To what extent can prescriptions be meaningfully exchanged between primary care terminologies? A case study of four western European classification systems

Michiel Meulendijk1, Marco Spruit2, Armel Lefebvre2, Sjaak Brinkkemper2
1Department of Public Health and Primary Care, Leiden University Medical Center, Albinusdreef 2, 2333 ZA Leiden, The Netherlands
2Department of Information and Computing Sciences, Utrecht University, Princetonplein 5, 3584 CC Utrecht, The Netherlands
E-mail: m.c.meulendijk@lumc.nl

Abstract: The diversity of terminologies used in primary care causes significant challenges regarding semantic interoperability. Attempts to address these challenges usually focus on the creation of metaterminologies, with the peculiarities of national variations of terminologies being overlooked. In this study the extent to which primary care data can be meaningfully exchanged between nationally implemented terminologies is assessed using a rule-based approach. To determine this, a model comprising primary care terminologies and including axioms to define their relations was developed. Generic metrics were designed to determine the completeness and accuracy of any two arbitrary vocabularies within an ontological model. These metrics were used on an implementation of the model to determine the data quality that is preserved when expressing similar data in different primary care terminologies. The results show that values of terminologies which are closely related can express each other's concepts relatively well. The authors conclude that the current state of accuracy and completeness between primary care terminologies does not allow for sufficiently meaningful semantic interoperability, but that their approach of mapping lower-level terminologies to each other next to an ontological approach may yield better results than relying solely on the latter.

1 Background

There is a great need for a more sophisticated level of semantic interoperability in primary care. Errors made during the prescription process often lead back to physicians’ use of incomplete health records and miscommunication with caretakers in secondary care [1–3]. This lack of semantic interoperability causes information exchange problems, which in turn can lead to medical errors; a loss of information when a drug prescription is communicated from a prescriber to a pharmacist could result in harmful changes to a patient's drug regimen (e.g. an incorrect change in dosage).

To improve semantic interoperability in primary care, the extent to which different national classification systems are able to express each other's concepts is worth exploring. In this paper, we match and map common primary care terminologies to each other in four western European countries, and validate the results.

1.1 Medical terminology mapping

To share data between concepts from different terminologies, two approaches can be distinguished. Either merging terminologies to form a single metaterminology or mapping the concepts of the different terminologies to one another [4].

The main attempts to create metaterminologies in the medical domain are Systematized Nomenclature of Medicine, Clinical Terms (SNOMED-CT) and Unified Medical Language System (UMLS). SNOMED-CT is a ‘comprehensive, multilingual clinical healthcare terminology’, consisting of coded concepts with descriptions [5]. UMLS aims to ‘bring together many health and biomedical vocabularies and standards to enable interoperability between computer systems’ [6]. It consists of a metathesaurus containing uniquely identifiable concepts with their equivalents in commonly used terminologies such as ICD10 and International Classification of Primary Care. Both SNOMED-CT and UMLS allow for the representation of relationships between concepts, complete with cardinalities and synonyms.

There have been many approaches to map values from one terminology to their closest relatives in another in the medical domain. After the second revision of the popular primary care terminology ICPC, a mapping to the more extensive ICD10 terminology was published [7]. In several instances, mappings between international standards and their national adaptations were created to facilitate data exchange [8, 9].

Both approaches have been used extensively to solve problems of semantic interoperability in the medical domain, under more for the mapping of drugs [10], biomedical data [11], and surgical complications [12], generally with positive results.

1.2 Data quality

The reliability of ontology-based systems is commonly recognised to be partially dependent on the data quality of the ontologies used as their input. As ‘the validity and quality of the ontology data directly affects the validity and quality of the system using the ontology’, assessing the data quality through metrics is a logical and essential step in creating effective and efficient systems [13].

Over the years, a wide variety of variables, and metrics to measure them have been developed. In their study on quality dimensions in ontological foundations, Wand and Wang [14] discerned between accuracy and precision, reliability, timeliness, and currency, completeness, and consistency. In their more recent approach, Batini and Scannapieco [15, p. 16:6] simplified these to accuracy, completeness, time-related dimensions, consistency, and minor parameters of lesser importance (i.e. accessibility and source quality).

