Intelligent Data Mining for Medical Quality Management

Wolf Stühlinger

TILAK – Tiroler Landeskrankenanstalten Ges.m.b.H.
Quality Management

Anichstrasse 35, A-6020 Innsbruck, Austria
wolf.stuehlinger@tilak.or.at

Oliver Hogl, Michael Müller

FORWISS – Bavarian Research Center
for Knowledge-Based Systems
Knowledge Acquisition Research Group
Am Weichselgarten 7, D-91058 Erlangen, Germany
oliver.hogl@forwiss.de

Abstract. In the healthcare sector cost pressure is growing, quality demands are rising and the competitive situation amongst suppliers is mounting. These developments confront hospitals more than ever with the necessities of critically reviewing their own efficiency under both medical and economical aspects. At the same time growing capture of medical data and integration of distributed and heterogeneous databases create a completely new base for quality and cost management. Statistics provide a powerful apparatus for the analysis of these data. However, the bandwidth of potential results is limited to manually built hypotheses. Innovative explorative methods of data analysis, known as data mining, automatically seek interesting statements about the data and are thus able to reveal unusual patterns too. Against this background we applied intelligent data mining methods to patient data from several clinics and from years 1996 to 1998. With the Knowledge Discovery Assistant (KDA) domain experts like medical quality managers are able to formulate relevant questions in the Knowledge Discovery Question Language (KDQL) which abstracts from database and data mining terminology and narrows the gap between domain experts and data mining specialists. For several questions in the field of medical service controlling interesting data mining results were discovered in this project. Questions were targeted on diagnoses, medical treatments, complications, and documentation. Within separable groups of patients, e.g. those with identical diagnoses, conspicuous subgroups were to be identified who – due to their age or sex – received a treatment other than expected from the given diagnosis without a medical indication being given or documented.

Keywords. Knowledge discovery, medical quality management, medical service controlling, language for business questions, data mining tool.

1 Introduction

Reforms in the healthcare sector have caused a continually rising cost pressure during the last years. At the same time quality demands on hospital and other suppliers of medical services are increasing. Along with an aggravating competitive situation the demand for enforced cost and service controlling in all fields of the healthcare sector, in diagnostics as well as in therapeutics and administration [Vog98, TGM98] is growing, aiming at the exploitation of efficiency potentials. On the other hand, the introduction of integrated hospital information systems (HIS) and the step by step conversion to electronic patient data files enable the capture of large amounts of data and thus a comprehensive documentation of historical diagnostic and therapeutic information on cases. Electronic documentation goes along with standardization efforts. One example for that are the diagnostic keys ICD9 and ICD10 [MB99]. Henceforth, distributed, heterogeneous, operative databases can be integrated, consolidated in a data warehouse or hospital information system, and made available within clinics or beyond. Rising costs and quality pressure on the one hand and new technologies of data processing on the other hand create both the necessity and the opportunity of a data based quality management in the health care sector, e.g. in the form of medical service controlling. Controlling in the business management sense comprises coordination, planning, and monitoring of business goals. Medical service controlling applies corresponding instruments to medical services. Fig. 1 shows a taxonomy of issues in the field of medical quality management.

Objective of a study launched by TILAK and FORWISS was the inspection of supposedly quality relevant criteria for medical service controlling in patient data as well as the discovery of new criteria. In the course of this, several questions on diagnoses, medical treatments, complications, and documentation quality were to be examined. This includes the search for indices in the data enabling the detection of quality relevant differences in care patterns. Above that, we attempt to discriminate normal clinic stays from stays with complications and good from less good documentation. Furthermore, temporal comparisons were to be drawn and generally conspicuous patterns in the available patient data revealed. Applying this knowledge efficiency potentials can be found, countermeasures taken, and compulsory quality standards or guidelines created.

