning of the atomic unit and of the resulting more complex unit should not be distorted with the composition. In addition, some compositions may yield illogical units, for example, combining the two terms "AIDS" and "flu" into "AIDS flu." Specific grammatical rules need to be defined to prevent such combinations.

In the second approach, contents (e.g. terms) are added as they are encountered in the various data sources (e.g. conference abstracts). Compared to the previous approach, one does not attempt to systematically anticipate and enumerate all possible terminological combinations, for example, by listing all the possible types of fractures (simple, complex, hair-line) for each possible bone. Instead, one would add terms corresponding to the most common/frequent ones (e.g. found more frequently in the data sources analyzed), and then add more complex terms as, and when, they are needed or encountered in the data sources. The main benefit of this approach is that it avoids the unnecessary generation of large numbers of terms occurring through combinatorial explosion and the enumeration of nonsensical combinations.

In the Q-Codes, the second approach has been adopted. Terms (simple and complex) are added as, and when, they are needed, and when they are encountered in the relevant data sources, such as conference abstracts.

Desideratum 2: Concept Orientation

According to most, if not all, researchers in medical informatics and in knowledge representation in general, the fundamental unit of symbolic processing is the concept. A concept can be defined as a mental representation of an object within a specific domain; an object refers to anything perceived or conceived (ISO 9000:2015). Thus, in a given specific domain, a concept embodies and conveys a precise meaning (D. A. Evans, Cimino, Hersh, Huff, & Bell, 1994; David A. Evans, 1988; Lindberg, Humphreys, & McCray, 1993; Rassinoux, Miller, Baud, & Scherrer, 1996; Volot et al., 1993). Concepts are lexically realized as terms (simple one-word terms or more complex multi-word terms). The collection of terms in a given domain is part of its vocabulary. Concept orientation means that terms must correspond to at least one meaning (nonvagueness) and no more than one meaning (unambiguity). Concerning the issue of unambiguity, some authors, such as Moorman and colleagues (Moorman, van Ginneken, van der Lei, & van Bemmel, 1994) argue that ambiguity can be allowed as long as the unequivocal meaning is preserved based on the term's usage in a given context.

A total of 182 single- and multi-word terms comprise the Q-Codes (Jamoulle, Grosjean, et al., 2017). Each one of these 182 terms was given only one definition (Jamoulle & Resnick, 2016). This, in turn, meets both the nonvagueness and the unambiguity criteria for Concept Orientation.

Desideratum 3: Concept Permanence

The desideratum of Concept Permanence follows directly from that of Concept Orientation (desideratum 2 above). Concept Permanence requires that once a concept has been created, its meaning is immutable, i.e. it cannot be changed or violated (Cimino, 1998). This condition of semantic immutability holds even if the concept's preferred name changes or if the concept is marked as inactive, deprecated or archaic (Cimino, 1998). For instance, consider a concept with the name "pacemaker". In this case, the concept's meaning does not change even if it is renamed to "implantable pacemaker". Conversely, consider a concept with name "non-A non-B hepatitis". Here, one cannot simply rename it to "hepatitis C" as "non-A non-B hepatitis" is not a synonym of "hepatitis C". In this situation, we cannot assert with certainty that someone with "non-A non-B hepatitis" is definitely suffering from "hepatitis C". Thus, such a renaming entails an alteration in the meaning.

In addition to semantic immutability, Concept Permanence also demands that concepts are not deleted if they are inactive or deprecated. Instead, they should be flagged as such.

Since the latest version of the Q-Codes taxonomy was recently completed (Jamoulle, Grosjean, et al., 2017), none of the current terms have become old or inactive. As the taxonomy evolves from one version to the next, extensive records will be kept so that older or inactive terms will be flagged and not removed. These records will help to ensure that the Q-Codes will continue to meet the Concept Permanence desideratum.

Desideratum 4: Non-Semantic Concept Identifier

Each concept should be assigned a unique identifier. In the simplest case, the concept's name also serves as its unique identifier. However, the main drawback of this strategy is that it hinders subsequent modifications to the concept's name.

Another approach is to assign a hierarchical code which reflects the position of the term in the hierarchy (Cimino, 1998). One advantage to this approach is that, with some knowledge of the hierarchy, the codes can become readable by humans, and thus, hierarchical relationships can be understood (Cimino, 1998).

