*Medical Informatics Europe '97*
*C. Pappas et al. (Eds.)*
*IOS Press, 1997*

# A Formal Model of Diabetological Terminology and Its Application for Data Entry

C. Birkmann[*], T. Diedrich[§], J. Ingenerf[*], J. Rogers[+], W. Moser[*], R. Engelbrecht[*]
[*]medis - Institut für Medizinische Informatik und Systemforschung, GSF - Forschungszentrum für Umwelt und Gesundheit, Ingolstädter Landstraße 1, D-85764 Neuherberg, Germany
[+]Medical Informatics Group, Depart. of Computer Science, University of Manchester, Manch. M13 9PL, UK
§DFI - Diabetes Forschungsinstitut Düsseldorf, Auf'm Hennekamp 65, D-40225 Düsseldorf, Germany

**Abstract.** This paper summarises the developmental activities for an electronic patient record system in diabetology based on GALEN technologies. It focusses on the modelling of primarily terminological medical knowledge of this subspecialty and describes its application for predictive data entry.

## 1. Introduction

A formal model of terminological knowledge and a corresponding structured data entry system as part of an electronic patient record system (EPR) are being developed for the subspecialty diabetology and will lateron be integrated into a GP-diabetes management system. The final clinical workstation is intended to support clinicians in their care and management of diabetes and will allow to enter and manipulate administrative and medical patient data. The current developments based on GALEN technology are part of the EU-funded Galen-In-Use project, which establishes Terminology Servers for medical applications based on a computer-supported model of medical terminology.

The GALEN-based structured data entry system is to demonstrate that the GALEN technology improves the quality and consistency of the captured data and further enhances the maintainablility and scalability of medical record systems. It will thereby profit from a semantically sound representation of medical terminology. The data entry system will enable the user to enter data in a structured and standardised manner and in some cases as free text also. The GALEN approach inherently supports the generation of dynamic, context-relevant forms and thus enables intuitive, quick, and easy entry of detailed clinical data. In a later phase of the project the diabetes management system will furthermore demonstrate the added functionality the GALEN approach offers: multilingual appearance, report generation, retrieval of data, and data driven decision support.

## 2. Data Entry in Electronic Patient Record Systems

The demand for electronic patient record systems in Germany is increasingly met. Therefore, it is now important to offer high-quality electronic systems with improved medical content and functionality. Data entry by commercially available electronic patient record systems is not as easy and consistent as desirable. Research in electronic patient record systems attempts to find practicable and convincing solutions to improve data entry [1, 2]. The GALEN project aims to permit easy and consistent data entry by tackling the problem of terminology [3].

Terminology is considered by GALEN as software. Terminological services (e.g. the conversion between natural languages) will be provided based on a centrally shared, semantically valid model of clinical terminology. This terminological model which forms the heart of the Terminology Server supports computer-based systems and allows detailed clinical information to be captured, represented, manipulated, and displayed.

## 3. Diabetes Model

### 3.1 GALEN Technology

The diabetological terminology is represented with the representation language of the GALEN project, the Galen Representation And Integration Language (GRAIL), resulting in the diabetes model. GRAIL allows to represent medical concepts compositionally with elementary clinical concepts, attributes combining these elementary concepts, and resulting complex concepts [4]. In addition a constraint mechanism with several levels of sanctioning allows to construct only sensible complex concepts [4].

The Terminology Server combines separate modules each of which serves a different aspect of the provided terminological services [3]. The concept module organizes terminological knowledge, whereas the extrinsic module links non-terminological to terminological knowledge [3]. The other modules are the code conversion module and multilingual module which are mapping concepts to codes or natural languages and vice versa [3].

### 3.2 Modelling Terminological Knowledge

The diabetes model has been developed systematically extending the concept module and specifically for the diabetological data entry system. It is based on the terminological knowledge of the standardised DIABCARD dataset version 2.0 [5]. This dataset contains medical information concerning signs and symptoms, findings, diagnoses, and the therapy of diabetes and other relevant diseases. Tool was the Knowledge Management Environment integrated Terminology Server (KiT) [6]. The non-terminological knowledge of the dataset has been added to the extrinsic module or the application model which is integrated in the concept module.

The modelling process was characterised by different ways of proceeding: Basic structures and constructs had to be proposed, consolidated or revised. A consistent model based on the source characterised above was developed rapidly. Special solutions were found for quaint details required by the source and not easily fitted into the preexisting terminological model.

The terminological diabetes model includes a dictionary of concepts and a grammar. Nearly 1200 items of the data source have been modelled including their relationship to each other as provided by the data source.

