Example : machine learning for adverse drug events prevention

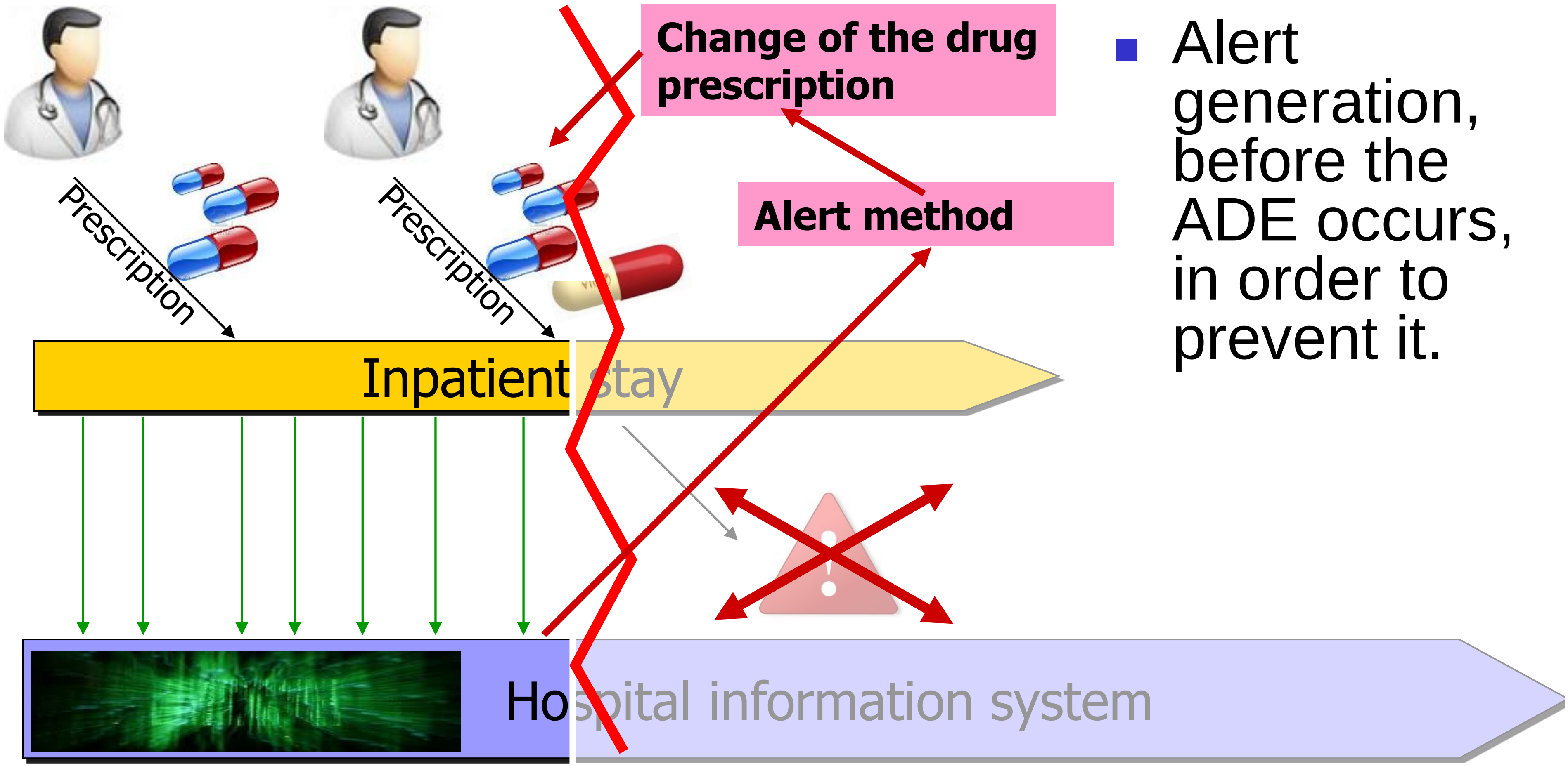
Adverse drug events

- ADEs = Adverse Drug Events

- Several definitions. Institute of Medicine (2007):
  - “An injury resulting from the use of a drug”
  - “An injury due to medication management
    rather than the underlying condition of the patient”

- Epidemiological data:
  - 98,000 deaths per year in the US
  - An ADE would occur in 5-9% of inpatient stays

Prospective prevention of ADEs



- Alert generation, before the ADE occurs, in order to prevent it.

Adverse drug events prevention by rule-based AI: CDSS

ADE prevention using rule-based CDSS

- CPOE:
  - computerized physician order entry
  - process of electronic entry of medical practitioner instructions for the treatment of (hospitalized) patients
- CDSS:
  - Clinical decision support system
  - Health information technology system that is designed to provide physicians and other health professionals with clinical decision support
  - Often based on level 1 artificial intelligence (rules)
- CPOE + CDSS = the “obvious” solution for adverse drug events prevention?

CDSS & CPOE: Over-alerting, alerte-fatigue => poor clinical efficiency!

- **Over-alerting**: too numerous and inappropriate alerts
- Alerts interrupt the clinicians’ workflow and induce **alert-fatigue**
  - Too many alerts
  - => time and mental energy consumption
  - => a mental state whereby users start ignoring critical alerts along with those that may be clinically insignificant
- May prevent CDSS from improving patient safety
- Alert override:
  - **up to 96% of alerts** are overridden by prescribers
  - But alert override is **often inappropriate**, and is sometimes followed by actual ADEs

Machine learning for ADE prevention

Idea driven by

Pr Regis Beuscart, head of the PSIP Project

Funded by the European Research Council, 7th framework program (agreement N°216130)

Administrative data

88 years old woman

Diagnoses

I10 Arterial hypertension
Z8671 Personal history of myocardial ischemia
I620 Non-traumatic subdural hemorrhage

Medical procedures

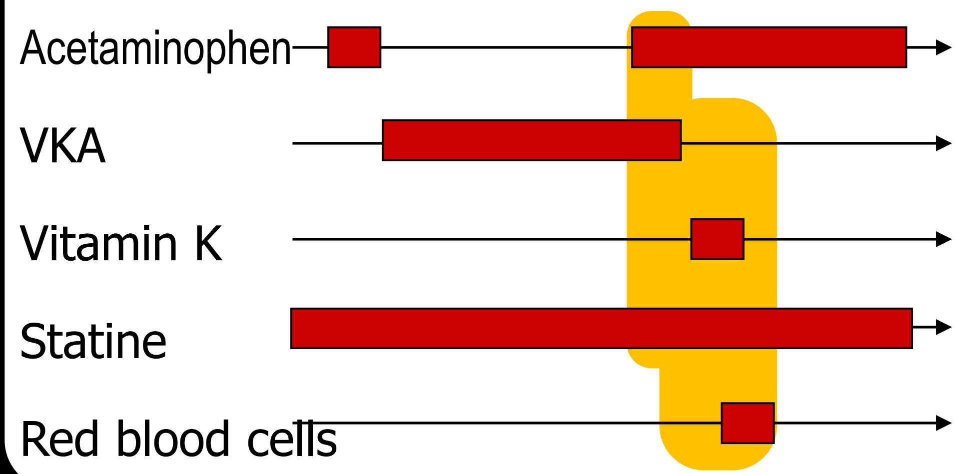
ABJA002 Drainage of an acute subdural hemorrhage, by craniotomy
FELF001 Transfusion