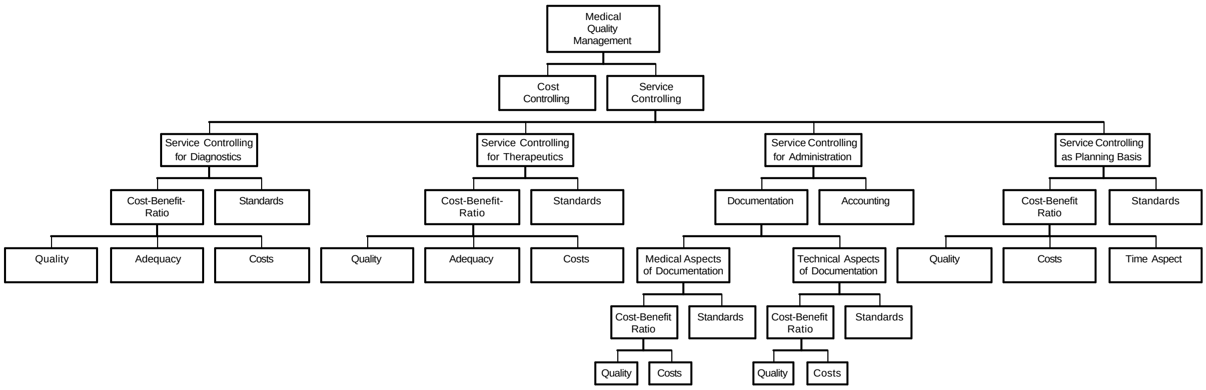


Fig. 1: Taxonomy of issues

Statistics provide a powerful apparatus for the analysis of this data. The bandwidth of potential results however is limited to manually built hypotheses. Innovative explorative methods of analysis, known as data mining, automatically seek interesting statements about the data and are thus able to reveal unusual connections too [Fay96]. Intelligent data mining incorporates advantages of both knowledge acquisition from data and knowledge acquisition from experts. This technology integrates domain experts intensely into the knowledge discovery process by acquiring domain knowledge and using it to focus the analyses as well as to filter the findings [HM96, Mül99].

Business management controlling already deploys data mining methods [Bis96]. Using them, conspicuous deviations between calculated and actual values of critical indices are searched for in hierarchy trees. This way often costly manual navigation is taken away from the controller and valuable time can be saved.

In the field of medicine an increasing deployment of data mining techniques for the analysis of medical data can be noted. This concerns e.g. the classification of tumors or the analysis of epidemiologic data [Bro98]. Relevant data to medical service controlling has been analyzed successfully using classic descriptive analyses of statistics.

The study represents a first approach to the application of data mining in patient data in the field of medical service controlling. For the first time innovative data mining methods in combination with expert knowledge have been used to analyze huge sets of electronically available patient data under several aspects in order to discover evidence to potentials for improvement with regard to efficiency and quality of clinical processes. It could be shown that substantial implicit knowledge is hidden in the available data, which cannot be discovered with conventional methods of analysis straight away.

In this paper we first illustrate the process of data based medical quality management. Chapter 3 describes the sets of data which have been used for the analyses. Chapter 4 introduces KDQL, a language for the representation of business questions which are used to focus the analyses and structure the findings. After that, we show important results which have been discovered by the analyses. In chapter 6 we present the Knowledge Discovery Assistant (KDA), a tool for intelligent data mining in this domain. Finally, we show which benefits can be achieved by our data based approach to medical quality management and give an outlook to future work.

2 The Process of Data Based Quality Management

After the exhaustion of the potentials of standard business reports and the use of interactive analytical methods on multidimensional data the focus is shifting to the automated discovery of interesting patterns in data. In the following, we propose a process model for the data based quality management as shown in fig. 2. Several parties within a hospital are involved into this process: The database manager builds and maintains data structures for the acquisition of data in clinical processes which are usually kept in a data warehouse or, more specifically, in a hospital information system. The medical data manager bridges the gap between the management, controlling, and quality management bodies of an organization and medical as well as social staff on the job. He manages the entirety of centrally acquired financial and medical data, presents them in a user-oriented manner and composes reports which are required for the information of the employees. On the other hand the medical data manager supports medical and nursing staff with the documentation of data, which are necessary for the management of a ward as well as the accounting of treatments. The data collected can be evaluated on three levels: standard reports, online analytical processing (OLAP), and data mining. Statistics and visualization are the essential basis for each of them.

