

Der Zusammenhang zwischen dem diagnostischen Wert eines Labortests und der Qualität des analytischen Verfahrens

U. E. Spichiger

Zusammenfassung

Aus den verfügbaren Daten (vergl. Tab. 1 und 2) geht hervor, daß verschiedene Laboratorien mit den eingeführten Analysenverfahren zur Untersuchung der Magnesiumkonzentration im Serum generell ungenügend vergleichbare Resultate liefern, um hochwertige diagnostische Aussagen zu machen. Die enge Zusammenarbeit mit einem gut geführten Labor kann zur Verbesserung der Aussagekraft führen und macht die saubere Erhebung der analytischen Fehler möglich.

Der Individualitätsindex für die Magnesium Gesamtkonzentration im Serum zeigt, daß es sich lohnt individuelle Differenzen zu beurteilen. Allgemeine Referenzbereiche können als grobe Richtlinien dienen, erlauben jedoch nur eine unpräzise diagnostische Aussage.

Summary

The relationship between the analytical performance of different methods for the analysis of total magnesium ions and the diagnostical value of the results are discussed with concern to the diagnostical specificity and sensitivity of a test. The allowable analytical error for a medically reliable interpretation of data was defined by the college of American Pathologists (CAP) in 1976. The available statistical data derived from quality control procedures document the generally insufficient analytical performance of different analytical methods with respect to the interpretation of the small differences in magnesium levels observed for individuals. The collaboration with a single clinical chemical laboratory as well as individual monitoring is recommended to minimize the analytical error and to favour a reliable interpretation of the analytical results. As demonstrated by Harris's individuality index the reference interval may only serve as a rough limit of "not extremely deprived individuals".

Résumé

Quant aux méthodes différentes pour l'évaluation du dosage de l'ion de magnésium totale les dimensions relatives des variations analytiques et des variations biologiques sont des facteurs essentiels à prendre en compte pour l'interprétation des résultats d'un laboratoire. Ces facteurs sont discutés en visage de la sensibilité et de la spécificité diagnostique. L'erreur analytique admise pour une interprétation médicale assez sensible était défini par le Collège des Pathologistes Américains (CAP) à 1976. Les valeurs statistiques évalués à base des études de contrôle de qualité documentes l'application généralement insuffisante des méthodes analytiques relative aux variations biologiques étroites. La collaboration avec un seul laboratoire de chimie cliniques et l'évaluations des taux individuelles sont recommandés pour réduire la dispersion des valeurs analytiques et pour favoriser l'interprétation efficace. L'index d'individualité de Harris représente le valeur diagnostique des limites de l'intervalle de référence. En cas du dosage de l'ion de magnésium totale les valeurs de référence ne servent que des limites pour la diagnose des individus "pas extrêmement déprivés".

Die Beurteilung eines einzelnen Analysenresultates kann sehr einfach am Referenzbereich des zuständigen Laboratoriums vollzogen werden. Anamnese, klinischer Befund und zahlreiche andere diagnostische Messwerte sind jedoch oft nicht mit dem Laborresultat in Einklang zu bringen. Ebenso stellt sich häufig die Frage, weshalb Resultate nicht signifikant ausfallen, obwohl subjektiv beurteilt, und durch andere Parameter auch

objektivierbar, eine Veränderung eingetreten ist.

Die Aussagekraft eines einzelnen Laborresultates steht in engem Zusammenhang mit:

1. dem Vergleichskollektiv, das den Referenzbereich liefert,
2. der Lage der Diskriminations-/ Entscheidungsschwelle (Cutoff value) zwischen verschiedenen Kollektiven und der damit verbundenen Ausprägung der diagnostischen Parameter,
3. dem gesamten analytischen und präanalytischen Fehler der La-

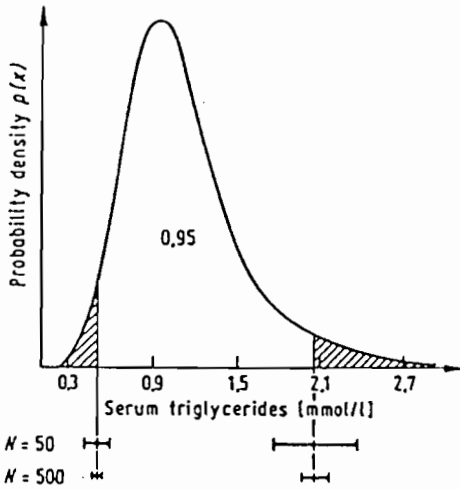
borresultate und der biologischen Variation der Kollektive.

1. Vergleichskollektiv und Referenzintervall

Die IFCC-Definition des Referenzintervalls beruht auf einem statistischen Konzept:

Nach Auswahl des geeigneten Kollektivs durch sorgfältige medizinische Vorabklärungen wird das Referenzintervall eines Laborparameters, ungeachtet der Art der Häufigkeitsverteilung, als das zentrale 95% Intervall aller Resultate bezeichnet (Abb. 1).

