Registration in practice: Comparing free-text and compositional terminological-system-based registration of ICU reasons for admission

Nicolette de Keizer [a], Ronald Cornet [a], Ferishta Bakhshi-Raiez [a], Evert de Jonge [b]

[a] *Dept of Medical Informatics,* [b] *Dept of intensive Care, Academic Medical Center, Universiteit van Amsterdam, The Netherlands*

Abstract

Reusability of patient data for clinical research or quality assessment relies on structured, coded data. Terminological systems (TS) are meant to support this. It is hardly known how compositional TS-based registration affects the correctness and specificity of information, as compared to free-text registration. In this observational study free-text reasons for admission (RfA) in intensive care were compared to RfAs that were composed using a compositional TS. Both RfAs were registered in the Patient Data Management System by clinicians during care practice. Analysis showed that only 11% of the concepts matched exactly, 79% of the concepts matched partially and 10% of the concepts did not match. TS-based registration results in more details for almost half of the partial matches and in less details for the other half. This study demonstrates that the quality of TS-based registration is influences by the terminological system's content, its interface, and the registration practice of the users.

Keywords: Terminological system, information storage and retrieval, medical records, evaluation

1. Introduction

Most potential advantages of electronic patient records, such as availability of patient data for decision support and the re-use of patient data for clinical research or quality assessment [1], rely on structured, coded data, not free text [2]. Structured data entry (SDE) [3] and terminological systems (TS) [4] are means to support this process of capturing patient data in a structured and standardized way. SDE is a method by which clinicians record patient data directly in a structured format based on predefined fields for data entry. Terminological systems provide terms denoting concepts and their relations from a specific domain [5] and can be used within predefined fields for data entry.

Nowadays most terminological systems do have a computer-based implementation. Terminological systems can either enumerate all concepts (pre-coordination), or allow post-coordination, i.e. enabling to compose new concepts by qualifying pre-coordinated concepts with more detail. Generally it takes longer to select and post-coordinate concepts corresponding to a patient's findings, diagnoses, or tests from long lists of standard terms drawn from terminological system than to enter a summary in free text. Worse, the standard codes and terms provided by a terminological system may constrain clinical language [6]. Although the disadvantages of capturing structured, coded data might be outweighed by more informative data and automatic processing of data, evidence on the effect of structured and TS-based registration of patient data on the correctness and specificity of these data compared to free-text is hardly available. Many studies compared the content coverage (correctness and specificity) of a TS by retrospectively coding a set of diagnoses [7]. Studies in which the feasibility of automated coding has been investigated also usually use an experimental design in which free text from a medical record is coded retrospectively by some natural language processing algorithm (e.g. [8,9]). Cimino et al [10] use an observational, cognitive-based approach for differentiating between successful, suboptimal, and failed entry of coded data by clinicians. They used the Medical Entities Dictionary (MED) which only included pre-coordinated concepts. To our knowledge no observational field studies exist in which free-text recording in a medical record is compared with prospectively recorded compositional TS-based diagnoses.

The aim of this observational study is to evaluate how clinicians in every day care practice register reasons for admission (RfA) by using compositional TS-based systems. TS-based registration was compared to free-text registration with regard to correctness and specificity of recorded RfA.

The outcome of this study depends on three factors: the terminological system's content, its interface, and the registration practice of the users. In this study, we aim at distinguishing the effect of content from the effect of the user interface and the user. If structured

TS-based registration of diagnoses results in (at least) the same information as free-text diagnoses, TS-based registration is preferred, as retrieval will be much easier and thereby re-use of the data will be much more feasible. If TS-based registration results in information loss we need to investigate the reasons for this to search for possibilities to improve the terminological system and its use.

