VINZENZ VONPAUL KLINIKEN gGMBH

MARIENHOSPITAL STUTTGART

Are regulation driven-performance criteria still acceptable?

Institut für Laboratoriumsmedizin

Matthias Orth

matthias.orth@vinzenz.de Tel. + 49 711 6489 2760

www.laborstuttgart.de

Regulation and performance criteria

Experiences with "RiliBÄK" as example of regulation driven-performance criteria

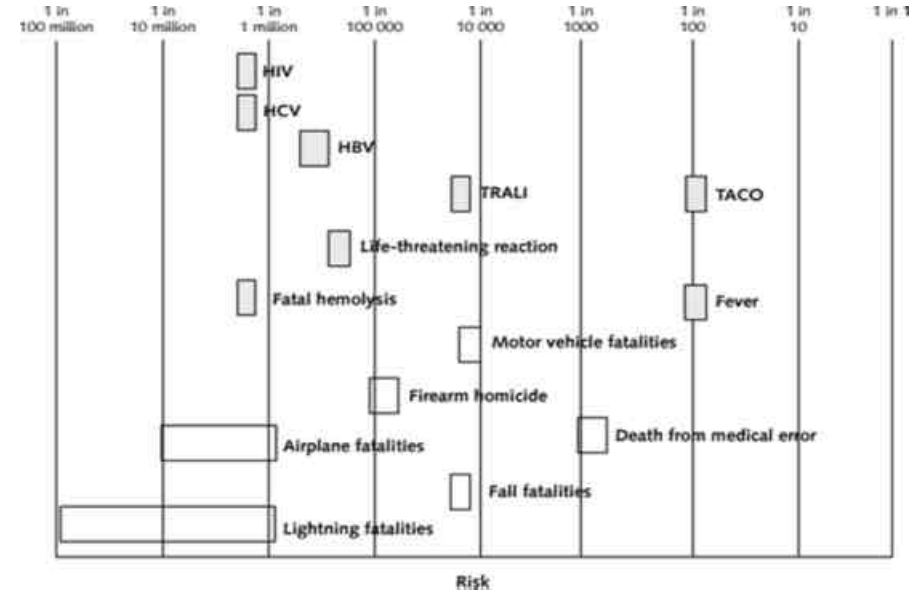
new challenge to laboratory medicine: Health Technology Assessment (HTA)

Performance criteria in companion diagnostics and DTC/DAT

Conclusions

Is our focus at the right target?

Risks of Red Blood Cell Transfusion



Carson JL et al. Ann Intern Med. 2012;157:49-58

How do others rate our performance in laboratory medicine services?

Table 1. Percentage and Aggregate Number of Ratings

Laboratory Service Category*	Excellent, % (No.)	Good, % (No.)	Average, % (No.)	Below Average, % (No.)	Poor, % (No.)
Quality/reliability of test results	45.6 (1939)	42.9 (1823)	9.8 (416)	1.3 (56)	0.4 (15)
Staff courtesy	50.4 (2069)	37.1 (1523)	9.7 (398)	2.2 (89)	0.6 (25)
Accessibility of pathologist	51.7 (1823)	34.1 (1201)	11.5 (406)	2.1 (74)	0.7 (23)
Accessibility of laboratory manager	46.5 (1524)	36.0 (1178)	13.7 (449)	2.7 (87)	1.2 (38)
Phlebotomy services	37.8 (1313)	43.7 (1515)	14.4 (501)	3.1 (106)	1.0 (35)
Test menu adequacy	36.7 (1427)	46.9 (1826)	14.0 (543)	1.7 (68)	0.7 (26)
Accessibility of laboratory staff	47.3 (1913)	36.5 (1475)	12.4 (500)	2.7 (111)	1.1 (44)
Courier services	38.0 (1039)	41.1 (1124)	15.6 (428)	3.2 (87)	2.1 (57)
Routine test TAT	33.5 (1389)	44.7 (1855)	17.0 (704)	3.4 (142)	1.4 (56)
Laboratory management responsiveness	40.4 (1380)	40.1 (1372)	14.4 (492)	3.6 (123)	1.5 (51)
Inpatient stat test TAT	36.7 (1177)	41.7 (1338)	15.0 (480)	4.4 (142)	2.2 (71)
Critical value notification	44.3 (1833)	39.3 (1624)	11.4 (470)	3.1 (128)	1.9 (79)
Clinical report format	33.7 (1396)	46.0 (1905)	15.5 (644)	3.1 (127)	1.7 (71)
Outpatient stat test TAT	33.6 (1170)	40.3 (1407)	17.4 (605)	6.2 (216)	2.6 (89)
Esoteric test TAT	17.1 (629)	38.0 (1398)	32.9 (1212)	8.9 (328)	3.1 (116)

