ROBERT KOCH INSTITUT

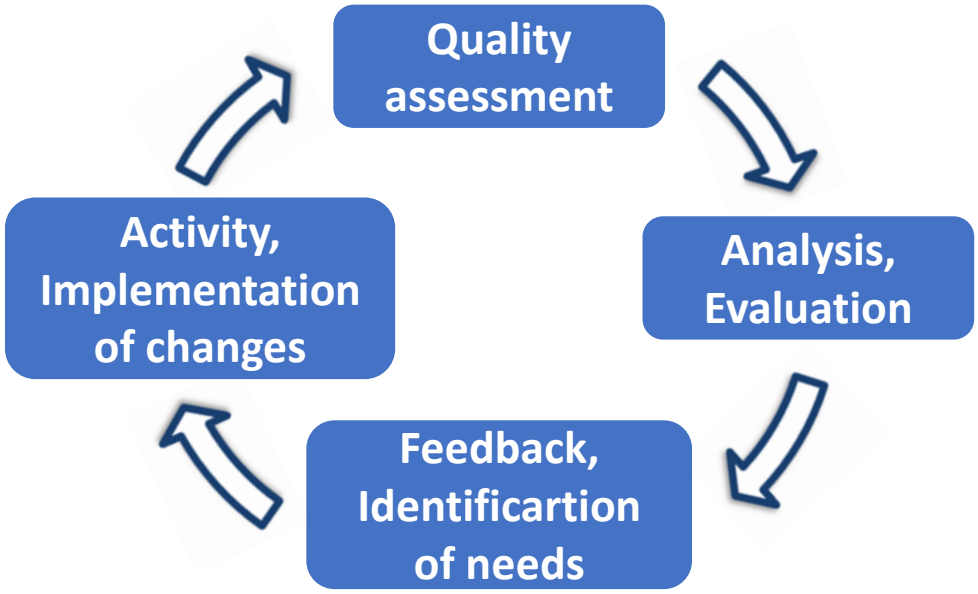
STATISTICAL APPROACH FOR OPTIMIZATION OF EQA STUDIES OF MOLECULAR AND SEROLOGICAL VIRAL DIAGNOSTICS

20th October 2017

Leonid Rumer, C. Domingo, Oliver Donoso Mantke, Yuliya Dobrydneva, Matthias Greiner and Matthias Niedrig

Without diagnostics, medicine is blind.

And yet, diagnostics receive much less attention than vaccines and drugs.



SAMPLE TYPES OF PROFICIENCY PANEL

Laboratory	Types of samples in proficiency panel: Positive samples and their dilutions: 10^0 10^{-1} …10^{-n}	Negative samples: Not targeted viruses	Negative samples: True negative	Score
Lab 1 …… Lab N				

Correct result : 1 or 2 points
Incorrect result : 0 points
Score : Total sum of the points

- Totally about 15 samples in the panel
- Information is requested on:

Methods
– Example: different versions of PCR (TaqMan, Sybr Green, Nested etc.)
Technical factors
– Example: kit for sample treatment, in-house protocol of sample treatment