Free-text reports

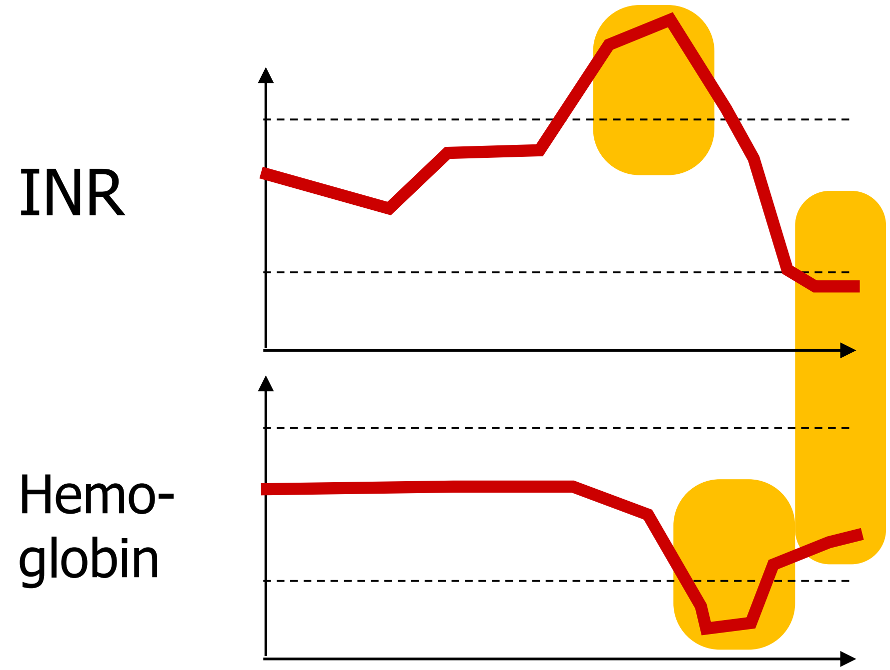
Discharge letter

Surgical report

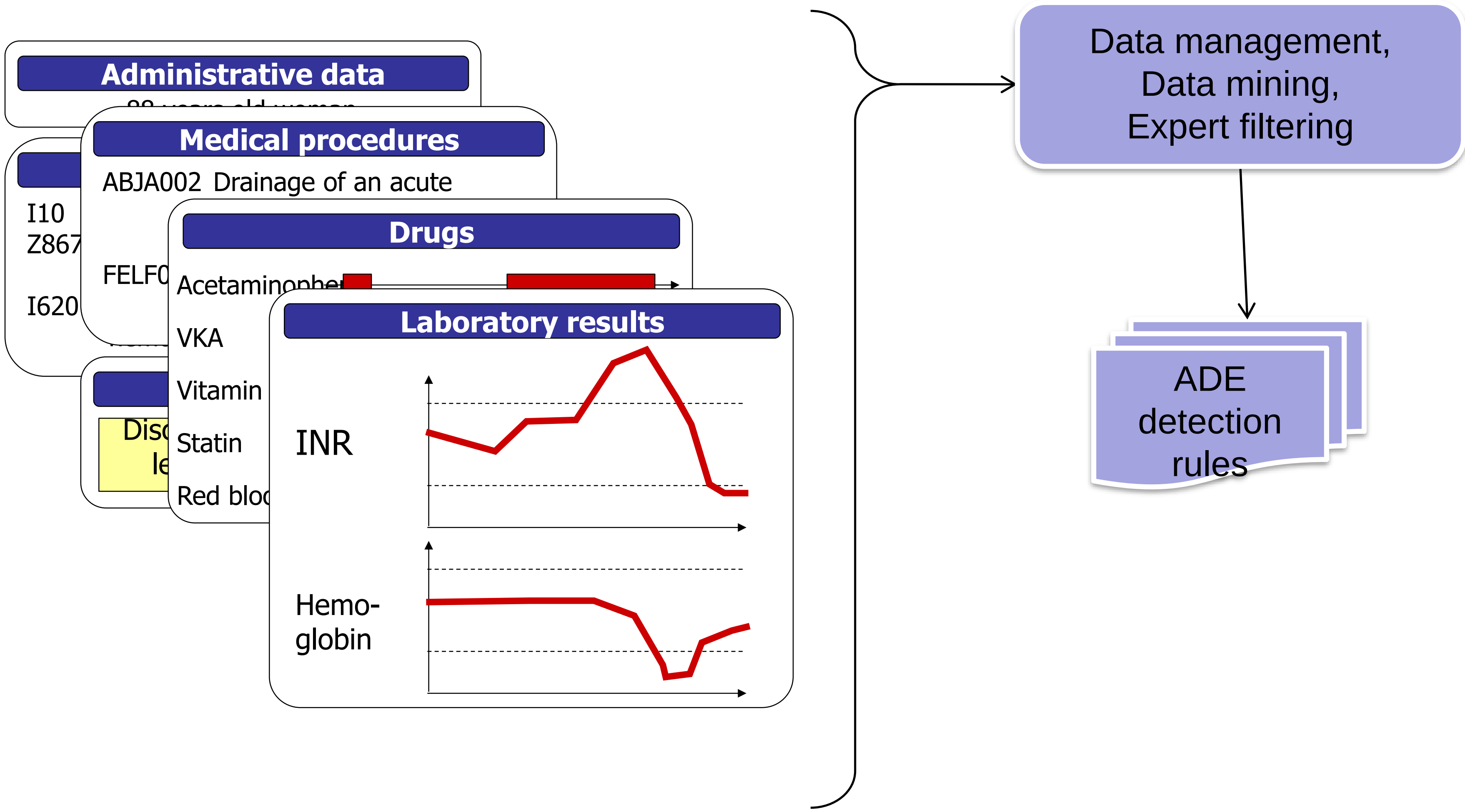
Drugs



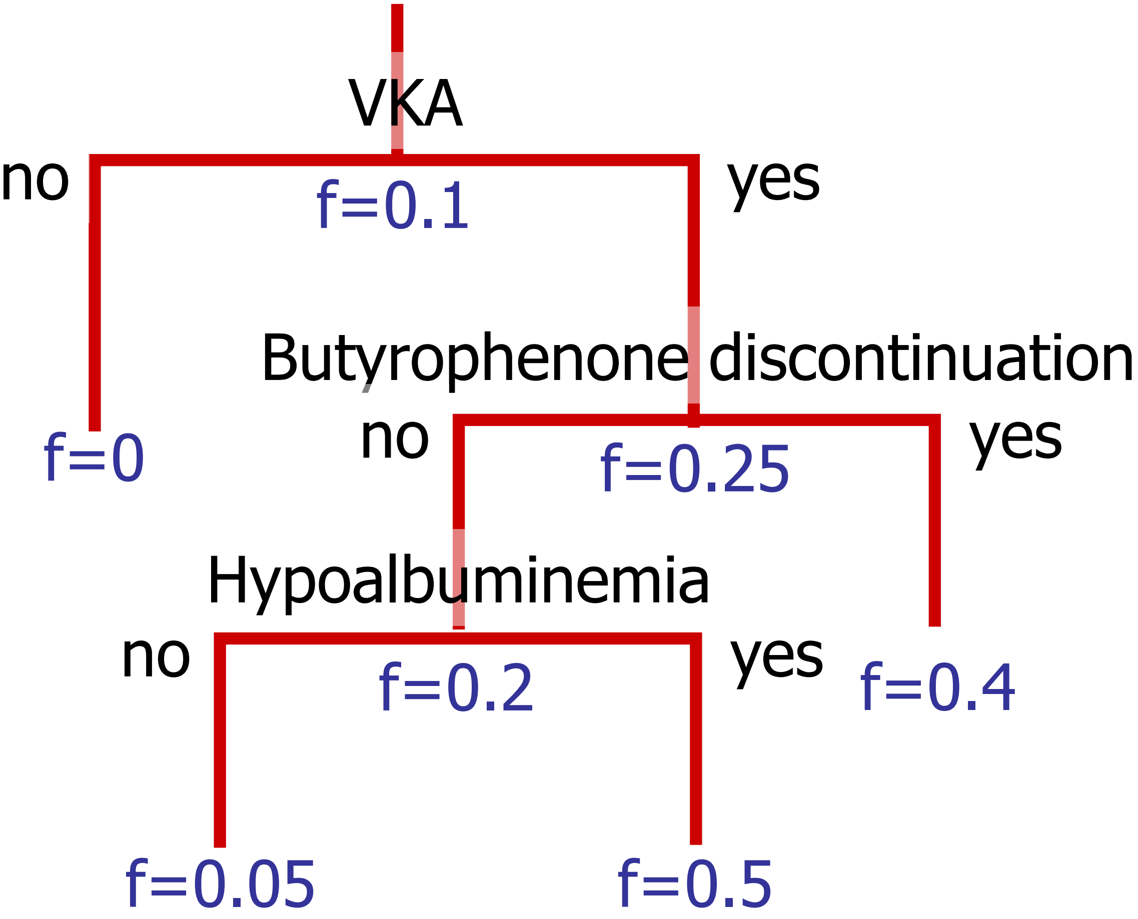
Laboratory results



Available data: ~175,000 inpatient stays from 6 hospitals (F, Dk, Bu)

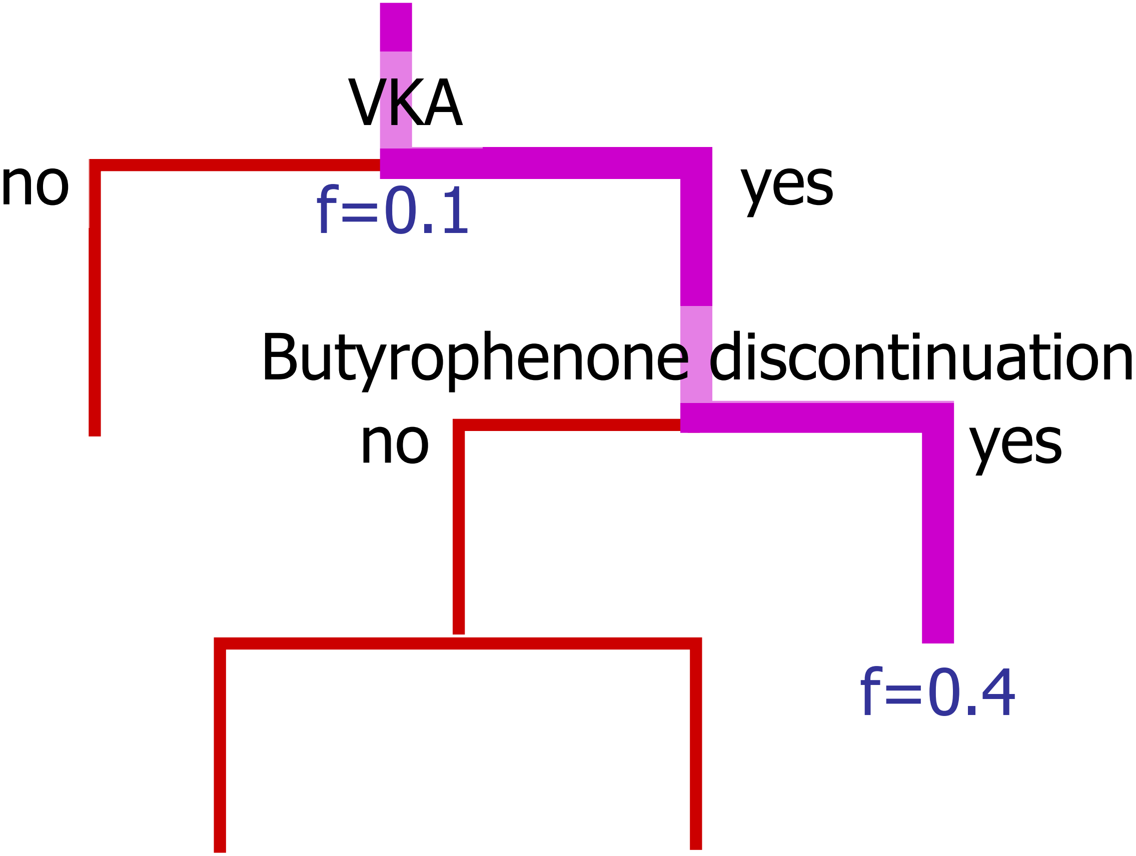


