

# Developing risk prediction methods in health care

*And also a depersonalized searchable database for health care data*

CHALLENGES: Health care, prevention

PRODUCTIVE SECTOR: Clinical informatics

## PROBLEM DESCRIPTION

Risk prediction methods aim to substitute time and money consuming clinical inferring methods.

## CHALLENGES AND GOALS

The research focused on the following topics:

- How to build a searchable database from everyday health data?
- What kind of risk prediction methods could be obtained from such a database?
- How to find unexpected signals in databases?

## MATHEMATICAL AND COMPUTATIONAL METHODS

A wide range of mathematical models were applied. High dimensional regression and Bayesian statistics were used in risk prediction methods, cryptography were used to depersonalize the data

GYFkód	B	DBK	DB	DRB	D	BK	K	RBK	DRK	DBR
0643	4,91%	6,54%	8,59%	12,47%	7,98%	3,07%	1,23%	2,25%	2,04%	5,73%
1501	3,41%	11,22%	8,05%	22,20%	8,05%	1,71%	2,93%	1,46%	2,44%	3,17%
0242	6,52%	10,87%	7,88%	1,63%	10,33%	4,62%	1,36%	0,27%	0,54%	3,26%
0901	7,21%	10,34%	12,54%	2,82%	9,72%	2,82%	0,94%	0,31%	19,12%	2,19%
0502	4,67%	7,67%	5,67%	18,33%	7,33%	1,67%	5,67%	2,67%	1,67%	4,33%
0140	13,95%	6,46%	5,78%	3,74%	4,42%	11,56%	7,14%	0,34%	0,68%	0,34%
0801	10,54%	1,36%	0,68%	1,70%	0,34%	10,88%	8,84%	14,29%	0,68%	0,00%
0115	4,41%	12,87%	9,56%	2,57%	10,29%	6,25%	2,94%	1,84%	1,47%	1,47%
01A6	12,40%	0,78%	1,16%	0,39%	0,39%	12,79%	11,63%	4,26%	0,39%	0,00%
1901	18,22%	10,08%	4,65%	12,79%	3,10%	5,43%	1,94%	3,88%	1,16%	1,55%
1801	4,71%	5,10%	7,06%	12,55%	5,49%	3,14%	0,00%	5,10%	4,31%	5,88%
2001	3,25%	16,67%	10,16%	8,94%	6,50%	2,44%	0,81%	4,47%	4,88%	5,69%
1401	1,38%	11,52%	6,91%	8,29%	13,82%	0,92%	0,92%	0,92%	5,53%	6,91%
1309	8,16%	13,27%	13,78%	3,06%	11,73%	5,10%	4,08%	2,04%	1,02%	2,04%
0702	4,42%	18,78%	11,60%	13,81%	11,05%	1,10%	0,55%	0,55%	1,10%	3,31%
0112	6,40%	13,95%	15,70%	3,49%	11,63%	1,74%	1,16%	0,00%	1,16%	1,16%
1601	18,71%	0,00%	0,00%	0,58%	0,00%	15,20%	7,02%	5,85%	0,00%	0,00%
1304	33,93%	0,00%	0,00%	0,00%	0,00%	24,40%	3,57%	2,38%	0,00%	0,00%
0154	5,42%	0,00%	3,61%	5,42%	1,20%	4,22%	9,04%	17,47%	3,01%	0,00%
1701	6,25%	12,50%	16,25%	6,25%	6,88%	3,13%	2,50%	0,00%	0,00%	3,13%
1101	11,25%	10,63%	8,75%	7,50%	4,38%	10,00%	8,13%	1,88%	0,63%	2,50%
0701	14,56%	1,90%	0,63%	1,90%	1,90%	6,33%	22,78%	1,90%	0,63%	2,53%
0403	5,30%	16,56%	15,89%	8,61%	7,28%	3,31%	1,99%	0,00%	9,93%	7,95%
1201	9,40%	20,81%	9,40%	0,67%	4,03%	16,78%	2,68%	0,67%	0,00%	0,00%
13B2	11,41%	7,38%	8,05%	6,71%	7,38%	5,37%	5,37%	6,04%	1,34%	6,04%
0940	4,76%	4,08%	9,52%	1,36%	8,16%	2,04%	3,40%	0,68%	12,93%	2,72%
2002	2,74%	17,81%	10,96%	5,48%	16,44%	4,11%	0,68%	0,68%	6,16%	4,11%
2237	12,14%	12,86%	11,43%	2,86%	8,57%	4,29%	1,43%	0,00%	1,43%	2,14%
0402	12,86%	1,43%	0,00%	1,43%	1,43%	8,57%	8,57%	0,71%	0,00%	0,00%
0116	12,21%	12,21%	14,50%	3,82%	11,45%	6,87%	1,53%	0,00%	1,53%	0,00%
0301	14,06%	3,91%	10,16%	9,38%	7,81%	3,91%	3,91%	3,13%	0,00%	4,69%
0106	11,48%	2,46%	1,64%	1,64%	0,82%	14,75%	8,20%	5,74%	2,46%	1,64%
0101	4,92%	7,38%	7,38%	17,21%	7,38%	3,28%	4,92%	10,66%	0,00%	1,64%
0306	9,92%	7,44%	19,83%	7,44%	7,44%	6,61%	5,79%	2,48%	0,00%	3,31%
0902	2,48%	8,26%	11,57%	4,96%	12,40%	0,00%	0,83%	0,00%	33,06%	1,65%
1011	20,66%	0,00%	1,65%	0,83%	0,83%	9,09%	4,13%	10,74%	0,00%	0,00%
1503	2,70%	12,61%	12,61%	17,12%	13,51%	4,50%	2,70%	2,70%	1,80%	2,70%
1202	4,72%	17,92%	19,81%	4,72%	14,15%	5,66%	0,00%	0,00%	1,89%	3,77%
0509	16,98%	0,00%	0,94%	0,00%	0,00%	6,60%	5,66%	19,81%	0,00%	2,83%
Végösszei	8,73%	8,64%	7,98%	7,03%	6,84%	5,92%	4,07%	3,28%	3,16%	2,73%

