Zeitschrift: Mitteilungen aus Lebensmitteluntersuchungen und Hygiene = Travaux de chimie alimentaire et d'hygiène

Herausgeber: Bundesamt für Gesundheit

Band: 93 (2002)

Heft: 2

Artikel: In house-Validierung mikrobiologischer Prüfverfahren

Autor: Hübner, Philipp / Gautsch, Sylvia / Jemmi, Thomas

DOI: https://doi.org/10.5169/seals-981717

Nutzungsbedingungen

Die ETH-Bibliothek ist die Anbieterin der digitalisierten Zeitschriften auf E-Periodica. Sie besitzt keine Urheberrechte an den Zeitschriften und ist nicht verantwortlich für deren Inhalte. Die Rechte liegen in der Regel bei den Herausgebern beziehungsweise den externen Rechteinhabern. Das Veröffentlichen von Bildern in Print- und Online-Publikationen sowie auf Social Media-Kanälen oder Webseiten ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. Mehr erfahren

Conditions d'utilisation

L'ETH Library est le fournisseur des revues numérisées. Elle ne détient aucun droit d'auteur sur les revues et n'est pas responsable de leur contenu. En règle générale, les droits sont détenus par les éditeurs ou les détenteurs de droits externes. La reproduction d'images dans des publications imprimées ou en ligne ainsi que sur des canaux de médias sociaux ou des sites web n'est autorisée qu'avec l'accord préalable des détenteurs des droits. En savoir plus

Terms of use

The ETH Library is the provider of the digitised journals. It does not own any copyrights to the journals and is not responsible for their content. The rights usually lie with the publishers or the external rights holders. Publishing images in print and online publications, as well as on social media channels or websites, is only permitted with the prior consent of the rights holders. Find out more

Download PDF: 15.07.2025

ETH-Bibliothek Zürich, E-Periodica, https://www.e-periodica.ch

In house-Validierung mikrobiologischer Prüfverfahren

Philipp Hübner[1], Sylvia Gautsch[2] und Thomas Jemmi[3]

[1]Kantonales Laboratorium Zürich, Zürich
[2]Kantonales Laboratorium Basel-Stadt, Postfach, CH-4056 Basel
[3]Bundesamt für Veterinärwesen, Postfach, CH-3003 Bern

Eingegangen 9. Januar 2002, angenommen 12. März 2002

Einleitung

Am 1. November 2000 hat die Schweizerische Gesellschaft für Lebensmittelhygiene (SGLH) einen Workshop zur Validierung von mikrobiologischen Prüfverfahren durchgeführt. Das anlässlich dieses Workshops verteilte Diskussionspapier bildet die Grundlage für die vorliegende Publikation, welche keinen gesetzlich bindenden Charakter hat und das Ziel verfolgt, einen Leitfaden (guideline) zur In house-Validierung von Prüfverfahren aufzustellen, der sowohl den Laboratorien als auch den Auditoren dienlich ist.

Die Validierung soll eine Methode anhand von Laboruntersuchungen möglichst gut charakterisieren und deren Eignung für den vorgesehenen Zweck belegen. Im Gegensatz zu chemischen Methoden können bei mikrobiologischen Prüfverfahren Messunsicherheiten oft nicht genau bestimmt werden, weil lebende Materie gemessen wird, welche sich dynamisch verhält in Bezug auf Vermehrung, Absterben, Schädigung und Inaktivierung. Ein Nachweis von Mikroorganismen, vor allem im quantitativen Bereich, ist daher stets nur eine Annäherung an den wahren Wert. Die Fehler von mikrobiologischen Nachweismethoden bewegen sich bei quantitativen Prüfverfahren in der Grössenordnung einer halben 10er Potenz oder grösser (Abb. 1). Aussagen über Validierungsparameter sind daher stets unter diesen Gesichtspunkten zu betrachten.

