

# AUS DEM LABOR

Laborwerte – vergleichbare Ergebnisse?

## Messgrössen und Messverfahren – Konstanz und Vergleichbarkeit

Vergleichbare Ergebnisse versprechen Inserate für Laborgeräte. Was wir eigentlich wollen, sind identische Resultate bei identischen Proben. Oder praktischer gefragt: Sind heute Resultate aus verschiedenen Laboratorien und verschiedenen Geräten ohne weiteres austauschbar? Sicher eher als früher, als auch im professionellen Labor eine Vielzahl verschiedener Geräte und Methoden vorherrschten.

Aber so locker, wie die Werbung oft tönt, dass alles ungeheuer präzise und vergleichbar sei, ist die Wirklichkeit denn doch nicht. Es hängt natürlich von der Situation und den Ansprüchen ab: Wenn Abweichungen von  $\approx 10\%$  akzeptabel sind, kann man der Reklame glauben. In vielen Fällen bleibt es aber beim alten Rat-schlag an den Arzt, das Labor nicht ohne Not zu wechseln, und an das Labor, die Methode nicht ohne Not zu wechseln. Und bei der



Dr. phil. Peter Hagemann  
Schaffhausen

Empfehlung an den Arzt, Zweifelsfälle partnerschaftlich mit seinem Labor abzuklären.

### Messgrössen der Routine

Es fragt sich zunächst, wie gross die Streuung überhaupt sein darf. Der Kanadier Tonks hatte bereits vor 50 Jahren einen Beitrag zum Thema geliefert und dabei ein Viertel des Referenzintervalls als tole-

**TAB. 1** Gemessene Streuung und tolerierte Variation  
√ wenn in mindestens zwei (von drei) Wertebereichen eingehalten, – nicht verfügbar (3)

Messgrösse	Cholesterin	Kreatinin	Glucose	HDL	LDL	Phosphat	Triglyceride	Harnsäure
tolerierte Variation	$\pm 4\%$	$\pm 4\%$	$\pm 4,5\%$	$\pm 4\%$				
Architect	√	√	√		√	√	√	√
Coulter AU	√	√	√	√	–	√		√
Vitros	√	√	√	√	–		√	√
Cobas	√	√	√	√	√	√	√	√
Advia	√		√	√		√	√	√
Konelab	√	√	√	√	√	–	√	

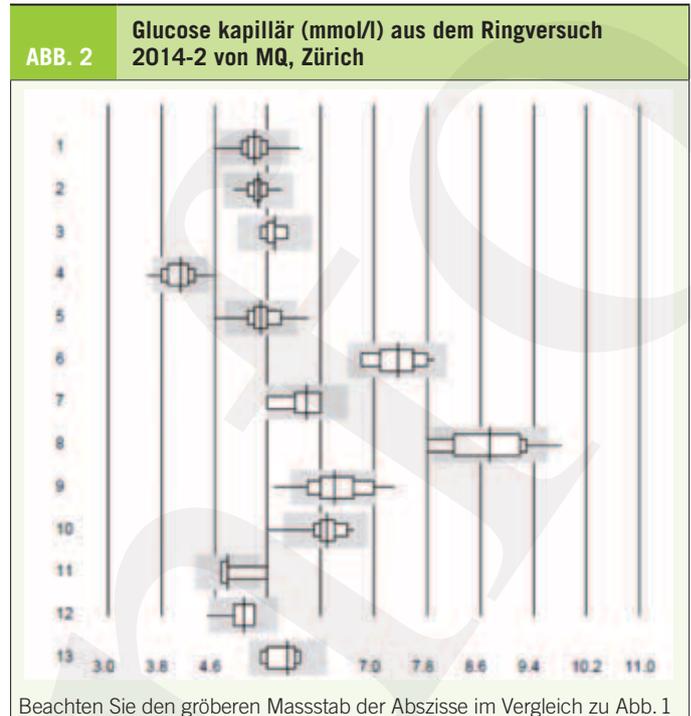
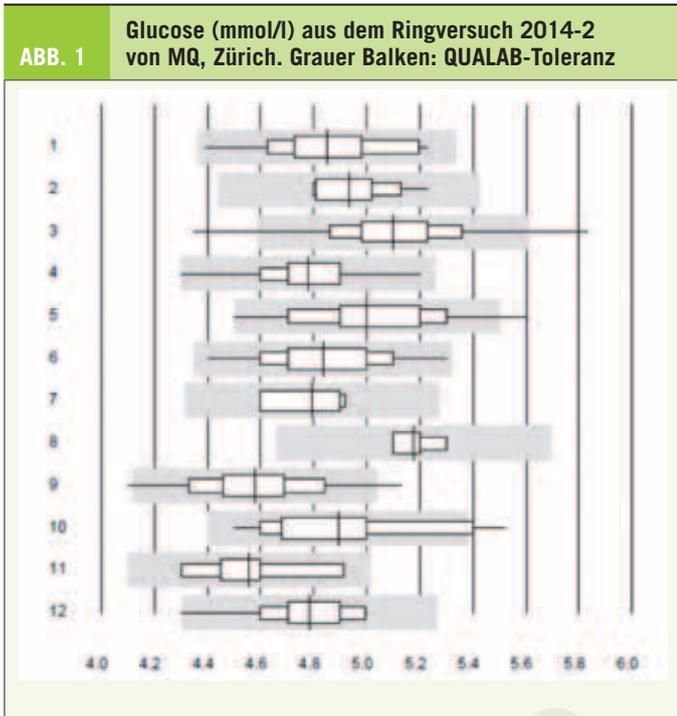