Der Zusammenhang zwischen dem diagnostischen Wert eines Labortests und der Qualität des analytischen Verfahrens



0.025 Fraktile: 0.51 mmol/l

0.975 Fraktile: 2.05 mmol/l

Abb. 1: Statistische Auswertung von Laborresultaten eines Referenzkollektivs entsprechend den Empfehlungen (1987) der IFCC und ICSH [10]. Bestimmung des zentralen 0.95 Referenzintervalls und Darstellung des Einflusses der Stichprobenzahl. Geglättete Häufigkeitsverteilung der Serum-Triglycerid Konzentrationen in [mmol/l].

Dies bedeutet, daß die Elimination von je 2.5 % der Resultate auf jeder Seite der Häufigkeitsverteilung zum Referenzintervall führt [1].

Die Definition des Kollektivs mit den Einschränkungen, die allenfalls gemacht werden, bleibt offen. Es kann sich dabei sowohl um das Kollektiv von anamnestisch gesunden, nicht rauchenden, nicht Alkohol trinkenden Erwachsenen handeln, wie auch um ein Kollektiv von Personen, welche mit Placebo behandelt wurden. Auch ein Kol-

lektiv von Patienten der Notfallstation, welche keine Symptome des Herzinfarktes oder von akuter Angina pectoris aufweisen, kann für eine gezielte Fragestellung ein Referenzkollektiv darstellen. Da sich die Resultate innerhalb des Referenzkollektivs zwangsläufig mit den Resultaten des abzuklärenden Kollektivs, den Symptomträgern, überlappen, wird sich immer eine Grauzone zwischen beiden Kollektiven einstellen, innerhalb welcher die Interpretation fehleranfällig ist (Abb. 2).

2. Die diagnostischen Parameter und die Lage der Entscheidungsschwelle

Das Risiko einer Fehlinterpretation läßt sich abschätzen, wenn die nachfolgenden diagnostischen Parameter und Kriterien zur Beschreibung der diagnostischen Aussagekraft eines Testes bekannt sind:

Terminologie der diagnostischen Parameter [2] zur Bewertung der **diagnostischen Aussage** (FN: falsch negative Resultate):

Praevalenz:

Fraktion der Symptomträger (D) innerhalb des abzuklärenden Kollektivs. Wahrscheinlichkeit $P(D)$ für das Auftreten des Symptoms.

diagnostische Sensitivität:

Fraktion der echt positiven Resultate (TP) im Kollektiv der Symptomträger ($TP/(TP+FN)$).

diagnostische Spezifität:

Fraktion der echt negativen Resultate (TN) im gesamten Referenzkollektiv ($TN/(TN+FP)$) während der Studie ($TN/(TN+FP)$).

Unspezifität:

Fraktion aller falsch positiven Resultate (FP) innerhalb des Referenzkollektivs ($FP/(FP+TN)$).

Charakterisierung und Bewertung eines Testes:

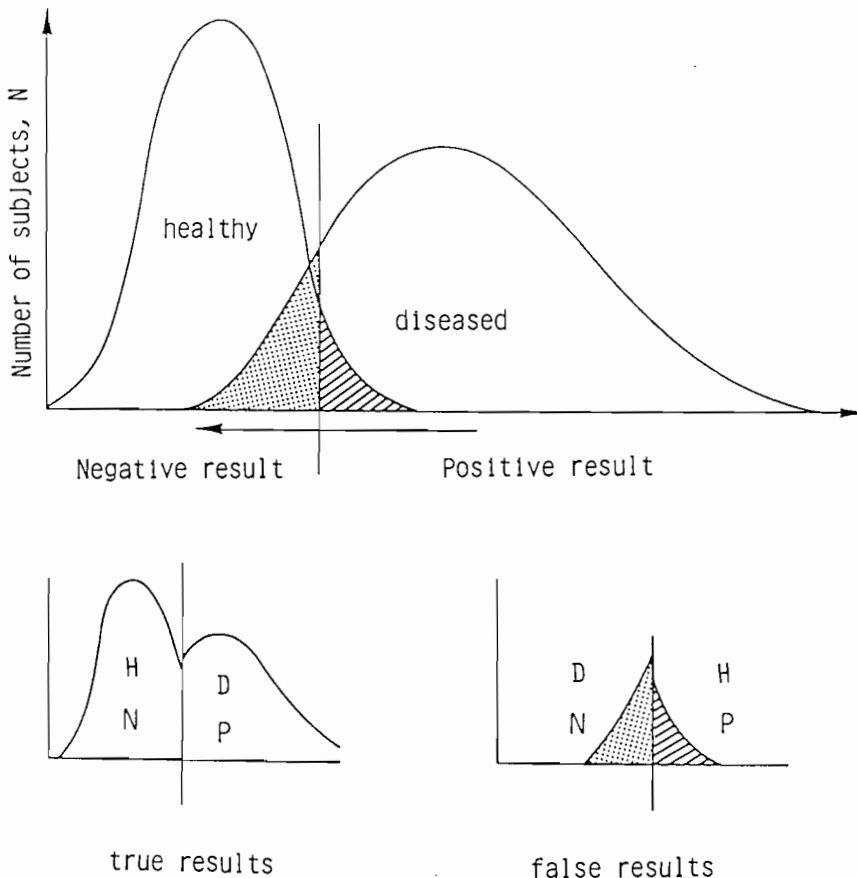


Abb. 2: Hypothetische Häufigkeitsverteilung der Testresultate von Referenzindividuen (H) und Symptomträgern (D). Getrennte Darstellung der Summe aller echt negativen (H, N) und echt positiven Resultate (D, P), sowie aller falsch negativen (D, N) und falsch positiven (H, P) Testresultate.