In the medical domain, completeness, correctness, concordance, plausibility, and currency are commonly used to assess data quality [16]. In their literature, study on data quality in electronic health records, Weiskopf and Wang found that especially completeness and correctness (or accuracy) are regarded as insightful factors. Metrics most frequently employed to assess these dimensions are element presence (in the case of completeness) and comparison with a gold standard (for both correctness and completeness).

Considering the research question, the information exchange model validated in this paper will be tested for its accuracy and completeness. Even though some of the other parameters
mentions are essential to the quality of data, they can only be ensured through successful implementation. Consistency of data throughout systems depends heavily on information architectures and communication protocols. The same goes for time-related dimensions and accessibility. Source quality in the primary care domain is partly guaranteed by the non-profit organisations maintaining the ontologies, and partly by the authority of whoever implements and uses them.

Defining accuracy is usually done by defining its antonym: inaccuracy implies that a ‘real-world state different from the one that should have been represented’ is displayed [14]. From this follows that accuracy is the correct representation of a real-life phenomenon. Flaws in data, for example, due to input errors, are occurrences of inaccuracy; specifying a drug dosage as 100 mg instead of 10 mg is an accuracy error with potentially far-reaching implications. Lack of precision or ambiguity of concepts can cause data to be inaccurate as well; a recording of a bone fracture without specifying which bone is broken may be too inaccurate to be meaningful.

An ontology is generally considered to be complete if it ‘represents every fact of the real world’ within its domain [15, p. 16.7]. Meeting this criterion requires an ontology to have a narrowly described domain, specifying which attributes are characterized to be included. Missing data does not automatically mean that this criterion is not met; however, since values may be expected to be missing if their attributes are optional; drug prescriptions, for example, may include specific brands if alternatives have been proven to be ineffective for a specific patient, but usually do not.

2 Research design

2.1 Rationale

The diversity of clinical terminologies causes significant challenges regarding semantic interoperability in primary care. Their different perspectives, levels of detail, national customisations, and use in various phases of the prescribing process make information exchange difficult. Doubts regarding concepts’ completeness and accuracy when transferring data appear.

This leads to the following research question, which will be explored in this paper: to what extent can primary care data be meaningfully exchanged between nationally implemented terminologies?

2.2 Approach

To investigate this issue, we strive to create a model of primary care concepts. This model will consist of the concepts present in a selection of nationally implemented terminologies. The relations between these concepts will be determined, specifying how concepts’ attributes are interconnected and how relations’ cardinalities exist. On the basis of this model, a set of rules will be proposed through which concepts from any one classification system can be expressed in those of another. Finally, testing an implementation of this rule-based model with sample data will address the issues put forward in the research question.

In his 2013 paper, Kierkegaard describes the current situation in Europe regarding the adoption of e-prescriptions [17]. He distinguishes between leading, trailing, and passive countries. To compare health terminologies to one another, the countries included in this paper should have reached a certain maturity with regard to their health care technology. Consequently, two leaders and two trailers were selected as case studies. The countries whose primary care situations will be explored are technology leaders the Netherlands (NL) and Denmark (DK), and trailing countries Germany (GM) and the UK.

After a preliminary investigation, a process involving terminology matching and mapping is performed to map values from one terminology to its equivalents in another, insofar as this is possible [18]. Bellahsene et al. [19, p. v] define matching as ‘the task of finding semantic correspondences between elements of two schemas’ and mapping as the task that ‘derives the relationships between elements and structures in heterogeneous schemas’; for mapping to be successful, matching should be completed beforehand. This process results in a comprehensive model detailing to what extent information exchange between these terminologies can be facilitated.

Finally, the model is evaluated against the data quality parameters of completeness and accuracy.