Artificial intelligence
Example of decision tree



- VKA= vitamin K antagonists (anticoagulant)
- INR= international normalized ratio. Evaluates VKA activity
- INR>5 => risk of hemorrhage
- The tree attempts to explain INR>5

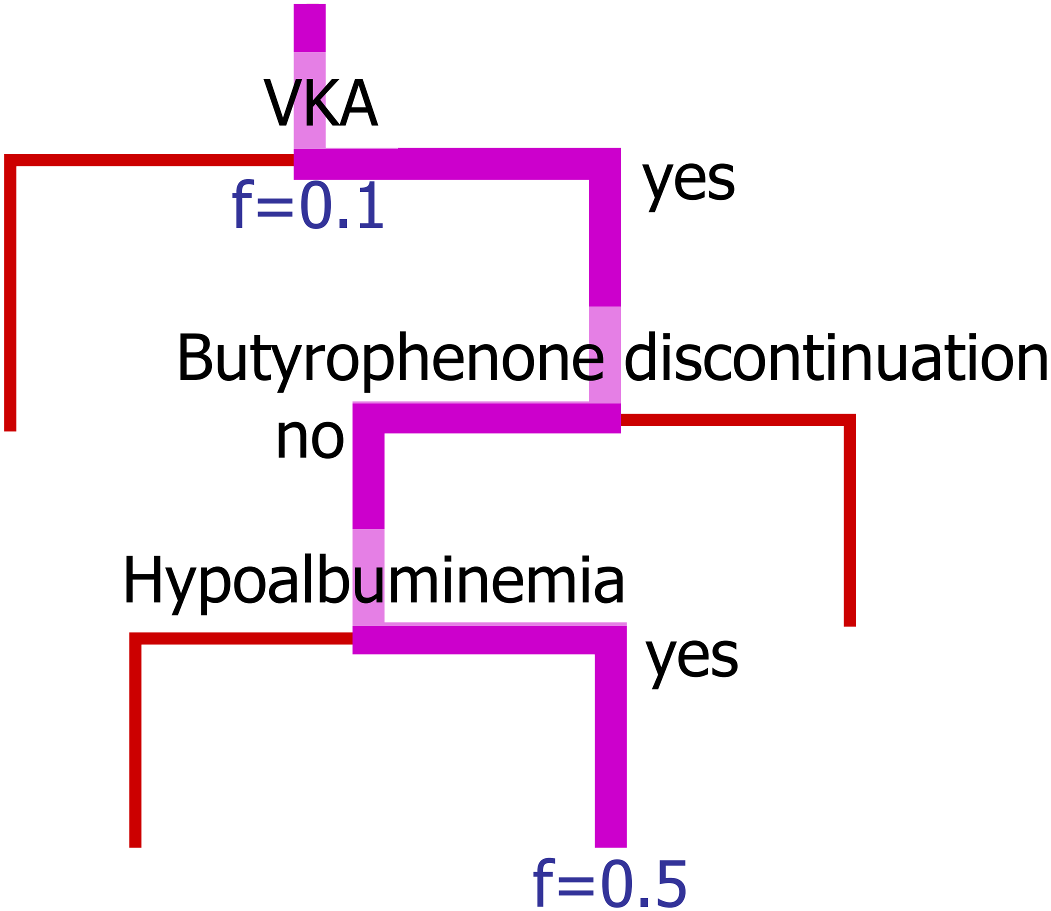
Artificial intelligence
Example of decision tree