Der Zusammenhang zwischen dem diagnostischen Wert eines Labortests und der Qualität des analytischen Verfahrens

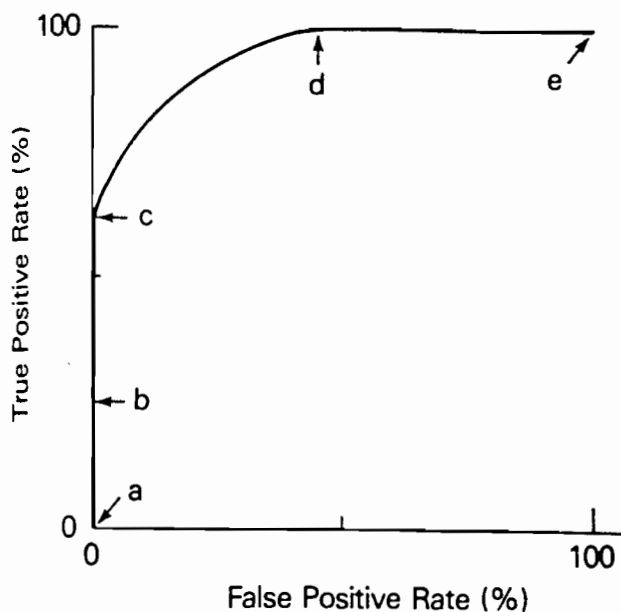
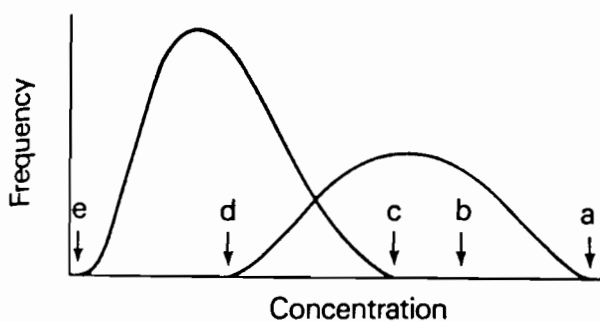


Abb. 3: Receiver Operating Characteristic Curve (ROC) [2]. Darstellung der Häufigkeit (oben) bzw. der Summenhäufigkeit (unten), des prozentualen Anteils von echt positiven Resultaten (Sensitivität) in Abhängigkeit des prozentualen Anteils an falsch positiven (Unspezifität) bei vorgegebener Praevalenz. Die Lage der Entscheidungsschwelle (a–e) verändert den Wert der diagnostischen Sensitivität und Unspezifität; Lage des Schwellenwertes bei a, b oder c: der Anteil an falsch positiven Ergebnissen ist Null, der Anteil an echt positiven steigt von 0 auf ca. 60 %; d, die echt positiven Fälle werden zu 100 % erfaßt, während gleichzeitig der Anteil an falsch positiven auf ca. 40 % ansteigt; e, die echt positiven Fälle werden alle erfaßt, aber gleichzeitig werden alle symptomlosen Fälle zu Kranken gemacht.

positiver prädiktiver Wert:

(+) PV or PV_{pos} [%]

Fraktion der echt positiven Resultate (TP) unter allen positiven Resultaten $((FP+TP) \times 100)$ in Prozenten.

negativer prädiktiver Wert:

(-) PV or PV_{neg} [%]

Fraktion aller echt negativen Resultate (TN) unter allen negativen Resultaten in Prozenten. $(TN+FN) \times 100$

ROC-Kurve ist. Die kleinste Fehler-rate bei gegebener Praevalenz liegt im Bereich der maximalen Neigung der Tangente. Wird die Entscheidungsschwelle über die Skala aller Resultate der beiden Kollektive verschoben, so ändern sich die diagnostischen Parameter eines Testes, abhängig von der Praevalenz des Symptoms (Abb. 4).

Effizienz:

Fraktion aller korrekt klassierten Personen innerhalb aller untersuchten Personen [%]. $((TN+TP)/(TN+TP+FN+FP)) \times 100$

Die ROC-Analyse (ROC: Receiver Operating Characteristic Curve, vergl. Abb. 3 [2]) erlaubt ein schnelles optisches Erfassen des Verhältnisses zwischen echt positiven (Sensitivität) und falsch positiven Resultaten (Unspezifität) eines Testes beim Verschieben der Entscheidungsschwelle über die ganze Werteskala der beiden Kollektive [3, 4, 5]. Sie erlaubt den Vergleich des diagnostischen Wertes von verschiedenen Testen [2]. Die ROC-Analyse zeigt ebenfalls den Zusammenhang zwischen dem diagnostischen Risiko und der Praevalenz eines Symptoms im untersuchten Kollektiv auf [3, 4]. Die Trennschärfe ist höher je stärker gewinkelt die

Daraus folgt: Je stärker das Referenzkollektiv durch eine gezielte diagnostische Fragestellung eingeschränkt wird, je effizienter ein Kollektiv voruntersucht ist, desto sicherer ist die Aussage. Die diagnostischen Parameter sowie die diagnostische Effizienz eines Analysenverfahrens verändern sich mit der Lage der Entscheidungsschwelle [3].