A second advantage is that a hierarchical code can be used to search for all members of a particular class. For instance, searching for "QC1" can allow a user to find all of the terms that belong to "QC1 age group" (e.g. "QC11 infant", "QC12 child", "QC13 adolescent", etc.). However, the difficulty with this method of searching for terms in the same class arises when a term appears in more than one class or place in the hierarchy (Cimino, 1998).

During the creation of the Q-Codes, each term was assigned a hierarchical code (e.g. "QR31 qualitative study"). With these hierarchical codes, one can see the relationships between the terms. For example, "QR3 research method" is the broader term encompassing the narrower term "QR31 qualitative study".

Although these hierarchical codes are easy for humans to understand and use, there is a major limitation to systems such as this. As noted by Cimino (Cimino, 1998), hierarchical coding systems can "run out of room." In fact, the hierarchical coding system utilized by the Q-Codes has currently "run out of room."

In the hierarchical coding system employed by the Q-Codes, only nine separate terms are possible for the first level under any top-level category. For example, at the first level under the top-level category "QR research", nine separate terms labeled QR1, QR2, QR3 to, and including, QR9 are possible. Next, terms at the first level (e.g. QR1) can have only nine terms (e.g. QR11, QR12, QR13, etc.). This is also the case for any term on the second and successive levels. However, there is no limit to the number of levels for each top-level category.

This limitation can be solved in at least three ways. The first solution is to expand the numbering at each level of the hierarchical code. For example, the first level terms can be labeled QR01 to and including QR99, thus, allowing for 99 terms on this level. However, in theory, this delays the point in time at which the expanded hierarchical coding system will, again, "run out of room."

A second solution is to represent the hierarchies with links between parents and children (Cimino, 1996b). For example, there would be a link between "research method" (the parent) and "qualitative study" (the child), instead of using the unique identifiers "QR3" and "QR31" respectively.

A third solution involves providing tree addresses for each term, like MeSH and the Gabrielli Nomenclature (Cimino, 1996b). Cimino (Cimino, 1996b) points out that tree addresses provide arbitrary "length and breadth." In future versions of the Q-Codes taxonomy, efforts will be made to expand the numbering at each level of the hierarchical code.

Desideratum 5: Polyhierarchy

Hierarchical structures can have at least two forms. One such form is a polyhierarchical structure. A polyhierarchical structure refers to a tree structure in which a term has more than one parent or broader term (American Society for Indexing, n.d.). This means that some of the terms appear in more than one place in the hierarchy (Coletti & Bleich, 2001).

According to Cimino (Cimino, 1998), there seems to be universal consensus that medical informatics resources (such as vocabularies) should have a hierarchical structure. This facilitates searching and locating concepts either by traversing the tree-like hierarchy, or by grouping similar concepts together.

Such a hierarchical organization is also useful for disambiguation. For instance, if a concept named "cell" is located under "anatomic entity", then one can infer that this concept has a different intended meaning than if it appears under "power source" (Cimino, 1998). Here, the parent concepts ("anatomical entity" and "power source") help in making the meaning of child concept ("cell") unambiguous (c.f.: desideratum 2). The majority of current standard vocabularies are strict hierarchies (Cimino, 1998). In this case, each child concept can have only one concept as its parent.

The main strength of strict hierarchies is that they are more amenable for computational purposes. A structure in which each child has a unique parent is far more efficient and easier to process than one where multiple parents are allowed. On the other hand, polyhierarchies might provide a more realistic and accurate conceptualization of a domain, as in the case of the concept "hepatorenal syndrome", which needs two parents, "liver diseases" and "renal disease".

Concerning the Q-codes, we adopted a strict hierarchy, i.e. a single parent per child concept. For example, the concept "QR31 qualitative study" has the concept "QR3 research method" as its only parent. We favored this arrangement over a polyhierarchy as this structure complements that of the International Classification of Primary Care (ICPC) (Jamoulle et al., 2018).