3 Prescribing in primary care

3.1 Concepts

The central process in primary care is the patient’s consultation of his or her physician. In its traditional arrangement, this process leads to the practitioner determining a diagnosis and prescribing medicine. One of the most commonly used approaches to structuring the prescription process is the Subjective, Objective, Assessment, and Plan method, which results in a set of concepts such as a diagnosed disease and a prescribed drug. The right-hand side of Fig. 1 displays the concepts resulting from a prescribing process structured by SOAP. These concepts correspond with those identified by the Dutch College of General Practitioners [20, p. 56].

The main action a General Practitioner (GP) performs in the first, subjective, step is recording the patient’s grievances, which leads to a complaint. After, optionally, performing one or more measurements, this is diagnosed as a disease. In the prescription process, this disease is treated by a drug, with a specified dosage. If a patient suffers drug allergies, the choice of medications available for treatment is reduced. Similarly, other diseases a patient suffers from, or comorbidities, may lead to exclusion of one or more drugs. These contraindications may be indicated by performed measurements, and result in excluding drugs or adjusting their dosages.

3.2 Terminologies

Within primary care, a multitude of terminologies exist to classify the concepts described in the prescribing process. Both national and international terminologies are in use to structure diagnoses, drug prescriptions, and measurements. Their use differs based on their purpose, geographical region, caretakers’ preference, and history. Table 1 shows how countries have implemented the international terminologies or, if they have not, which national system they use instead.

4 Relationships between terminologies

4.1 Matching

The process of matching encompasses the finding of values with similar meanings in two terminologies. It is a precursor for the mapping process, in which the terminologies are related in a schema, with regard to structures and cardinalities. In this paper, many terminology mappings were already available, and the matching process was limited to the few classification systems which had not been matched yet. The process of matching, in these cases, was performed through manual assignment. The authors would list all values from a certain terminology that needed matching, and would iterate over every other value from a different terminology looking for an equivalent. As the terminologies that needed matching were relatively small, this approach was feasible.

For most disease terminologies, mappings (and thus matchings) were readily available. The only exception to this was the Dutch ICPC1-NL variation; the sub-values were manually related to equivalents in ICPC2, where possible. Table 2 summarises the routes taken to exchange terms of any one disease terminology to another one.

As most dosage terminologies were not standardised, no mappings existed between them. By dividing the lists of common British and German SIG codes into the Dutch format for frequency, form, and instructions, they could be matched with Dutch equivalents. Use of the Dutch T25 table as an intermediate terminology in all cases makes presentation of a table for dosage routes similar to Table 2 superfluous.
4.2 Mapping

To create a model for information exchange between existing medical terminologies, the relations between them should be defined. As shown in the previous section, all concepts as shown in Fig. 1 can be modelled by at least one terminology (with the exception of complaint). The complaint is a (very) short statement by the patient about his or her illness; as it is subjective and uninterpreted, it cannot be linked to predefined terminologies. On the basis of Fig. 1, it can be concluded that the remaining concepts can be expressed as one of these four: disease, drug, dosage, and measurement. Contraindications, being either drug allergies or comorbidities, can be expressed as (restrictions on) drugs or as diseases, respectively. As explained above, complaints cannot be expressed through any terminology.

In the model, all terminologies containing similar concepts are related to one another. Where possible, existing mappings were used to determine the degree to which concepts of one terminology could be meaningfully expressed in another. Where existing mappings were not available, they were created specifically for this paper. With the exception of ICPC1-NL, all terminologies depicting diseases had mappings available. Between the dosage terminologies no mappings existed; only a limited mapping of the Dutch system to Logical Observation Identifiers Names and Codes (LOINC) has been published.