3. Die analytischen Fehler und die biologische Variation

Die **Trennschärfe**, mit der sich zwei Kollektive unterscheiden lassen, und die Lage der Entscheidungsschwelle stehen in engem Zusammenhang mit der Summe aller Fehler, welche ein Resultat beeinflussen. Der gesamte Fehler setzt sich aus dem systematischen Fehler und dem zufälligen, der Streuung, zusammen.

Die **gesamte Streuung** läßt sich in die folgenden Beiträge gliedern:

- 3.1. Die gesamte **analytische** Variation, von welcher eine Drift bzw. ein konstanter Fehler der Daten abzutrennen ist.
 - 3.2. Die **biologische** Variation: die interindividuelle und die intraindividuelle Variation.
 - 3.3. die technische und die biologische, **präanalytische** Variation.
- Mit der folgenden Formulierung wird der gesamte Fehler eines Laborresultates als Summe einzelner Komponenten erfaßt [6]:

$$S_{RW} = \sqrt{s_B^2 + s_{AT}^2 + s_{Rest}^2}$$

S_{RW}	Streuung der Referenzwerte
s_B	Biologische Streuung
s_{AT}	Analytische Streuung von Tag zu Tag (Unpräzision von Tag zu Tag)
s_{Rest}	Reststreuung (z. B. aus Entnahme und Vorbereitung des Untersuchungsmaterials)

3.1. Der gesamte Analysenfehler setzt sich zusammen aus dem syste-

Der Zusammenhang zwischen dem diagnostischen Wert eines Labortests und der Qualität des analytischen Verfahrens

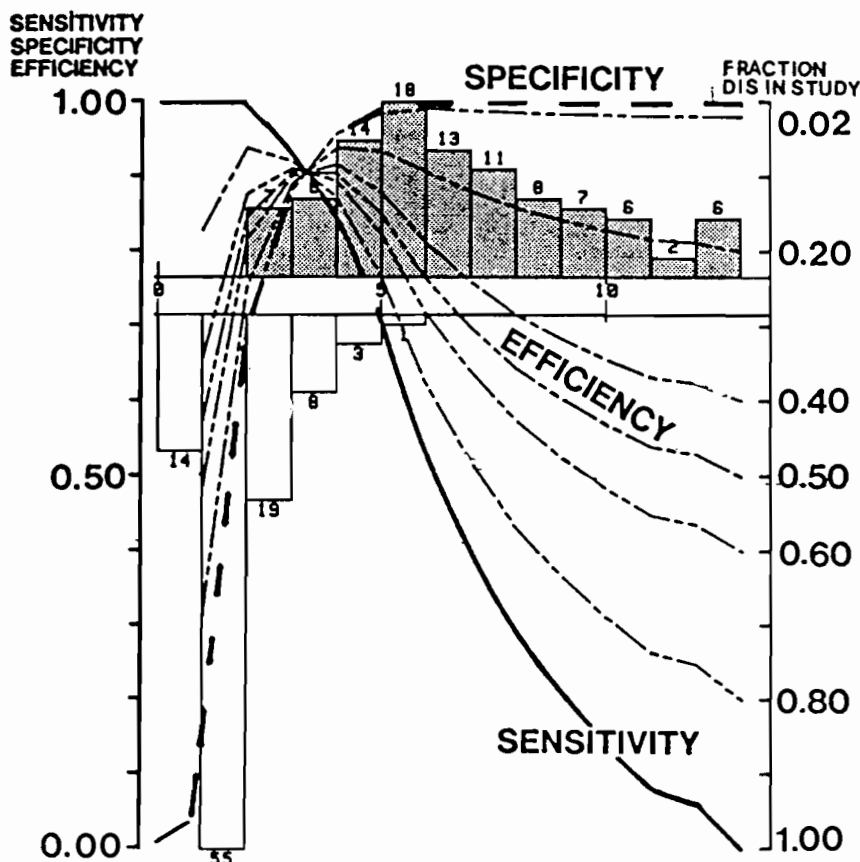


Abb. 4: Plot der diagnostischen Parameter in Abhängigkeit der Lage der Entscheidungsschwelle [3]. Stabdiagramm der Resultate von 200 AMI (akuter Myocardinfarkt) verdächtigen Patienten: 100 erwiesen sich als Referenzpersonen, 100 als AMI-Erkrankte (hypothetische Verteilung mit einer Praevalenz von 0.5). Die Ausprägung der diagnostischen Parameter verändert sich beim Verschieben der Diskriminationsschwelle mit der Praevalenz („fraction DIS in study“).

matischen Fehler (Richtigkeit, bzw. accuracy) und dem zufälligen Fehler, der Reproduzierbarkeit bzw. Repeitierbarkeit (reproducibility, repeatability) [7] (vergl. Abb. 5).

