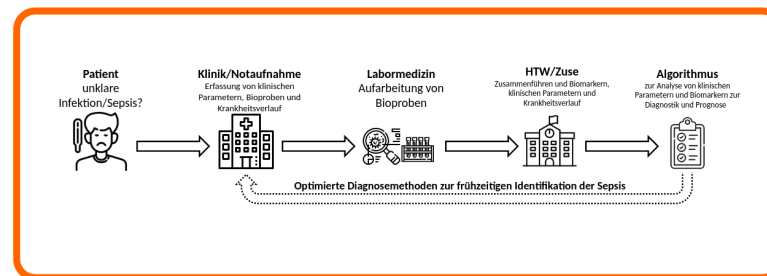


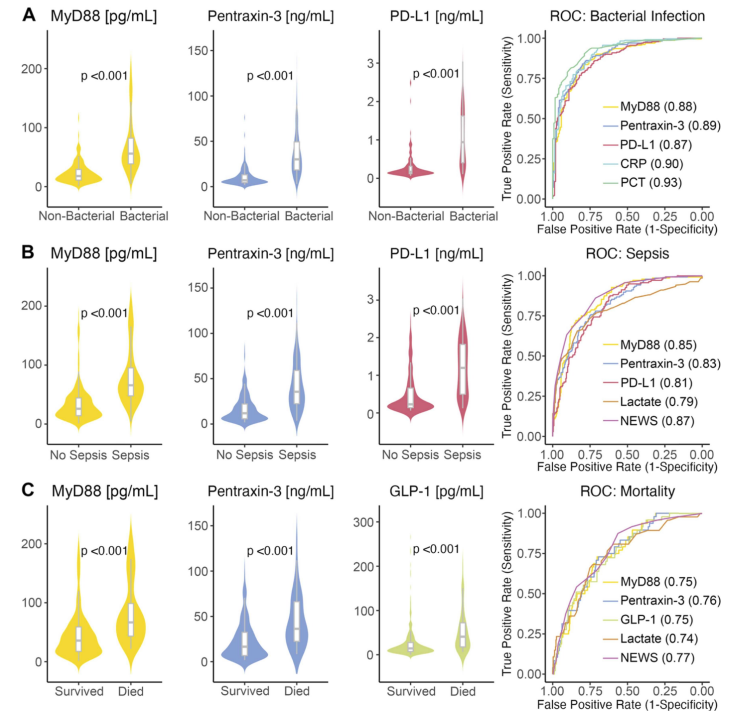
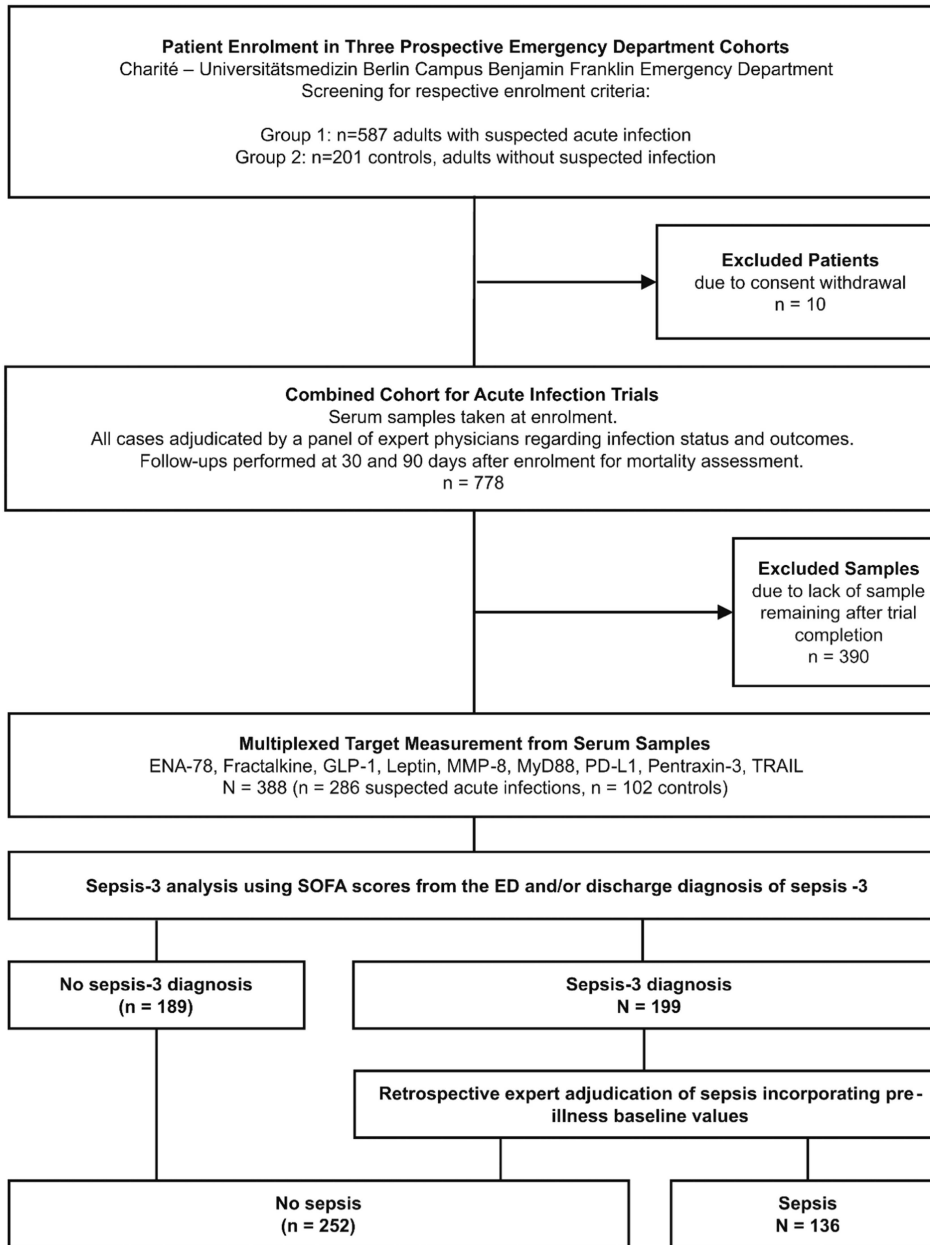
## Logical Data Analysis