A number of elementary concepts relevant for general medical applications as well as specific diabetological applications has been added to the preexisting concept structure. Clean taxonomies were kept and created, existing structures extended, and new structures either created in analogy to existing ones or frequently from scratch after profound analysis. Examples of new elementary concepts are kinds of body parts (fig. 1), roles (e.g. AntigenicRole, AllergicRole), units (e.g. mg/dl, IU), drugs, and enzymes.

- BodyPart
  - NAMEDSensoryOrgan
    - EyeOrgan
  - NAMEDSensoryOrganSubPart.
    - NAMEDEyeOrganSubPart
      - Macula
      - OcularLens
      - OpticPapilla
      - Retina

**Fig. 1 Newly created taxonomy showing elementary concepts of the eye anatomy.**
The levels of indenting denote the subsumption hierarchy. The taxonomy presents is-a-kind-of relations.

**Tab. 1 Attributes introduced and their usage**

| attribute | usage |
|---|---|
| hasDiagnosisProbability | how reliable is the diagnosis of a pathological phenomenon? |
| hasDiseaseStage | which stage is a pathological phenomenon in? |
| hasDosage | in which dosage is a drug applied? |
| hasImmediateComplication / hasLateComplication | which pathological phenomenona occurring after a short / long period of time are complications of pathological phenomenona? |

New attributes were added. These enable the represention of details of e.g. pathological phenomena and drug therapies (tab. 1).

Complex concepts (tab. 2) are frequently necessary to represent the items of the data source. Modelling conventions once established simplify modelling. The sanctioning mechanism is carefully applied in order to prevent nonsense compositions. A secondary structure is frequently superimposed on the primary conceptual hierarchy to rearrange the conceptual hierarchy as needed in the diabetological context. Central medical concepts are frequently modelled as general and heterogeneous GRAIL concepts.

The data source contains non-terminological knowledge which is represented in the extrinsic module. Linking extrinsic knowledge to terminological knowledge is done by application attributes (fig. 2). These are hierarchically organised in the application model, subsumed

**Tab. 2 Examples of complex concepts**

| item of dataset | complex concept |
|---|---|
| goal of therapy | Feature which <<br>isFeatureOf [ Feature GeneralisedProcess GeneralisedSubstance ]<br>hasQuantity ( Quantity which <<br>hasMagnitude PrimitiveValueType<br>hasUnit Unit > )<br>isGoalOf TherapeuticAct >. |
| visual symptom | Vision which <<br>hasAbnormalityStatus nonNormal<br>hasPersonReporting Patient >. |
| cataract | ( OcularLens which<br>hasOpacity ( Opacity which<br>hasAbsoluteState opaque ) ) name Cataract. |
| visual acuity | ( Vision which<br>isActedOnSpecificallyBy Glasses ) name VisualAcuity. |

by the ApplicationAttribute, designed conceptually, and arranged semantically thus permitting navigation.

( DrugPreparation which isMadeOf Drug )
extr. hasBrandName DrugBrandName;
extr. hasGenericName DrugGenericName.

**Fig. 2 Application attributes (underlined) and their usage.**
Extr. - extrinsically.

## 4. Application

The application will adopt the medical record architecture developed by the German MEDWIS-project DIADOQ which implements a procedure-based and problem-oriented approach to the EPR [7]. The application will enable structured data entry for current patient data. It will also record the measures together with their results and medical reasons of examinations and therapies. In its content the application like the diabetes model is based on the DIABCARD dataset version 2.0.

Starting point and presently focus of the application is the German Diabetes Passport (fig. 3, center; adapted from the German by the authors). Edited by the *Deutsche Diabetes-Gesellschaft (DDG)* the Diabetes Passport contains the measures internationally recognised as minimally necessary to prevent diabetic complications. It aims to achieve the goals of the St. Vincent Declaration. The Diabetes Passport presents a miniature EPR ideally suited to demonstrate the benefits of the GALEN technologies.

The application allows to enter clinical data of the Diabetes Passport right away. These data are values for therapeutical goals and actual values of several clinical parameters used to monitor diabetics. For the laboratory parameters the user can choose a unit from a list of adequate units (fig. 3, left). Explanations can be displayed for the clinical parameters and how these are measured. Further dialogues enable the user to enter detailed medical data (fig. 3, right) to be stored in the patient's overall medical record. A summary of these detailed data is transferred to the corresponding input field of the Diabetes Passport. Additional functionality is offered: The entered actual values are compared with the corresponding normal range and the given therapeutic goal. The therapeutic goal in some cases is preset with a value, usually the upper limit of the normal range. The present application as outlined with e.g. normal ranges and unit selection (fig. 3) is supported by the present diabetes model (3.2).

