

Analysis of guideline compliance - a data mining approach

Vojtěch SVÁTEK, Antonín ŘÍHA,
Jan PELEŠKA, Jan RAUCH

EuroMISE Centre – Cardio

University of Economics, Prague

*Computer Science Institute of the Czech Academy of
Science*

Computerized Guidelines and Protocols, April 13, 2004

Guideline compliance analysis via data mining (1)

- Guideline compliance analysis carried out
 - traditionally by **clinicians**:
mostly statistical methods
 - more recently also by **computer scientists**:
possible use of modern techniques, such as
knowledge modelling or data mining
- We speak about the latter

Guideline compliance analysis via data mining (2)

- Discovered non-compliance may lead to
 - feeding back to **guideline authors**
(incomplete, inconsistent or outdated documents)
 - feeding back to **field clinicians**
(errors in clinical practice)
 - feeding back to **knowledge engineers**
(errors in document formalisation)
- Depending on **interpretation...**

Guideline compliance analysis via data mining (3)

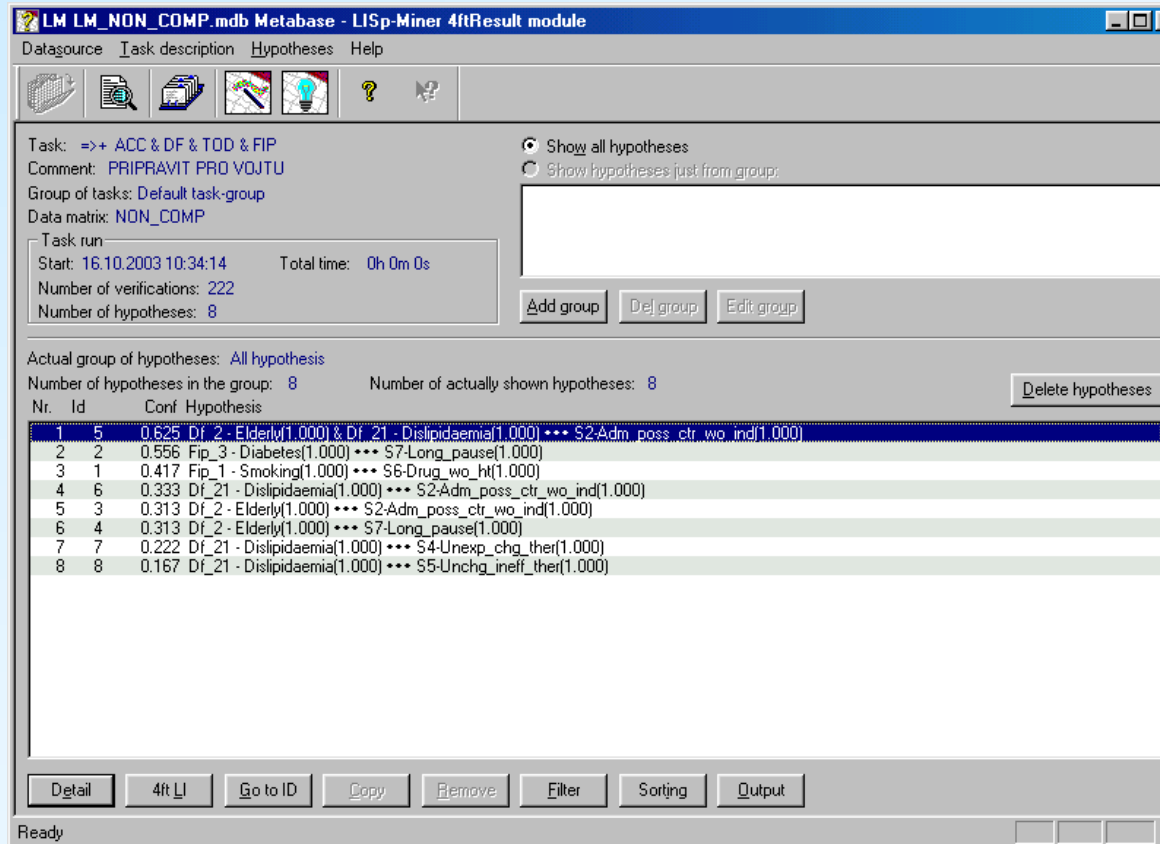
- Our approach
 - detect non-compliance for **individual patient records** using an operational model of the guideline (in OCML or Prolog)
 - apply a **data-mining** tool to
 - determine frequent types of non-compliance
 - find associations among the non-compliance patterns and other patient data
 - **frequent** patterns together with their **associations** submitted to medical expert for interpretation

Guideline compliance analysis via data mining (4)

- Experiment in the area of hypertension
 - 48 patient records
 - compared to an (ad hoc) **OCML** model of the WHO hypertension guidelines
 - the model covered 10 generic non-compliance patterns (NCPs)
 - 61 instances of these patterns discovered in data
 - subsequent analysis by the **LISp-Miner** tool
 - associating 10 NCPs with 39 other binary attributes
 - 8 associations found for a reasonable parameter setting
 - interpretation by a physician (data-donor)
 - mostly missing background knowledge or outdated guidelines

LISp-Miner user environment (1)

List of association hypotheses



The screenshot displays the LISp-Miner 4ftResult module interface. The window title is "LM LM_NON_COMP.mdb Metabase - LISp-Miner 4ftResult module". The interface includes a menu bar (Data source, Task description, Hypotheses, Help) and a toolbar with icons for database, search, task, hypothesis, help, and refresh.