2. Materials & Methods

2.1 PDMS and Terminological system DICE

This study took place in an adult Intensive Care Unit with 24 beds in 3 units, with more than 1500 yearly admissions. Since 2002, this ward uses a commercial Patient Data Management System (PDMS), Metavision. This PDMS is a point-of-care Clinical Information System, which runs on a Microsoft Windows platform, uses a SQL server database and includes computerized order entry; automatic data collection from bedside devices such as a mechanical ventilator; some simple clinical decision support; and (free-text) clinical documentation of e.g. reasons for admission and complications during ICU stay. As part of the National Intensive Care Evaluation (NICE) project [11], a national registry on quality assurance of Dutch ICUs, for each patient a minimal dataset among which the reason for admission is extracted from the PDMS. Since April 1st 2005 a pilot study is running in which the compositional terminological system DICE [12] is integrated with the PDMS (see Figure 1) to evaluate its usability for structured registration of reasons for ICU admission. The main reasons for the development of DICE were the need for a terminological system that supports a) registration of intensive-care-specific reasons for admission, commonly either a severe acute medical condition or observation after a large surgical condition b) semantic definitions of concepts, enabling selection of patients by aggregating diagnoses on different features, and c) assignment of multiple synonymous Dutch and English terms to these concepts.

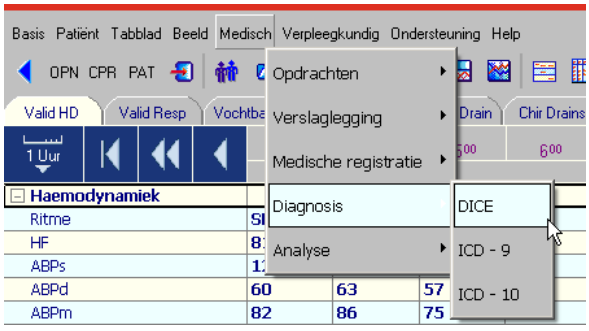

Figure 1: Activation of TS-based registration within the Patient Data Management System

DICE implements frame-based definitions of diagnostic information for the unambiguous and unified classification of patients in Intensive Care medicine. DICE defines more than 2400 concepts including about 1500 reasons for admission and uses 45 relations. DICE is implemented as a SOAP-based Java terminology service together with clients for knowledge modeling and browsing [13]. DICE is used to add controlled compositional terms to clinical records. The implementation of DICE offered the physicians two ways to search for the appropriate diagnosis concept: (a) a short list containing the most frequently occurring diagnoses, (b) entry of (a part of) its preferred or synonymous term. Once a concept is selected, DICE uses post-coordination to provide concepts with more detailed information, as shown in Figure 2. The user interface of the client by which concepts are browsed stimulates but does not enforce users to specify additional qualifiers of a concept, e.g. a Coronary Artery Bypass Graft (CABG) can be further qualified by the number of bypasses; the types of bypasses and whether it was a re-operation or not.
At the start of the pilot physicians got a 15-minutes training on the use of DICE. During the pilot, registration of DICE-based reasons for admission as part of the NICE minimal dataset was voluntary. This means that after the first 24 hours of ICU admission a physician could add a controlled term from DICE into the PDMS to describe the reason for ICU admission. As the reason for admission is an essential part of the clinical documentation the regular registration of free-text-based reasons for admission into the PDMS was continued during the pilot for each patient at the time of admission.

2.2 Data collection and analysis

For all patients admitted between April 1st 2005 and December 1st 2005 the free-text reasons for admission

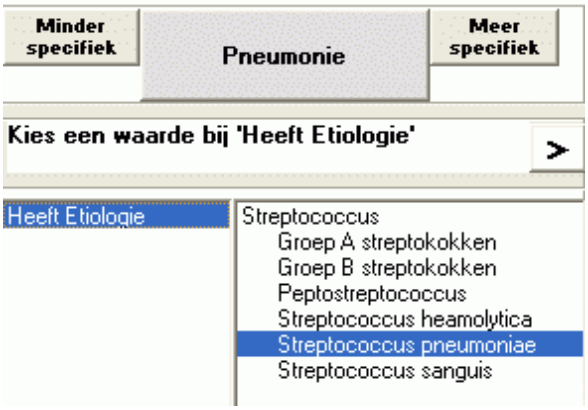

Figure 2: User interface presenting options for post-coordination

and (if available) the structured DICE-based reasons for admission were extracted from the PDMS. As free-text recording of reasons for admission is mandatory, for all patients admitted to the IC a free-text description was available. Since DICE-based registration of reason for admission was voluntary it could be possible that "difficult or complex" reasons for admissions were not registered with DICE. To investigate this possible selection bias the free-text reasons for admission were compared between the groups with and without structured DICE-based reasons for admission.

