How do others rate our performance in laboratory medicine services?

<table>
<thead>
<tr>
<th>Laboratory Service Category*</th>
<th>Excellent, % (No.)</th>
<th>Good, % (No.)</th>
<th>Average, % (No.)</th>
<th>Below Average, % (No.)</th>
<th>Poor, % (No.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quality/reliability of test results</td>
<td>45.6 (1039)</td>
<td>32.4 (769)</td>
<td>11.8 (281)</td>
<td>5.6 (130)</td>
<td>0.4 (10)</td>
</tr>
<tr>
<td>Staff courtesy</td>
<td>50.4 (1269)</td>
<td>37.1 (952)</td>
<td>9.7 (238)</td>
<td>1.7 (43)</td>
<td>0.1 (1)</td>
</tr>
<tr>
<td>Accessibility of pathologist</td>
<td>51.7 (1293)</td>
<td>34.1 (810)</td>
<td>11.6 (273)</td>
<td>2.1 (50)</td>
<td>0.1 (1)</td>
</tr>
<tr>
<td>Accessibility of laboratory manager</td>
<td>46.5 (1152)</td>
<td>36.0 (878)</td>
<td>11.7 (283)</td>
<td>2.7 (63)</td>
<td>0.1 (1)</td>
</tr>
<tr>
<td>Phlebotomy services</td>
<td>37.8 (933)</td>
<td>43.7 (1053)</td>
<td>14.4 (351)</td>
<td>3.1 (73)</td>
<td>1.0 (25)</td>
</tr>
<tr>
<td>Test menu adequacy</td>
<td>36.7 (912)</td>
<td>46.9 (1116)</td>
<td>14.0 (343)</td>
<td>1.7 (43)</td>
<td>0.7 (17)</td>
</tr>
<tr>
<td>Accessibility of laboratory staff</td>
<td>47.3 (1113)</td>
<td>36.5 (873)</td>
<td>12.4 (293)</td>
<td>2.7 (63)</td>
<td>1.1 (3)</td>
</tr>
<tr>
<td>Counter services</td>
<td>38.0 (939)</td>
<td>41.1 (924)</td>
<td>15.6 (373)</td>
<td>3.2 (73)</td>
<td>2.1 (50)</td>
</tr>
<tr>
<td>Routine test TAT</td>
<td>33.3 (809)</td>
<td>44.7 (1065)</td>
<td>17.0 (407)</td>
<td>3.4 (81)</td>
<td>1.4 (32)</td>
</tr>
<tr>
<td>Laboratory management responsiveness</td>
<td>40.4 (1028)</td>
<td>40.1 (972)</td>
<td>14.4 (349)</td>
<td>3.6 (84)</td>
<td>1.5 (35)</td>
</tr>
<tr>
<td>Inpatient stat test TAT</td>
<td>36.7 (917)</td>
<td>41.7 (1038)</td>
<td>15.0 (360)</td>
<td>4.4 (103)</td>
<td>2.2 (56)</td>
</tr>
<tr>
<td>Critical value notification</td>
<td>44.3 (1133)</td>
<td>39.3 (992)</td>
<td>11.4 (275)</td>
<td>3.2 (73)</td>
<td>1.2 (25)</td>
</tr>
<tr>
<td>Clinical report format</td>
<td>33.7 (918)</td>
<td>46.0 (1170)</td>
<td>11.5 (273)</td>
<td>3.1 (73)</td>
<td>1.7 (41)</td>
</tr>
<tr>
<td>Outpatient stat test TAT</td>
<td>33.6 (917)</td>
<td>40.3 (924)</td>
<td>17.4 (423)</td>
<td>6.2 (153)</td>
<td>2.6 (63)</td>
</tr>
<tr>
<td>IS/T local test TAT</td>
<td>17.1 (429)</td>
<td>38.0 (939)</td>
<td>32.9 (781)</td>
<td>4.9 (118)</td>
<td>3.1 (73)</td>
</tr>
</tbody>
</table>

* TAT indicates turnaround time.