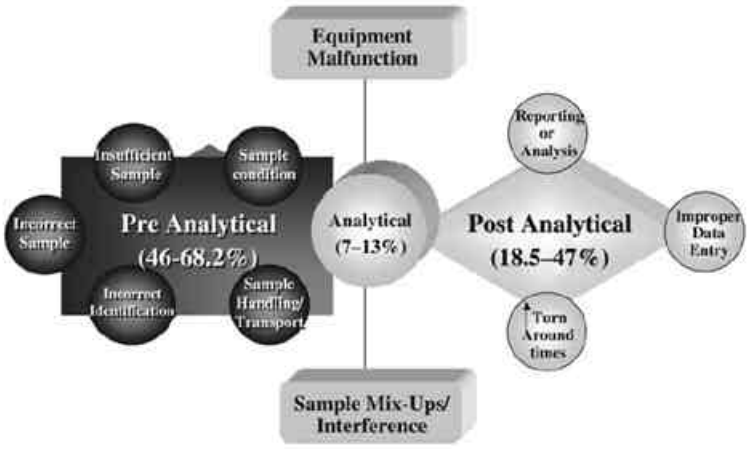
* TAT indicates turnaround time.

4329 respondents

responsibility for processes out of the laboratory

Arch Pathol Lab Med. 2009;133:38–43

Is our focus at the right target?



Plebani M. Clin Chem Lab Med 2006;44:750

Quality criteria to be covered by regulation

performance criteria for daily routine quality controls

performance criteria for EQAS

performance criteria for tests with numeric as well as for alpha-numeric results

use of reference method values and/or method specific values for EQAS

optional: minimum time interval / maximum frequency for ordering a specific test

FDA
clinical validity (accuracy with which test identifies, measures, or predicts presence or absence of a clinical condition or predisposition in a patient)

CLIA
safety and effectiveness of the test system. does not address the clinical validity of any test


Bekanntmachungen

Richtlinie der Bundesärztekammer zur Qualitätssicherung laboratoriumsmedizinischer Untersuchungen

Gemäß dem Beschluss des Vorstands der Bundesärztekammer vom 11.04.2014 und 20.06.2014

Experiences with "RiliBÄK" as example of regulation driven-performance criteria

Legal background behind RiliBÄK

EU IVD directive

German Medical Devices Act ("Medizinproduktegesetz")

German Medical Devices Operator Ordinance ("Medizinproduktebetreiberverordnung")

German Medical Association ("Bundesärztekammer")

RiliBÄK
every professional employing laboratory tests in human healthcare is obliged to comply to all regulations specified in RiliBÄK

part A (the description of a quality management system closely resembling DIN EN ISO norm 15189 as a framework for structural quality) **(GROSS ERROR)**

part B with extensive appendices covering analytical performance goals in internal as well as in external quality programs in tabulated form for 84 selected quantitative and 50 semiquantitative tests in hematology, hemostaseology, clinical chemistry, TDM, endocrinology, serology in different matrices (such as serum, plasma, whole blood, urine, cerebrospinal fluid) as well as for genetical and microbiological tests and sperm analysis **(RANDOM and SYSTEMATIC ERROR)**

acceptable % root mean square of measurement deviation (RMSMD)

validity range for RMSMD and EQAS

maximum allowable deviation in EQAS

reference method (RMW) or method-specific consensus value (SW) in interlab tests

1 lfd. Nr.	2 Analyt	3 Zulässige rela-tive Abwei-chung des Ein-zelwertes bzw. des relativen quadratischen Mittelwertes	4 Gültigkeitsbereich der Spalten 3 und 5			5 Zulässige relative Ab-weichung beim Ring-versuch	6 Referenzmethode beim Ring-versuch
			Typ	Bis	Einheit		
23	Ferritin	13,5 %	10	600	µg/l	25,0 %	SW
24	FSH	14,0 %	4	70	U/l	23 %	SW
25	Gamma-Glutamyl-Transferase (γ-GT) EC 2.3.2.2	11,5 %	20 0,33	300 5	U/l µkat/l	21,0 %	RMW
26	Glucose	11,0 %	40 2,2	400 22	mg/dl mmol/l	15,0 %	RMW
27	Hämatokrit	5,0 %	10 0,1	60 0,6	% l/l	9,0 %	SW
28	Hämoglobin	4,0 %	2 1,2	20 12,4	g/dl mmol/l	6,0 %	RMW
29	Hämoglobin A 1c (HbA1c)	10,0 %	30	140	mmol/mol Hb	18,0 %	RMW
30	Harnsäure	7,0 %	2 119	13 773	mg/dl µmol/l	13,0 %	RMW
31	Harnstoff	10,5 %	15 2,5	200 33	mg/dl mmol/l	20,0 %	RMW
32	Humanes Choriongonadotropin (hCG)	14,0 % 17,0 %	> 100 2	1 500 ≤ 100	IU/l IU/l	30,0 %	SW