Systematische Fehler können in der Form eines proportionalen Fehlers oder eines Bias zum gesamten analytischen wie auch praeanalytischen Fehler (z. B. Art der Blutentnahme) beitragen und die Lage der Laborresultate gegenüber der Entscheidungsschwelle verschieben.

Der **gesamte analytische Fehler TE** läßt sich als Summe der zufälligen und systematischen Fehler wie folgt darstellen [8]:

$$TE = RE + SE + CE + PE \quad (1)$$

$$RE: \text{„random error“} = \pm t_{(n,d)} \cdot s_a$$

$$(t_{(n \rightarrow \infty, \alpha=0.05)} = 1.96)$$

$$SE: \text{Methodenvergleich:}$$

$$\text{„systematic error“} = (a + bX_i) - X_i = \hat{Y}_i - X_i$$

$$CE: \text{„constant error“} = \text{bias} \pm t_{(n,\alpha)} \cdot (s_a \sqrt{N})$$

$$PE: \text{Wiederfindungs-Experiment:}$$

$$\text{„proportional error“} = (\text{Rec \%} - 100) \cdot X_i$$

wobei a: Achsenabschnitt; b: Steigung der Regressionsgeraden; N: Stichprobenzahl; X, Y: Meßwerte; Xi: unabhängige Variable; \hat{Y}_i : Schätzwert der abhängigen Variablen; REC % Wiederfindung in Prozenten.

Wird ein maximal zulässiger Analysenfehler definiert, so muß die ana-

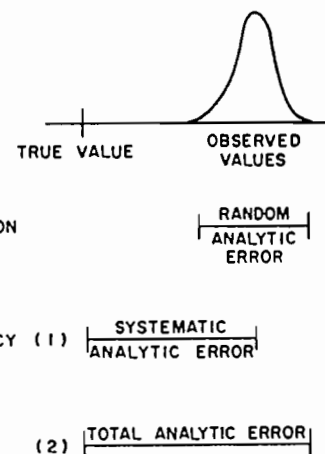


Abb. 5: Definition des totalen analytischen Fehlers durch die abstrakten Begriffe Praezision und Richtigkeit [8]. Praezision: zufälliger Fehler, Streuung der Resultate; Richtigkeit: systematischer Fehler, Bias (Verschiebung der Resultate, um einen konstanten Betrag); totaler analytischer Fehler: systematischer + zufälliger Fehler.

lytische Variation enger werden, je größer die Wahrscheinlichkeit für das gleichzeitige Vorliegen von systematischen Fehlern ist [9] (vergl. Abb. 6). Langzeitdriften werden oft nicht beachtet, sie bedingen besondere Prüfprogramme. Kurzzeitige Qualitätskontrollprogramme unterschätzen meist den analytischen Fehler.

Zur Abschätzung des **maximal zulässigen analytischen Fehlers** wird die analytische Variation mit der biologischen in Zusammenhang gebracht.

Der statistisch maximal zulässige analytische Fehler wurde 1976 am CAP Meeting (Comittee of American Pathologists) in Aspen wie folgt festgelegt [10]:

$$s_a \leq 1/2 s_b \quad (2)$$

Wie dies in einer IFCC-Publikation [11] mathematisch ausgeführt wird, wird damit der zulässige analytische Fehler so eingeschränkt, daß er, innerhalb des Referenzbereichs, die biologische Variabilität um weniger als 11.8 % erweitert:

$$s_{\text{tot}} = \pm 1.118 \sqrt{s_b^2} \quad (3)$$

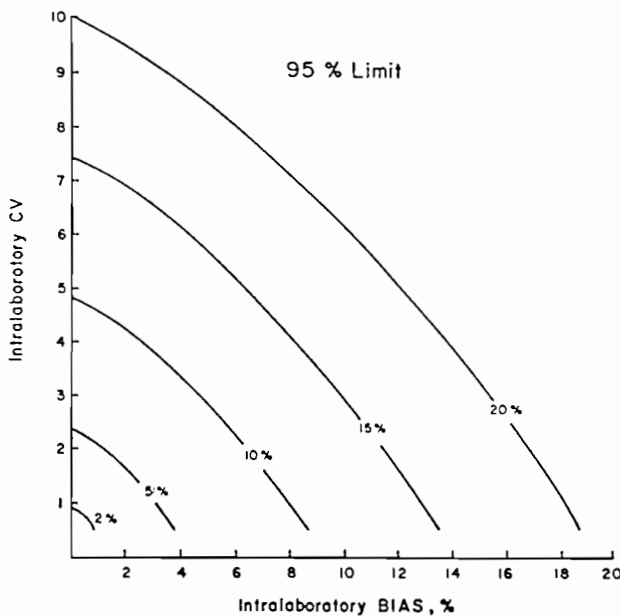


Abb. 6: Begrenzung von Bias und Variationskoeffizient (CV), um den tolerierbaren totalen analytischen Fehler bei 95 % aller Resultate innerhalb der Grenzen von 2, 5, 10, 15 und 20 % zu halten [9].