In the following, we concentrate on three tasks of the cross industry standard process for data mining (CRISP-DM, [CC+99]): business understanding, modeling and evaluation. The business understanding task focuses on understanding the project objectives and requirements from a quality management perspective, and then converting this knowledge into a data mining problem definition. The modeling task contains the search for patterns of interest in a particular representational form. Having built a model, it is indispensable to thoroughly evaluate the model and review the steps executed to construct the model, to be certain it properly achieves the business objectives. Particularly the business understanding and the evaluation task, which precede and follow the modeling task, require a tight cooperation and interaction between experts from data mining and quality management. The results of data mining, e.g. evidence for capacity planning, quality relevant factors, guidelines or even new or revised medical knowledge, are kept in a central statement database from where different representations of data mining results can be generated in various formats, e.g., for the presentation in the hospital intranet or in printed reports. This information can be deployed to answer questions of quality managers which in turn can help to solve tasks given by the hospital administration. In addition to that, medical staff of the hospital can get evidence of the quality of their efforts.

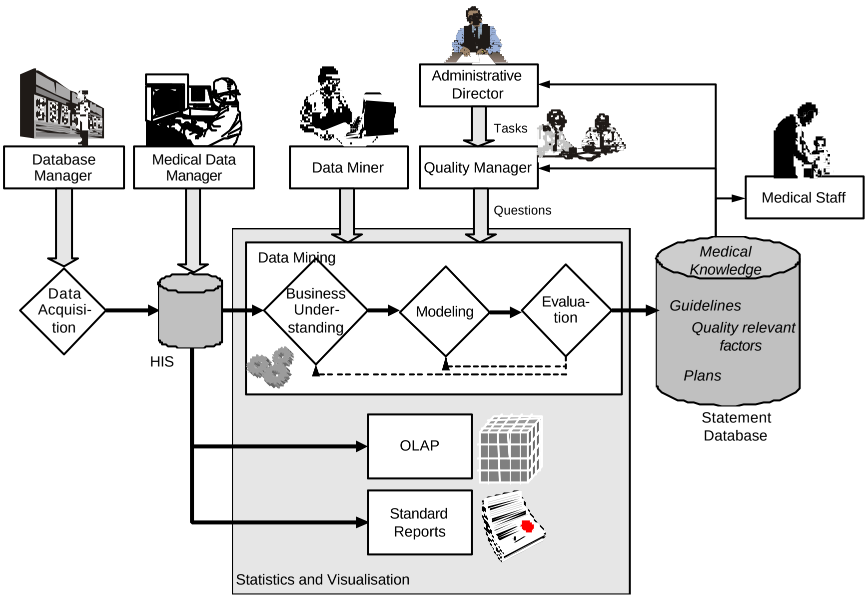


Fig. 2: A process model for data based quality management

3 Patient Data

Roughly 55.000 to 60.000 data sets of treatment cases were available from the clinics administered by TILAK for years 1996 to 1998 each. These data were given in relational form. The object types `patient` and `treatment case` are to be distinguished, where `patient` and `treatment case` are in a 1:n-relation. Amongst others, a patient is described by *age*, *sex*, and *native country*. Attributes of a treatment case in the block "diagnoses" are *primary diagnosis* and *additional diagnoses*, in the block "medical treatments" *type* and *number of treatments* and in the block "accounting information" *overall length of stay* and *treating wards*. Diagnoses are coded in accordance with the hierarchically structured diagnosis key ICD9 with 3 to 5 digits. Medical treatments were coded in a TILAK specific four digit treatment key. Associated with a diagnosis are standard values such as *upper* and *lower limit of length of stay*, which have been introduced by Austrian clinics as countrywide guidelines.

For analyses data have been denormalized, i.e., all attributes describing a treatment case have been combined in a single relation. Analyses have been carried out within patient groups of selected clinics (see table 1). In the eye clinic as well as in the dermatological, gynaecological, neurological, and urological clinic it can be assumed that inter-clinic fluctuation is low. The homogeneity of the patients groups establishes a reliable basis for the analyses.