Selection of quality control material based on RiliBÄK specifications (!) (range, target value assignment)

Calculation of root mean square of measurement deviation (RSMD)

$$\%RSMD = \frac{\sqrt{k^2 (SD_{cc}^2) + Bias^2}}{TV}$$

SD_{cc}: Standard deviation
Bias: Difference of observed mean from Target value (TV)
K: statistical coverage factor (1 for metric, 3 to calculate specification)
TV: target value

Procedure for non-tabulated tests with new control samples (new control cycle)

Process for repeated failures of column 3 at the end of control cycles („event“ according to §2 Medical Products Safety Plan Ordinance)

Open discussion whether different analytical performance standards might be acceptable between real laboratory tests and point of care tests

Mueller, C., A. Scholer, et al. (2004). N Engl J Med **350**: 647-54
Straseski, J. A., M. E. Lyon, et al. (2011). Clin Chem **57**: 1566-73

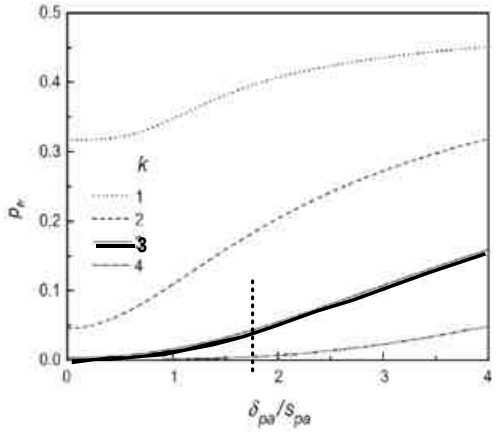


Figure 2 False rejection probability as a function of the systematic error, expressed in multiples of the standard deviation, both observed during a pre-analytical period. Parameter is the coverage factor k.

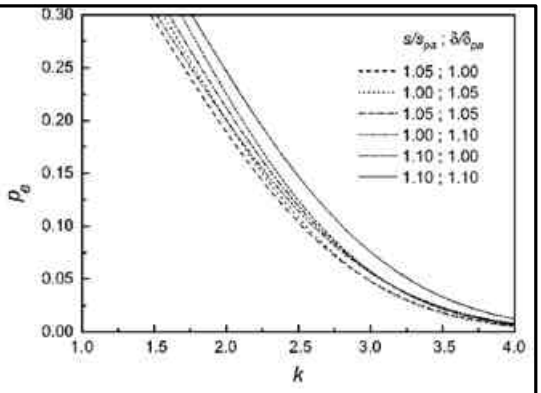


Figure 3 Error detection probability vs. coverage factor *k* for different size and type of error, $\delta_{pa}/s_{pa} = 1.7$.

empirical: $\delta_{pa}/s_{pa} = 1.7$

Macdonald, R. (2006). LaboratoriumsMedizin **30**: 111-7

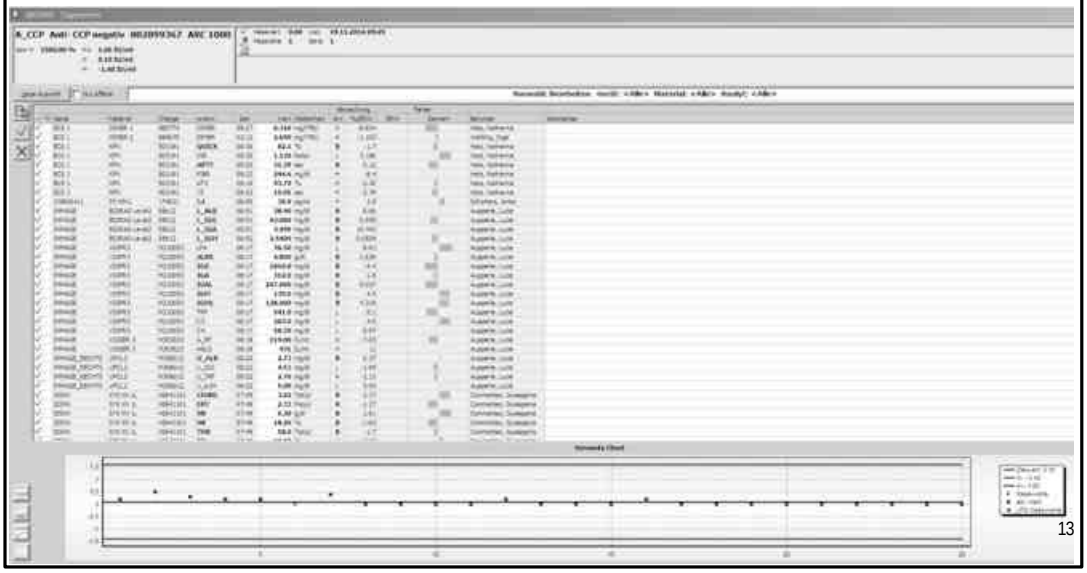
Tabelle B 2-1 – Interne Qualitätssicherung