Definitionen:

- biologische Variation: s_b
- analytische Variation: s_a
- gesamte beobachtete Variabilität: s_{tot}

3.2. Die biologische Variation kann, entsprechend der diagnostisch relevanten Fragestellung, auf 2 Arten definiert werden: Für **individuelle** wiederholte Untersuchungen entspricht die biologische Variation (s_b) der intraindividuellen Variation (s_{intra}), d. h. der Variation eines Analyten, der über mindestens 20 Tage in einem definierten Specimen eines einzelnen Individuums bestimmt wird.

$$s_b = s_{intra} \quad (4)$$

Für das **Gruppen-Screening** errechnet sich die biologische Variation zwischen verschiedenen Individuen aus der Summe der intraindividuellen (s_{intra}^2) und interindividuellen Varianz (s_{inter}^2), angenommen daß sich die biologischen Varianzen addieren.

$$s_b = \pm \sqrt{s_{intra}^2 + s_{inter}^2} \quad (5)$$

In vielen Fällen ist die Annahme, daß sich die biologische Varianz addiert,

nicht erfüllt. Dies bedeutet, daß die biologische Variation und damit auch der zulässige Analysenfehler für ein Gruppenscreening überschätzt werden. Es resultiert ein geschätzter Referenzbereich, welcher weiter ist als der experimentell untersuchte.

Die Signifikanz von Unterschieden

Der Sinn der Verknüpfung zwischen dem analytischen Fehler und der biologischen Variabilität wird

deutlich, wenn nach der Signifikanz von Unterschieden zwischen einzelnen Resultaten gefragt wird.

Allgemein wird der erforderliche Unterschied zweier Testresultate (Δc), welcher eine signifikante Aussage erlaubt, aufgrund der gesamten Variation, welche ein Resultat belastet, statistisch mit Hilfe der Parameter der Student t-Verteilung, wie folgt definiert und berechnet:

$$\Delta c > \pm \sqrt{2 \cdot (t^2 s_D^2 + t^2 s_a^2)} \quad (6)$$

Das Signifikanzniveau wird durch den tabellierten t-Wert, mit der Wahrscheinlichkeit $P_{(n, \alpha)}$ bestimmt.

Zur Berechnung der minimalen, signifikanten Differenz zwischen 2 Resultaten (Δc) **desselben Individuums** ist nur die intraindividuelle Streuung relevant. Diese auf dem 95%-Niveau signifikante Differenz Δc für $t_{(n=50, \alpha=0.05)} = 2$ wird berechnet nach:

$$\Delta c > \pm 2.83 \sqrt{s_{intra}^2 + s_a^2} \quad (7)$$

Je nach Fragestellung kann die Aussage erweitert werden, indem in der obigen Gleichung s_{intra}^2 durch s_{inter}^2 bzw. $(s_{intra}^2 + s_{inter}^2)$ ersetzt wird. Der

letzte Ausdruck führt oft zu überschätzten, signifikanten Differenzen. Das Verhältnis der intraindividuellen zur interindividuellen Variation wird als sogenannter **Individualitätsindex** bezeichnet. Der Quotient erlaubt es abzuschätzen, wie aussagekräftig der Vergleich eines Resultates eines einzelnen Individuums mit der interindividuellen Variation bzw. dem Referenzbereich ist [12]. Der Individualitätsindex wird auch als „**Harris ratio**“ bezeichnet.

$$I = \pm \sqrt{s_{intra}^2 / s_{inter}^2} \quad (8)$$

Ist $I > 1.4$, dann besteht bei einem Vergleich des individuellen Resultates mit dem Referenzintervall eine hohe Wahrscheinlichkeit für das Auftreten von falsch positiven Resultaten.

Ist $I < 0.6$, dann kann durch den Vergleich mit dem Referenzintervall keine praezise diagnostische Aussage gemacht werden. Es besteht eine hohe Wahrscheinlichkeit für falsch negative Resultate.

Wie sehen nun die Verhältnisse für die Analyse der Magnesium Gesamtkonzentration im Serum aus?

Ein Überblick über die zur Zeit in den USA und in der Schweiz übliche analytische Variation wird in Tab. 1 gegeben. Die analytische Variation steigt an der unteren Grenze des physiologischen Bereichs, unter 0.6 mmol/l, auf über 10 % Variationskoeffizient (Vk) an. Werden verschiedene Laboratorien und Analysenmethoden auf verschiedenen Konzentrationsniveaus miteinander verglichen, steigt die analytische Variation bis auf 40 % des Mittelwertes an. Gemäß Tab. 1 wird für die nachfolgenden Berechnungen eine durchschnittliche analytische Variation innerhalb eines einzelnen Laboratoriums von 6 % Vk vorausgesetzt.

Die durchschnittliche intra- und interindividuelle biologische Variation (s_b) des Magnesiumspiegels im Blutserum von gesunden Individuen beträgt nach C. G. Fraser [15], über 2 bis 22 Wochen bestimmt [Vk %]:

individuelle Variation:

Der Zusammenhang zwischen dem diagnostischen Wert eines Labortests und der Qualität des analytischen Verfahrens

Tab. 1: Magnesium Gesamtkonzentration im Serum: „State of the Art“

Analytische Variation (Vk): Langzeit-Variation gemäß verschiedenen Qualitätskontroll-Untersuchungen.