For each admission having both a free-text reason for admission and one or more DICE-based reasons for admission, these reasons for admission were compared by two independent researchers, both experienced in DICE and intensive care medicine. Each pair consisting of one free-text and one or more DICE-based reason for admission was scored as either being an exact match, partial match or mismatch. A match was considered exact when the DICE-based reason for admission was semantically equivalent to the free-text registration. For example the abbreviated free-text "AVR" was considered an exact match with the DICE concept "aortic valve replacement". A concept pair was considered as partially matching when one concept subsumed the other (e.g. "3-fold CABG" and "CABG") or when the concepts were siblings with equal anatomical and pathological properties (e.g. "hepatitis A" and "hepatitis B"). A concept pair is considered a mismatch in all other cases. For each partial match the two researchers independently assessed which concepts, attributes or relations were missing or were additionally represented in the DICE-based reason for admission. Comments on missing details in the DICE-based registration were classified either as a) "not registered but available in DICE", b) "value of relation is missing in DICE", e.g. although a CABG can be qualified by type of graft (LIMA, RIMA, PIMA and venous) the value "LIMA-lad" is missing or c) "relation is missing in DICE", e.g. "bleeding of the cerebellum, right side" can not completely be registered by DICE since the relation "laterality" is missing.

Different scores of the researchers were solved based on consensus and if necessary by asking an intensivist as an independent third party.

Figure 3 presents an example of a partial match. The free text "AVR-bio + CABG" coming from the clinical documentation part of the PDMS is displayed at the top of the screen dump. In the middle of Figure 3 the DICE-based reasons for admission are presented and at the bottom the scoring of agreement, in this case a partial match, is presented. A "+" indicates that the