Lfd. Nr.	Messgröße/Untersuchung	Häufigkeit der (Kontroll-)untersuchung
1.	6-Acetylmorphin	täglich
2.	ABO-Merkmale	wöchentlich
3.	Amphetamine	täglich
4.	Barbiturate	täglich
5.	Benzodiazepine	täglich
6.	Borrelia burgdorferi, Antikörper gegen	täglich
7.	Buprenorphin	täglich
8.	Candida albicans, Antikörper gegen	täglich
9.	Cannabinoide	täglich
10.	Chromatographische Analyse mit Identifizierung der Wirksubstanz (STA)	täglich
11.	Cocain und Metabolite	täglich
12.	Direkter Coombstest	wöchentlich
13.	dsDNA, Autoantikörper gegen	täglich
14.	Echinococcus, Antikörper gegen	täglich
15.	Elektrophorese mit Immunreaktion	monatlich
16.	Entamoeba histolytica, Antikörper gegen	täglich
17.	Erythrozytenantigene, Antikörper gegen (Suchtest)	täglich
18.	extrahierbare, nukleäre Antigene, Autoantikörper gegen	täglich
19.	Glatte Muskulatur, Autoantikörper gegen	täglich
20.	Glutaminase, Antikörper gegen	täglich
21.	HBc-Antigen, Antikörper gegen	täglich
22.	HBe-Antigen, Antikörper gegen	täglich
23.	HBs-Antigen, Antikörper gegen	täglich
24.	Hepatitis A-Virus, Antikörper gegen	täglich
25.	Hepatitis C-Virus, Antikörper gegen	täglich

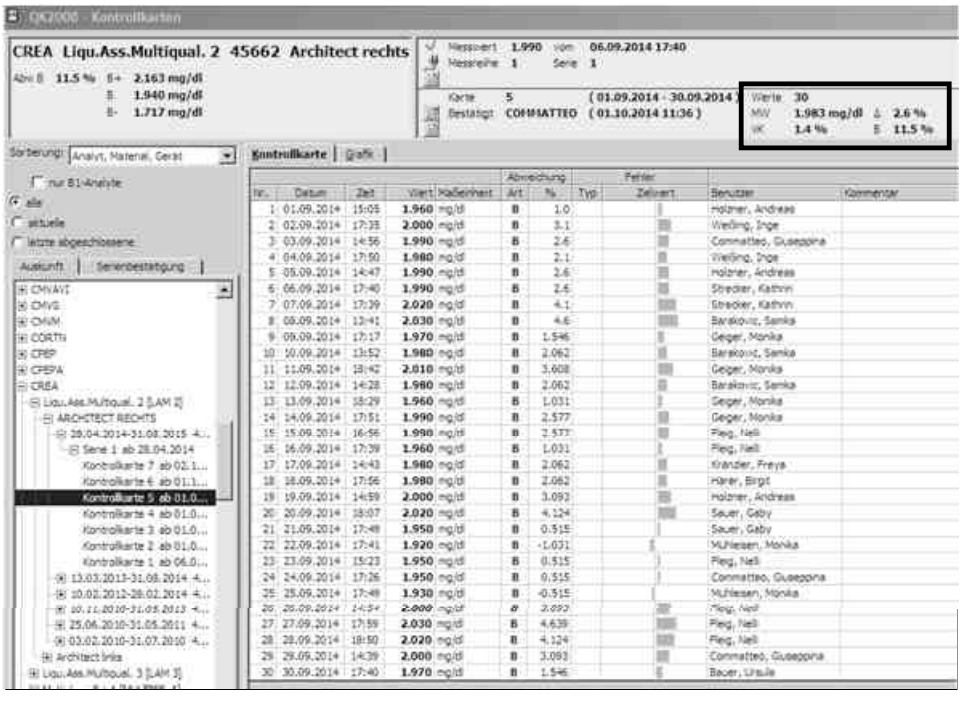
Tabelle B 2-2 – Externe Qualitätssicherung (Ringversuche)