Rule #1 (4th leaf):

VKA
& butyrophenone discontinuation
→ P=0.4

Artificial intelligence
Example of decision tree

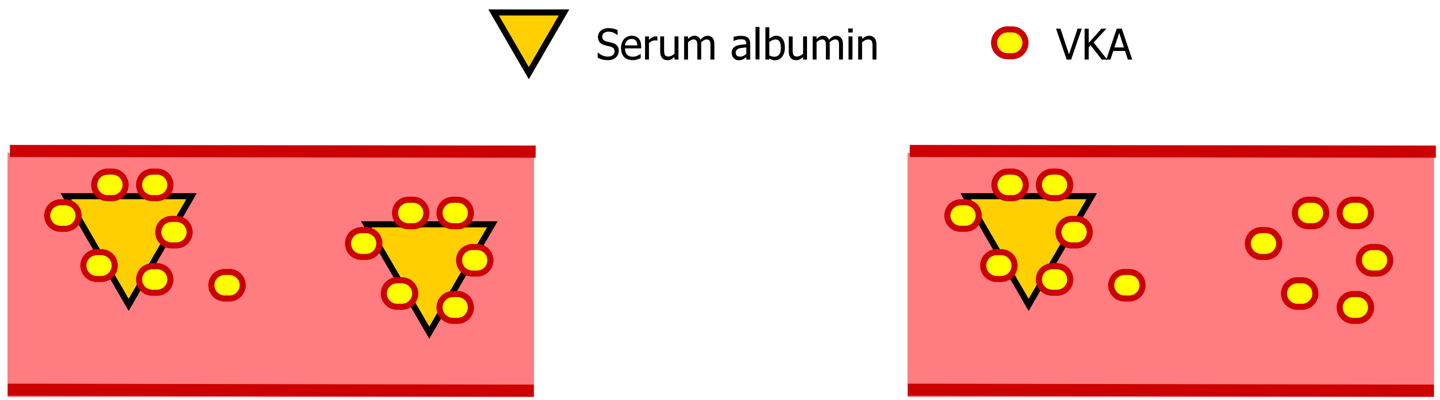


Rule #2 (3rd leaf):

VKA
& no butyrophenone discontinuation
& hypoalbuminemia
→ P=0.5

Artificial intelligence
Expert validation of rules

Albumine = plasmatic protein to which VKA bind. Only the non-bound part is biologically active.



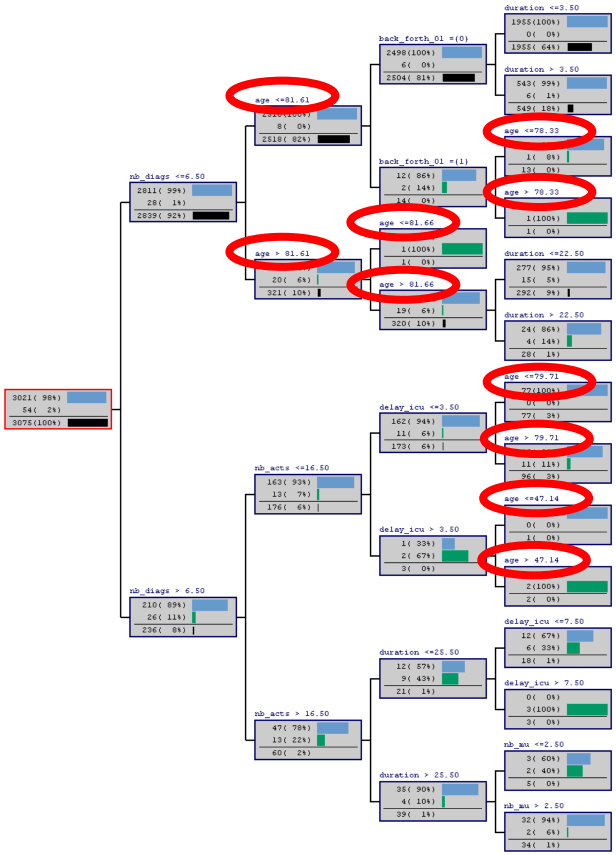
<u>Normal state:</u>
99% of the VKA bind to albumin.
Only 1% of VKA are biologically active.
The intake is based on it.

<u>Hypoalbuminemia:</u>
decrease of the bound fraction,
increase of the non-bound fraction
=> too high INR (with constant intake)

=> Need for validation, explanations, reorganization!

Actual result: risk of death

Explanatory variables: ~20 administrative variables, ~500 laboratory results, ~500 drugs. Result: old people die more than young people…