**Fig. 1** Process-data diagram depicting concepts relevant during a primary care consultation

**Table 1** Overview of terminologies implemented in four European countries

<table>
<thead>
<tr>
<th>Disease</th>
<th>Drug</th>
<th>Dosage</th>
</tr>
</thead>
<tbody>
<tr>
<td>UK</td>
<td>NL</td>
<td>GM</td>
</tr>
</tbody>
</table>

**Table 2** Overview of the routes when matching one disease terminology’s terms to those of another. \( \rightarrow \) indicate to which intermediate terminologies were mapped

<table>
<thead>
<tr>
<th>From</th>
<th>ICPC1-NL</th>
<th>ICPC2(-DK)</th>
<th>ICD10</th>
<th>ICD10-GM</th>
<th>READ-UK</th>
</tr>
</thead>
<tbody>
<tr>
<td>ICPC2(-DK)</td>
<td>ICPC1-NL</td>
<td>ICPC2a</td>
<td>ICD10b</td>
<td>ICD10b</td>
<td>ICD10b</td>
</tr>
<tr>
<td>ICPC1-NLa</td>
<td>ICD10b</td>
<td>ICD10b</td>
<td></td>
<td>ICD10b</td>
<td>READ-UKa</td>
</tr>
<tr>
<td>ICD10</td>
<td>ICPC2b</td>
<td>ICPC2b</td>
<td>ICD10c</td>
<td>ICD10c</td>
<td>READ-UKd</td>
</tr>
<tr>
<td>ICD10-GM</td>
<td>ICD10c</td>
<td>ICD10c</td>
<td></td>
<td>ICD10c</td>
<td>READ-UKc</td>
</tr>
<tr>
<td>READ-UK</td>
<td>ICD10d</td>
<td>ICD10d</td>
<td>ICD10d</td>
<td>ICD10-GMc</td>
<td></td>
</tr>
</tbody>
</table>

Numbers refer to the matching available in references. aManual assignment by the researchers, lacking available resources, b[7], c[8], d[21].

© The Institution of Engineering and Technology 2017
The results of this modelling approach can be seen in Figs. 2 and 3. Both international standards (e.g. ICPC2), regardless of whether they have been implemented, and actual national adaptations (e.g. ICPC1-NL) have been included. The relations’ cardinalities depict to what degree the terminologies’ concepts can be expressed in those of another. A 1–1 relationship would infer that all concepts of either terminology have exactly one match in the other. A 1–0..1 relationship defines that any concept in the first terminology has either one or no equivalents in the second terminology. Similarly, a 1–0..* relationship means that any concept of the first terminology has any number (including zero) of equivalent concepts in the other. The cardinalities are a comprehensive result of the matching process, which linked concepts from different terminologies to one another. This means that two terminologies between which a 0..*–0..* relationship exist may contain concepts for which 1–1 relations exist. These specific matching details will be addressed in the implementation phase and the creation of the inference rules, which will be described later.

4.2.1 Diseases: A total of six terminologies able to classify diseases have been included in the model. All of them describe concepts as code-description combinations, whereby the code acts as a unique identifier.

The international standards ICPC2 and ICD10 were originally mapped when the former was released. As ICD10 holds much more specific diseases, only a small number of its concepts have equivalents in ICPC2. On the other hand, most ICPC2 concepts can meaningfully be expressed in ICD10 codes; for some unspecified concepts such as L76 – fracture (other), manual intervention would be required to successfully select a corresponding ICD10 code out of a list of possibilities. Some subjective complaints such as P05 – feeling old, cannot be mapped to ICD10 concepts.

Three of the national terminologies studied, the Dutch, German, and Danish, are based on international standards. The Dutch ICPC1-NL is an amended version of the original ICPC1 terminology, enhanced for primary care practise with sub-concepts (such as 790.01 – diabetes mellitus type 1). Some of these sub-concepts were added to the international version when ICPC2 was released, enabling the successful mapping of a small number of ICPC1-NL sub-concepts to ICPC2 main concepts. The Danish version of ICPC2 is merely a translation of the original terminology, meaning that all concepts can be meaningfully transferred between the two. The German version of ICD10 was altered to better correspond to primary care needs. It excludes some diseases that are very rare in western Europe (such as bubonic plague) and chapters detailing morbidity and mortality causes. This results in most, but not all, concepts being mapped between the two terminologies.