	1996	1997	1998
Dermatology	2605 cases	1955 cases	1758 cases
Eye Clinic	3923 cases	3847 cases	2879 cases
Gynaecology	5682 cases	5372 cases	3937 cases
Neurology	3721 cases	3808 cases	2121 cases
Urology	5024 cases	4544 cases	4236 cases

Table 1: Definition of homogeneous groups of treatment cases

4 Question-Driven Data Mining for Medical Quality Management

In the field of medical service controlling (see fig. 1) we focused on questions for diagnostics, therapeutics, and for the administration. In addition to that, we drew temporal comparisons in order to review the effects of measures taken in the past.

We used the new language KDQL (Knowledge Discovery Question Language) which is designed to enable business users in the medical quality management domain to represent business questions in order to focus data mining queries and to structure data mining results. KDQL abstracts from database terminology, e.g., attribute names, and data mining terminology, e.g., names of data mining algorithms. Questions in KDQL are refined using a domain model and transformed into concrete data mining queries. In the following, we distinguish the design of five different components of business questions like:

Are there indications on complications regarding the length of stay in various patient groups in the neurological clinic?

In this example, the five components are instantiated as follows:

- Question Type: *are there*
- Question Object: *indications on complications*
- Question Arguments: *length of stay*
- Question Group: *in various patient groups*
- Question Context: *in the neurological clinic*

The KDQL syntax is an approach to a controlled language, which can easily be used and understood by business users of data mining systems. Thus, most structures have been adopted from the syntax of natural language questions. We define KDQL in an extended BNF grammar, where "[]" represents 0 or one occurrence and words in sans serif represent keywords.

⟨Question⟩ ::=
⟨QuestionType⟩
[⟨QuestionGroup⟩]
[⟨QuestionContext⟩] ?

In KDQL, "⟨QuestionType⟩" determines the type of possible answers. As different question types require different types of question objects and question arguments, the corresponding formats will vary. Details on the specification of the question arguments will be given below. The optional statement "⟨QuestionGroup⟩" enables the user to select criteria for the formation of groups and "⟨QuestionContext⟩" allows to focus on a view on the available data.

The Question Type. We distinguish between confirmative questions such as

> *Is there a difference between risk and non-risk patients regarding their length of stay in the clinic?*

which seek to confirm a hypothesis and explorative questions such as

> *Which indications on complications regarding the length of stay in various patient groups are there in the neurological clinic?*

which trigger the search in hypotheses spaces. These questions are introduced by a question element which can be a single word, e.g., “who?”, “when?” etc., or a composed phrase, e.g., “which clinic?”. KDQL provides support for the most essential types of questions. As mentioned above we distinguish an explorative and a confirmative question type:

⟨QuestionType⟩ ::=
  ⟨confirmativeType⟩ | ⟨explorativeType⟩

Confirmative questions are yes/no questions while exploratory questions use an interrogative pronoun:

⟨confirmativeType⟩ ::=
  is there ⟨QuestionObject⟩ | are there ⟨QuestionObject⟩

⟨explorativeType⟩ ::=
  which ⟨QuestionObject⟩ | who | when | where | how | why |

The Question Object. The question object represents in a grammatical sense the direct object of a question. In this early stage of design, we make out two different types of question objects:

⟨QuestionObject⟩ ::=
  ⟨PatternType⟩ | ⟨DomainConcept⟩

Abstract concepts like “complications” or “quality of documentation” are domain-dependent whereas “dependency”, “difference”, “mutuality”, and “change” represent domain-independent pattern types:

⟨PatternType⟩ ::=
  ⟨Dependency⟩ | ⟨Difference⟩ | ⟨Mutuality⟩ | ⟨Change⟩

⟨DomainConcept⟩ ::=
  ⟨Complications⟩ | ⟨QualityOfDocumentation⟩ | ...