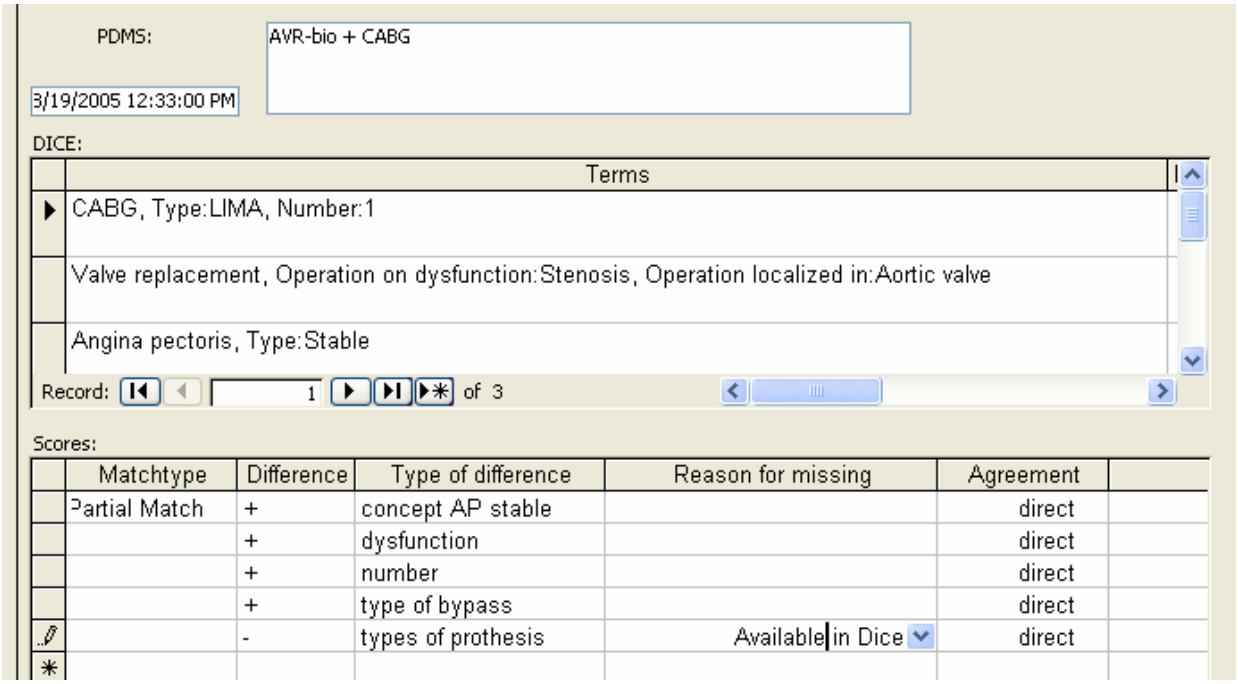

Figure 3: Scorings example of the agreement between free-text "AVR-bio + CABG" and the accompanying set of DICE-based reasons for admission. The bottom part represents the match type, the difference ("+" means DICE has additional detail," –" means DICE misses detail), the type of difference, the reason for missing (type of prosthesis is available in DICE) and if the two researchers directly agreed on the differences or after discussion.

DICE based registration includes more detail than the free-text registration on type of CABG, number of bypasses, dysfunction of the aortic valve and the Angina Pectoris diagnosis. The "-" indicates that the free-text registration includes details on the type of valve prothesis which is not registered in the DICE-based registration, although this qualifier is available in DICE. In this example all differences between the free-text and DICE-based reasons for admission were scored by both researchers which is indicated by "direct" agreement.

In this paper a TS-based diagnosis is regarded as correct when it exactly or partially matches the free-text diagnosis. Specificity of (correct) diagnoses is expressed by as "equal" (exact match), "more specific", "less specific" or "more and less specific" depending on differences in detail of the TS-based diagnoses compared to the free-text diagnoses.

3. Results

During the study period 799 admissions to the ICU took place. For all these admissions a free-text reason for admission was available and for 359 (45%) of these admissions a DICE-based reason for admission was available. Those admissions for which a DICE-based registration was missing do not represent other reasons for admissions than those for which a DICE-based registration was available. One free-text reason for admission could be described by more than one DICE-based reason for admission, e.g. "CABG + AVR" is one free-text reason for admission encoded by two DICE concepts "CABG" and "Aortic valve replacement". The 359 free-text reasons for admission were described by 457 DICE-based reasons for admission. Half of them were registered as pre-coordinated concepts such as "Pneumonia", half of them were registered using post-coordination, e.g. "Pneumonia; has aetiology Staphylococcus aureus".

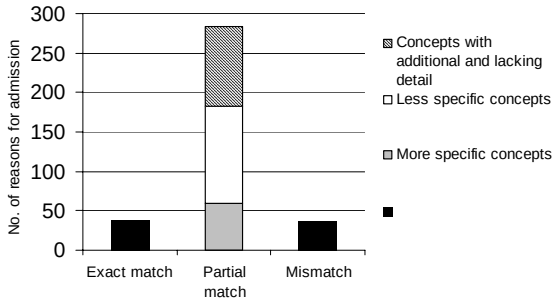

Figure 4: Distribution of exact match, mismatch and partial match (including whether the DICE based reason for admission included more and/or less specific detail).

Figure 4 shows that we found 38 (11%) exact matches, 284 (79%) partial matches and 37(10%) mismatches.