RESULTS OF THE EQA FOR MOLECULAR DETECTION OF DENV

Lab N	PCR-technique	Sample no. #2	#9	#12	#4	#14	#5	#13	#6	#10	#11	#3	#7	Score	Classification
		DENV-1	DENV-1	DENV-1	DENV-1	DENV-1	DENV-3	DENV-3	DENV-2	DENV-4	JE/ YF/ WN/ TBE	CHIK	Negative		
		Copy no [GE/mL]													
		7.0E+05	7.0E+04	7.5E+03	7.0E+02	7.0E+01	3.0E+04	3.0E+03	1.0E+05	1.0E+05	Neg	Neg	Neg		
8	Hemi-nested	++	++	++	++	++	++	++	++	++	–	–	–	22	Optimal
7	TaqMan	++	++	++	++	(–)	++	++	++	++	–	–	–	22	Optimal
13	SYBR-Green	++	++	++	++	(–)	++	++	++	++	–	–	–	22	Optimal
17a	TaqMan	++	++	++	++	(–)	++	++	++	++	–	–	–	22	Optimal
12	TaqMan	++	++	++	+	++	+	(–)	++	++	–	–	–	20	Improve
21	SYBR-Green[a]	++	++	++	++	(–)	++	++	++	++	(+)	–	–	20	Improve
2a	Nested	++	++	++	(–)	(–)	++	++	++	++	–	–	–	20	Optimal
2b	TaqMan[a]	++	++	++	(–)	(–)	++	(–)	++	++	–	–	–	18	Improve
4b	Nested	++	++	++	(–)	(–)	++	(–)	++	++	–	–	–	18	Improve
28a	Nested[b]	++	++	++	(–)	(–)	++	(–)	++	++	–	–	–	18	Improve
.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....
11	Nested	++	++	(–)	(–)	(–)	+	(–)	+	(–)	(+)	–	–	10	Improve
35a	Nested	++	++	(–)	(–)	(–)	(–)	(–)	(–)	(–)	–	–	–	10	Improve
34	TaqMan	+	+	(–)	(–)	(–)	(–)	(–)	(–)	+	–	–	–	9	Improve
23a	SYBR-Green	+	(–)	(–)	(–)	(–)	(–)	(–)	+	(–)	–	–	–	8	Improve
32	Hemi-nested	++	(–)	(–)	(–)	(–)	(–)	(–)	(–)	(–)	–	–	–	8	Improve
33	TaqMan	+	+	+	(–)	(–)	(–)	(–)	+	+	–	–	–	8	Improve
26	Nested[b]	(–)	(–)	(–)	(–)	(–)	(–)	(–)	(–)	(–)	–	–	–	6	Improve
35b	Nested	(–)	(–)	(–)	(–)	(–)	(–)	(–)	(–)	+	–	–	(+)	5	Improve
Correct positive/total results (%)		43/46 (93.5)	41/46 (89)	23/46 (50)	14/46 (30.4)	8/46 (17.4)	32/46 (69.5)	17/46 (37)	38/46 (82.6)	32/46 (69.5)	40/46 (87)	44/46 (95.6)	44/46 (95.6)		

YOUDEN'S INDEX

- **The index was suggested by W.J. Youden [1] as a way of summarizing the performance of a diagnostic test.**
- **Its value ranges from -1 to 1, and has a zero value when a diagnostic test gives the same proportion of positive results for groups with and without the disease, i.e the test is useless.**
- **A value of 1 indicates that there are no false positives or false negatives, i.e. the test is perfect.**
- **The index gives equal weight to false positive and false negative values, so all tests with the same value of the index give the same proportion of total misclassified results.**