We suggest the introduction of primitives for domain-independent concepts such as pattern types as well as primitives for domain-specific concepts, e.g., for the medical quality management domain. Question objects have different requirements on the format of the question arguments. To give just one example we show the specification of the pattern type ⟨Difference⟩:

⟨Difference⟩ ::=
  difference in ⟨subs_list⟩ |
  difference regarding ⟨ref_list⟩ in ⟨subs_list⟩ |
  difference regarding ⟨pat_list⟩

⟨subs_list⟩ refers to the list of subsets in which the difference is analyzed and ⟨ref_list⟩ denotes the list of references which can be used to narrow down the analysis.

The Question Arguments. Question arguments specify the set of arguments which the question object is based upon, i.e., attribute groups, attributes, attribute value groups, attribute values, records, and record groups. They can be connected by the Boolean operators *and*, *or*, and *either…or*. Since arguments can take different roles in the context of a question a deeper differentiation is necessary. Arguments can be used either as references or as specifications of subsets of data. In addition to that, the pattern list enables users to ask questions involving patterns on patterns. Changes, one example for a pattern type, can not only exist between attributes but also between patterns themselves. Thus, e.g., the analysis of changes on other patterns is possible:

Is there a change on the difference regarding the overall length of stay between risk-patients and non-risk-patients between patients from 1998 and 1999?

However, in order not to let questions get too complex only one level of patterns connected with *and* will be allowed:

⟨pat_list⟩ ::=
⟨dependency⟩ and ⟨dependency⟩ | ⟨difference⟩ and ⟨difference⟩ |
⟨mutuality⟩ and ⟨mutuality⟩ | ⟨change⟩ and ⟨change⟩

The Question Group. For many applications, the formation of groups of database records such as patients is necessary because only within a group certain analyses can be performed in a meaningful way. The question group is specified by naming the attribute that is to be used for the formation of the group:

⟨QuestionGroup⟩ ::=
within all patients | within various groups |
within groups defined by ⟨Attribute⟩

The Question Context. Optionally, in many cases not the whole set of data available but limited views may be of interest for the business user. Examples for such views are the selection of a specific clinic from which the data is to be considered in the analysis process or the selection of a specific year. For the definition of the question context again the Boolean connections *and*, *or*, and *either…or* are allowed:

⟨QuestionContext⟩ ::=
in all data | in various data views |
in a data view defined by ⟨Attribute⟩ ⟨AttributeValue⟩

Thus, the KDQL representation of the natural language question

Is there a dependency of the overall length of stay and the primary diagnosis within groups with the same age in the neurological clinic 1999?

is the following:

Is there dependency of
Overall Length Of Stay and *Primary Diagnosis*
within groups defined by *Age*
in a data view defined by
Clinic Neurology and *Year 1999* ?

Most questions expressed by business users will not be initially translatable because they contain concepts which are not part of the data mining concept world such as attribute groups, attribute value groups or certain question objects, e.g., “complications”. In order to make those questions processable by data mining methods they are refined using various taxonomies and transformed into data mining queries.

5 Results for the Medical Service Controlling

This chapter presents results which have been produced driven by concrete questions (subchapters 5.1 to 5.4) as well as results which have been generated driven by data only (subchapter 5.5). For reasons of clarity the data mining results described above are presented in another processing step as diagrams.

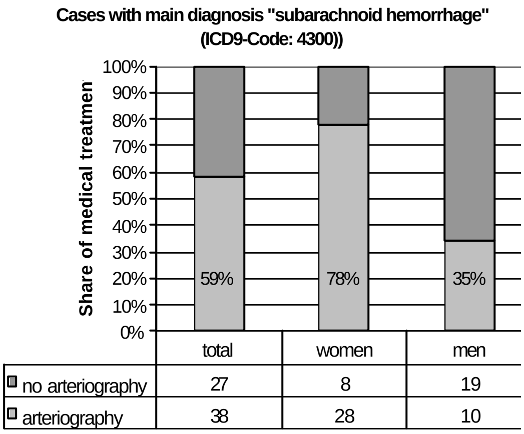
5.1 Medical Service Controlling for Diagnostics

To the question

In which conspicuous subgroups within differentiable patient groups, such as patients with identical diagnoses, have patients e.g. due to their age or sex received a treatment other than expected from the given diagnosis without a medical indication being given or documented?

for patients with primary diagnosis “subarachnoid hemorrhage” (ICD9-Code: 4300) in the neurological clinic in 1997 the following result has been found (see diagram 1). It becomes obvious that women in contrast to men are much more frequently treated with arteriography without any medical indication being given or documented.