There exist more true logical statements than atoms in the universe

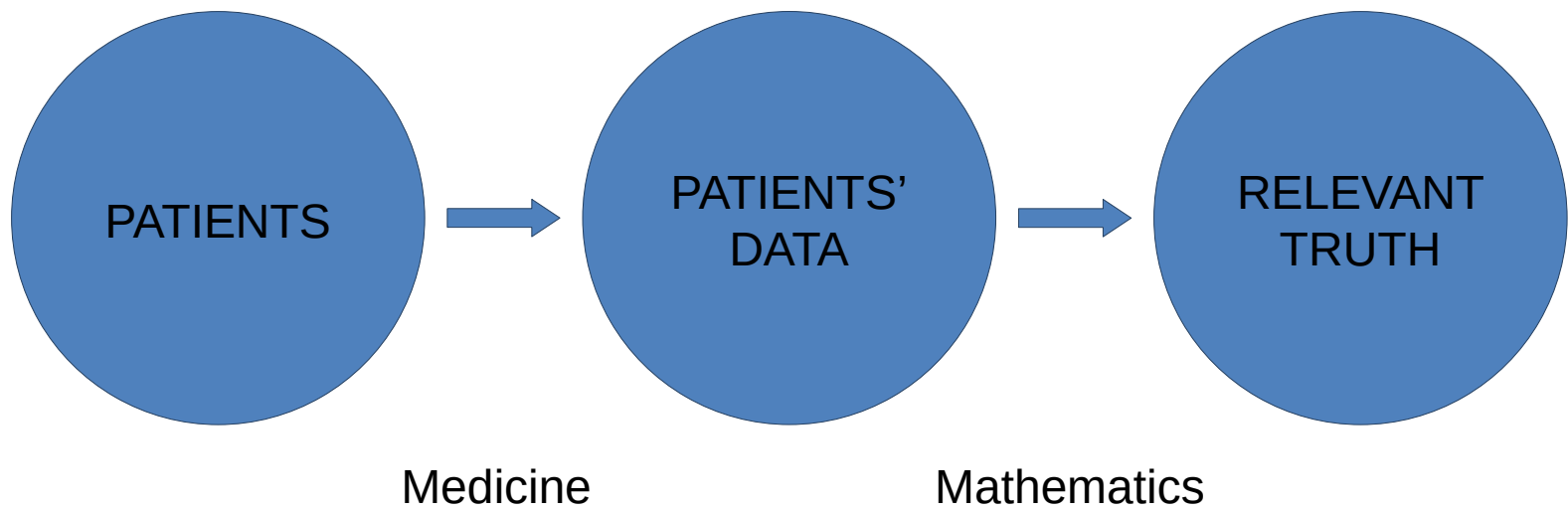
Marcus Weber\*, Kai Kappert, Marco Reidelbach, Ambros Gleixner, Konstantin Fackeldey, Wolfgang Bauer

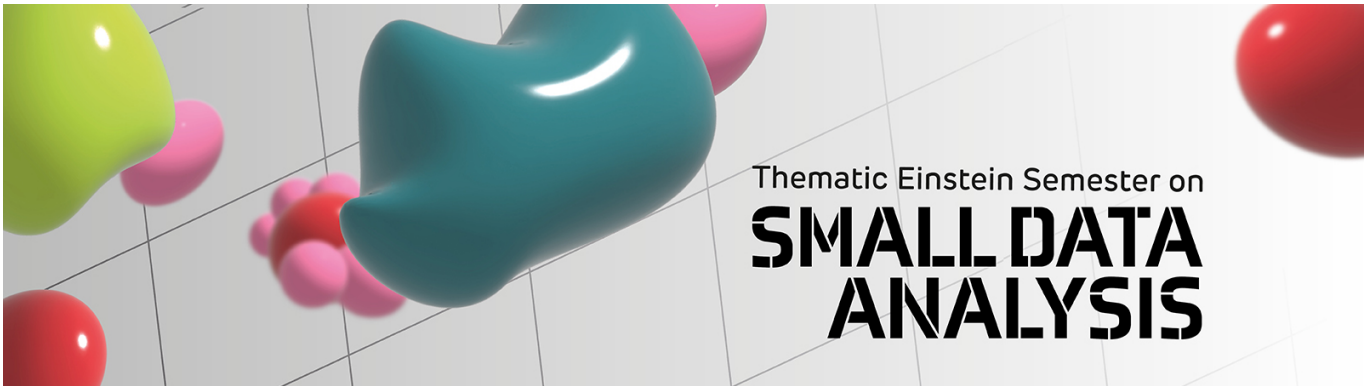


# SEPSIS



Wolfgang Bauer, Noa Galtung, Peter Geserick, Katharina Friedrich, Marcus Weber, Rajan Somasundaram, Eva Diehl-Wiesenecker, Kai Kappert: Pentraxin-3, MyD88, GLP-1, and PD-L1: Performance assessment and composite algorithmic analysis for sepsis identification *Journal of Infection*, Volume 91, Issue 3, 2025