4329 respondents

responsible for processes out of the laboratory

http://www.ifcc.org/ifcc-education-division/emd-committees/c-clm/

Arch Pathol Lab Med. 2009;133:38–43
Analytical quality criteria to be covered

- 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: quality specifications for calculated tests

Westgard multirules
Variation of test results

1. preanalytic variation
2. analytical variation (imprecision and bias)
3. biological variation within a single subject

1999 Stockholm consensus conference statement

hierarchy of models to set analytical quality specifications

1. Evaluation of the effect of analytical performance on clinical outcomes in specific clinical settings
2. Evaluation of the effect of analytical performance on clinical decisions in general:
   a. Data based on components of biological variation
   b. Data based on analysis of clinicians’ opinions
3. Published professional recommendations:
   a. From national and international expert bodies
   b. From expert local groups or individuals
4. Performance goals set by:
   a. Regulatory bodies
   b. Organisers of EQA schemes
5. Goals based on the current state of the art:
   a. As demonstrated by data from EQA or Proficiency Testing schemes
   b. As found in current publications on methodology
Identical performance criteria in real labs, POCT and DCT?

Diagnosis and monitoring: \( CV_{\text{analytical}} < 0.5 \) \( CV_{\text{within-subject}} \)

Screening: \( CV_{\text{analytical}} < 0.5 \sqrt{CV_{\text{within-subject}}^2 + CV_{\text{between-subject}}^2} \)


Challenges of HTA/outcome studies for diagnostic procedures

Qualifying performance testing in the medical laboratory by HTA is a yet unresolved challenge

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 specificity values only

"Evidence on current practice indicates that clinical practice has changed to such a degree that the original research question is no longer relevant to UK practice" Czoski-Murray, C., M. Lloyd Jones, et al. (2012). Health Technol Assess 16(50):i-xvi, 1-159.
Challenges of a general acceptance of the Stockholm criteria

Recommendations not widely introduced because data not available for many tests or concept not applicable (e.g. graphical presentation of titers, numerical + alphanumerical results, extreme analytical ranges)

In particular in immunoassays and mass-spectrometry, data highly dependent on method / matrix

Most data on biological validation on "simple Clinical Chemistry tests"

Skipping too many (complex) tests by giving no recommendations at all

2014 Milan consensus conference statement

hierarchy of models to set analytical quality specifications

1. Evaluation of the effect of analytical performance on clinical outcomes in specific clinical settings (very few analytes)

2. Evaluation of the effect of analytical performance on clinical decisions in general:
   a. Data based on components of biological variation (scrutinizing data)

3. Other goals
   a. From national and international expert bodies
   b. From expert local groups or individuals
   c. Regulatory bodies
   d. Organisers of EQA schemes
   e. As demonstrated by data from EQA or Proficiency Testing schemes
   f. As found in current publications on methodology

Pre-Analytical and Post-analytical Performance Goals - TBD

Measurement of „true“ value and correct medical interpretation of test result „(selecting the correct language)“

Test result has deviation from „true“ value (total analytical error TAE or permissible uncertainty (pU))

pU consists of dispersion of results („random error“) and systematic deviation from „true“ value, called „bias“

Preanalytic effects lead to
• Gross errors (e.g. sample mixup)
• Unsuitable results (e.g. wrong timing of TDM or in provocation test)
• Systematic in- or decrease of result caused by instability of analyte or by interference (hemolysis), unpredictable instability by recentrifugation of gel tubes or barricor tubes

components of error (random and systematic (bias) error) of
(A) a single result of measurement,
(B) the mean of four replicate measurements and
(C) the mean of infinite number of measurements, which eliminates the random error component

*Bioanalysis* (2014) 6(21), 2855–2875
bias may be indistinguishable from imprecision

In patient samples, uncertainty methods estimate the confidence we can have in the measurement result for the purpose of diagnosis. Proficiency testing and measurement uncertainty are related through the traceability chain to the reference standard.

target value is defined for the proficiency testing sample, which is used for calculating error.

Orth: Analytical Performance Specifications

CCLM https://doi.org/10.1515/cclm-2017-0341
Legal framework for performance criteria

**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

---

On 19 September 2014, the current version of the “Guideline of the German Medical Association on Quality Assurance in Medical Laboratory Examinations” was published. It featured an introduction by the German Medical Association.