The READ terminology, finally, is a separate classification having been developed and widely used in the UK. The official issue of READ codes by the British Health and Social Care Information Centre contains a mapping to ICD10. As it contains a very comprehensive terminology, most ICD10 codes can be successfully mapped to READ concepts, and vice versa.

4.2.2 Drugs and dosages: Unlike diseases, drugs are consistently classified internationally using the ATC terminology. While some countries employ additional national classification systems for diverse purposes, they are complementary to the international standard. The availability of drugs throughout different countries may differ; however, causing interoperability issues when transferring drug records between countries’ systems.

For the notation of drug dosages international standards do not exist. Countries have often implemented national approaches, which are usually created bottom-up and which differ in their degree of standardisation. Dosage notations exist of the frequency with which certain specified units should be taken, optionally augmented with instructions (such as take two pills once a day, at mealtimes).
The NL is the only country with a fully implemented dosage notation system. It uses the format \(XXTAB\), where \(X\) and \(Y\) are (ranges of) numbers and \(T, A, B\) and \(b\) are standardised strings. \(XT\) denotes the frequency (e.g. \(2D = \text{twice per day}\), \(Y\) is the form (e.g. \(1DR = \text{one drop}\)), and \(B\) are the optional instructions (e.g. \(ALCO = \text{be careful with alcohol in combination with this drug}\)).

While other countries support de facto standards for the \(T, A, B\) and \(b\) values, no recognised national standards are implemented. The UK uses a combination of Latin and English abbreviations as dosage notations (e.g. \(gt = \text{drop}\) and \(gr = \text{grain}\)). In GM, Latin and German terms are commonly in use (e.g. \(qs, quantum\ satis = \text{as many as required}, fein = \text{fine}\)), though there are no widely accepted de facto standards to denote frequency.

DK, finally, does not support any standardised format for dosage notations, except for specifying the number of units to be used daily (e.g. once a day two (tablets)).

### 4.2.3 Measurements:
Mappings between the two major laboratory terminologies, LOINC and Nomenclature for Properties and Units (NPU), do not exist. Only a limited, incomplete mapping of the Diagnostics table by the Dutch College of General Practitioners to LOINC could be accessed. Since all laboratory terminologies exist of, at least, thousands of entries, creating a custom mapping for this paper was not feasible. Consequently, no assessment of laboratory terminologies could be made, and in the remainder of this paper the measurement concept is therefore omitted.

### 5 Implementation
Information exchange between the terminologies is constrained by inference rules, which can be implemented through a rule engine. Objects used as input for the conversion process such as diseases or dosages are registered with the rule engine. If the rules are being fired, all objects are iterated through and applicable rules’ consequences executed. The conditions of the rules generally require one terminology to have filled-out values for a certain object. If these conditions are met, rules’ consequences consist of amending the object by adding equivalent values of other terminologies to it. The implementation does neither dictate the order in which objects should be iterated through, nor the order in which applicable rules should be run. The process simply continues until no more rules’ conditions are satisfied. A more thorough explanation of how this inference engine operates and how these types of rules were implemented can be found in our 2015 paper [24].

Figs. 4 and 5 below graphically depict the decision rules’ execution flow. Both processes are iterated for every object in working memory, and its rules are run parallel to each other. The decision nodes in both diagrams represent the conditions; if a condition is not met, either another one is tested or the process is terminated. For the sake of readability, routes leading to terminations are displayed as dashed lines. Consequences are modelled as actions and may incorporate formulae or database queries, which are not explicitly shown.

### 6 Evaluation

#### 6.1 Completeness
As described before, a key determinant of data quality is its completeness, which is the extent to which data’s expected attributes are present in a terminology. In a medical scenario, data would be consistently complete if it contained exactly the same information regardless of which terminology was used to describe it. For example, when describing dosages, all additional instructions \(b\) should be systematically included in all terminologies.