Desideratum 6: Formal Definitions

According to Cimino (Cimino, 1998), many researchers in medical informatics and knowledge representation have formulated the requirement that controlled vocabularies should include Formal Definitions. It should, however, be mentioned that a formal definition here does not mean that the concepts should necessarily be expressed in a particular formalism, such as First Order Logic or RDF'S/OWL. Instead, according to this desideratum, the Formal Definition of a concept is expressed as the different relationships in which the concept participates with other concepts. For example, the concept "hay fever" participates in hyponymy ("a type of") relationship with the concept "fever". This relationship is also referred to as "parent-child", "super-class/sub-class" or "generalization-specialization". The same concept "hay fever" participates with the concept "allergen" in a "caused by" relationship.

Concerning the Q-codes, we have in total 172 relationships, including eight top-level categories with a total of 44 first-level children. These 44 first-level terms have a total of 109 children. These 109 second-level terms have a total of 21 third-level children. For example, the concept "QR31 qualitative study" participates with the concept "QR3 research method" in the "is-a" relationship (e.g. "QR31 qualitative study" "is-a" "QR3 research method"). However, it is important for these relationships to be in a form which can be manipulated symbolically (i.e., with a computer), as opposed to narratives like those seen in a dictionary (Cimino, 1998). In the case of the Q-codes, these relationships are, indeed, represented symbolically, (i.e. in the form of "is-a" links, or "parent-child" relationships), thus, making it easier to manipulate them by a computer.

Desideratum 7: Reject "Not Elsewhere Classified"

This desideratum discourages the use of the category "Not Elsewhere Classified" (called "rag-bag" in this case) to represent terms that cannot be classified under any other category,

i.e. these terms cannot be classified elsewhere. The main motivation for such a rag-bag category is that no vocabulary can guarantee domain completeness at any one time (Cimino, 1998). Thus, the rag-bag category facilitates the representation of terms, which cannot be classified elsewhere given the current state of the taxonomy. According to (Cimino, 1998), one critical issue with having a rag-bag category is that terms classified under this category lack a formal definition (c.f. desideratum 6). Here, the terms in this category can only be defined via exclusion, based on knowledge of the rest of the terms in the taxonomy. Furthermore, as the vocabulary evolves, the meaning of these "Not Elsewhere Classified" or rag-bag terms could change accordingly. This gives rise to the phenomenon of semantic drift, which hinders the analysis of historical data.

In the case of the Q-Codes, the rag-bag category serves one major purpose, sorting of conference abstracts for the discovery and classification of terms. To assist in this process, the rag-bag category is comprised of four subcategories: (i) unable to code, unclear; (ii) acronym; (iii) out of scope of Family Medicine; and (iv) consider new code. Abstracts in the "unable to code, unclear" subcategory, contain no discernible terms related to contextual concepts in GP/FM.

An abstract/title containing an abbreviation or acronym is placed in the "acronym" subcategory. However, these abbreviations/acronyms are often unclear and not well defined.

Some conference abstracts contain concepts unrelated to GP/FM, and thus, they are placed in the "out of scope of Family Medicine" subcategory. These abstracts usually discuss hospital-based studies.

Finally, the abstracts in the "consider possible new code" subcategory contain newly discovered terms or concepts that are currently not present in the taxonomy. These new terms or concepts are defined and examined for relationships to the terms already present in the Q-Codes taxonomy. Finally, these new terms are added to the taxonomy at the appropriate level, as defined by their relationships to other terms in the taxonomy.

Thus, this category and its subcategories have been useful in discovering terms during the creation of the Q-Codes taxonomy. Currently, however, the rag-bag category is useful for suggesting possible terms as additions to the future versions of the taxonomy. The rag-bag category does not hold any terms at each successive release of a version of the Q-Codes. Therefore, this category has no significance for the various end-users and applications of the Q-Codes taxonomy.

Desideratum 8: Multiple Granularities

When a vocabulary is being constructed for a particular application, there is implicitly a preconception about the level of granularity ("details") at which the concepts should be expressed. Granularity can be at a single level or at multiple levels. As the name suggests, with single granularity, all concepts are presented along a single level. Conversely, multiple granularity allows concepts to be represented with progressively finer-grained precision: "Diabetes Mellitus", "Type II Diabetes Mellitus", and "Insulin-Dependent Type II Diabetes Mellitus" (Cimino, 1998). This desideratum asserts that terminologies with multiple granularity should be preferred over those with single granularity. The main issue with vocabularies that attempt to operate at a single level of granularity is their inadequacy for applications requiring finer-grained information. In addition, they will also be considered too overwhelming and cumbersome in applications requiring coarser-grained information (Cimino, 1998).