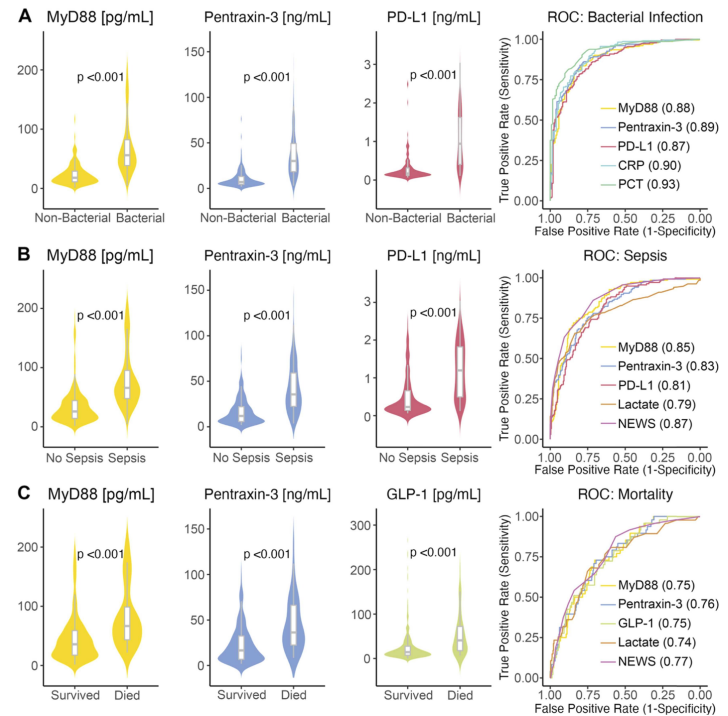
Medicine

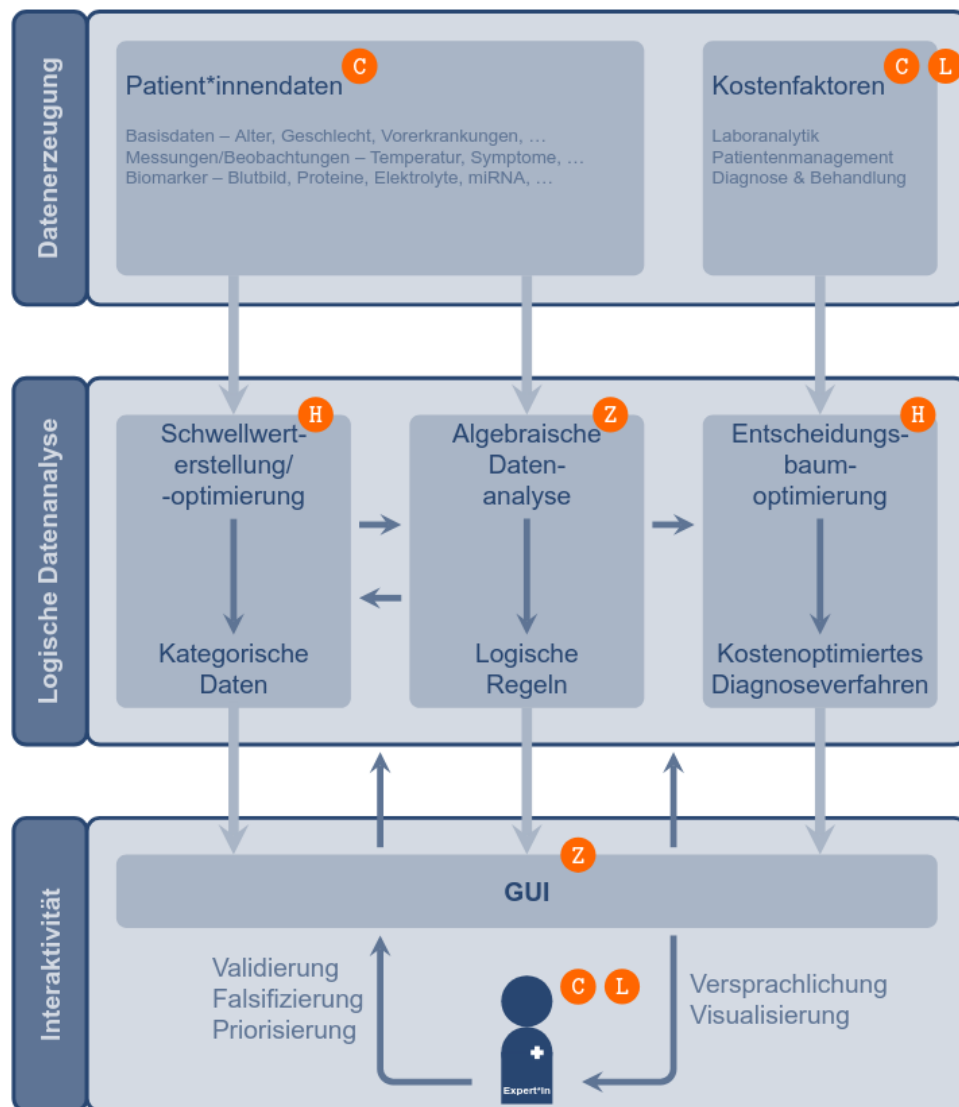
Mathematics

Thematic Einstein Semester on  
**SMALL DATA  
 ANALYSIS**

Not every true logical statement about sepsis is relevant. The relevance of true logical statements can only be accessed by using experts' or external knowledge not necessarily provided by the given dataset itself.

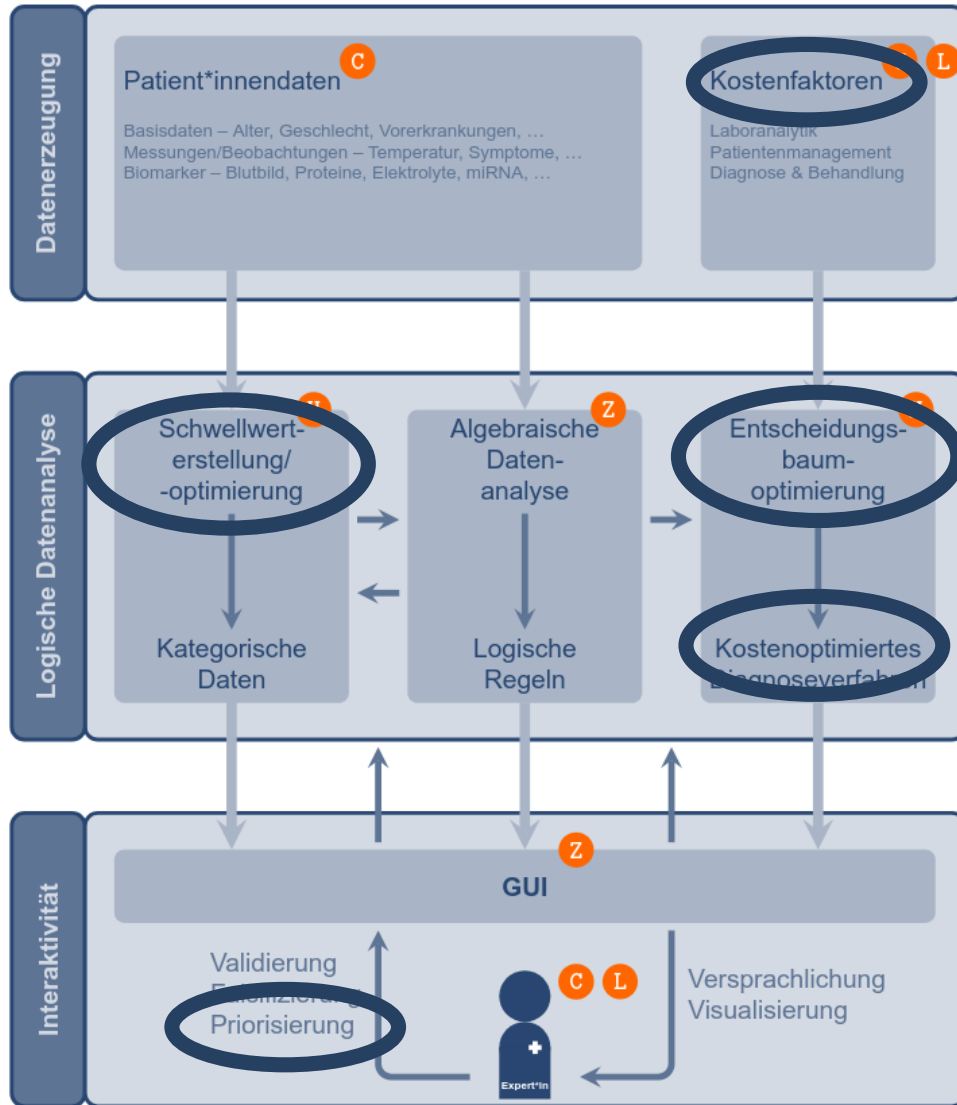
Not every rare observation is an exception. The relevance of observations can only be accessed by using experts' insights with regard to the study.