Abb. 1 und 2 mit Erlaubnis von MQ, Zürich

rierbare Streuung vorgeschlagen (1). Heute wird in diesem Sinne das Tabellenwerk der Arbeitsgruppe von Carmen Ricos (2) herangezogen, das Auskunft über intraindividuelle und interindividuelle Streuungen der häufigsten Messgrößen gibt.

Einen konkreten Vergleich von acht Messgrößen, bestimmt auf sechs Analysesysteme in insgesamt 63 teilnehmenden Laboratorien hat 2014 eine internationale Arbeitsgruppe unter Leitung von Linda Thienpont (Gent) vorgelegt (3). Der ausführliche Titel nimmt das Fazit aus der Sicht des Labors vorweg: „Inadequate Standardization among 6 Routine Laboratory Assays.“ In Tabelle 1 sind die Resultate zusammengefasst.

Die tolerierte Variation wurde hauptsächlich aus den Versuchsergebnissen abgeleitet (all-method trimmed mean) und entspricht nicht unbedingt den erwünschten Werten aufgrund der intraindividuellen biologischen Streuung (2). Ebenso wenig können die etablierten Begriffe Impräzision und Unrichtigkeit aus der Arbeit entnommen werden. Zwischen den Analysesystemen differierte die Kalibration mitunter erheblich, wobei einzelne Systeme deutlich „robuster“ sind als andere. Erheblich grösser als erwartet bzw. beim Hersteller ermittelt war ausserdem das Ausmass der Impräzision bei einzelnen Messgrößen, also sozusagen die im Labor selbst gemachte zusätzliche Streuung.

Ähnliche Resultate liegen für Zellzählgeräte vor (4). Allerdings wurden fünf Analytoren nur im eigenen Labor, nicht zwischen verschiedenen Laboratorien verglichen. Die tolerierte Streuung betrug für Leukozyten 7,5 %, für Thrombozyten 12,5 %. Die Autoren betonen wie Stepman et al. (3) die zentrale Bedeutung von Kalibration und ausserdem von Training und Qualitätskontrolle.

Das Beispiel aus einem der beiden schweizerischen Qualitätskontrollzentren umfasst im Unterschied zu Stepman et al. (3) vorwiegend Geräte aus dem Praxislabor (Abb. 1). Sie schneiden nicht schlechter ab als die mechanisierten Systeme. Die hierzulande paritätisch ausgehandelte zulässige Streuung (grauer Balken) wird – welche Überraschung – fast durchwegs eingehalten. Diskutabel erscheint mir bei uns, dass wir uns bekanntlich mit einem Vergleich

innerhalb von Analysesystemen begnügen (Impräzision) und die Unrichtigkeit nicht beurteilen. Hersteller und Benutzer von System 8 zum Beispiel dürften mit sich äusserst zufrieden sein und sehen keinen Anlass zu Verbesserungen.

### „Heikle“ Messgrößen

Vitamin D ist ein Beispiel für die meisten Fragestellungen punkto Ernährung, Intoxikation oder Mangel, gemessen als 25-OH-Cholecalciferol. Die Antikörper sind oft nicht spezifisch genug und reagieren auch mit Metaboliten, der Analyt wird mitunter unvollständig vom Bindungsprotein getrennt, Matrixeffekte (z.B. Lipide) spielen eine Rolle. Die Studie von Enko et al. (5) wurde in sieben europäischen Laboratorien an 133 Patientenproben mit sieben Analysesystemen bzw. -methoden durchgeführt. Es ergaben sich zwei Cluster von Resultaten, die nicht überlappten. Die höchste Korrelation zwischen zwei Methoden betrug  $r=0,96$ , die geringste  $r=0,46$ . Einer der geprüften Tests wurde später vom Markt genommen. Die Autoren bemerken etwas süffisant, dass sämtliche physiologischen und pathologischen Differenzen geringer waren als die analytischen Differenzen. Ausserdem wundern sie sich, wie denn die kritische Grenze der WHO von  $20 \mu\text{g/l}$  ( $50 \text{ nmol/l}$ ) eingehalten werden soll. Fazit ist, unbedingt bei einer einmal gewählten Methode bleiben.