	Vk	Mittelwert
Serumpool:		
Median	12 %	0.5 mmol/l
berechnet aus QC Programmen	6.4 %	0.82 mmol/l
in den USA (alle Methoden [10]):	5.5 %	1.44 mmol/l
Plasmapool USZ (53–57 Tage):		
(Calmagit, Photometric	6.05 %	0.81 mmol/l
(eigene Untersuchungen [13])	5.5 %	0.83 mmol/l
CH, nationales QC-Programm,		
alle Laboratorien [14]:	< 40 %	div.
(alle Methoden, CSCQ)		

$Vk_{intra} = 2.2 \%$
interindividuelle Variation:
 $Vk_{inter} = 5.9 \%$

Für alle praxisbezogenen Berechnungen wird die Streuung s ersetzt durch den von den Dimensionen unabhängigen Variationskoeffizienten Vk . Aus den biologischen Streuungskomponenten läßt sich die totale biologische Variation gemäß Gleichung (5) wie folgt berechnen:

$$Vk_b = \pm \sqrt{Vk_{intra}^2 + Vk_{inter}^2} = 6.4 \%$$

Die CAP-Empfehlung, wonach die analytische Variation kleiner als 50 % der biologischen sein soll, wird unter diesen Voraussetzungen nicht erfüllt. Die analytische und die interindividuelle biologische Variation sind bestenfalls annähernd gleich groß. Zur Beurteilung der Resultate wiederholter Untersuchungen einer einzelnen Person sollte die analytische Variation gemäß Gleichung (2) bei 1.1 % Vk liegen und bei maximal 3.15 % Vk für das Gruppen-Screening.

Konzentrationsunterschiede zur signifikanten Unterscheidung von 2 Testwerten auf dem 95%-Signifikanzniveau

Für **individuelle** Abklärungen wird die erforderliche Differenz gemäß Gleichung (6) und (7) unter Annah-

me einer analytischen Variation von 6 % Vk und einer intraindividuellen Variation von 2.2 % Vk berechnet. Die erforderliche Differenz gemäß CAP-Empfehlungen [10], um die Hypothese, daß sich 2 Werte nicht unterscheiden, mit 95 % Wahrscheinlichkeit zu verwerfen, beträgt bei einem Serum-Spiegel von $c_{Mg} = 0.8 \text{ mmol/l}$

$$\Delta c_{Mg} = \pm 0.14 \text{ mmol/l}$$

vorausgesetzt, daß die Werte durch ein einzelnes, zuverlässig geführtes Laboratorium ermittelt werden. Die individuellen Unterschiede werden durch den analytischen Fehler überdeckt. Die vom CAP publizierten Berechnungen ergeben aufgrund der festgestellten analytischen Fehler zwischen Laboratorien, signifikante Differenzen von $\Delta c_{Mg} = \pm 0.40 \text{ mmol/l}$ [10].

Für das Gruppen-Screening gelten folgende Voraussetzungen für die Berechnung:

analytische Variation:	$Vk = 6 \%$
totale biologische Variation:	$Vk = 6.4 \%$
interindividuelle Variation:	$Vk = 5.9 \%$
totale interindividuelle Variation, mit dem totalen biologischen Vk berechnet:	$Vk = 8.7 \%$
mit dem interindividuellen Vk berechnet:	$Vk = 8.4 \%$

Tab. 2: Kritische Werte zur Beurteilung der Resultate der Magnesium Gesamtkonzentration im Serum, berechnet nach den Empfehlungen der CAP [10]

Erlaubte analytische Variation (gemäß Gleichung [2]):

$Vk_a \leq 1.1 \%$	für individuelle Tests
$Vk_a \leq 3.15 \%$	für Gruppen-Screening:
2.8 %	(mit totalem Vk_b geschätzt)
	(mit Vk_{inter} geschätzt)

Erforderliche signifikante Differenz (gemäß Gleichung (6,7)) für $t_{(n=50, \alpha=0.05)}$ bei einem Serum-Spiegel von $c_{Mg} = 0.8 \text{ mmol/l}$:
für die **individuelle** Bewertung:

$$\Delta c_{Mg} = \pm 0.14 \text{ mmol/l}$$

für das Gruppen-Screening bei einem Vk von 8.7 %:

$$\Delta c_{Mg} = \pm 0.19 \text{ mmol/l}$$

Der **Individualitätsindex** (berechnet nach Gleichung (8)):

$$I = 0.37$$

(gemäß Gleichung (5))

Die erforderliche Differenz beträgt gemäß CAP-Empfehlungen (Gleichung 6 und 7) bei einem Serum-Spiegel von $c_{Mg} = 0.8 \text{ mmol/l}$:

Bei einem Vk von 8.7 %:

$$\Delta c_{Mg} = \pm 0.19 \text{ mmol/l}$$

Bei einem Vk von 8.4 %:

$$\Delta c_{Mg} = \pm 0.185 \text{ mmol/l}$$

Der Unterschied zwischen den erforderlichen Differenzen ist irrelevant. Da die analytische Variation dieselbe Ausprägung besitzt wie die gesamte biologische Streuung, können feine biologische Unterschiede nicht erfaßt werden.