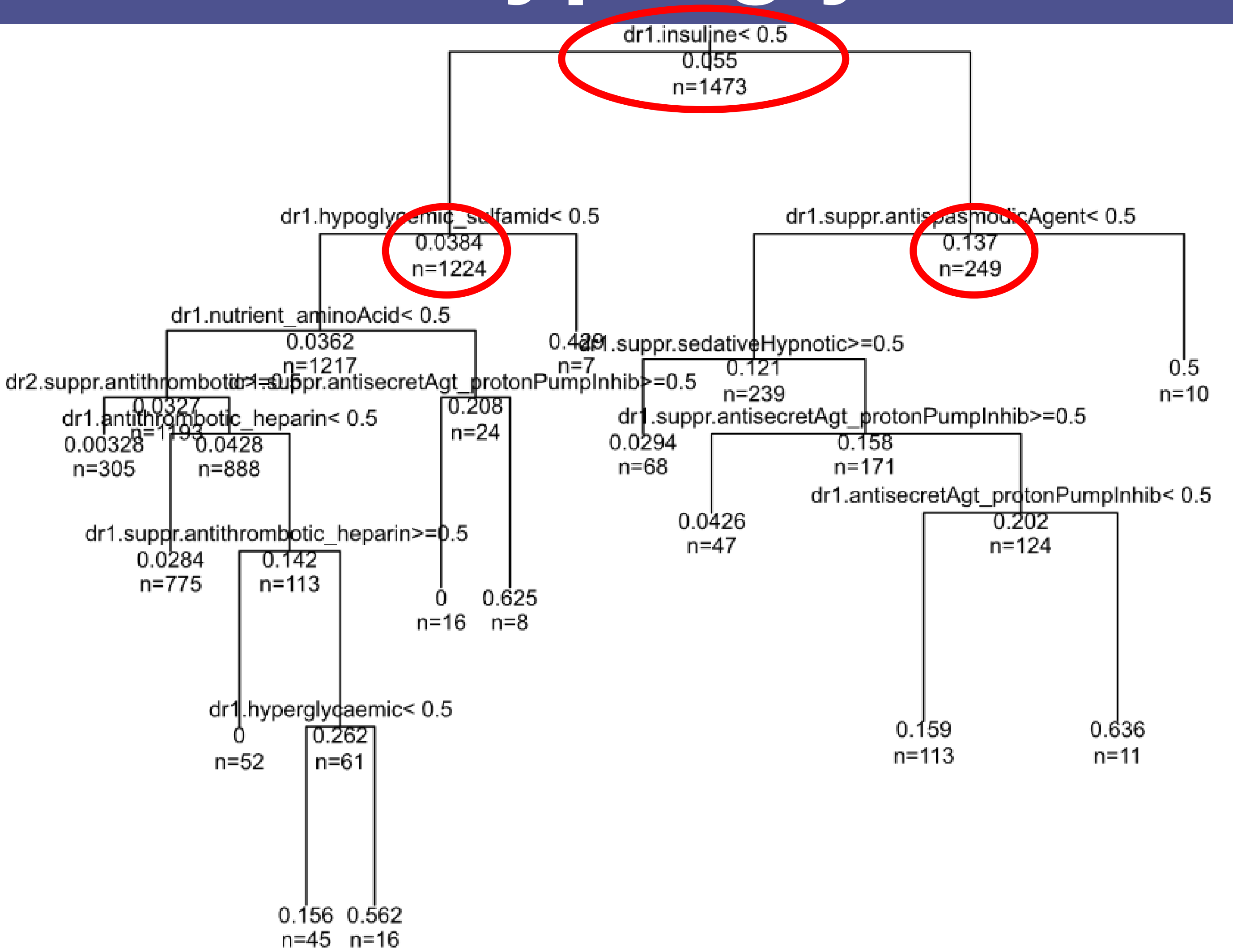


Base de connaissances : 15 règles

N°	N°SEGMENT	Condition d'appartenance	Support	Conclusion	AE_1 (C7=0)	AE_2 (C7=1)
1	10	**SI** age <=81.66 et age > 81.61 et nb_diags <=6.50	0.0003	AE_2 (C7=1)	0.0000	1.0000
2	16	**SI** duration <=3.50 et back_forth_01 ={0} et age <=81.61 et nb_diags <=6.50	0.6358	AE_1 (C7=0)	1.0000	0.0000
3	17	**SI** duration > 3.50 et back_forth_01 ={0} et age <=81.61 et nb_diags <=6.50	0.1785	AE_1 (C7=0)	0.9891	0.0109
4	18	**SI** age <=78.33 et back_forth_01 ={1} et age <=81.61 et nb_diags <=6.50	0.0042	AE_1 (C7=0)	0.9231	0.0769
5	19	**SI** age > 78.33 et back_forth_01 ={1} et age <=81.61 et nb_diags <=6.50	0.0003	AE_2 (C7=1)	0.0000	1.0000
6	20	**SI** duration <=22.50 et age > 81.66 et age > 81.61 et nb_diags <=6.50	0.0950	AE_1 (C7=0)	0.9486	0.0514
7	21	**SI** duration > 22.50 et age > 81.66 et age > 81.61 et nb_diags <=6.50	0.0091	AE_1 (C7=0)	0.8571	0.1429
8	22	**SI** age <=79.71 et delay_icu <=3.50 et nb_acts <=16.50 et nb_diags > 6.50	0.0250	AE_1 (C7=0)	1.0000	0.0000
9	23	**SI** age > 79.71 et delay_icu <=3.50 et nb_acts <=16.50 et nb_diags > 6.50	0.0312	AE_1 (C7=0)	0.8854	0.1146
10	24	**SI** age <=47.14 et delay_icu > 3.50 et nb_acts <=16.50 et nb_diags > 6.50	0.0003	AE_1 (C7=0)	1.0000	0.0000
11	25	**SI** age > 47.14 et delay_icu > 3.50 et nb_acts <=16.50 et nb_diags > 6.50	0.0007	AE_2 (C7=1)	0.0000	1.0000
12	26	**SI** delay_icu <=7.50 et duration <=25.50 et nb_acts > 16.50 et nb_diags > 6.50	0.0059	AE_1 (C7=0)	0.6667	0.3333
13	27	**SI** delay_icu > 7.50 et duration <=25.50 et nb_acts > 16.50 et nb_diags > 6.50	0.0010	AE_2 (C7=1)	0.0000	1.0000
14	28	**SI** nb_mu <=2.50 et duration > 25.50 et nb_acts > 16.50 et nb_diags > 6.50	0.0016	AE_1 (C7=0)	0.6000	0.4000
15	29	**SI** nb_mu > 2.50 et duration > 25.50 et nb_acts > 16.50 et nb_diags > 6.50	0.0111	AE_1 (C7=0)	0.9412	0.0588