Within the scope of etiology questions would rise, whether there is a third influence factor causing the conspicuousness e.g. treatment by different physicians. If it becomes apparent that the same results of treatment have set in, the more cost efficient method of treatment should be preferred in the future.



	total	women	men
no arteriography	27	8	19
arteriography	38	28	10

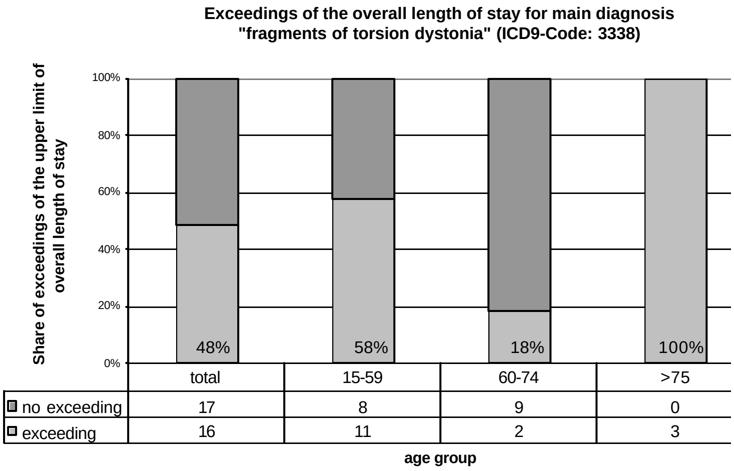
Diagram 1: Dependency of medical treatments on patients’ sex

5.2 Medical Service Controlling for Therapeutics

To the question

Which conspicuous subgroups can be identified within differentiable patient groups such as patients with identical diagnoses, who - discernible by the frequent occurrence of complications - should be treated differently, but have not been treated specifically?

for patients with the diagnosis “fragments of torsion dystonia“ (ICD9-Code: 3338) in the neurological clinic in 1996 the following result has been discovered (see diagram 2). Here it is obvious, that for “fragments of torsion dystonia“ the upper limit of length of stay is exceeded more frequently by patients between 15 and 59 years old than by those between 60 and 74 years old. In the scope of etiology it has to be investigated, if it is in fact the age which can be held responsible for the longer duration of stay. If that is the case, for patients of this age a modified treatment should be considered in the future. If not, the knowledge on the upper limits of length of stay has to be completed for this diagnosis by the corresponding exception.



	total	15-59	60-74	>75
no exceeding	17	8	9	0
exceeding	16	11	2	3

Diagram 2: Share of exceedings of the upper limit of length of stay in dependency of age

5.3 Medical Service Controlling for the Administration

To the question

What can be said on indices reflecting the quality of diagnoses and medical treatments?

for patients in the urological clinic in 1996 the following result has been found:

IF *primary diagnosis category was "phimosis and preputial changes"(ICD9-Code: 605)*
THEN *diagnoses were documented only by three digits (98%).*

This is conspicuous, because five more specific in form of four digit codes would have been available for documentation: paraphimosis (6050), phimosis (6051), preputial hypertrophy (6052), frenulum breve (6053), preputial block (6054), phimosis and preputial hypertrophy (6059). It is thus to be questioned, why this precise differentiation of diagnoses has not been made use of.

The rule

IF *primary diagnosis category was "male infertility" (ICD9-Code: 606)*
THEN *diagnoses were documented by only three digits throughout (100%).*

by contrast does not show this conspicuousness, because in this case only one unspecific four digit code would have been offered.