The Q-codes are a multi-granular taxonomy, as they contain up to three possible levels for each top-level category. In turn, this multi-level granularity allows the Q-Codes taxonomy to be used for many purposes, including indexing the grey literature. Finally, this granularity provides the user with the ability to choose general or specific terms according to their needs.

Desideratum 9: Multiple Consistent Views
According to this desideratum, if a vocabulary is intended for use in multiple applications, then there is a need to provide multiple, consistent views of the vocabulary, as dictated by the various applications (Cimino, 1998; van Ginneken, van der Lei, & Moorman, 1992). For example, if a vocabulary with multiple granularity (c.f. desideratum 8) is to be used in an application that requires coarse-grained concepts, such as "Diabetes Mellitus", then finer-grained concepts, such as "Insulin-Dependent Type II Diabetes Mellitus", could be collapsed into the coarser-grained concept and marked as a synonym. An alternative approach to providing multiple consistent views is to enable users to show/hide specific levels depending on their needs. In a more extreme case, an application may restrict the user to only one level of the hierarchy, while hiding the remaining levels.

In the case of the Q-Codes, users are able to display the different levels of the taxonomy according to their needs (see www.hetop.eu/q). First, the user sees the eight top-level categories. After choosing one of these categories (e.g. "QC patient category"), the second level, with all of its terms, becomes visible, revealing in this case QC1, QC2, QC3 up to and including QC6. The user can continue to climb down to the next level. At any level, the user can choose a particular term to view its definition, links to literature and other information, and relationships to other terms in the taxonomy. Thus, the user can see any one level while hiding the other levels, and, for any one term, he/she can see the associated broader and narrower terms.

Desideratum 10: Representing Context
Traditionally, vocabularies have been created without consideration for the specific contexts in which they are intended to be used. While this strategy helps in reducing implicit assumptions about the vocabulary and enables it to "stand alone", it also leads to confusion when determining whether the vocabulary concepts can be used in specific contexts (Cimino, 1998). Thus, this desideratum asserts that vocabularies should contain explicit information about the contextual usage of concepts, i.e. explaining how/when these concepts should or should not be used (Rector, Glowinski, Nowlan, & Rossi-Mori, 1995).

In the Q-codes, we are provided a glimpse of its purpose by the titles given to some of the terms, such as "QS4 primary care provider" and "QS41 family doctor". Thus, these titles indicate that the Q-Codes taxonomy is used for General Practice or Family Medicine. However, little is provided in the definitions of the terms as to how and when to use them.

The purpose and use of the Q-Codes taxonomy has been further documented in the literature (Jamoulle et al., 2018, 2017; Jamoulle & Resnick, 2016). As a "stand alone" taxonomy, the Q-Codes have been used for e-learning courses in GP/FM (Jamoulle, Grosjean, et al., 2017; Jamoulle & Resnick, 2016). In combination with other terminologies, such as ICPC-2, it has been proposed that the Q-Codes assist with indexing and retrieval of grey literature about GP/FM (Jamoulle et al., 2018, 2017; Jamoulle & Resnick, 2016). Beyond this, not much, if anything, has been provided in this literature about how and when individual terms should and should not be used. In the future, notes (annotations) will be added to the definitions, describing the proper use of the terms.

Desideratum 11: Graceful Evolution
All vocabularies are bound to evolve over time. However, experience shows that in most cases, changes are brought about for the convenience of the vocabulary's creators, and such changes tend to be problematic for the users (Cimino, 1996a). Thus, this desideratum states that those parties responsible for maintaining the vocabulary should ensure its graceful evolution. This can be achieved by assuring that all changes as well as the reason/request for these changes be properly documented and logged.

The Q-codes are currently in their infancy, being used for indexing grey literature only for one to two years. It is expected that they will continue to grow and evolve. We are currently exploring methods to document all changes and reasons/requests for changes, as they continue this growth and evolution.