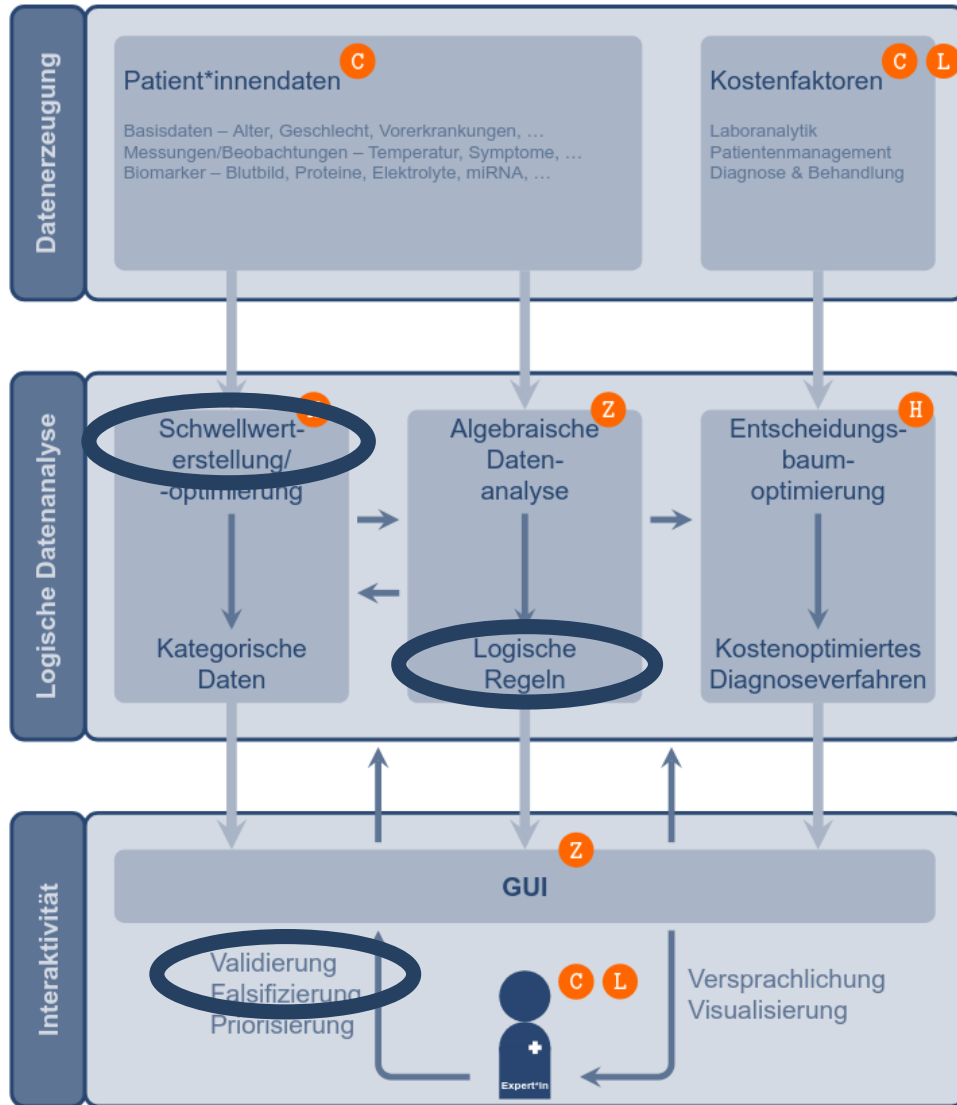
DaKoDi,  
BMFTR  
Feb. 2026

RELEVANCE

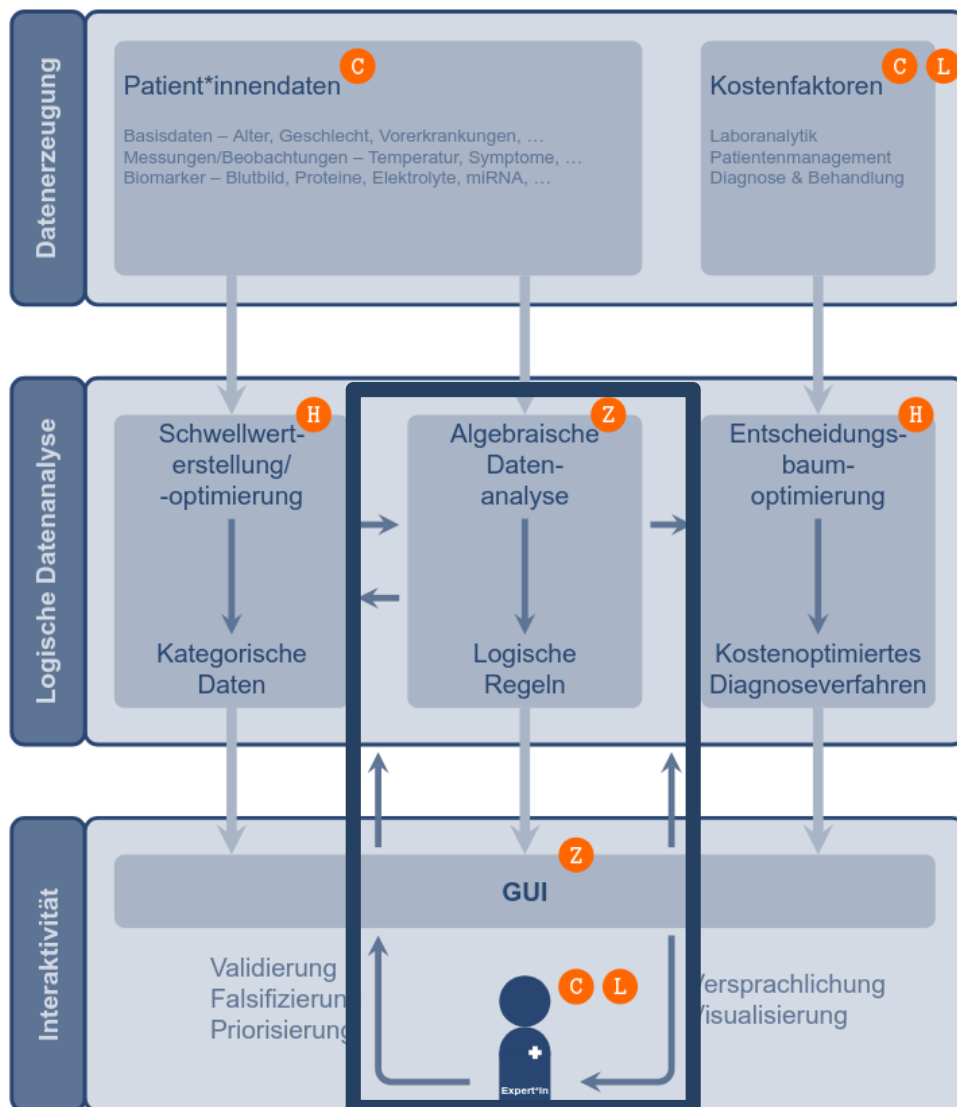


DaKoDi,  
BMFTR  
Feb. 2026

# EXCEPTIONS



DaKoDi,  
BMFTR  
Feb. 2026



DaKoDi,  
BMFTR  
Feb. 2026

protein	threshold
(E)NA-78	> 786.51
(F)ractalkine	> 17307.21
(G)LP-1	> 22.71
(L)eptin	> 20818.6
(M)MP-8	> 59704.08
M(y)D88	> 51.88
(P)D-L1	> 844.84
Pentra(x)in-3	> 28200.36
(T)RAIL	> 92.8

		C	D	E	F	G	H	I	J	K
		Fractalkine	GLP-1	Leptin	MMP-8	MyD88	PD-L1 (epito	Pentraxin-3	TRAIL	sepsis
		0	0	1	1	1	0	1	0	1
		1	1	0	1	1	1	0	0	1
		1	0	0	0	0	1	1	0	1
		1	0	0	0	1	1	1	0	0
		1	0	0	0	1	1	1	0	0
		0	0	1	0	0	0	0	0	1
		0	1	0	1	0	0	0	1	0
9	121	1	1	0	0	0	0	0	0	0
10	125	0	0	0	0	0	0	0	0	0
11	126	0	1	0	0	0	0	0	1	0
12	127	0	0	0	0	0	0	1	1	1
13	129	0	1	0	1	0	0	0	1	0
14	133	0	1	1	0	0	1	0	1	1
15	134	1	1	0	0	0	1	0	1	1
16	135	1	1	0	0	1	1	0	1	1
17	137	0	0	0	0	1	0	0	0	0
18	144	0	1	0	0	1	0	0	1	0
19	145	0	1	1	0	1	1	0	1	1
20	148	0	1	0	1	0	1	1	1	0
21	149	0	0	0	0	0	0	0	0	0
22	154	0	0	0	0	0	1	1	0	1
23	155	0	1	0	0	1	0	1	0	0
24	156	0	1	1	0	0	0	1	0	0
25	157	0	0	0	1	1	0	1	0	0
26	158	1	0	0	1	0	0	0	0	0
27	159	0	0	1	0	1	1	1	1	0

$$A \cdot (B + 1) = A \cdot (B + 1) \cdot (s + 1)$$

AND                      XOR

If A has a high concentration and B does not have a high concentration, then the patient does not have sepsis.

$$0 = A(B + 1)s$$

$$A \cdot (B + 1) = A \cdot (B + 1) \cdot (s + 1)$$

$$0 = A(B + 1)s$$

*A true logical statement which is expressed by an equation of Boolean expressions ( $S_1 = S_2$ ) with identical truth values for each observation is equivalent to a Boolean expression ( $S_1 + S_2 = 0$ ) which is always false.  $S_1 + S_2$  applied as a selection criterion does not select an observed measurement.*

true logical statement = a selection expression which does not select any observed patient

1	A	B	C	D	E	F	G	H	I	J	K
ID	ENA-78	Fractalkine	GLP-1	Leptin	MMP-8	MyD88	PD-L1 (epitor)	Pentraxin-3	TRAIL	sepsis	
2	101	0	0	1	1	1	0	0	1	0	1