Der **Individualitätsindex** für die Magnesium Gesamtkonzentrationen im Serum beträgt, berechnet nach Gleichung (8):

$$I = \sqrt{2.2^2/5.9^2} = 0.37$$

Die Beurteilung individueller Meßwerte mit Hilfe des Referenzbereichs ist somit nicht aussagekräftig. Es sind viele falsch negative Resultate zu erwarten, da die individuellen Variationen wesentlich enger sind als die interindividuellen.

Zusammenfassend (vergl. Tab. 2) ergeben sich folgende **Empfehlungen** für die Messung der totalen Magnesiumkonzentration im Serum und die Bewertung der Resultate.

1. Die diagnostische oder wissenschaftliche Fragestellung soll klar

Der Zusammenhang zwischen dem diagnostischen Wert eines Labortests und der Qualität des analytischen Verfahrens

formuliert werden. Kollektive, welche gegeneinander abgegrenzt werden sollen, sind möglichst eindeutig und eng begrenzt zu definieren.

- Die biologischen Parameter sollen bestimmt oder in der Literatur gesucht werden [15]. Individuelle Variationen können meist ohne großen zusätzlichen Aufwand in einer Vorperiode der Untersuchung am ausgewählten Kollektiv selbst bestimmt werden.
- Die notwendige analytische Richtigkeit und Präzision und der erlaubte Fehler sollen berechnet und die signifikanten Unterschiede für das relevante Signifikanzniveau (t-Tabelle) geschätzt werden. Alle Laboratorien können heute über die unvermeidlichen analytischen Fehler während der Untersuchungsperiode aufgrund der Resultate der Qualitätskontrollprogramme Auskunft geben.
- Präanalytische Fehler sollen ausgeschlossen (Art der Entnahme des Specimens) und das geeignetste Specimen gewählt werden.
- Geeignete Kombinationen von Analysen können oft zur Erhöhung der Aussagekraft führen (Profil-Analysen), während redundante Labortests mit einer stark

streuenden Variationskomponente die Unschärfe der Entscheidung und die Anzahl falscher Resultate erhöhen. Zur Beurteilung des diagnostischen Wertes von Labortests wird die Effizienz bzw. die Spezifität bei gleicher Sensitivität vice versa, mit Rücksicht auf die Prävalenz eines Symptoms im untersuchten Kollektiv verglichen.

Literatur

- Solberg, H. E.: International Federation of Clinical Chemistry (IFCC) and International Committee for Standardization in Haematology (ICSH). *J. Clin. Chem. Clin. Biochem.* **25** (1987) 645–656.
- Mark, H., Zweig, E., Robertson, A.: *Clin. Chem.* **28**, 6 (1982) 1272–1276.
- Gerhardt, W., Keller, H.: *Clinical and Laboratory Investigations. Scand. J. Clin. Lab. Invest.* **46** (Suppl.) (1986) 181.
- Metz, C. E., Goodenough, D. J., Rossmann, K.: *Radiology* **109** (1973) 297–303.
- van der Helm, H. J., Hische, E. A. H.: *Clin. Chem.* **25**, 6 (1979) 985–88.
- Stamm, D.: *DG Klinische Chemie Mitteilungen* **21**, 2 (1990) 69–77.
- Carey, R. N., Garber, C. C.: Evaluation of Methods. In: Kaplan, L. A., Pesce, A. J. (eds.): *Clinical Chemistry*. St. Louis 1984: C. V. Mosby Company, pp. 338–59.
- Westgard, J. O., Carey, R. N., Wold, S.: *Clin. Chem.* **20**, 7 (1974) 825–833.
- Ehrmeyer, S. S., Laessig, R. H.: *A. J. C. P.* **89**, 1 (1988) 14–18.
- College of American Pathologists (CAP). CAP Conference, Aspen 1976. In: *Elevitch, F. R.*, (ed.): *Proceedings of the Aspen Conference on Analytical Goals. Clin. Chem.* 1977.
- Fraser, C. G.: *JIFCC*, **2**, 2 (1990).
- Harris, E. K.: *Clin. Chem.* **21**, 1 (1975) 1457–64; *Clin. Chem.* **22**, 8 (1976) 1343–50.
- Spichiger, U. E.: Thesis, ETH Nr. 8830, 1989.
- Zehnder, R.: Centre de Contrôle de Qualité Suisse, La Chaux-de-Fonds. Mündliche Mitteilung anlässlich eines Vortrages an der Jahresversammlung der SGKC, Porrentruy 1989.
- Fraser, C. G. (ed.): *Interpretation of Clinical Chemistry Laboratory Data*. Oxford 1986, Blackwell Scientific Publications.

(Correspondence to: Dr. Ursula Spichiger-Keller, Eidgenössische Technische Hochschule, Laboratorium für Organische Chemie, Universitätsstr. 16, CH-8092 Zürich)