Lfd. Nr.	Messgröße/Untersuchung	Häufigkeit der Teilnahme/ Intervall pro
1.	ABO-Merkmale	Quartal
2.	Borrelia burgdorferi, Antikörper gegen	Halbjahr
3.	Candida albicans, Antikörper gegen	Halbjahr
4.	Cannabinoide	Quartal
5.	CD4-T-Zellen	Halbjahr
6.	CD8-T-Zellen	Halbjahr
7.	Chromatographische Analyse mit Identifizierung der Wirksubstanz (STA)	Halbjahr
8.	Cocain und Metabolite	Quartal
9.	Differenzierung, Blutausstrich	Quartal
10.	Direkter Coombstest	Quartal
11.	dsDNA, Autoantikörper gegen	Halbjahr
12.	Echinococcus, Antikörper gegen	Kalenderjahr
13.	Entamoeba histolytica, Antikörper gegen	Kalenderjahr
14.	Erythrozytenantigene, Antikörper gegen (Suchtest)	Quartal
15.	Glutaminase, Antikörper gegen	Halbjahr
16.	HBc-Antigen, Antikörper gegen	Halbjahr
17.	HBe-Antigen, Antikörper gegen	Halbjahr
18.	HBs-Antigen, Antikörper gegen	Halbjahr
19.	Hepatitis A-Virus, Antikörper gegen	Halbjahr
20.	Hepatitis C-Virus, Antikörper gegen	Halbjahr
21.	HIV, Antikörper gegen	Halbjahr
22.	IgE-Antikörper, allergenspezifischer-Einzelallergentest Methodenspezifische Kontrolle im Rotationsverfahren mit 6 Leitallergenen aus folgenden Gruppen: a) saisonales Inhalationsallergen, b) ganzjähriges Inhalationsallergen, c) Nahrungsmittelallergen, d) Insektengiftallergen	Halbjahr

•instant assessment of analytical control samples and detection of critical deviations by operator
•automatic calculation of RMSMD is integrated into all major lab information systems
•drawback of RMSMD: no information whether systematic (i.e., bias) or random (i.e. imprecision) issues have caused the violation



13



14

Control	Assay	Unit	Level	N	Mean	SD	CV (%)	All Site Mean	% RMSD	% RMSD Acceptance Limit	Sigma-Metric	%TEa
Technopath Multichem IA Plus	2nd Gen Testosterone (2P13-20)	ng/mL	Level 1	32	0.34	0.024	7.0	0.3	11.7	20.5	2.4	23
			Level 2	32	5.29	0.252	4.8	5.0	8.2	20.5	3.5	23
	Alpha Fetoprotein (3P36)	ng/mL	Level 1	31	7.2	0.21	3.0	7.1	3.1	17.0	7.1	21.9
			Level 2	31	88.0	2.10	2.4	86.5	3.0	17.0	8.5	21.9
			Level 3	31	201.1	5.44	2.7	196.5	3.6	17.0	7.2	21.9
	CA 19-9 XR (2K91)	U/mL	Level 1	30	17.3	2.5	14.7	17.5	14.6	14.0	3.1	46.3
			Level 2	30	35.1	4.1	11.6	35.9	11.6	14.0	3.8	46.3
			Level 3	30	136.8	9.7	7.1	139.6	7.2	14.0	6.2	46.3
	CEA (7K68)	ng/mL	Level 1	31	3.1	0.3	9.8	3.1	9.8	14.0	2.5	24.7
			Level 2	31	17.9	0.8	4.5	18.1	4.6	14.0	5.3	24.7
			Level 3	31	46.7	2.2	4.5	49.3	4.6	14.0	5.2	24.7
	Cortisol (8D15)	ng/mL	Level 1	31	3.5	0.2	4.8	3.4	5.5	18.5	5.3	28
			Level 2	31	14.2	0.4	3.0	13.9	3.6	16.0	6.7	28
			Level 3	31	33.7	1.2	3.7	33.6	3.7	16.0	7.5	28
	Estradiol (7K72)	pg/mL	Level 1	32	54.3	4.6	8.4	54.9	8.4	22.0	3.1	26.9
			Level 2	32	170.3	6	3.5	172.7	3.8	22.0	7.3	26.9
			Level 3	32	419.9	16.8	4.0	431.3	4.7	22.0	6.1	26.9
	Progesterone (7K77)	ng/mL	Level 1	34	1	0.1	10.6	1.0	10.6	22.0	2.5	26
			Level 2	34	9.9	0.65	6.6	9.7	6.9	17.0	3.6	26
			Level 3	34	23.5	1.26	5.4	23.4	5.4	17.0	4.6	26
	Ferritin (8D15)	ng/mL	Level 1	30	16.2	1.0	6.4	16.2	6.4	13.5	3.3	21
			Level 2	30	181.7	10.1	5.6	180.4	5.6	13.5	3.6	21
			Level 3	30	330.6	18.1	5.5	334.0	5.5	13.5	3.7	21
	FSH (7K75)	U/L	Level 1	16	5.2	0.1	2.1	5.3	2.6	6.2	9.0	21.2
			Level 2	16	17.1	0.3	1.9	17.7	3.6	11.2	9.5	21.2
			Level 3	16	38.9	1.2	3.2	40.0	4.6	10.9	5.8	21.2
	Free T3 (7K83)	pg/mL	Level 1	30	3	0.3	9.5	3.0	9.6	14.5	2.4	24
			Level 2	30	4.2	0.3	6.5	4.1	6.9	14.5	3.5	24
			Level 3	30	9.3	0.5	5.1	9.1	5.7	14.5	4.4	24
	Free T4 (7K65)	ng/dL	Level 1	30	0.6	0	0.0	0.0	0.0	15.0	3.4	24
			Level 2	31	1.7	0.1	5.3	1.7	5.4	15.0	4.4	24
			Level 3	31	2.9	0.2	5.3	2.9	5.4	15.0	4.5	24
	TSH (7K62-22)	µIU/mL	Level 1	30	0.107	0.004	3.5	0.099	8.9	14.0	4.5	23.7
			Level 2	30	4.343	0.129	3.0	4.155	5.5	13.5	6.4	23.7