3	108	1	1	0	1	1	1	1	0	0	1
4	110	0	0	0	0	0	0	0	0	0	0
5	113	1	0	0	0	1	1	1	1	0	0
6	115	1	0	0	0	1	1	1	1	0	0
7	116	0	0	1	0	0	0	0	0	0	1
8	119	0	1	0	1	0	0	0	0	1	0
9	121	1	1	0	0	0	0	0	0	1	0
10	125	0	0	0	0	0	0	0	0	0	0
11	126	0	1	0	0	0	0	0	0	1	0
12	127	0	0	0	0	0	0	0	1	1	1
13	129	0	1	0	1	0	0	0	0	1	0
14	133	0	1	1	0	0	1	0	1	0	1
15	134	1	1	0	0	0	1	1	0	1	1
16	135	1	1	0	0	1	1	0	1	1	1
17	137	0	0	0	0	1	0	0	0	0	0
18	144	0	1	0	0	1	0	0	1	0	0
19	145	0	1	1	0	1	1	0	1	0	1
20	148	0	1	0	1	0	1	1	1	1	0
21	149	0	0	0	0	0	0	0	0	1	0
22	154	0	0	0	0	0	1	1	0	0	1
23	155	0	1	0	0	1	0	1	0	1	0
24	156	0	1	1	0	0	0	0	1	0	0
25	157	0	0	0	1	1	0	1	1	0	0
26	158	1	0	0	1	0	0	0	0	0	0
27	159	0	0	1	0	1	1	1	1	1	0

Having 10 binary digits at hand, i.e., 9 proteins plus (s)epsis, we can in principle create  $2^{10} = 1024$  different patterns. Out of these possible patterns, we only have observed 194. We have not observed 830 possible patterns. How many selection criteria exist which select a **subset** of these 830 non-observed patterns? It is  $2^{830}$ , because every non-observed pattern can either be part of such a subset or not. This means, our given data provides the possibility to formulate  $2^{830}$  “empty” selection criteria leading to true logical statements about the relationship of the 10 Boolean variables. This number is much bigger than the number of atoms in the universe.

protein	threshold
(E)NA-78	> 786.51
(F)ractalkine	> 17307.21
(G)LP-1	> 22.71
(L)eptin	> 20818.6
(M)MP-8	> 59704.08
M(y)D88	> 51.88
(P)D-L1	> 844.84
Pentra(x)in-3	> 28200.36
(T)RAIL	> 92.8

$$FTs(y + 1), \quad (F + 1)xTs \in \mathcal{I}$$

protein	threshold
(E)NA-78	> 786.51
(F)ractalkine	> 17307.21
(G)LP-1	> 22.71
(L)eptin	> 20818.6
(M)MP-8	> 59704.08
M(y)D88	> 51.88
(P)D-L1	> 844.84
Pentra(x)in-3	> 28200.36
(T)RAIL	> 92.8

$$FTs(y + 1), \quad (F + 1)xTs \in \mathcal{I}$$

$$FTs(y + 1) \cdot E \in \mathcal{I}$$

$$FTs(y + 1) + (F + 1)xTs \in \mathcal{I}$$

$\mathcal{I}$  is an ideal inside the set of Boolean polynomials.