*Cancer related health care codes and their frequency in the database*

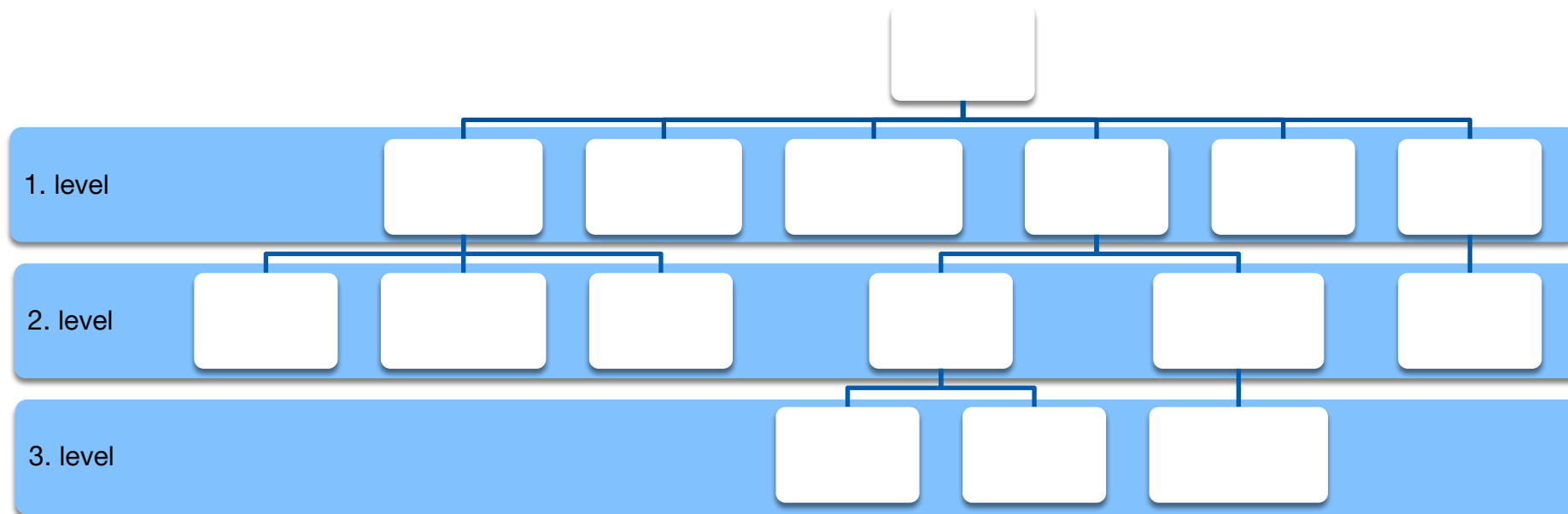
# Developing risk prediction methods in health care

*And also a depersonalized searchable database for health care data*

## Results and Benefits

The online, depersonalized database is on and working. With the depersonalization method, the life pathways of individual patients is still searchable. This allows that once any disease is developed and registered in the database, the antecedent data can be searched. This definitely help researchers to develop better risk models.

The Hospitaly Kft. developed a **depersonalized, searchable** database based on national health care data.



Possible treatments of colorectal cancer patients, as searchable in our database