Desideratum 12: Recognizing Redundancy

Redundancy is a phenomenon whereby the same information can be stated in multiple different ways (Cimino, 1998). For instance, consider a case of "pneumonia in the lower lobe of the left lung" in a patient. Now, this information is to be entered in an electronic patient record, which is based on medical vocabulary. However, if the vocabulary does not have a corresponding concept for "pneumonia in the lower lobe of the left lung", then the user may code it under the concept "Pneumonia" and include an additional label/modifier "left lower lobe". If at some later time, the concept "Left Lower Lobe Pneumonia" is indeed added to the vocabulary, then there will be two ways to code the same concept in the vocabulary: the old way, under "Pneumonia" and with another term indicating location; and the new way, directly under the concept "Left Lower Lobe Pneumonia". Such redundancies are to be avoided, but are inevitable when the vocabulary evolves. Thus, this desideratum asserts that a mechanism to recognize redundancies should be put in place. The two desiderata of Formal Definitions and Representing Context (c.f.: desiderata 6 and 10 respectively) can help in detecting redundancies.

At the present time, there are no redundant terms in the Q-codes. However, by satisfying desideratum 6 (Formal Definitions), one measure has been put in place to detect the presence of redundancies. As mentioned above, efforts will be made to institute desideratum 10 (Representing Context), further increasing the chances that redundancies are detected.

Conclusion

The analysis of the Q-Codes taxonomy demonstrates that it meets eleven of the twelve desiderata. For the remaining desideratum, Representing Context, notes or annotations will be added to the definitions of the terms, describing how and when they should be used. In addition, extensive records, noting any future changes, will be kept to guarantee that the Q-Codes continue to meet these desiderata, as they grow and evolve.

From these results, slight improvements can be made to the Q-Codes component of, and thus, the 3CGP terminology. This, in turn, improves the ability to index the grey literature with this terminology, providing greater access to and preventing loss of needed information.

For future work, the ICPC-2 component of 3CGP will be evaluated for the presence of the twelve desiderata. From this evaluation, any forthcoming recommendations will be provided for improvements to 3CGP, and thus, to indexing and retrieval of grey literature, leading to future research in GP/FM.

REFERENCES

American Society for Indexing. (n.d.). About Taxonomies & Controlled Vocabularies [A Special Interest Group of the American Society for Indexing]. Retrieved November 14, 2018, from http://www.taxonomies-sig.org/about.htm

Bellefontaine, S. P., & Lee, C. M. (2014). Between Black and White: Examining Grey Literature in Meta-analyses of Psychological Research. *Journal of Child and Family Studies*, *23*(8), 1378–1388. https://doi.org/10.1007/s10826-013-9795-1

Cimino, J. J. (1996a). Formal descriptions and adaptive mechanisms for changes in controlled medical vocabularies. M*ethods of Information in Medicine*, *35*(3), 202–210.

Cimino, J. J. (1996b). Review paper: coding systems in health care. *Methods of Information in Medicine*, *35*(4–5), 273–284.

Cimino, J. J. (1998). Desiderata for controlled medical vocabularies in the twenty-first century. *Methods of Information in Medicine*, *37*(4–5), 394–403.

Coletti, M. H., & Bleich, H. L. (2001). Medical subject headings used to search the biomedical literature. *Journal of the American Medical Informatics Association: JAMIA*, *8*(4), 317–323.

Côté, R. A., & Robboy, S. (1980). Progress in medical information management. Systematized nomenclature of medicine (SNOMED). *JAMA*, *243*(8), 756–762.

Denda, K. (2002). Fugitive Literature in the Cross Hairs. *Collection Management*, *27*(2), 75–86. https://doi.org/10.1300/J105v27n02_07

Evans, D. A., Cimino, J. J., Hersh, W. R., Huff, S. M., & Bell, D. S. (1994). Toward a medical-concept representation language. The Canon Group. *Journal of the American Medical Informatics Association: JAMIA*, *1*(3), 207–217.

Evans, D. A., Rothwell, D. J., Monarch, I. A., Lefferts, R. G., & Cote, R. A. (1991). Toward representations for medical concepts. *Medical Decision Making: An International Journal of the Society for Medical Decision Making*, *11*(4 Suppl), S102-108.