15

AALT Method Decision Chart

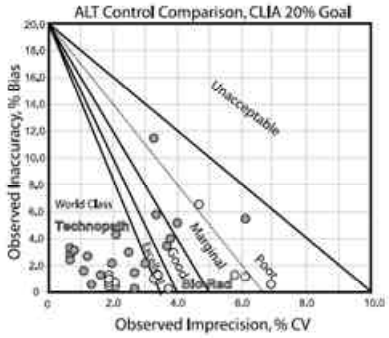


Figure 5. Six Sigma Method Performance comparison of BioRad (○) and Technopath (●) controls for AALT.

Chloride Method Decision Chart

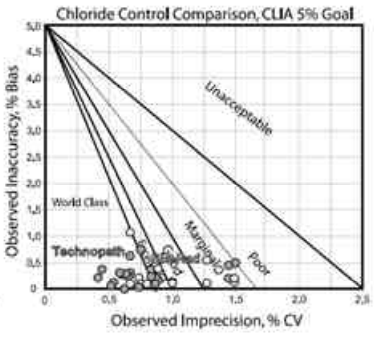


Figure 6. Six Sigma Method performance comparison of BioRad (○) and Technopath (●) controls for chloride.

16

Bekanntmachungen

Richtlinie der Bundesärztekammer zur Qualitätssicherung laboratoriumsmedizinischer Untersuchungen

Gemäß dem Beschluss des Vorstands der Bundesärztekammer vom 11.04.2014 und 20.06.2014

Revision 2014

Experiences with "RiliBÄK" as example of regulation driven-performance criteria

- Alkaline Phosphatase: RMSD reduced from 13% to 11%; EQAS reduced from 21 to 18%
- CA 19-9 replaced by CA 15-3
- FSH added
- Lipase deleted
- pCO2: goals made more complex (2 levels)
- FT4: goals simplified (1 level)
- Transferrin: RMSD reduced from 9.5% to 8.0%; EQAS reduced from 15% to 12%
- FT3: RMSD reduced from 14.5% to 13.0%; EQAS reduced from 24% to 20%
- Vancomycin: EQAS reduced from 21% to 18.0%

Deutsches Ärzteblatt | Jg. 111 | Heft 38 | 19. September 2014

outcomes studies: health technology assessment (HTA)

multidisciplinary process firmly rooted in research and the scientific methods that summarizes information about medical, social, economic and ethical issues related to the use of a health technology

It is expected that with HTA the risk of implementing measures that negatively affect patient outcomes is reduced

In general HTA is being performed by formally independent institutions employing scientific methods

in most countries, the paramount aim of HTA is at decision-making in health politics and healthcare budgeting

concept behind HTA is a prohibition of use of a certain technology unless permission is granted (positive list)

current focus on HTA for introducing new technologies is severe threat to innovations in laboratory medicine as well as in using established laboratory tests!