Furthermore in this area in the eye clinic in 1997 it has been discovered, that for primary diagnosis category "Inflammation of eyelids" (ICD9-Code: 373) only unspecific single medical treatments have been documented (93%). In addition to that it has been found, that for certain diagnoses such as "Diabetes melitus" (ICD9-Code: 250) no single medical treatment was documented at all (92%). In the scope of etiology it has to be investigated, whether these differences in documentation can be attributed to a lacking differentiation of the ICD9-key for these diagnoses.

5.4 Temporal Comparisons

After these static considerations changes of the uncovered phenomena between different years in all areas are of interest. In the complications area for example changes regarding the length of stay have been analysed. Thus in the eye clinic the following has to be noticed:

- For primary diagnosis category "cataract" (ICD9-Code: 366) the share of cases falling below the lower limit of length of stay rose by around 15 %.
- For the primary diagnosis category "disorders of refraction and accommodation" (ICD9-Code: 367) the share of cases falling below the lower limit of length of stay fell by ca . 8 %.
- The share of cases falling below the lower limit of length of stay fell for most primary diagnosis categories.

In the first case it has to be investigated, whether improved methods of treatment of a cataract have shortened the length of stay in this case.

5.5 Data-Driven Discoveries

In addition to the above question-driven data evaluation, we also performed data-driven analyses which produced another set of interesting findings. Thus, strong associations, e.g., within diagnoses and medical treatments have been discovered.

The discovery

IF *one of the single medical treatments [SMT] was "continuous ventriculometry",*
THEN *further SMTs were "respirator therapy" and "burr-hole trepanation". (77%)*

in the neurological clinic in 1997 gives evidence which could increase the reliability of plans. Furthermore conspicuousnesses regarding the length of stay with several influence factors, such as

IF *one of the SMTs was "combined strabotomy",*
THEN *the overall length of stay is between 2 and 6 days. (92%).*

and

IF *primary diagnosis category was "benign neoplasias of skin" (ICD9-Code: 216),*
THEN *the case fell below the lower limit of length of stay, because it was an out-patient treatment. (100%)*

in the eye clinic in 1996. The first rule can aid an increased reliability of plans for the management of resources and the second allows conclusions that the lower limit of length of stay for primary diagnosis category "benign neoplasias of skin" is not adequate. Furtermore, influences of attributes have been analysed which were not in the centre of consideration in the focused actions. One example for that is the following rule from the neurological clinic in 1997:

IF *the type of discharge was "death",*
THEN *one of the SMTs was " respirator therapy". (49%)*

The discovery of dependencies between medical parameters, such as diagnosis and therapy, and personal information, such as age and sex, e.g.,

IF *one of the additional diagnoses was „essential hypertension"(ICD9-Code: 401),*
THEN *the patient was female. (92%)*

and

IF *primary diagnosis was „ keratoconus"(ICD9-Code: 3716),*
THEN *the patient was between 15 and 59 years old. (100%)*

in the eye clinic in 1997 can support the verification or completion of medical knowledge.

6 The Knowledge Discovery Assistant

KDQL and the evaluation methods described above are for the most part implemented in the Knowledge Discovery Assistant (KDA), a tool for intelligent data mining, which is being adapted to and evaluated in the medical quality management domain. The goal was to develop an easy-to-use graphical user interface which enables user interaction without deeper knowledge about data mining methods and databases. Fig. 5 shows the main window of the KDA. The main window is horizontally divided