..., -15, -10, -5, 5, 10, 15, ...

<5>

...,  $x^2y+xy$ ,  $7zx+7z$ ,  $x$ ,  $y$ ,  $xy$ ,  $xy+y$ , ...

$\langle x+1, x, y \rangle$

$\sigma \in \mathcal{I}$  which exactly selects all unobserved patterns generates  $\mathcal{I}$

1	A	B	C	D	E	F	G	H	I	J	K
ID	ENA-78	Fractalkine	GLP-1	Leptin	MMP-8	MyD88	PD-L1 (epitor)	Pentraxin-3	TRAIL	sepsis	
2	101	0	0	1	1	1	0	0	1	0	1
3	108	1	1	0	1	1	1	1	0	0	1
4	112	1	0	0	0	0	1	1	1	0	1
5	113	1	0	0	0	1	1	1	1	0	0
6	115	1	0	0	0	1	1	1	1	0	0
7	116	0	0	1	0	0	0	0	0	0	1
8	119	0	1	0	1	0	0	0	0	1	0
9	121	1	1	0	0	0	0	0	0	1	0
10	125	0	0	0	0	0	0	0	0	0	0
11	126	0	1	0	0	0	0	0	0	1	0
12	127	0	0	0	0	0	0	0	1	1	1
13	129	0	1	0	1	0	0	0	0	1	0
14	133	0	1	1	0	0	1	0	1	0	1
16	135	1	1	0	0	1	1	0	1	1	1
17	137	0	0	0	0	1	0	0	0	0	0
18	144	0	1	0	0	1	0	0	1	0	0
19	145	0	1	1	0	1	1	0	1	0	1
20	148	0	1	0	1	0	1	1	1	1	0
21	149	0	0	0	0	0	0	0	0	1	0
22	154	0	0	0	0	0	1	1	0	0	1
23	155	0	1	0	0	1	0	1	0	1	0
24	156	0	1	1	0	0	0	0	1	0	0
25	157	0	0	0	1	1	0	1	1	0	0
26	158	1	0	0	1	0	0	0	0	0	0
27	159	0	0	1	0	1	1	1	1	1	0

E	F	G	L	M	y	P	x	T	s
1877	90658	22	561	60876	81	479	52470	408	yes
1	1	0	0	1	1	0	1	1	1
E	F	G+1	L+1	M	y	P+1	x	T	s

```

11 -> y*x*T*s + x*T*s
12 -> F*x*T*s + x*T*s
13 -> y*P*T*s + P*x*T*s + G*M*y + G*y*x + E*y*T + L*y*T + M*y*T + y*x*T + P*x*T
+ G*M*s + M*y*s + G*x*s + y*x*s + E*T*s + L*T*s + M*T*s + P*T*s + x*T + M*s +
x*s
14 -> M*P*T*s + y*P*T*s + G*x*T*s + M*x*T*s + P*x*T*s + G*T*s + P*T*s + x*T*s
15 -> G*P*T*s + F*L*y + G*M*y + L*M*y + F*L*P + L*M*P + L*y*P + G*y*x + E*y*T +
F*y*T + L*y*T + G*P*T + F*x*T + G*x*T + M*x*T + P*x*T + G*M*s + M*y*s + G*x*s +
y*x*s + E*T*s + F*T*s + G*T*s + L*T*s + M*T*s + y*T*s + P*T*s + L*P + M*s + x*s
16 -> F*P*T*s + y*P*T*s + F*T*s + y*T*s
17 -> E*P*T*s + L*P*T*s + y*P*T*s + G*x*T*s + L*x*T*s + M*x*T*s + E*T*s + F*T*s
+ G*T*s + M*T*s
18 -> M*y*T*s + G*M*y + G*y*x + G*M*s + M*y*s + G*x*s + y*x*s + M*T*s + M*s +
x*s
19 -> L*y*T*s + M*y*T*s + L*T*s + M*T*s
20 -> G*y*T*s + G*T*s
21 -> F*y*T*s + F*T*s
22 -> E*y*T*s + M*y*T*s + y*P*T*s + E*T*s + M*T*s + P*T*s
23 -> F*M*T*s + M*y*T*s + E*P*T*s + L*P*T*s + E*x*T*s + G*x*T*s + L*x*T*s +
G*T*s + P*T*s + x*T*s
24 -> E*M*T*s + G*M*T*s + M*y*T*s + y*P*T*s + E*x*T*s + M*x*T*s + G*T*s + M*T*s
+ P*T*s + x*T*s
25 -> G*L*T*s + G*M*T*s + L*M*T*s + L*P*T*s + E*x*T*s + G*x*T*s + M*x*T*s +
G*L*M + E*F*y + E*G*y + F*G*y + E*L*y + E*M*y + E*F*P + E*y*P + F*y*P + L*y*P +
E*F*x + F*G*x + G*M*x + F*y*x + G*y*x + L*y*x + M*y*x + F*P*x + G*P*x + L*P*x +
M*P*x + y*P*x + G*M*T + E*y*T + F*y*T + G*y*T + L*y*T + M*y*T + E*P*T + L*P*T +
y*P*T + E*x*T + F*x*T + L*x*T + P*x*T + E*F*s + E*G*s + E*L*s + E*M*s + G*M*s +
L*M*s + E*y*s + F*y*s + L*y*s + M*y*s + L*P*s + M*P*s + y*P*s + E*x*s + F*x*s +
G*x*s + M*x*s + y*T*s + F*y + G*y + E*P + G*P + L*P + y*P + E*x + G*x + P*x +
y*T + E*s + M*s + y*s + y
26 -> F*L*T*s + F*M*T*s + F*y*T*s + E*P*T*s + L*P*T*s + M*P*T*s + y*P*T*s +
E*x*T*s + L*x*T*s + M*x*T*s + P*x*T*s + L*T*s + M*T*s + y*T*s
27 -> E*L*T*s + G*L*T*s + E*M*T*s + G*M*T*s + L*M*T*s + E*y*T*s + G*y*T*s +
M*y*T*s + E*P*T*s + F*P*T*s + M*P*T*s + y*P*T*s + E*x*T*s + M*x*T*s + P*x*T*s +
L*T*s + y*T*s + P*T*s
28 -> F*G*T*s + E*P*T*s + L*P*T*s + E*x*T*s + G*x*T*s + L*x*T*s + P*T*s + x*T*s
29 -> E*G*T*s + G*M*T*s + G*P*T*s + E*x*T*s + M*x*T*s + P*x*T*s

```

## Gröbner Basis

$$S_1 \cdot s + S_2$$

**Positive Rules.** Imagine  $S_2 \notin \mathcal{I}$ , i.e.,  $S_2$  selects a subset of patients. The requirement is  $S_1 \cdot s + S_2 \in \mathcal{I}$ . We can conclude that the selection criterion  $S_1 \cdot s$  must select the same subset of patients like  $S_2$ , because it has to “cancel out with  $S_2$ ”, i.e.  $S_1 \cdot s = S_2 \notin \mathcal{I}$  if applied to the dataset. This means that the patients in this special subset must have sepsis. To give an example:  $GMys + Gyxs + GMy + Gyx$  is an element of the Gröbner Basis of  $\mathcal{I}$  with  $S_1 = GMy + Gyx$  and  $S_2 = GMy + Gyx$ .  $S_2$  selects 22 patients and therefore  $S_2 \notin \mathcal{I}$ . This means, that the 22 patients selected by  $S_2 = Gy(M + x)$  have sepsis. Patients who have a high concentration of GLP-1 and of MyD88 and either a high concentration of MMP-8 or Pentraxin-3 have sepsis.