Revision of the “Guideline of the German Medical Association on Quality Assurance in Medical Laboratory Examinations – Rili-BAEK”

---

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)**
Special Part B1: Quantitative tests in medical laboratories

1. Principles of quality assurance
2. Minimum requirements are listed that need to be met to assess the quality of quantitative results of examinations in medical laboratories.
3. All quantitative tests performed by medical laboratories are subject to IQC.
4. All measurands listed in table B1 a to c are subject to EQA.

Special Part B1: 2. Carrying out quality assurance

1. Internal quality assurance
1. Carrying out individual measurements of control samples
2. Evaluating the results of the individual measurements of control samples
3. Calculating and evaluating the root mean square of the error of measurement after completing a control cycle.
4. Establishing internal laboratory limits of permissible error for measurands that are not listed in Table B1
5. Point-of-care testing with unit-use reagents
6. Measurands with small test frequencies
7. Documentation

2. External quality assurance (round robin test)
principle: root mean square of the error of measurement

\[ a^2 + b^2 = c^2 \]

Pythagoras of Samos
(570 BC – 510 BC)

error limits

old RiliBÄK
Total error: \(2 \times 5\% + 6\% = 16\%\)

New RiliBÄK Root mean square of the error of measurement:
\[ \sqrt{5^2 + 6^2} = 7.8\% \]
Calculation of root mean square of measurement deviation (RSMD)

\[ \Delta_{\text{max}} = \sqrt{k^2 \cdot s_{ep}^2 + \delta_{ep}^2} \]

k=3, coverage factor for calculating the internal laboratory deviation limits

\( s_{ep} \), empirical standard deviation of the control sample measurements used in the calculations during the pre-evaluation period

\( \delta_{ep} \), systematic deviation of measurement of the control sample measurements used in the calculations during the evaluation period (ep)

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


*J Lab Med* **2015; 39**: 26–69

---

**Orth: Analytical Performance Specifications**

---

**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:** \( d_{pa} / s_{pa} = 1.7 \)
<table>
<thead>
<tr>
<th>No.</th>
<th>Measurand</th>
<th>Permissible relative deviation of a single result or the relative root mean square, respectively (From)</th>
<th>4</th>
<th>Rili-BAEK applicable concentration intervals of columns 3 and 5 (To)</th>
<th>Unit</th>
<th>5</th>
<th>Permissible relative deviation in EQA</th>
<th>6</th>
<th>Type of target value in EQA</th>
</tr>
</thead>
<tbody>
<tr>
<td>38</td>
<td>Lactate</td>
<td>11.0%</td>
<td></td>
<td>9</td>
<td>90</td>
<td>mmol/L</td>
<td>18.0%</td>
<td></td>
<td>NV</td>
</tr>
<tr>
<td>39</td>
<td>Lactate dehydrogenase</td>
<td>9.0%</td>
<td></td>
<td>100</td>
<td>700</td>
<td>U/L</td>
<td>18.0%</td>
<td></td>
<td>RMV</td>
</tr>
<tr>
<td></td>
<td>(LDH) EC 1.1.1.27</td>
<td></td>
<td></td>
<td>1.67</td>
<td>11.7</td>
<td>μkat/L</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>40</td>
<td>Leucocytes</td>
<td>6.5%</td>
<td></td>
<td>2</td>
<td>30</td>
<td>10^9/L</td>
<td>18.0%</td>
<td></td>
<td>RMV</td>
</tr>
<tr>
<td>41</td>
<td>Lithium</td>
<td>6.0%</td>
<td></td>
<td>0.3</td>
<td>3.5</td>
<td>mmol/L</td>
<td>15.0%</td>
<td></td>
<td>RMV</td>
</tr>
<tr>
<td>42</td>
<td>Magnesium</td>
<td>7.5%</td>
<td></td>
<td>0.3</td>
<td>3.5</td>
<td>mmol/l</td>
<td>5.0%</td>
<td></td>
<td>RMV</td>
</tr>
<tr>
<td>43</td>
<td>Sodium</td>
<td>3.0%</td>
<td></td>
<td>110</td>
<td>180</td>
<td>mmol/l</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>44</td>
<td>pCO₂</td>
<td>7.5%</td>
<td></td>
<td>≤35</td>
<td></td>
<td>mmHg</td>
<td>12.0%</td>
<td></td>
<td>NV</td>
</tr>
</tbody>
</table>