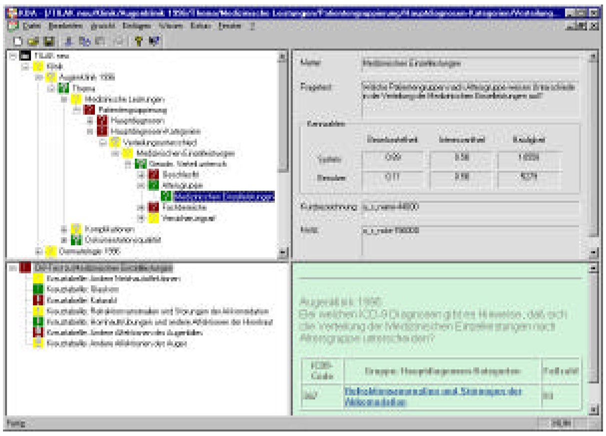


Fig. 5: The KDA's GUI

Fig. 6: The HTML presentation of results

into a window for the management of questions in the upper half and a window for the management of answers in the lower half. For both, questions and answers, a structured view as well as an inspector window is provided. The structured view for questions contains the question structures which have been built during the refinement process. The color of the question mark symbols in the structured view indicates, whether a question has been answered, and its size gives a first impression of its global interestingness. The inspector window gives a detailed view on the selected question, which includes in addition to the question text and its short name various attributes such as the interestingness of the question and a measure of the extent to which the question has been answered. For each selected question the associated results are shown. This list can be sorted by a single interestingness facet like novelty or usefulness or by a global interestingness value and the results can be accessed as shown above.

To allow the deployment of the results for distributed users, a highly structured HTML-hypertext with a navigation tree based on the structures of questions can be generated, as shown in fig. 6. In addition to that, a report in printable rich text format can be produced.

7 Conclusions and Future Work

We have shown, that intelligent data mining in addition to conventional analyses and statistical studies in patient data can deliver further evidence for medical service controlling. In detail, practical use can be derived from the discoveries for all parties involved in the data based medical quality management process: For the clinic administration the analysis of costly complications and identification of "normal", improvement of the reliability of plans and the setting of standards and enhancement of adequacy of the scoring system can be supported. The quality manager is aided in refining his quality management knowledge, reviewing and improving quality and adequacy of therapies, discovering self-contained, sound quality management indices as well as reviewing and improving the compliance with medical guidelines. But also for the medical data manager new insight can be won in terms of reviewing and improving data correctness as well as the quality of documentation. Above that and in addition to quality relevant information, medical knowledge can be broadened, aiding the work of physicians and nursing staff in the clinics.

Further works concern primarily the statistical consolidation of the results, support for the conversion of conclusions as well as the development of a accordingly adapted automated data mining process and the introduction of this technology as a component of a comprehensive and steady set of controlling instruments.

References

[Bro98] Brossette, S.: *Data mining and epidemiologic surveillance*, PhD thesis, Birmingham (AL), University of Alabama at Birmingham, 1998.

[CC+99] Chapman, P.; Clinton, J. et. al. : *The CRISP-DM Process Model*, Discussion Paper, http://www.crisp-dm.org/, March 1999.

[FPS+96] Fayyad, U.; Piatetsky-Shapiro, G. et al.: *From Data Mining To Knowledge Discovery: An Overview*, in: Fayyad, U.; Piatetsky-Shapiro, G. et al. (editors): *Advances in Knowledge Discovery and Data Mining*, AAAI Press, Menlo Park, California, 1996, pp. 1-34.

[HM96] Hausdorf, C.; Müller, M.: *A Theory of Interestingness for Knowledge Discovery in Databases Exemplified in Medicine*, in: Lavrac, N.; Keravnou, E.; et al. (editors): *First International Workshop on Intelligent Data Analysis in Medicine and Pharmacology (IDAMAP-96)*, Budapest, Hungary, August 1996, pp. 31-36.

[KL93] H. Kunath, U. Lochmann: *Klassifikation als Voraussetzung für Qualitätssicherung*, ecomed, Landberg/Lech, 1993.

[MB99] Meads, S.; Brown, F.: *ICD-10 Coding Fundamentals : A Comprehensive Coding Guide for Health Care Professionals*, Practice Management Information, Los Angeles, 1999

[Mül98] Müller, M.: *Interessantheit bei der Entdeckung von Wissen in Datenbanken*, PhD thesis, University Erlangen-Nürnberg, Erlangen, 1998.

[TGM98] Tweet, A.; Gavin-Marciano, K.: *The Guide to Benchmarking in Healthcare: Practical Lessons from the Field*, American Hospital Publications, Chicago, 1998.