**Negative Rules.** Imagine  $S_1 \cdot (1 + S_2) \notin \mathcal{I}$ , i.e., the intersection of the set of patients selected by  $S_1 \notin \mathcal{I}$  and the complement of the set selected by  $S_2$  with  $(1 + S_2) \notin \mathcal{I}$  is not empty. In this case,  $S_1 \cdot s + S_2$  can only select an empty set of patients, if the patients selected by  $S_1 \cdot (1 + S_2)$  do not have sepsis. To give an example:  $FyTs + FTs$  is an element of the Gröbner Basis of  $\mathcal{I}$  with  $S_1 = FyT + FT$  and  $S_2 = 0$ , i.e., with  $S_1(S_2 + 1) = FT(y + 1)$ . This selection criterion selects 24 patients who do not have sepsis. Patients with a high concentration of Fractalkine and TRAIL and a low concentration of MyD88 do not have sepsis.

# Relevance?

```

11 -> y*x*T*s + x*T*s
12 -> F*x*T*s + x*T*s
13 -> y*P*T*s + P*x*T*s + G*M*y + G*y*x + E*y*T + L*y*T + M*y*T + y*x*T + P*x*T
+ G*M*s + M*y*s + G*x*s + y*x*s + E*T*s + L*T*s + M*T*s + P*T*s + x*T + M*s +
x*s
14 -> M*P*T*s + y*P*T*s + G*x*T*s + M*x*T*s + P*x*T*s + G*T*s + P*T*s + x*T*s

```

```

11 -> neg: 7 1 * x * T * (y + 1)
12 -> neg: 7 1 * x * T * (F + 1)
13 -> pos: 28 1 * (E*y*T + G*M*y + G*y*x + L*y*T + M*y*T +
y*x*T + P*x*T + x*T) neg: 64 1 * (y + 1) * (E*P*x*T + E*x*T
+ E*T + G*M*P*x*T + G*M*x*T + G*M + G*P*x*T + G*x*T + G*x + L*P*x*T + L*x*T +
L*T + M*T + M + P*T + x*T + x)
14 -> neg: 23 1 * T * (G*x + G + M*P + M*x + y*P + P*x + P +
x)

```

```
: # STEP3
for i=1:length(GB1)
    nf=normal_form(GB1[i], gens(GB2));
    statement_1=derivative(nf,s_index);
    if (nf!=0 && (statement_1 !=0))
        statement_2 = normal_form(s*statement_1 + nf, gens(GB2));
        statement_3 = normal_form(statement_1*(1+statement_2), gens(GB2));
        print(i); print(" -> ");
        if (statement_2 !=0)
            counter=0;
            for j = 1:size(expr_list,1)
                if (evaluate(statement_2, chartable[j,:]*1)==1);
                    counter=counter+1;
                end
            end
            printstyled(" pos: ", color=:green);
            print(counter, " ");
            printstyled(factor(statement_2), color=:green);
        end
        if (statement_3 !=0)
            counter=0;
            for j = 1:size(expr_list,1)
                if (evaluate(statement_3, chartable[j,:]*1)==1);
                    counter=counter+1;
                end
            end
            printstyled(" neg: ", color=:red);
            print(counter, " ");
            printstyled(factor(statement_3), color=:red);
        end
        print("\n");
    end
end
```

## Exception?

E	F	G	L	M	y	P	x	T	s
1877	90658	22	561	60876	81	479	52470	408	yes
1	1	0	0	1	1	0	1	1	1
E	F	G+1	L+1	M	y	P+1	x	T	s

17 -> 148 (#1)

$1 * F * L * y * P * x * T * (s + 1) * (M + 1) * (G + 1) * (E + 1)$

$L * M * x * T * s + L * M * x * T + L * y * x * T + L * P * x * T * s + L * P * x * T + L * x * T$

```

# STEP4
for i = 1:size(uq_expr_list,1)
    res=normal_form(uq_expr_list[i],gens(GB1));
    counter=0;
    printstyled(i, color=:red);
    printstyled(" -> ", color=:red);
    for j = 1:size(expr_list,1)
        if (uq_expr_list[i]==expr_list[j])
            counter=counter+1;
            printstyled(indices[j], color=:red);
            print(" ");
        end
    end
    printstyled(" (#", color=:red);
    printstyled(counter, color=:red);
    printstyled(")", color=:red);
    print("\n");
    printstyled(factor(uq_expr_list[i]), color=:green);
    wrtnorm=0;
    if (wrtnorm==1)
        printstyled("=\n", color=:green);
        printstyled(uq_expr_list[i], color=:green);
    end
    print("\n");
    printstyled(res, bold=:true);
    print("\n\n");
end

```

### Step 1: generating the Gröbner Basis of ideal $\mathcal{I}$

If the generator  $\sigma$  of  $\mathcal{I}$  is computed, we can generate a Gröbner Basis of this ideal. In the course of the algorithm (in Step 4), polynomials will be determined which will also become further generators and extend the ideal  $\mathcal{I}$ .

### Step 2: generating the Gröbner Basis of ideal $\mathcal{J}$

In the course of the algorithm (in Step 3), polynomials of the Gröbner Basis of  $\mathcal{I}$  will be selected and denoted as “relevant”. These polynomials generate a further ideal  $\mathcal{J}$  with its Gröbner Basis. We start with an empty ideal  $\mathcal{J} = \{0\}$ .

### Step 3: inner loop to identify “insights”

Transform the remainder polynomials of the Gröbner Basis of  $\mathcal{I}$  with regard to the ideal  $\mathcal{J}$  into positive and negative rules about sepsis. Then go through this list of rules and figure out, which ones are relevant. The respective polynomials of the Gröbner Basis will be added to the list of generators which generate the ideal  $\mathcal{J}$ . Continue with Step 2, if relevant rules have been found. Otherwise go to Step 4.