Task: =>+ ACC & DF & TOD & FIP
Comment: PRIPRAVIT PRO VOJTU
Group of tasks: Default task-group
Data matrix: NON_COMP

Task run
Start: 16.10.2003 10:34:14 Total time: 0h 0m 0s
Number of verifications: 222
Number of hypotheses: 8

Buttons: Add group, Del group, Edit group

Actual group of hypotheses: All hypothesis
Number of hypotheses in the group: 8 Number of actually shown hypotheses: 8
Delete hypotheses

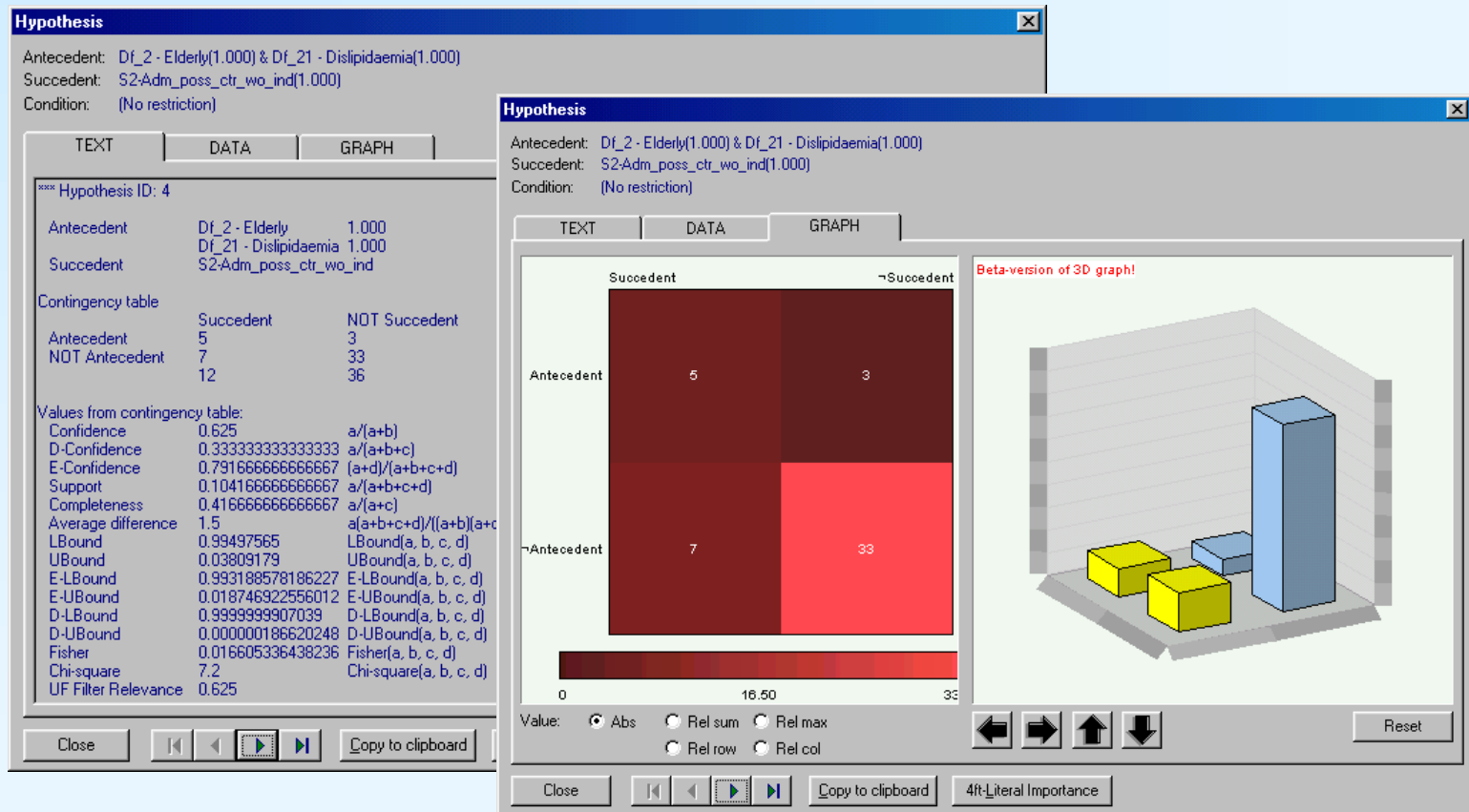
Nr.	Id	Conf	Hypothesis
1	5	0.625	Df_2 - Elderly(1.000) & Df_21 - Dislipidaemia(1.000) *** S2-Adm_poss_ctr_wo_ind(1.000)
2	2	0.556	Fip_3 - Diabetes(1.000) *** S7-Long_pause(1.000)
3	1	0.417	Fip_1 - Smoking(1.000) *** S6-Drug_wo_ht(1.000)
4	6	0.333	Df_21 - Dislipidaemia(1.000) *** S2-Adm_poss_ctr_wo_ind(1.000)
5	3	0.313	Df_2 - Elderly(1.000) *** S2-Adm_poss_ctr_wo_ind(1.000)
6	4	0.313	Df_2 - Elderly(1.000) *** S7-Long_pause(1.000)
7	7	0.222	Df_21 - Dislipidaemia(1.000) *** S4-Unexp_chg_ther(1.000)
8	8	0.167	Df_21 - Dislipidaemia(1.000) *** S5-Unchg_ineff_ther(1.000)

Buttons: Detail, 4ft L, Go to ID, Copy, Remove, Filter, Sorting, Output

Ready

LISp-Miner user environment (2)

Details of a hypothesis



Future : bottom-up formalisation & ‘compliance’ analysis ?

- Step-by-step bottom-up formalisation (cf. the talk by Růžička...)
 - allows to proceed from original document to ‘literal’ operational model
 - can be carried out to large extent by knowledge engineer alone, except for addition of background knowledge
 - data-mining-based compliance analysis then may serve for **posterior identification** of missing background knowledge!

Thank you for your attention