Table1. Example of 5 exact matches, 5 partial matches and 5 mismatches

	Free-text diagnoses	**DICE-based diagnoses**
Exact matches	THOCR	Oesophageal cardiac resection, entrance: transhiatal
	SAB	Subarchnoid bleeding
	re-CABG x2 venous	CABG, Re-operation: true, Type:Venous graft, Number:2
	Staphylococcal sepsis	Sepsis, has etiology: Staphylococcus aureus
	Stomach bleeding	GI bleeding; localized in stomach
Partial matches	SAB	Subarchnoid bleeding; closing: coil
	Respiratory insufficiency	Respiratory insufficiency; due to: pneumonia
	CABGx3 and Ao-biovalve	CABG & valve replacement
	Respiratoire insufficiency bij benzodiazepine intoxicatie	Accidental intoxication with sedatives and hypnotics
	Large posterior infarction	Acute pulmonary oedema ; due to acute myocardial infarction
Mismatches	Abdominal bleeding	Renal insufficiency
	Hypercapnia with reduced consciousness	COPD
	Hyponatremia with cerebral oedema	Self intoxication
	Resp insufficiency after cardiogenic shock	Myocardial infarction
	Respiratory insufficiency due to pneumonia	Perforated gallbladder

According to our definition 90% ((38+284)/359) of all concepts were correct but for 79% of all concepts (all partial matches), there were some discrepancies in specificity. One-third of the partial matches add some details as well as miss some details compared to the free-text reason for admission. Twenty-two percent of the partial matches was more specific and forty-four percent of the partially matches was less specific compared to the free-text reason for admission. Table 1 shows some examples of exact matches, partial matches and mismatches.

In total 582 comments were given on the 284 partially matched reasons for admission. Two hundred sixty (45%) comments were given on *additional* concepts, attributes or relations registered in the DICE-based registration of reasons for admission that were not described in the free-text reason for admission. On the other hand 325 (55%) comments were given on *missing* concepts, attributes or relations in the DICE-based registration of reasons for admission compared to the free-text reasons for admission.

Figure 5 shows the distribution of the 325 reasons why the DICE-based reasons for admission were missing detail. The majority (65%) of the details presented in free text but missing in the DICE-based registration was available in the DICE terminological system, but was not used by the clinicians.

The largest group of reasons for admission consisted of patients who were admitted to the ICU after cardiac surgery such as CABG and heart valve operations (n=112). In this patient group we found 95% correct concepts: 6(5%) exact matches, 100(90%) partial matches and 6(5%) mismatches. Among the partial matches the DICE-based registration of cardiosurgical reasons for admission contains more detail in 48% of the cases compared to the free-text registered ones. The main reason for missing detail in the remaining 52% cases is caused by the lack of a relation to describe the area of the heart to which the new graft is located, e.g. “CABG, LIMA-LAD” can be coded in DICE as “CABG, Type: LIMA” but without “LAD”.

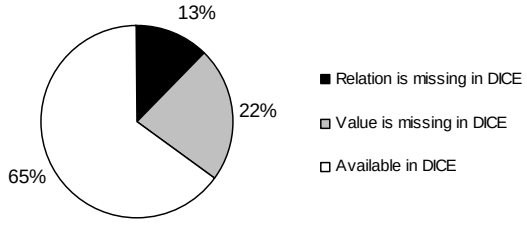

Figure 5: Reasons for missing detail in DICE-based registration of reasons for admission.

Table 2. Match scores for reasons for admission (RfA) on or not on the list of most frequently occurring reasons for admission.

	RfA on short list	RfA not on short list	All registered RfA
Mismatch	21 (7%)	17 (23%)	38 (11%)
Partial match	233 (82%)	51 (69%)	284 (79%)
Exact match	31 (11%)	6 (8%)	37 (10%)
Total	285 (100%)	74 (100%)	359 (100%)

As described above the DICE user interface supports two ways to search for the appropriate diagnostic concept: using a short list or entering (a part of) a term. Table 2 shows the scores for reasons for admission split up for those that could be selected from the short list of frequently occurring reasons for admissions and those that were not on this list. Twenty percent (n=74) of all reasons for admission was not on the short list of frequently occurring reasons for admission.

Reasons for admission that could be selected from the short list were scored differently from those reasons for admission that were not represented on this list (Chi-Square $p<0.001$). Significantly more mismatches were scored among the reasons for admission that were not on the short list.