The model described above was tested for its completeness by measuring the extent to which its terminologies can contain each other’s values. Orme et al. [13] introduce many metrics for determining ontologies’ complexity and cohesion. These measurement methods were used as a starting point to formalise metrics for measuring the degree of completeness between any two vocabularies in an ontological model.

For each term in a terminology, the number of equivalent terms in another terminology is determined through a lookup table containing the mapping between the two.

The sections below include tables containing the results of these equations for the disease and dosage terminologies. The first number in each cell shows how many concepts of one terminology have equivalents in another, the second is the same number expressed as a percentage of the first terminology. For terminologies between which direct mappings exist, this was done by determining how many unique values had equivalents, and relating that to the complete number of terminology entries. For terminology pairs that required intermediate terminologies, results from each intermediate step were used as input for the next as sub-queries.

#### 6.1.1 Diseases: In Table 3, the results of the completeness test are displayed for the disease concept. Only limited mappings between these terminologies exist, usually between the most adjacent ones in terms of extensiveness. Mappings between ICD10-UK and ICD10, ICD10 and ICD10-GM, and ICD10 and READ-UK were used. Noteworthy observations are that the more extensive terminologies (ICD10 and READ-UK) are able to express higher percentages of simpler terminologies (ICPC) than the other way around. Most diseases included in ICD10 can be expressed in READ-UK as well. Only a limited set of ICD10 sub-codes can be expressed in ICD2 and similarly in more distinguished terminologies. READ-UK concepts can be expressed rather well in ICD10 and vice versa. When attempting to express concepts from these more extensive terminologies in the lower-level ICPC terminology, only very few can be successfully exchanged.

#### 6.1.2 Dosages: Table 4 shows the results of this test for the dosage concept. As is evident from Fig. 3, the Dutch SIG terminology is able to systematically contain all dosage sub-concepts (albeit not all possible values). As such, it was chosen as an intermediate to convert values from one terminology to another. Being the most extensive dosage terminology, SIG-NL can only be partially expressed in others. About 11% of its values can be expressed in the British SIG terminology, 4% in the German SIG terminology, and virtually none in the Danish one. Conversely, many values of other terminologies (between 44 and 100%) can be expressed in the Dutch terminology. Amongst the values that are consistent throughout SIG-NL and SIG-UK are especially many (\(x + \text{t}\) and \(a\) inputs such as \(D = d\) (day) and \(C = \text{caps} (\text{capsule})\). SIG-GM does not support systematic recording of \(t\), but limited categorisation of \(a\) such as \(T = \text{pil}\) (tablets). The Danish terminology only supports the number of non-specific units per day, enabling only \(t\)-value \(D = d\) (day).

About 18 of the \(a\) and \(b\) values in the British dosage terminology can be expressed as their German equivalents. Exactly the same number of concepts in SIG-GM can be expressed in the British terminology. These include \(\text{tab} = \text{pil}\) (tablet) and \(\text{aq} = \text{aqua}\) (add water). SIG-UK does support \(v\) values, so \(D = d\) (day), the only Danish SIG code, can be expressed as well. As SIG-DK does not support \(v\) values, it cannot be expressed in the Danish terminology.

#### 6.2 Accuracy
A second key determinant of data quality is accuracy, interpreted as the degree to which real-world objects are correctly represented by the data. In the medical scenario explored in this paper, data would accurately reflect real-world objects if it would retain meaningful information regardless of which terminology was used to express it.

Two ways in which conversions can yield inaccurate results can be distinguished: either erroneous information is added (e.g. when converting \(\text{diabetes mellitus}\) to \(\text{diabetes mellitus type A}\) or correct information is omitted (e.g. when converting \(\text{testicular cancer}\) to \(\text{cancer}\)). While qualitative judgement is necessary to determine
whether or not a matching between values of two terminologies is partially inaccurate, cardinalities can be used as indicators. Given that terminologies contain no synonyms, any one-to-many relationship would mean that all conversions, except at most one, contain partial inaccuracies, i.e. do add superfluous information. After all, if a concept diabetes mellitus in a certain terminology has relations to three other terminologies in a different classification system, only one of these could be exactly similar; the other two would have to be sub-concepts of the disease (such as diabetes mellitus types A and B). Similarly, any many-to-one relationship between terminology values, except at most one, would result in loss of accuracy.