Date de création : 02/05/2008 16:39:16

Actual results: risk factors of hyperglycemia…

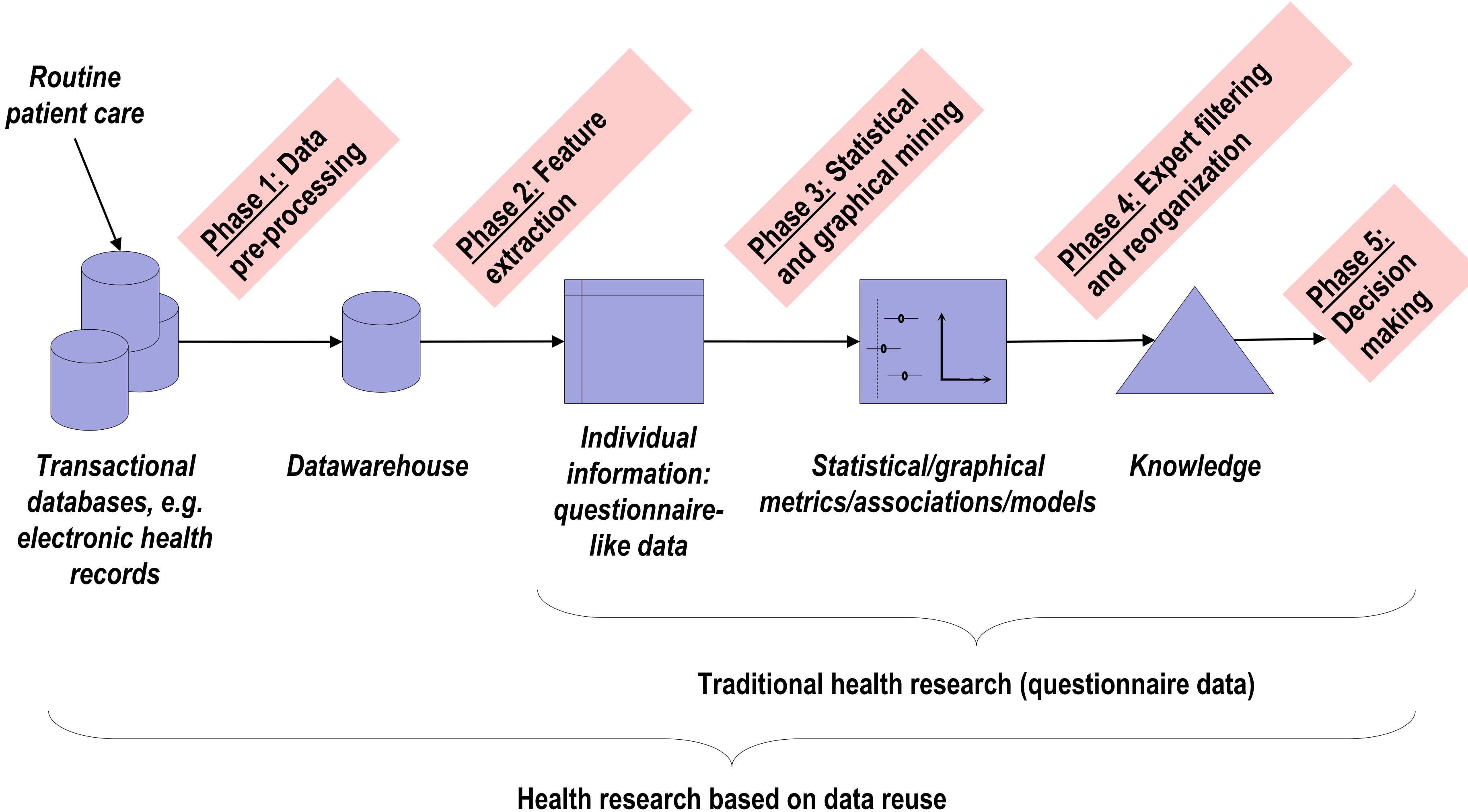


The failure of fully-automated machine learning

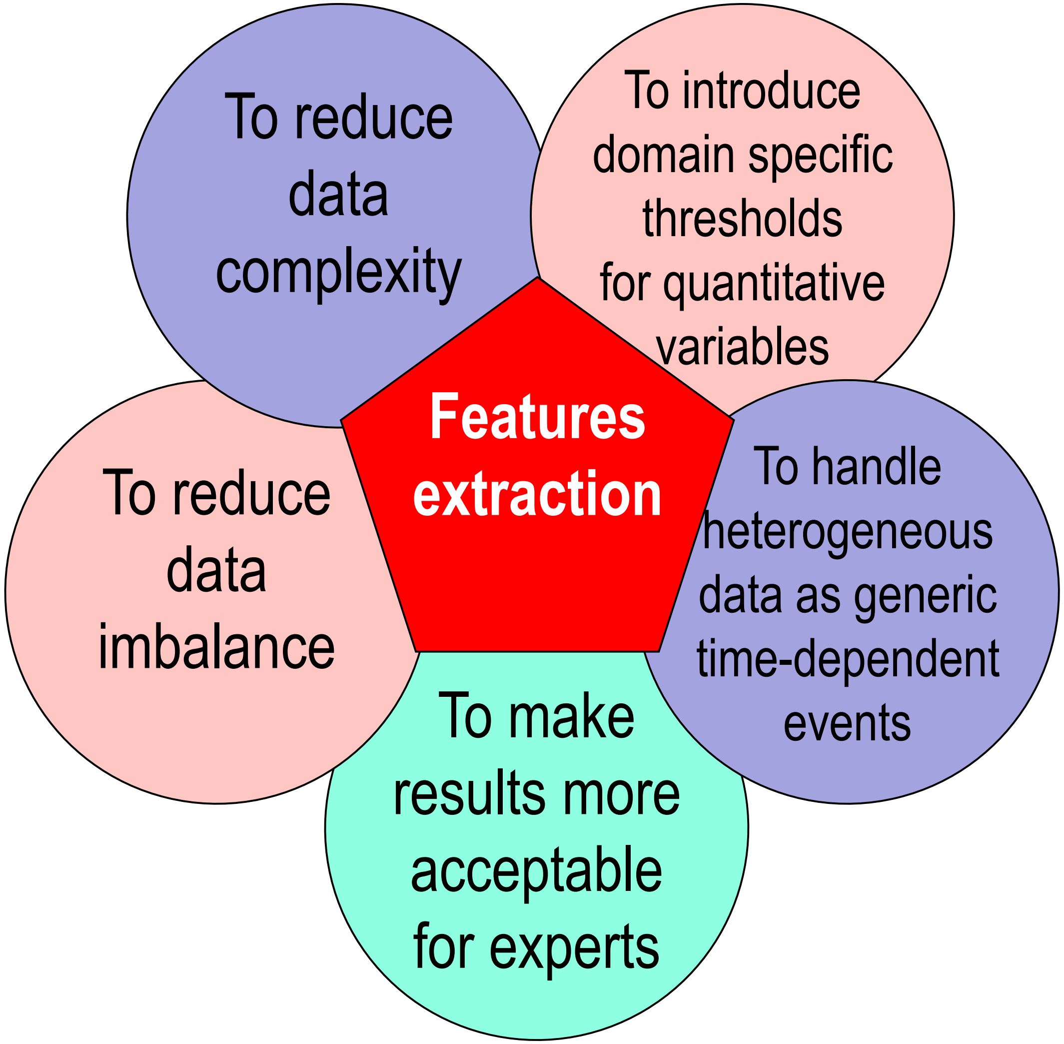
- Supervised data mining:
  - Good predictive power
  - Enables to filter, reorganize and explain knowledge
- "Black boxes", such as deep learning
  - Better predictive power
  - Does not enable to manage knowledge!
- However, some other steps are from far more crucial: feature extraction

Feature extraction

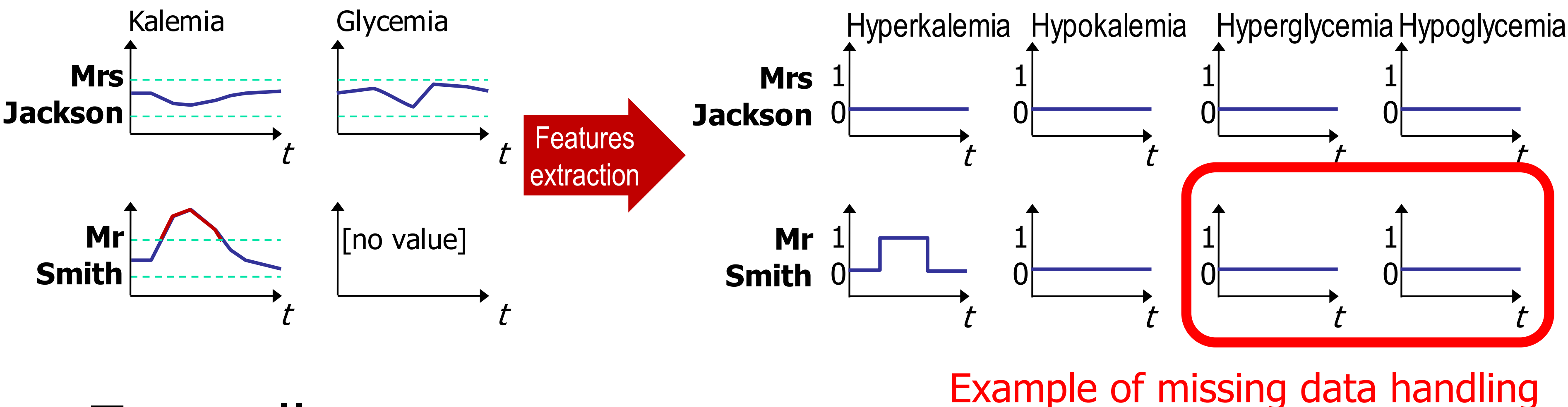
Structured data reuse process in healthcare



The objectives of feature extraction

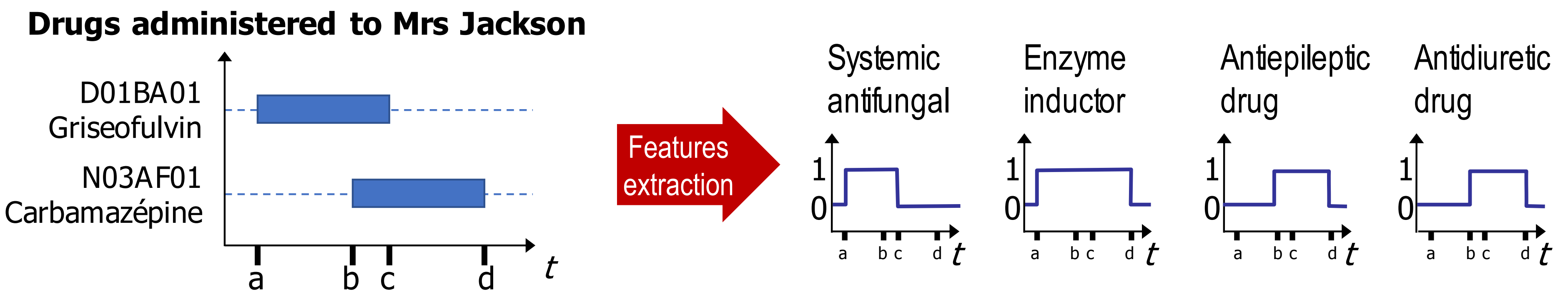


Feature extraction, example 1: laboratory results



- Formally:
  - Example of 2 patients, 2 parameters measured 5 times
  - Before: 1 table with 2 lines + 1 table with 10 lines
  - After: 1 table with 2 lines