General challenges of HTA

•evidence of efficiency gains and improvements in health remains valid when different definitions of health outcomes used?

•often crude measurements such as life expectancy, not considering quality of years of life gained

•besides parameter studied influence of numerous factors on health outcomes

•time lag between the introduction of a new technological solution and its impact on health outcomes

'false savings' because they may lead to increased costs or other unintended consequences in the long term (e.g. in screening tests with high rates of false positives followed by extensive diagnostic procedures or even invasive treatment measures)

' undervalued positive effects ' of new technologies when outcomes can be detected only after long periods of observation such as in screening programs of low grade types of cancer or of risk markers for slowly progressing diseases such as coronary heart disease

Challenges of HTA for diagnostic procedures

Qualifying performance testing in the medical laboratory by HTA is a yet unresolved challenge
Reid, M. C., M. S. Lachs, et al. (1995). JAMA **274**: 645-51

General concept of laboratory medicine which only delivers data to the attending physicians such as the presence or absence of a certain disease. Most meta-analyses for diagnostic test studies still pool diagnostic sensitivity and sensitivity values only Willis, B. H. and M. Quigley (2011). BMC Med Res Methodol **11**: 27

diagnostical and analytical per-formance goals of a certain labora-tory test might even have to be defined for different clinical situa-tions and have to be revised in specified intervals thereafter
Sandberg, S., and Thue, G. Scand J Clin Lab Invest. 1999;59:531

„Evidence on current practice indi-cates that clinical practice has changed to such a degree that the original research question is no lon-ger relevant to UK practice"Czoski-Murray, C., M. Lloyd Jones, et al. (2012). Health Technol Assess **16**(50): i-xvi, 1-159.

NHS, Diagnostics Assessment Programme manual 2011

The evaluation of diagnostics differs from the evaluation of treatments, diagnostic tests have few direct outcomes

Most outcomes follow from treatments that are either initiated or not initiated based on the results of the tests (Surrogate markers!)

Tests are frequently done in conjunction with other tests or measurements, and it is the composite of the series of tests that is used in clinical decision-making

Only very rarely do studies of diagnostic tests follow patients through treatment to final outcomes. Also, evaluation of diagnostics usually requires that the clinical management process is described and that the effects of that process are known or assumed. If the effects of treatment are not known, analyses can be performed, but the validity of the results will be less certain in ways that may not be completely specifiable. This increases the uncertainty with which decisions can be made on use of diagnostic technologies

In statistics, '**test accuracy**' means the proportion of test results that are correct. This is not a useful definition for the purposes of this document, because a test may be incorrect in more than one way and for more than one reason

HTA and testing intervals

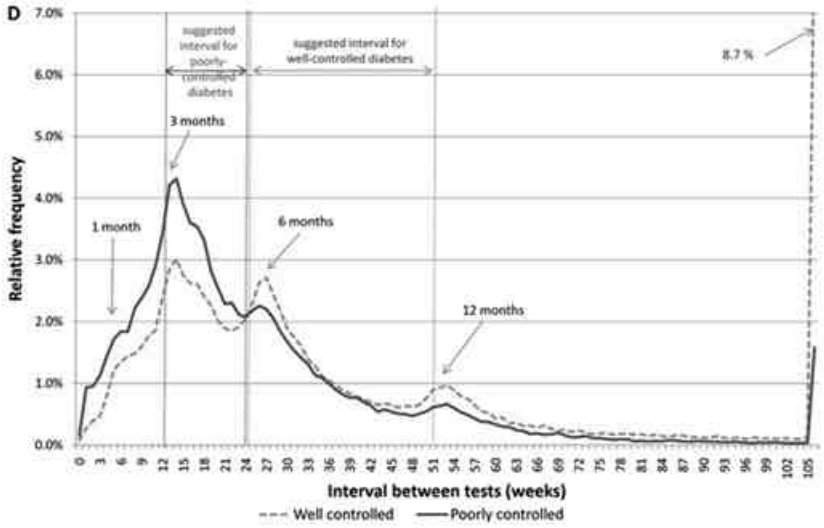
HTA adds further level of complexity to concept of quality indicators and performance goals in the medical laboratory that not only analytical quality indicators have to be agreed on for tests but also for