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

J Lab Med 2015; 39(1): 26–69

- 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
Six Sigma Performance of BioRad and Technopath controls for ALT (left) and Chloride (right)

- 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%

Performance criteria have to be revised in a timely and controlled process.
Quality performance specifications: Challenge of calculated results (e.g. anion gap, ratios, eGFR)

1. Error propagation in formulas consisting of test results, constants and estimated factors

2. Linearity of uncertainties

3. Probability density function of single pU factor? (rectangular, triangular, normal, U-form, asymmetrical)

4. Reliability of single pU?

5. Mathematical model to calculate total pU

Guide to the Expression of Uncertainty in Measurement (GUM)

Summary of procedures for evaluating and expressing uncertainty components

- Specify the measurand: (what is being measured and the mathematical functional relationship between the measurand and the input quantities upon which it depends)
- Identify sources of uncertainty
- Identify and correct for systematic error (bias) where possible
- Quantify uncertainty components: determine the standard uncertainty associated with each of the input quantities, including any uncertainty associated with the correction for systematic error. An uncertainty estimate obtained by the statistical analysis of serial observations OR uncertainty estimate obtained by other means (authoritative published report, a calibration certificate, personal experience or a numerical quantity associated with a certified reference material)
- Calculate the value of the measurand: that is, calculate the result of the measurement from the functional relationship which connects the various input quantities to the measurand
- Calculate the combined standard uncertainty of the measurand: that is, calculate the combined standard uncertainty of the measurand from the standard uncertainties (and covariances if present) associated with the various input quantities. These standard uncertainties are combined according to the rules based on the law for the propagation of uncertainties
- Calculate the expanded uncertainty of the measurand by applying an appropriate coverage factor, k. The expanded uncertainty is equal to the combined standard uncertainty of the measurand multiplied by k. For medical laboratory applications, k is typically given the value of 1.96 (or 2.0). This provides an expanded uncertainty which includes 95.0% (or 95.4%) of the values within the distribution of the measurand. The expanded uncertainty calculated in this manner provides a coverage interval on the assumption that the distribution of the measurand is normal

Monte Carlo Stimulation procedure ‘automatically’ takes into account any nonlinearities in the functional relationship

- graphical representation of the distribution of the measurand can be obtained directly from the MCS procedure
- **significant reduction in the mathematical skills required for most evaluations**
- MCS generally provides improved estimates for non-linear models
- MCS provides a coverage interval corresponding to a stipulated coverage probability (normal distribution, 95% for coverage factor of 1.96 or 95.4% for coverage factor of 2.0. For asymmetric distributions the shortest 95% coverage interval is quoted)

\[ u^2(AG) = u^2(Na^+) + u^2(K^+) + u^2(Cl^-) + u^2(HCO_3^-) \]

Thus: \[ u^2(AG) = (1.2)^2 + (0.10)^2 + (1.5)^2 + (1.2)^2 \]

\[ = 1.44 + 0.01 + 2.25 + 1.44 \]

\[ = 5.14 \]

And: \[ u(AG) = 2.26, \text{ mmol/L (AG standard uncertainty)} \]
Summary

- Lack of outcome-based performance criteria should trigger the use of other analytical performance goals lower in hierarchy if widely accepted both by medical professionals and from the health-economical network.
- Performance criteria have to be constituted and revised by medical professionals.
- Performance criteria should be established for the complete 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.
- Performance criteria should be mandatory for all tests performed in healthcare (exceptions have to be clearly defined!)
- Results from EQAS testing schemes can be used in a formalized process to revise performance goals.

Orth: Analytical Performance Specifications
matthias.orth@vinzenz.de