Evans, David A. (1988). Pragmatically-Structured, Lexical-Semantic Knowledge Bases for Unified Medical Language Systems. *Proceedings of the Annual Symposium on Computer Application in Medical Care*, 169–173.

Jamoulle, M., Cardillo, E., Ittoo, A., Vander Stichele, R., Resnick, M. P., Grosjean, J., … Vanmeerbeek, M. (2018). Indexing grey literature in General Practice: Family Medicine in the Era of Semantic Web. In D. Farace & J. Frantzen (Eds.), *GL19 Conference Proceedings* (Vol. 19). National Research Council of Italy Piazzale Aldo Moro 7, Rome: TextRelease,. Retrieved from http://www.textrelease.com/images/GL19_Jamoulle_et_al.pdf

Jamoulle, M., Grosjean, J., Resnick, M., Ittoo, A., Treuherz, A., Vander Stichele, R., … Vanmeerbeek, M. (2017). A Terminology in General Practice/Family Medicine to Represent Non-Clinical Aspects for Various Usages: The Q-Codes. *Studies in Health Technology and Informatics*, *235*, 471–475.

Jamoulle, M., & Resnick, M. P. (2016). *General Practice / Family Medicine multilingual terminology – English version*. Strépy-Bracquegnies, Belgium: Le livre en papier. Retrieved from http://www.publier-un-livre.com/fr/le-livre-en-papier/349-general-practice-family-medicine-multilingual-terminology-english-version

Lindberg, D. A., Humphreys, B. L., & McCray, A. T. (1993). The Unified Medical Language System. *Methods of Information in Medicine*, *32*(4), 281–291.

Mahood, Q., Van Eerd, D., & Irvin, E. (2014). Searching for grey literature for systematic reviews: challenges and benefits. *Research Synthesis Methods*, *5*(3), 221–234. https://doi.org/10.1002/jrsm.1106

Moorman, P. W., van Ginneken, A. M., van der Lei, J., & van Bemmel, J. H. (1994). A model for structured data entry based on explicit descriptional knowledge. *Methods of Information in Medicine*, *33*(5), 454–463.

New York Academy of Medicine. (2018, October). What is Grey Literature? | Grey Literature Database. Retrieved November 13, 2018, from http://www.greylit.org/about

Paez, A. (2017). Grey literature: An important resource in systematic reviews. *Journal of Evidence-Based Medicine*. https://doi.org/10.1111/jebm.12265

Pappas, C., & Williams, I. (2011). Grey Literature: Its Emerging Importance. *Journal of Hospital Librarianship*, *11*(3), 228–234. https://doi.org/10.1080/15323269.2011.587100

Rassinoux, A. M., Miller, R. A., Baud, R. H., & Scherrer, J. R. (1996). Modeling principles for QMR medical findings. *Proceedings: A Conference of the American Medical Informatics Association. AMIA Fall Symposium*, 264–268.

Rector, A. L., Glowinski, A. J., Nowlan, W. A., & Rossi-Mori, A. (1995). Medical-concept models and medical records: an approach based on GALEN and PEN&PAD. *Journal of the American Medical Informatics Association: JAMIA*, *2*(1), 19–35.

TextRelease. (n.d.-a). Grey Literature - GreySource, A Selection of Web-based Resources in Grey Literature [GreyNet International 2018]. Retrieved November 13, 2018, from http://www.greynet.org/greysourceindex/documenttypes.html

TextRelease. (n.d.-b). GreyNet International, Grey Literature Network Service [GreyNet International 2018]. Retrieved November 13, 2018, from http://www.greynet.org/home.html

Tillett, S., & Newbold, E. (2006). Grey literature at The British Library: revealing a hidden resource. *Interlending & Document Supply*, *34*(2), 70–73. https://doi.org/10.1108/02641610610669769

van Ginneken, A. M., van der Lei, J., & Moorman, P. W. (1992). Towards unambiguous representation of patient data. *Proceedings. Symposium on Computer Applications in Medical Care*, 69–73.

Volot, F., Zweigenbaum, P., Bachimont, B., Ben Said, M., Bouaud, J., Fieschi, M., & Boisvieux, J. F. (1993). Structuration and acquisition of medical knowledge. Using UMLS in the conceptual graph formalism. *Proceedings. Symposium on Computer Applications in Medical Care*, 710–714.