•testing intervals

and for

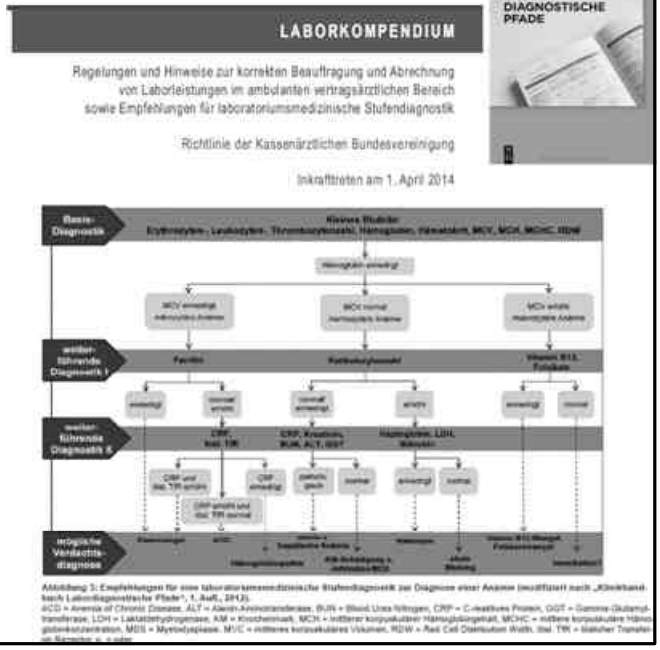
•sequence of tests (screening/confirmation testing)

Inappropriate Requesting of Glycated Hemoglobin (Hb A_{1c}) Is Widespread: Assessment of Prevalence, Impact of National Guidance, and Practice-to-Practice Variability



Driskell OJ et al. Clin Chem 2012; 58: 906

Diagnostic pathways



Challenges of HTA for diagnostic procedures

Disease prevalence in the population under question

Houben PH, et al. Scand J Prim Health Care. 2010;28:18-23

availability of other diagnostic methods

cost structure of the health system in this population

acceptance of monetary gain of certain medical procedure not equally accepted in different nations. E.g., concept of costs per QUALY is accepted in some countries (with wide differences among countries) but is highly defeated and even considered to be unethical in Germany

Hirth RA et al. Med Decis Making. 2000;20:332-42

Challenges of HTA for diagnostic procedures

In companion diagnostics, a certain test result of a (new) laboratory test is the prerequisite for prescription of drug

Khoury, J. D. and D. V. T. Catenacci (2014). Arch Pathol Lab Med

For regulation of the drug, the approval of the laboratory test is sine qua non

Substantial concern that HTA of new laboratory tests is shifted from laboratory medicine to drug companies (FDA, EMEA), (setting performance goals for a blood count should therefore also be done by drug companies?)

Challenges by DTC/DAT

EU

Direct to consumer testing (DTC)/ Direct Access Testing (DAT) :

no quality criteria at all have to be followed if laboratory tests are performed by non-health care professionals allowing a free movement of services under the consumer rights directive 2011/83/EU

Orth, M. and P. Luppa (2014). "„Direct to consumer testing" – boon or bane for the self-determined patient?" Dtsch Arztebl International **111**: in press

USA

A laboratory is defined to be a facility that performs certain testing on human specimens in order to obtain information that can be used for the diagnosis, prevention, or treatment of any disease or impairment of a human being

CLIA regulations and standards do not differentiate between facilities performing DAT and facilities performing provider ordered testing. All facilities must obtain appropriate CLIA certificate prior to conducting patient testing, *including* DTC/DAT

Conclusions

Current focus on HTA by health care policy makers may pose a severe threat to the introduction of new laboratory for patient use

Regulation-driven performance criteria for medical laboratory testing, even when based on analytical performance goals low in hierarchy - might be a promising alternative to HTA if widely-accepted both by medical professionals and from the health-economical network

Regulation-driven performance criteria have to be constituted by medical professionals
In case of referrals to DIN EN ISO norms, federal organization of health care system has to be respected

Performance criteria should be established for a wide array of laboratory tests and **updated on a regular basis** employing different analytical performance goals, in particular goals based on biological variation and the state of the art (i.e. technically achievable) outcome studies

These performance criteria should be mandatory for all tests performed in healthcare (exception have to be clearly defined!)

Results from from EQAS testing can be used in a formalized process to revise performance goals

Challenges of a general acceptance of the Stockholm criteria

Recommendations not widely introduced because such data were not available for many tests or the concept could not be applied to these tests (e.g. graphical presentation of titers, numerical + alphanumerical results, extreme analytical ranges)

In particular in immunoassays and mass-spectrometry, data highly dependent on method / sample material used. Challenge for laboratory and physicians who try to implement an improved assay when faced with data on (pre)analytical performances obtained with different methods or other sample types

Most data on biological validation were obtained on "simple Clinical Chemistry tests". Skipping too many (complex) tests by giving no recommendations at all and focusing on established tests might impede a fast progress in laboratory medicine, in particular for innovations