In 82% of the cases the two researchers directly agreed on the assigned scores, disagreement on the other 18% was easily resolved after short discussion.

4. Discussion

Terminological systems offer the possibility to structure and standardize medical data, which improves the re-usability of these data for clinical research and quality assessment. In this study we compared the correctness and specificity between prospectively collected TS-based reasons for admission and free-text-based reasons for admission. We focused on the recorded data as such without taking into account the clinical consequences of the correctness and specificity of these data. We analyzed 359 reasons for admission to a Dutch Intensive Care registered in the PDMS by clinicians during actual care practice by using free text as well as by using the DICE terminological system. According to our definition 90% of the concepts were correctly registered based on the terminological system DICE. Only 11% of the cases had a perfect match. However, a partial match could be measured in 79% and there

were only 10% mismatches. One should be aware that if we change our definition of correctness to only "concepts with a perfect match" a completely different conclusion appears.

Among the partial matches about half of the TS-based reasons for admissions had additional detail compared to the free-text reason for admission. A possible explanation of this result could be the functionality of the terminology service in which users are encouraged to further specify a medical concept by additional qualifiers. Sixty-five percent of the information that is lacking in the other half of the partial matches was available in DICE but was not specified by the users. Further training and an improved user interface can contribute to improving these recorded reasons for admissions. Medical concepts on the short list of frequently occurring reasons for admission, counting for 80% of all reasons for admission, do have a better score than those not on this list. This is not a surprising result as the frequently occurring reasons for admission have got more attention during the modeling process of the terminological system than those not on the list. The reasons for missing concepts, attributes or relations gave us good insight into possibilities for (simple) improvements in DICE. For example the concept CABG could be extended with an attribute to describe which area of the heart is supported by the new graft. However, although we used free-text reasons for admission as they were recorded in daily care practice as a kind of golden standard, we observed many cases in which the TS-based registration included more detail than the free-text reasons for admission. Further research is necessary to determine the relevance of the details present in free-text as well as in the TS-based registration.

One weakness of our study is that the moment on which the free-text reason for admission is registered is not exactly the same as the moment on which the DICE based reason for admission has been registered. Although both reasons for admission were registered in the first 24 hours of admission, changing insight into the patient's condition could be an explanation for the discrepancy (partial match or mismatch) between the free-text reasons for admission and the DICE-based reason for admission. We will investigate this in further research. Another weakness is the fact that TS-based registration and free-text registration have not necessarily been done by the same physician. However, when two different physicians recorded the reason for admission of a particular patient both physicians were directly involved in treating the patient and hence both knew the patient's condition very well. Finally, there are no clear registration rules regarding what constitutes a reason of admission of a patient. As mismatches seemed to be mainly caused by above mentioned limitations of the registration process rather than the terminological system, they have not been further investigated.

According to other studies in which the quality of structured and standardized registration of medical data was audited our study has a strong surplus value because this data comes from a real-practice situation and is not collected retrospectively in an experimental setting. Physicians in our observational study who recorded the reasons for admission treat the patients and were not informed that DICE-based reasons for admission would be compared to free-text reasons for admission. In studies such as [14-16] patient cases were selected, and structured, coded data were obtained by independent physicians or coders without a direct clinical relation with the patient.

The aim of our study corresponds most with [10] as both studies observe coding behavior of clinicians in actual practice. Although different methods are used (cognitive approach vs. document analysis) both studies compare TS-based registration with some kind of free text. We used written text while Cimino et al used video-taped spoken text. Cimino et al found a larger amount of exact matches than we did. Differences in definitions of match types partly explain this. Furthermore, the differences in results might be partly explained by the fact that in [10] TS-based registration took place at the same time as free-text registration and because of other methods used. Furthermore, in [10] not only diagnoses but also drug information is included. The main difference between the two studies, however, is that our study used a compositional TS instead of MED which only contains pre-coordinated concepts. The availability of post-coordination might have a large influence on the specificity of recorded diagnoses. Our study confirms the findings of Cimino et al. that correctness and specificity of TS-based registration depends on three factors: the terminological system's content, its interface and the registration practice of the users.