Feature extraction, example 2: administered drugs

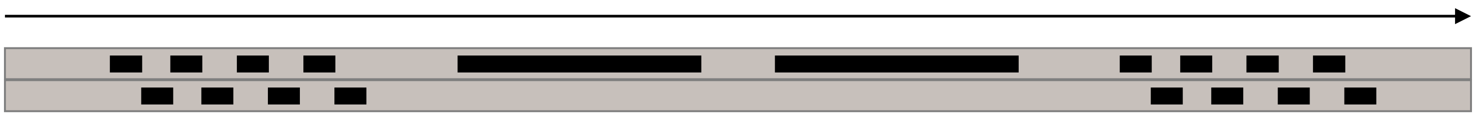


- Formally:
  - Example of 1 patient, 2 administered drugs
  - Before: 1 table with 1 line + 1 table with 2 lines
  - After: 1 table with 1 line

Book music

Street organ (Credit: Roman Bonnefoy - Creative Commons Attribution-Share Alike 3.0)

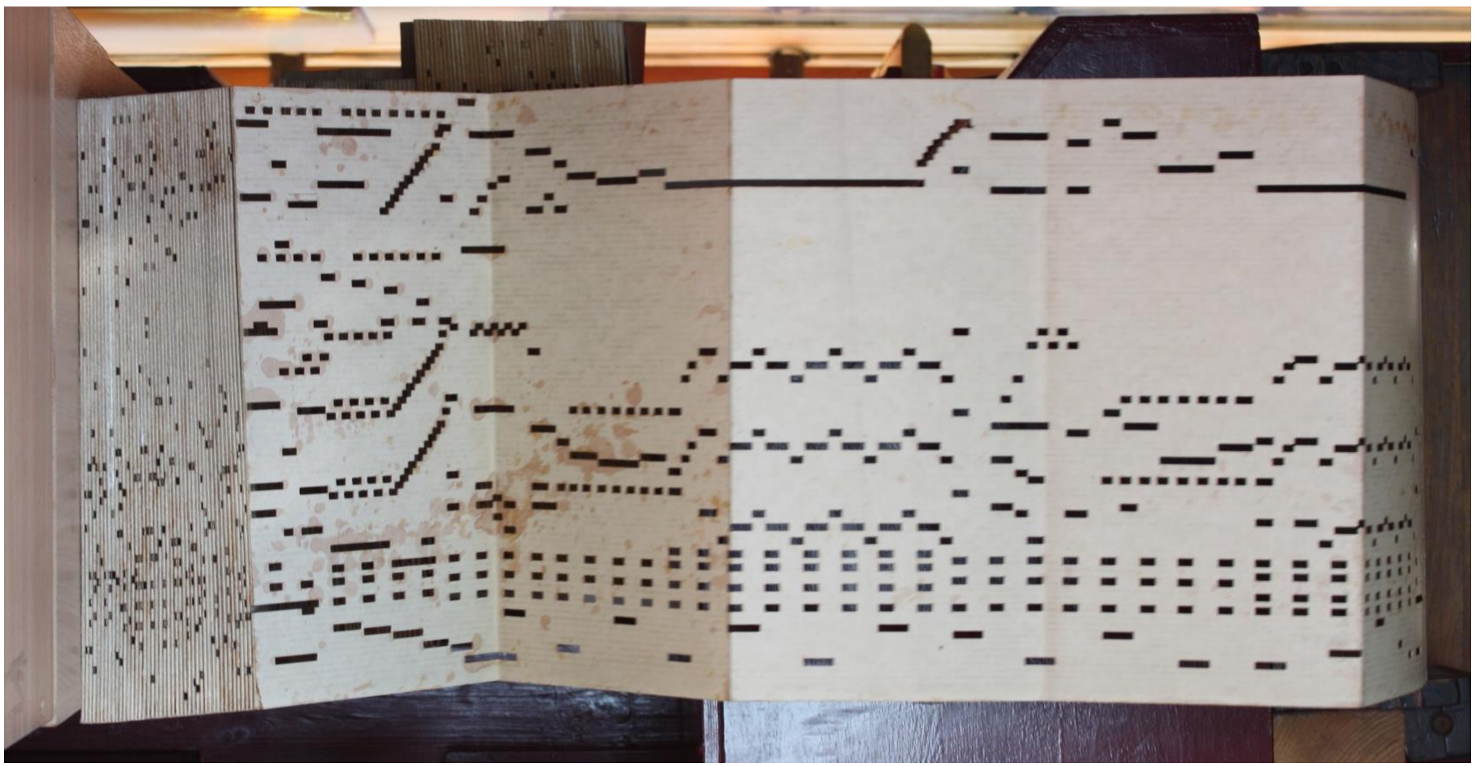
Book music



Time axis

Note (binary: hole=on, no hole=off)

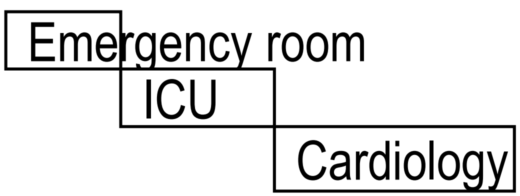
All notes are synchronized



Book music (resized picture - Credit: Richard Ash - Creative Commons Attribution-Share Alike 2.0)

e.g. representation of patient's flow

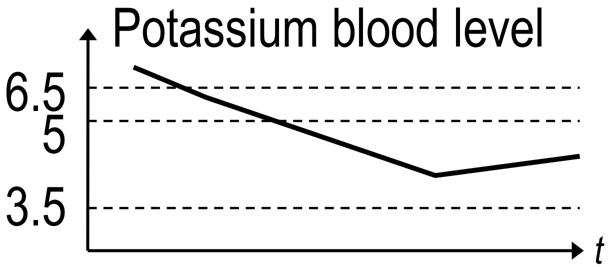
Data in the datawarehouse



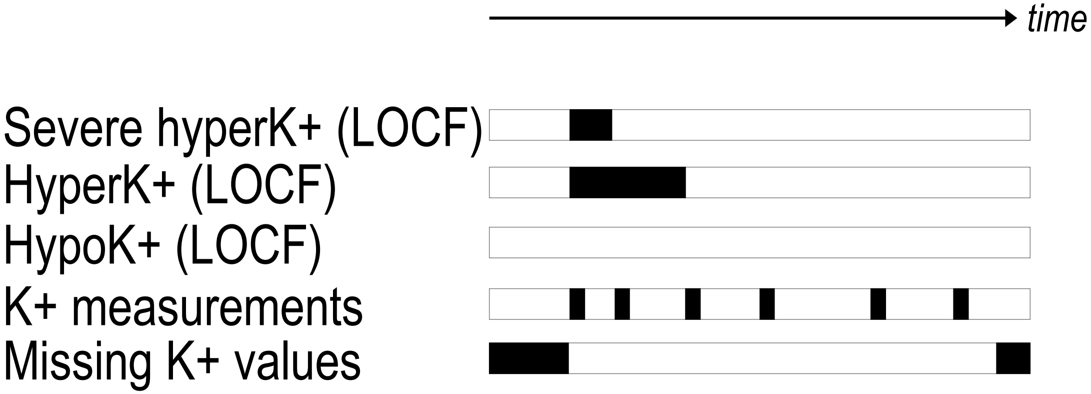
Example of "book music" data representation