Heikel ist es auch dort, wo kleine Differenzen entscheidend sind: Ein Beispiel ist das Glykohämoglobin. Eigentlich können die scharfen Grenzen der „American Diabetes Association“ für Prädiabetes ( $\leq 6,4\%$ ) bzw. Diabetes ( $\geq 6,5\%$ ) nicht einmal durch ein einzelnes Messgerät mit Sicherheit voneinander unterschieden werden. Geschweige denn durch zwei verschiedene Messgeräte.

Heikel ist schliesslich der Vergleich von Systemen für kapilläre Proben, wie sie im point-of-care-Bereich (POCT) häufig angewendet werden. Abbildung 2 zeigt die im Vergleich zu Abb. 1 grössere Divergenz zwischen den einzelnen Systemen, aber auch die grosse Zuverlässigkeit jedes einzelnen Fabrikats an und für sich: ein Archipel von proprietären Systemen, deren individuelle Streuung nota bene deshalb so gering ist, weil der Benutzer sie nicht manipulieren

kann. Besonders problematisch sind Vergleiche von POCT-Resultaten mit solchen der professionellen Analytik: Neben der Methode differiert das Probenmaterial, liegen ungleiche Empfindlichkeiten für Einflussgrößen und Störfaktoren etc. vor. Eine weitere Variable ist die klinische Situation, die dem Labor ohnehin nicht bekannt ist, und von der auch der Anwender nicht unbedingt weiss, ob sein POCT-Gerät dafür validiert ist (6).

### Massnahmen

In der referierten Arbeit (3) werden Korrekturfaktoren auf der Basis eines Mittels aller verwendeten Methoden vorgeschlagen. Das ist wohl ein pragmatischer Vorschlag, aber gleichzeitig eine opportunistische Scheinlösung und eine Absage an alle Prinzipien der Analytik und der Rückführbarkeit auf etablierte Standards. Lokal hat

es sich indes bewährt, z.B. die Glucosebestimmung in der Notfallstation mit derjenigen im Labor per Faktor zu harmonisieren. In einem Editorial im selben Heft (7) werden weitere Ansätze diskutiert, aber am Schluss bleibt der Ruf nach mehr Standardisierung und Kalibration. In manchen Fällen dürften uns erst neue Messgrößen und neue Testverfahren aus der Sackgasse führen (8). Industrie und Technik sind in diesem Punkt weiter: Schraube und Mutter von verschiedenen Herstellern passen zusammen.

### Dr. phil. Peter Hagemann

Labormedizinisches Zentrum Dr. Risch AG, 8200 Schaffhausen  
peter.hagemann@risch.ch

### Literatur:

1. Tonks DB. A study of the accuracy and precision of clinical chemical determination in 170 Canadian laboratories. Clin Chem 1963;9:217-33
2. Ricos C et al. Desirable Specifications, update 2014. Verfügbar unter [www.westgard.com/Desirable Biological Variation](http://www.westgard.com/Desirable%20Biological%20Variation)
3. Stepman HCM et al. Measurements for 8 Common Analytes in Native Sera Identify Inadequate Standardization among 6 Routine Laboratory Assays. Clin Chem 2014 60:855-63
4. Gu W et al. Operating Scheme and Experience of Executing internal Comparisons of Blood Cell Analysis in a Clinical Laboratory. Clin Lab 2014;60:1627-34
5. Enko D et al. 25-Hydroxy-Vitamin-D Status: Limitations in Comparison Across Different Assay Methods. Clin Lab 2014;60:1541-50
6. Karon B. Using Glucose Meters in Intensive Care Units. Clin Lab News 2015;41/1
7. Zegers I, Schimmel H. To Harmonize and Standardize: Making Measurement Results Comparable. Clin Chem 2014;60:911-3 (Editorial)
8. Meeusen JW, Lieske JC. Looking for a Better Creatinine. Clin Chem 2014;60:1036-9 (Perspective)

### Take-Home Message

- ◆ Routineresultate von verschiedenen Laborgeräten sind für moderne Analysensysteme im Rahmen von 10% Abweichung vergleichbar
- ◆ Das heisst aber eben nicht, dass verschiedene Verfahren bei derselben Probe identische Resultate ergeben, wie es der Arzt eigentlich erwartet
- ◆ Besonders heikel sind z.B. Vitamin D, scharfe Entscheidungsgrenzen wie bei Glykohämoglobin sowie Vergleiche von kapillären Systemen
- ◆ Wichtigste Massnahme ist nach wie vor, ein bewährtes Analysensystem nicht ohne Not zu wechseln