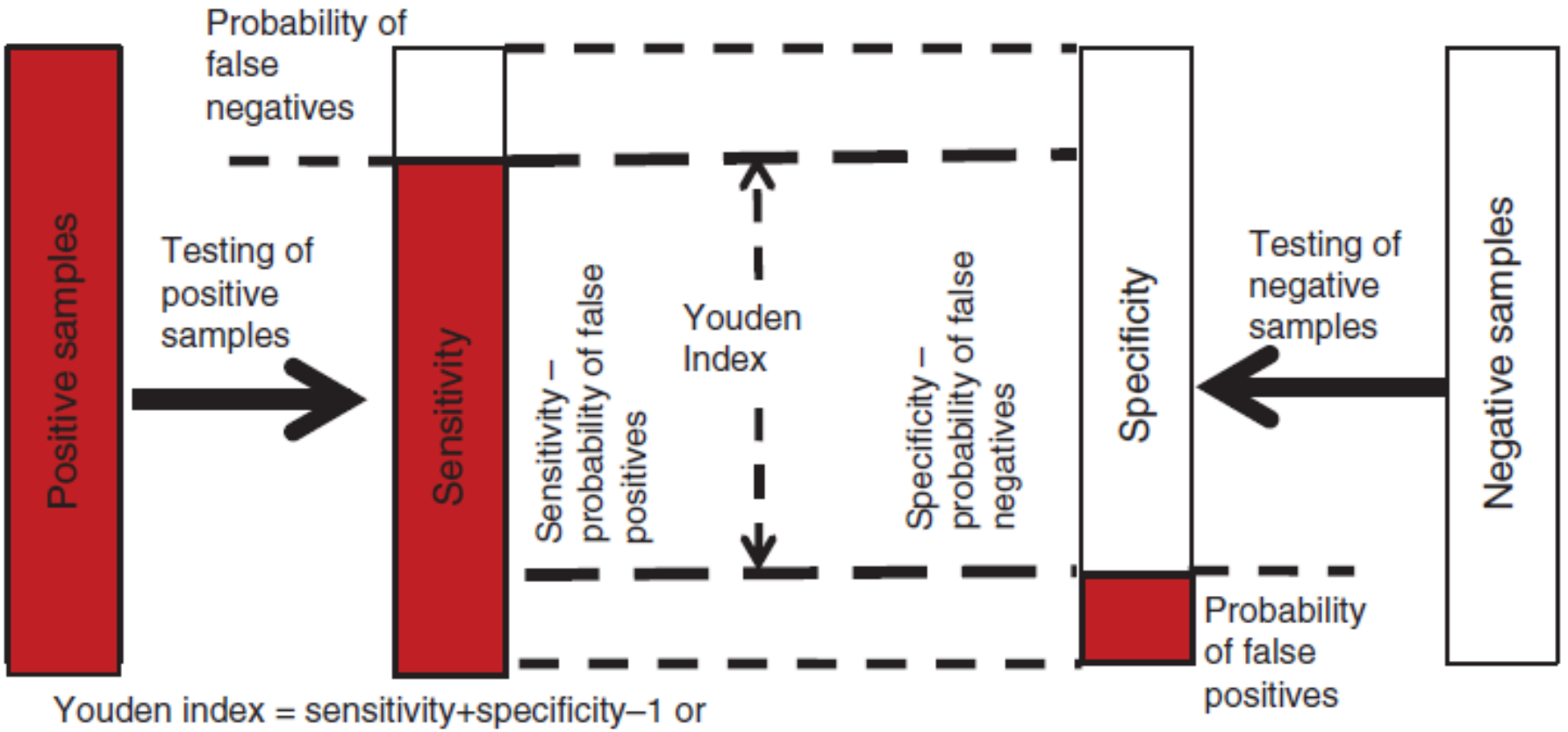
$$J = \text{sensitivity} + \text{specificity} - 1$$

with the two right-hand quantities being sensitivity and specificity. Thus the expanded formula is:

$$J = \frac{\text{true positives}}{\text{true positives} + \text{false negatives}} + \frac{\text{true negatives}}{\text{true negatives} + \text{false positives}} - 1$$

[1] Youden, W.J. (1950). Index for rating diagnostic tests. Cancer. 3: 32–35

ILLUSTRATION OF YOUDEN INDEX.



Youden index = sensitivity+specificity–1 or

= sensitivity–probability of false positives

= specificity–probability of false negatives

Sensitivity = probability of correct testing of positive sample

Specificity = probability of correct testing of negative sample

EQA PANEL

Number of correct responses	**7**	6	5	4	3	2	1	0
Number of incorrect responses	**0**	1	2	3	4	5	6	7
Probability of the outcome if the parameter is really 0.99[c]	**0.932**	0.0659	0.001997	3.36×10^{-5}	2.65×10^{-7}	2.06×10^{-9}	6.93×10^{-12}	1×10^{-14}
p-Value[d]	**1**	0.068	0.002031	0.000034	3.42×10^{-7}	2.07×10^{-9}	6.94×10^{-12}	1×10^{-14}
Two-sided confidence interval of the consistent outcome[e] (at significance level α=0.1)	**0.72–1**							
Left-sided confidence interval of the consistent outcome[e] (at significance level α=0.1)	**0.72–0.99**							

Application of exact binomial test for goodness-of-fita for the case of testing seven equal samples: estimation of consistency of a particular outcome with the target value 0.99 at significance level α = 0.1.

Minimal required sample sizes for discrimination of selected parameter values

Target level	Sample size (number of equal samples) for discrimination of the target level from the next to the target level[b]					
	0.995	0.99	0.95	0.90	0.85	0.80
0.999	855	299	32	16	10	8
0.995		1364	59	16	10	8
0.99			85	29	10	8
0.95				124	44	21
0.90					199	61
0.85						260

a Power analysis for one-tailed exact binomial test for goodness-of-fit was performed with significance level 0.1 and power 80 using G*Power [13].
b Next to the target level = target level–effect size of power analysis.