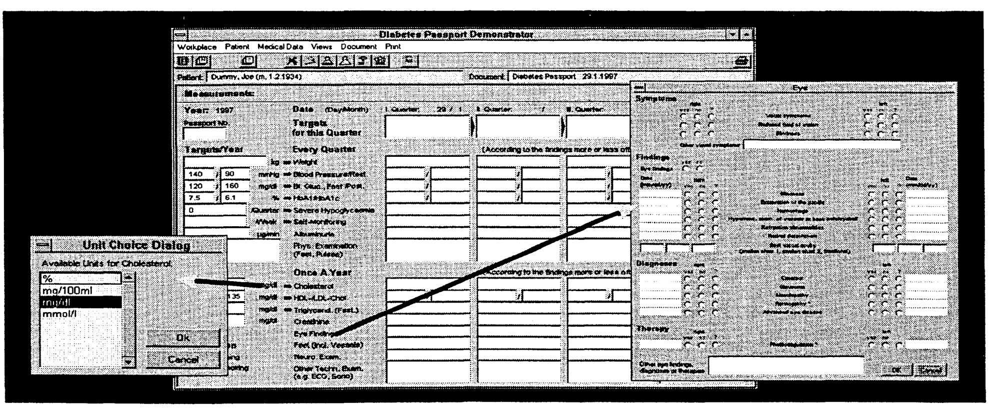


**Fig. 3 Aspects of the Application: Diabetes Passport, unit selection and detailed input of eye findings**

## 5. Discussion and Conclusion

The modelling activity has extended the terminological model to support a data entry system for diabetic management. About 2000 elementary and complex concepts and 20 attributes are added by the diabetes model. These are covering general and specialised medical knowledge. Numerous constraints added allow to automatically generate many further sensible composites. Considering the kind of modelled knowledge, the potentially generated concepts, and the clean taxonomies of the structured knowledge, we feel that our terminological diabetes model will be useful for other medical applications - general and diabetological.

Modelling in GRAIL is demanding, because representing terminological knowledge has to be precise and unambiguous. The difficulty of modelling terms depends on the regularity of the necessary structures and atoms, the coverage of the existing model, and the length of the items. Exquisite concepts are considerably more difficult to represent than common concepts. Modelling according to established conventions whenever possible is convenient and produces a regularly structured model more easily projected to the application.

Some information has been deliberately included in addition to the medical knowledge of the data source. It is intended to complete corners of the preexisting model or the represented medical context. Terms denoting two or more concepts are frequently modelled in every possible form. Concepts are frequently presented in two alternative nominalisations. Metalanguage like „other“ residual in the data source is referred to the application to be dealt with there.

The verified diabetes model will be quality checked according to various criteria: modelling structure; central reconciliation; verification against the application; clinical correctness of the information presented in the user interface; subjective satisfaction of the user; intuitive interaction with the application; reusability.

The system currently implemented as an offline version reflects the present diabetes model. It thus does not automatically adapt to changes in the model, which is a future aim. Further technical implementations to be realised in the near future are the application programming interface and filter of the EPR. As these technical features progress, the diabetes model will be modified according to the needs of the technical implementation. The application will be extended to cover all of the medical knowledge included in the DIABCARD dataset. It is furthermore planned to integrate technologies for speech recognition, report generation, and clinical guidelines.

## 6. References

[1] **van Ginnecken AM.** The structure of data in medical records. Yearbook of Medical Informatics. Stuttgart. 61 - 70. 1995.

[2] **Moorman PW, van Ginnecken AM, Siersema PD, van der Lei J, van Bemmel JH.** Evaluation of reporting based on descriptional knowledge. Yearbook of Medical Informatics. Stuttgart. 199 - 207. 1996.

[3] **Rector AL, Solomon WD, Nowlan WA, Rush TW.** A Terminology Server for medical language and medical information systems. Proceedings of IMIA WG6. Geneva. 1 - 16. May 1994.

[4] **Rector AL, Gangemi A, Galeazzi E, Glowinski AJ, Rossi-Mori A.** The GALEN CORE Model schemata for anatomy: Towards a re-usable application-independent model of medical concepts. Proceedings of Medical Informatics Europe. Lisbon. 186 - 189. 1994.

[5] **Engelbrecht R, Hildebrand C, Blecher M.** Improving patient care by the use of smart cards. Proceedings of AMICE 95. Amsterdam. 227 - 233. 1995.

[6] **Solomon D, Toft C.** The GRAIL KnoME. GALEN Documentation, July 1996.

[7] **Böhm V, Moser W, Diedrich T, et al.** A computer-based patient record system for diabetes outpatient clinics - the conceptual model and its implementation. In: The future of health information management. Eds.: Hoffmann U, et al. Munich: MMV. 27 - 33. 1996.