### Step 4: outer loop to identify “exceptions”

Go through every observed pattern of the dataset and determine the selection criteria constructed according to Tab. For every such polynomial compute the remainder with regard to ideal  $\mathcal{I}$ . Using these remainders one decides whether a rarely observed pattern is really an exception. All remainders, which belong to “exceptions” are collected and added to the list of generators of ideal  $\mathcal{I}$ . If exceptions have been identified, then continue with Step 1. Otherwise the algorithm stops here.

E	F	G	L	M	y	P	x	T	s
1877	90658	22	561	60876	81	479	52470	408	yes
1	1	0	0	1	1	0	1	1	1
E	F	G+1	L+1	M	y	P+1	x	T	s

cycle No.	polynomials to $\mathcal{J}$ (inner)	polynomials to $\mathcal{I}$ (outer)	exceptions
1	$FyTs + FTs, GMys + GyxS + GMy + Gyx$	$LyPxTs + LPxTs$	No. 2237
2	$LyTs + MyTs + LTs + MTs, ELys + EMys + ELs + EMs$	$yxTs + xTs$	No. 127, 545
3	$EyTs + MyTs + yPTs + ETs + MTs + PTs$		

type (pos/neg)	number (sepsis)	selection criterion
neg	45(1)	$T (y + 1) (L + M)$
neg	24(0)	$F T (y + 1)$
neg	49(0)	$T (y + 1) (E + M + P)$
pos	22(22)	$G y (M + x)$
neg	33(1)	$E (y + 1) (L + M)$
pos	1(1)	$G y T (L M + L x + M x + M)$
neg	1(1)	$G T (y + 1) (L M + L x + M x + M)$
pos	2(2)	$F G y T (M + x)$
neg	3(0)	$F G T (y + 1) (M + x)$
pos	1(1)	$G y T (E M + E x + M P + M x + M + P x)$
neg	3(0)	$G T (y + 1) (E M + E x + M P + M x + M + P x)$
pos	1(1)	$E G y (L M + L x + M x + M)$
neg	2(1)	$E G (y + 1) (L M + L x + M x + M)$

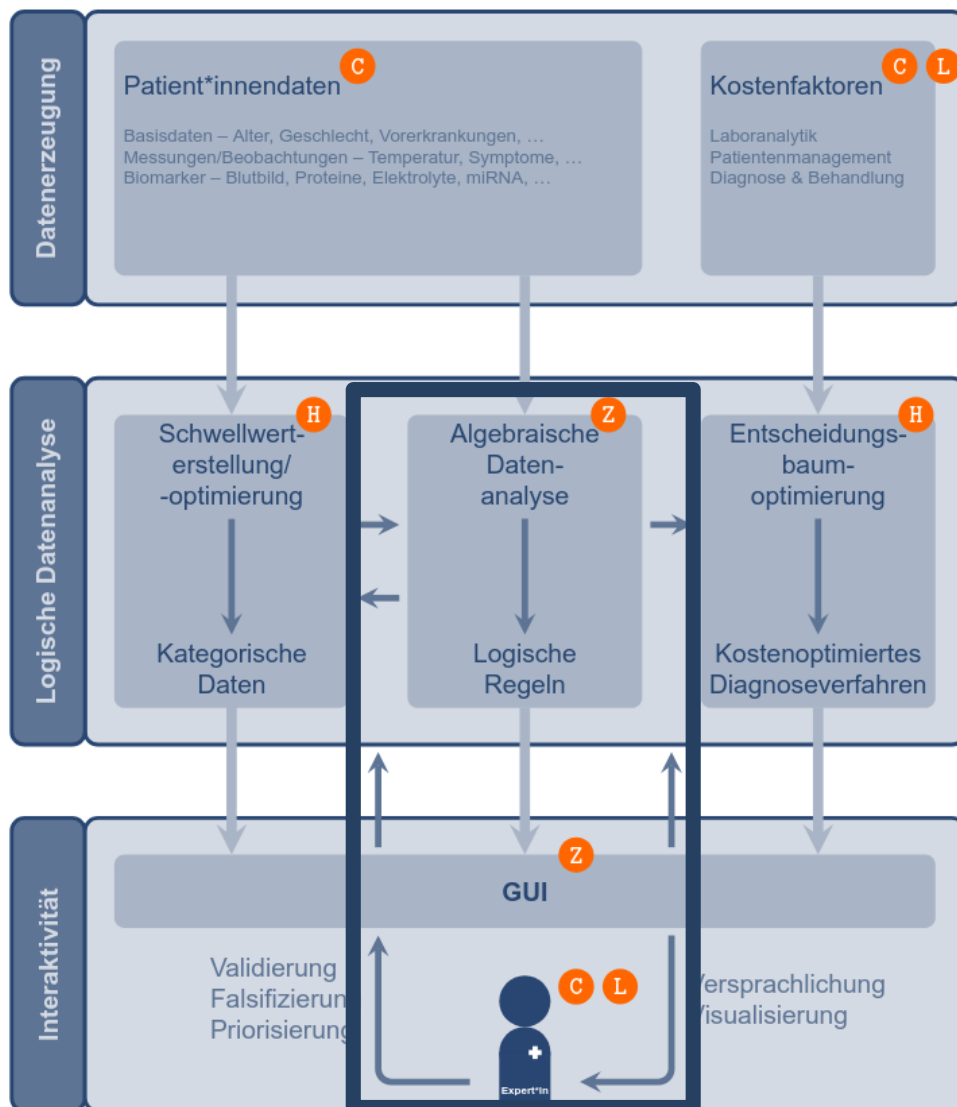
type (pos/neg)	number (sepsis)	selection criterion
neg	108(6)	$T(y + 1)$
neg	110(13)	$(y + 1)(E + M + P)$
neg	102(12)	$(y + 1)(L + M)$
neg	253(39)	$y + 1$
pos	67(62)	$Gy$
pos	58(40)	$y(M + x)$
pos	50(37)	$G(M + x)$
pos	137(98)	$y$

High TRAIL levels  $\Rightarrow$  Anti-inflammatory and apoptotic signaling  
 $\Rightarrow$  Reduced systemic inflammation

Low MyD88 levels  $\Rightarrow$  Attenuated TLR signaling  
 $\Rightarrow$  Lower pro-inflammatory cytokine production

High GLP-1 levels  $\Rightarrow$  Metabolic dysregulation and immune modulation  
 $\Rightarrow$  Associated with increased mortality

High MyD88 levels  $\Rightarrow$  Amplified TLR signaling  
 $\Rightarrow$  Excessive cytokine release and inflammation



DaKoDi,  
BMFTR  
Feb. 2026