6.2.1 Diseases: The differences between the Dutch version of ICPC1 and the later ICPC2 are relatively few. As ICPC2 was designed to be backwards compatible with its earlier version, all concepts which were included in the original ICPC1 can be expressed in ICPC2, and vice versa. Incompatibility problems arise when trying to express the Dutch sub-concepts in ICPC2; a detailed recording of S12.01 – mite bite can only be expressed in ICPC2 as the more generic S12 – insect bite or stinging, thus losing information. In fewer instances, ICPC2 concepts cannot be mapped to ICPC1-NL (sub) concepts without losing details, or at all; A11 – chest pain was introduced in ICPC2 and has no ICPC1-NL equivalent.

While a relatively high percentage of ICPC2 concepts can be expressed in ICD10 codes, the difference in perspective between the two classification systems hinders accuracy. While both terminologies strive to categorise diagnoses, ICPC2 also includes subjective complaints as they are experienced by patients. For example, ICPC2 contains a code-description pair H13 – plugged feeling in ear; the nearest equivalent in ICD10 is H93.8 – other specified disorders of ear. The latter code is neither a specified disorder nor an accurate description of the complaint. Vice versa, ICPC2 cannot accurately express many of the more detailed diagnoses included in ICD10. For example, the nearest ICPC2
equivalent for the ICD10 concept E12 – malnutrition-related diabetes mellitus is T90 – diabetes mellitus. In this case, the specific causes of the disease are not retained.

The Danish version of ICPC2 is merely a translation and has exactly the same expressive capabilities as the international standard. The same is not true of the German version of ICD10; it is as much an adaptation of the English standard as a translation.

Concepts recording causes for morbidity and mortality were deemed unnecessary in primary care practice, and therefore excluded; an example is B20.2 – cytomegalovirus following human immunodeficiency virus infection. Additional concepts deemed useful in primary care practice were added such as D70.0 – compromised immune system after radiation; no equivalent for this concept in the standard ICD10 issue exists.

The British READ terminology, because of its larger scope, supports a far larger number of concepts than the other terminologies. Consequently, all concepts present in the implemented ICPC and ICD10 versions can be expressed meaningfully in READ code-description combinations, while only a limited number of READ concepts can be reflected in other terminologies. Concepts as detailed as X40.0 – congenital
To:

<table>
<thead>
<tr>
<th>To:</th>
<th>ICPC1-NL</th>
<th>ICPC2-(DK)</th>
<th>ICD10</th>
<th>ICD10-GM</th>
<th>READ-UK</th>
</tr>
</thead>
<tbody>
<tr>
<td>From:</td>
<td>ICPC1-NL</td>
<td>738 (57.43%)</td>
<td>698 (54.32%)</td>
<td>626 (48.72%)</td>
<td>665 (51.75%)</td>
</tr>
<tr>
<td></td>
<td>ICPC2-(DK)</td>
<td>719 (99.04%)</td>
<td>686 (94.49%)</td>
<td>614 (84.57%)</td>
<td>652 (89.81%)</td>
</tr>
<tr>
<td></td>
<td>ICD10</td>
<td>2,832 (23.35%)</td>
<td>2,893 (23.85%)</td>
<td>9,975 (82.23%)</td>
<td>10,578 (87.20%)</td>
</tr>
<tr>
<td></td>
<td>ICD10-GM</td>
<td>2,737 (24.16%)</td>
<td>2,779 (24.53%)</td>
<td>9,975 (88.06%)</td>
<td>8,665 (76.50%)</td>
</tr>
<tr>
<td></td>
<td>READ-UK</td>
<td>5,658 (9.46%)</td>
<td>5,729 (9.58%)</td>
<td>51,622 (86.31%)</td>
<td>47,731 (79.80%)</td>
</tr>
</tbody>
</table>

Absolute values denote the number of terms in the vertical terminology that can be expressed in the horizontal terminology, values between brackets are the absolute values expressed as percentages of the vertical terminology.