INFORMATIONAL CAPACITIES OF CURRENTLY FEASIBLE SAMPLE SIZES ALLOCATED FOR EVALUATION OF ONE PARAMETER.

Number of equal samples	Highest value of the next to the target level[a,b,c] which can be discriminated from the target values			Maximal number of incorrect identifications consistent with the target values[d]			Probability of an outcome consistent with the target values 0.99, 0.995 and 0.999 for the labs with low parameter values[e]		
	0.999	0.995	0.99	0.999	0.995	0.99	0.5	0.7	0.8
1	0.001	0.001	0.001	0	0	0	0.5	0.7	0.8
2	0.201	0.201	0.201	0	0	0	0.25	0.49	0.64
3	0.449	0.448	0.448	0	0	0	0.125	0.343	0.512
4	0.585	0.585	0.585	0	0	0	0.063	0.240	0.41
5	0.669	0.669	0.669	0	0	0	0.031	0.168	0.328
6	0.725	0.725	0.73	0	0	0	0.016	0.118	0.262
7	0.765	0.765	0.765	0	0	0	0.008	0.082	0.21
8	0.795	0.795	0.795	0	0	0	0.004	0.058	0.168
9	0.818	0.818	0.818	0	0	0	0.002	0.040	0.134
10	0.837	0.837	0.837	0	0	0	0.001	0.028	0.107
11	0.852	0.852	n.d.	0	0	0	0.0005	0.020	0.086
12	0.864	0.864	n.d	0	0	1			
13	0.875	0.875	n.d.	0	0	1			
14	0.884	0.884	n.d.	0	0	1			
15	0.892	0.892	n.d.	0	0	1			
16	0.899	0.899	n.d.	0	0	1			
17	0.905	0.905	n.d.	0	0	1			
18	0.910	0.910	n.d.	0	0	1			
19	0.915	0.915	n.d.	0	0	1			
20	0.919	0.919	0.855	0	0	1			
21	0.923	0.923	0.86	0	0	1			
22	0.927	n.d.	0.865	0	1	1			
23	0.930	n.d.	0.87	0	1	1			
24	0.933	n.d.	0.875	0	1	1			
25	0.936	n.d.	0.881	0	1	1			
26	0.938	n.d.	0.885	0	1	1			
27	0.940	n.d.	0.89	0	1	1			
28	0.943	n.d.	0.894	0	1	1			
29	0.945	n.d.	0.897	0	1	1			
30	0.947	n.d.	0.901	0	1	1			
31	0.948	n.d.	0.904	0	1	1			
32	0.95	n.d.	0.906	0	1	1			

data not included

n.d. = not determinable

EXAMPLES OF CURRENTLY FEASIBLE OPTIMIZED TEST PANELS.

Panel #	Positive sample (copy number)	Truly negative sample (copy number)	Mixture of confounding viruses (copy number)	Secondary (sero)types: (if available) (copy number)					Total number of samples
				1	2	3	....	n	
1	7	7	1	–	–	–		–	15
2	7	7	1	1	1	1		1	15+n
3	11	3	1	–	–	–		–	15
4(optimized Dengue panel)	8	3	1	1	1	1		–	15
5	16	3	1	–	–	–		–	20
6 (realistic prospective)	29	3	1	–	–	–		–	33

CONCLUSIONS

- **The immediate goal of EQA is defined as to obtain a statistically reliable estimation for every laboratory whether its performance meets the proficiency standard, while the overall goal is to match every laboratory to its specific performance level.**
- **Youden index requires an estimate of sensitivity and specificity and incorporates the relationship of these performance parameters.**
- **Dependence of informational capacities of test panel from the panel size and content is quantitatively analyzed and the optimal design and informational capacities of both idealized panels (whose size is not restricted by financial factors) and currently feasible panels are considered.**
- **Our approach provides the basis both for rational design of currently feasible EQA test panels and for an increased panel size.**
- **Our approach provides the basis for the upfront planning of EQAs ensuring that the data will allow objective statistical evaluation and comparison of participant performances.**
- **It enables both the rational design of currently feasible test panels and to provide reasons for rational panel size increase.**

QUESTIONS?


Prevention
Quality Assurance
Quality Control
Detection


PLAN
DO
CHECK
ACT


Quality
Excellent
Good
Poor