e.g. representation of laboratory results

Data in the datawarehouse

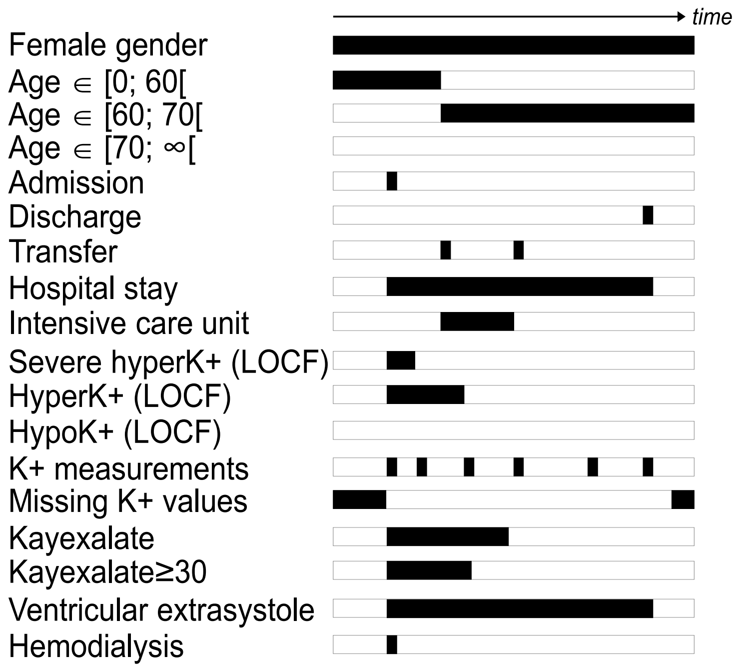


Example of "book music" data representation

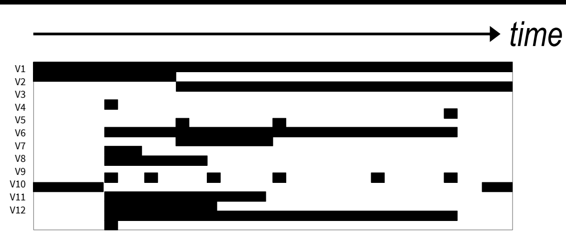


e.g. tabular representation

Book music (expanded view)



Book music (condensed view)



Example of tabular representation

id	Variable	From	To
1	female	1960-06-30	+inf
1	age0_60	1960-06-30	2020-06-29
1	Age60_70	2020-06-30	2030-06-29
…	…	…	…
1	transfer	2020-06-30	2020-06-30
1	transfer	2020-07-03	2020-07-03
1	inhospital	2020-06-29	2020-07-10
…	…	…	…
1	hyperk+	2020-06-29	2020-06-30
1	k+measure	2020-06-29	2020-06-29
1	k+measure	2020-06-29	2020-06-29
…	…	…	…

The place of domain-specific knowledge

- Literature review published par Meystre et al. in 2017
- Classifications performed by Arnaud Dezetrée & Adrien Lecoeuvre

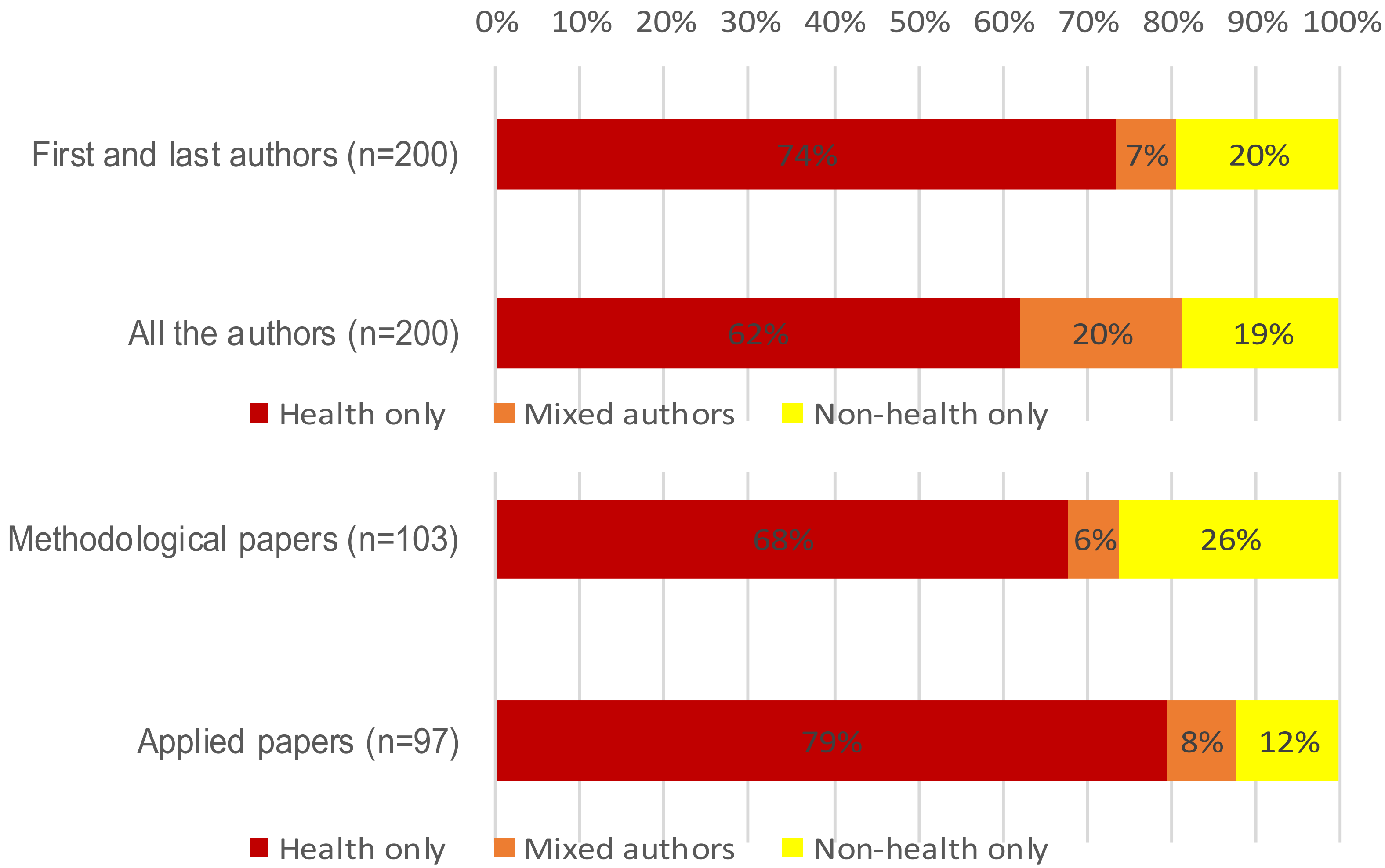
38

Clinical Data Reuse or Secondary Use: Current Status and Potential Future Progress

S. M. Meystre[a], C. Lovis[b], T. Bürkle[c], G. Tognola[d], A. Budrionis[e], C. U. Lehmann[f]

[a] Medical University of South Carolina, Charleston, SC, USA
[b] Division of Medical Information Sciences, University Hospitals of Geneva, Switzerland
[c] University of Applied Sciences, Bern, Switzerland
[d] Institute of Electronics, Computer and Telecommunication Engineering, Italian Natl. Research Council IEIIT-CNR, Milan, Italy
[e] Norwegian Centre for E-health Research, University Hospital of North Norway, Tromsø, Norway
[f] Departments of Biomedical Informatics and Pediatrics, Vanderbilt University Medical Center, Nashville, TN, USA

The place of domain-specific knowledge



Needs / wishes

- Defining methodological framework or methods for feature extraction

- Methods for statistical mining of time-dependent data