To:

<table>
<thead>
<tr>
<th>To:</th>
<th>SIG-NL</th>
<th>SIG-UK</th>
<th>SIG-GM</th>
<th>SIG-DK</th>
</tr>
</thead>
<tbody>
<tr>
<td>From:</td>
<td>SIG-NL</td>
<td>89 (11%)</td>
<td>37 (4%)</td>
<td>1 (0%)</td>
</tr>
<tr>
<td></td>
<td>SIG-UK</td>
<td>109 (44%)</td>
<td>18 (7%)</td>
<td>1 (0%)</td>
</tr>
<tr>
<td></td>
<td>SIG-GM</td>
<td>45 (55%)</td>
<td>18 (22%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td></td>
<td>SIG-DK</td>
<td>1 (100%)</td>
<td>1 (100%)</td>
<td>0 (0%)</td>
</tr>
</tbody>
</table>

Absolute values denote the number of terms in the vertical terminology that can be expressed in the horizontal terminology, values between brackets are the absolute values expressed as percentages of the vertical terminology.

7 Discussion

In this paper, the researchers have explored the extent to which primary care terminologies can meaningfully express each other’s concepts. To determine this, a model containing these terminologies and including axioms to define their relations was developed. Generic metrics were designed to determine the completeness and accuracy of any two arbitrary vocabularies within an ontological model. Tests investigating completeness and accuracy show that simpler disease terminologies such as ICD10 can be expressed relatively well in more extensive ones such as SNOMED-CT; mainly concepts which do not equate to diagnoses such as subjective complaints (e.g. plugged feeling in ear), cannot be mapped to ICD10 concepts. Vice versa, only a limited number of concepts of extensive terminologies can be mapped to simpler ones and still retain their meaning; the highly specific nature of many ICD10 entries (e.g. malnutrition-related diabetes mellitus) makes accurate matching with ICD concepts difficult.

With the exception of the NL, standardisations for dosage notations do not exist. When testing de facto notation standards, it appears that only the most common abbreviations have equivalents in most terminologies. Completeness among these classification systems is thus negligible. Only the most common terms have unambiguous counterparts in most terminologies. The accuracy for more complex or less frequently used terms is compromised; terms can often only be partially matched to others (e.g. add water, not milk, aqua, and water).

The inability to meaningfully exchange large numbers of concepts between primary care terminologies creates significant semantic interoperability challenges. These are commonly addressed by advocating the use of an overarching terminology such as SNOMED-CT or the UMLS Metathesaurus. Matching concept of primary terminologies to metaterminologies is not necessarily a flawless process, however. Schulz et al. [25] note structural deficiencies between ICD and SNOMED that impair matching between them. In an attempt to map the Swedish translation of ICD10 to SNOMED-CT, Vikström et al. [26] find that even if extensive mapping rules would exist: 'obstacles to high-quality mapping remain due to structure and content characteristics in both coding systems'. Thus, even if the majority of primary care concepts have defined meanings in a metaterminology, this does not guarantee complete semantic interoperability. Therefore, in this paper we pursued an alternative approach omitting the need for metaterminologies.

To achieve a next step in semantic interoperability of primary care data, national terminologies should be unambiguously mapped to metaterminologies, or replaced by subsets of metaterminologies. Commonly used, but unstandardised notation systems (such as SIG codes) should be formally standardised and incorporated in...
metaterminologies. Without standardisation, we have shown that a direct mapping approach between low-level terminologies can yield beneficial results, and can be used next to more traditional ontological approaches involving metaterminologies in order to achieve optimal mapping results.

8 References

[7] Olofsson, I., Oskam, S., Lamberts, H.: ‘ICPC in the Amsterdam transition project’, Academic Medical Center/University of Amsterdam, Department of Family Medicine, Amsterdam, 2005