5. Conclusions

This study shows that comparing free-text registration of reasons for admission with TS-based registration of reasons for admission only 11% of the concepts exactly matched and 79% of the concepts partially matched. TS-based registration added details in almost half of these partially matches and missed details in the other half. The methods used in this

study provide insight into possibilities for further improvement of the content coverage of DICE. However, 65% of the information not captured by the TS-based reasons for admission was available in DICE, indicating that user interaction with the system is more of an impediment than the contents of the TS. This study shows that availability of concepts and qualifiers in a TS does not guarantee that physicians will use them all. We expect that this result is generalizable to other terminological systems using post-coordination such as SNOMED CT. Further research is needed to investigate how physicians will be optimally supported in compositional TS-based registration.

References

1.van Ginneken AM. The computerized patient record: Balancing efforts and benefit. Int J Med Inf 2002:65(2):97-119.

2. Wyatt JC.Clinical data systems, part II: components and techniques. Lancet 1994; 344: 1609-14.

3. Moorman PW, van Ginneken AM, van der Lei J, van Bemmel JH. A model for structured data entry based on explicit descriptional knowledge. Methods Inf Med 1994; 33(5):454-63.

4. Rossi Mori A, Consorti F, Galeazzi E. Standards to support development of terminological systems for healthcare telematics. Methods Inf Med 1998; 37(4-5): 551-63.

5. de Keizer NF, Abu-Hanna A, Zwetsloot-Schonk JH. Understanding terminological systems I: Terminology and typology. Methods Inf Med 2000;39(1):16-21.

6. Prowsner SM, Wyatt JC, Wright P. Opportunities for and challenges of computerization. Lancet 1998; 352(9140): 1617-22.

7. Arts DG, Cornet R, de Jonge E, de Keizer NF. Methods for evaluations of medical terminological systems. A literature review and a case study. Methods Inf Med 2005;44(5):616-25..

8. Elkin PL, Brown SH, Bauer BA, Husser CS, Carruth W, Bergstrom LR, Wahner-Roedler DL. A controlled trial of automated classification of negation from clinical notes. BMC Med Inform Decis Mak. 2005;5;5(1):13

9. Lussier YA, Shagina L, Freidman C. Automating SNOMED coding using medical language understanding: a feasibility study. In proceedings AMIA Annual Symposium, Washington DC (2001):418-22.

10. Cimino JJ, Patel VL, Kushniruk AW. Studying the human-computer-terminology interface. JAMIA 2001;8(2):163-73.

11. de Keizer N.F, de Jonge E. National IC Evaluation (NICE) : a Dutch quality control system. J ICU management 2005; 3:62-64.

12. de Keizer NF, Abu-Hanna A, Cornet R, Zwetsloot-Schonk JHM, Stoutenbeek CP. Analysis and design of an intensive care diagnostic classification. Methods Inf Med 1999; 38: 102-112.

13. Cornet R, Prins AK. An architecture for standardized termiology services by wrapping and integration of existing applications. In proceedings AMIA Annual Symposium, Washington DC (2003):180-4.

14. Los RK, Roukema J, van Ginneken AM et al. Are structured data structured identically. Investigating the uniformity of pediatric patient data recorded using OpenSDE. Method Inf Med 2005;44:631-638.

15. Brown PJ, Warmington V, Laurence M, Prevost AT. Randomised crossover trial comparing the performance of Clinical Terms Version 3 and Read Codes 5 byte set coding schemes in general practice. BMJ. 2003 May 24;326(7399):1127.

16.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. Journal of the American Medical Informatics Association 1997;4:238-251

Acknowledgement

We would like to thank Antoon Prins for implementing the DICE application.

Adress for correspondence

N.F. de Keizer, Academic Medical Center, dept Medical Informatics, POBox 22700, 1100DE Amsterdam, The Netherlands. Email: n.f.keizer@amc.uva.nl