Die vorliegende Publikation schlägt Verfahren für die In house-Validierung (neu: single laboratory validation) von mikrobiologischen Prüfverfahren vor. Sie stützt sich auf die EAL-G18 Richtlinie «Accreditation for Laboratories Performing Microbiological Testing» (1), das MICROVAL-WG 3 Dokument «Technical rules for validation criteria» (2), die AFNOR Richtlinie «Validation AFNOR des métho-

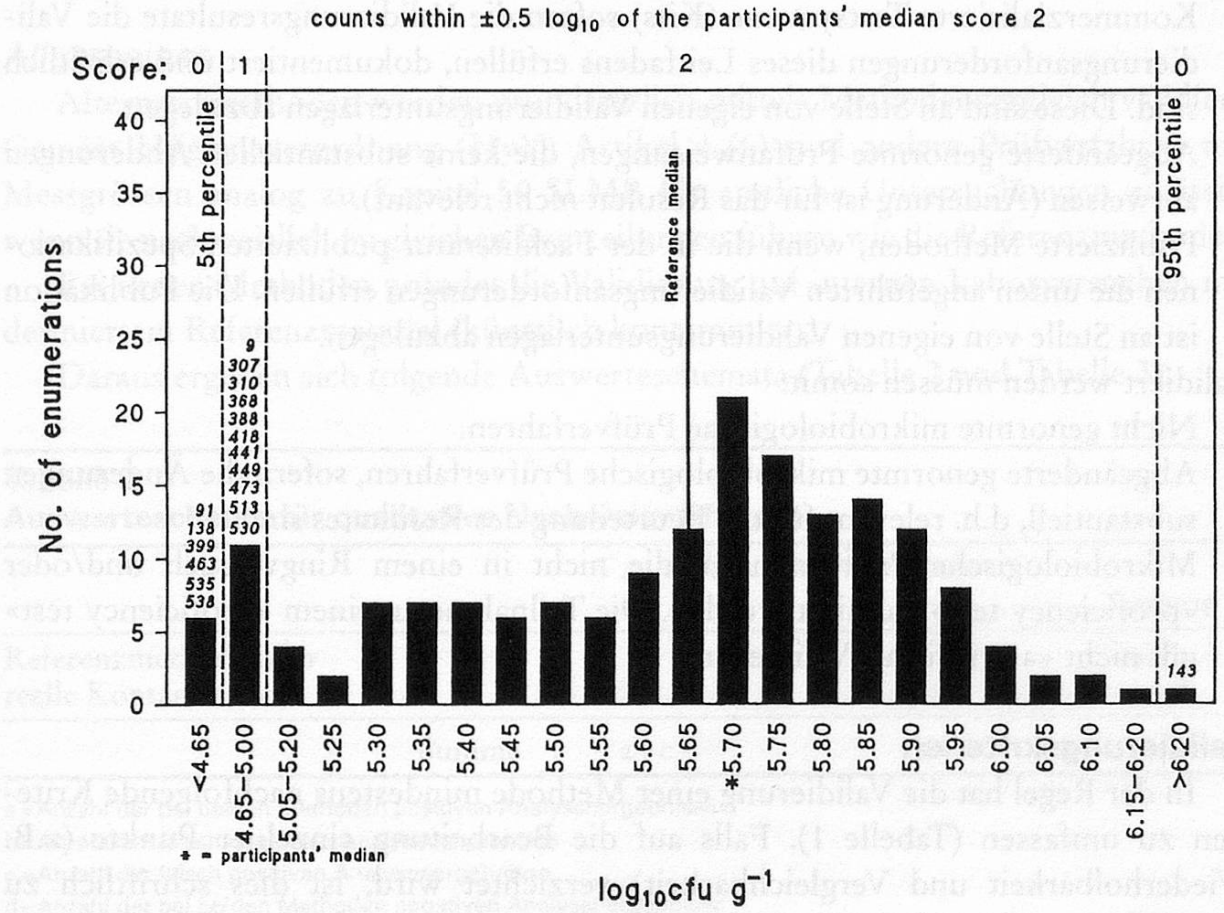


Abbildung 1 **Auswertung eines «proficiency test» für *Staphylococcus aureus* aus: PHLS Food EQA schemes, Distribution 096, Samples S0229 and S0230. Reproduziert mit Genehmigung durch PHLS**

des rapides d'analyses» (3), die «AOAC International Qualitative and Quantitative Microbiology Guidelines for Methods Validation» (4) und die provisorische Norm prEN ISO/FDIS 16140:2000(E) «Microbiology of food and animal feeding stuffs – Protocol for the validation of alternative method» (5).

In der Einleitung der prEN ISO/FDIS 16140:2000(E) steht explizit, dass wegen des (beträchtlichen) Ausmasses der in der Norm beschriebenen Methodenvergleiche dieses Vorgehen für die In house-Validierung durch individuelle Laboratorien nicht praktikabel sein könnte. Darauf basierend werden im Folgenden die spezifischen Anforderungen für eine In house-Validierung von mikrobiologischen Prüfverfahren durch individuelle Labors beschrieben.

Welche Methoden müssen validiert werden?

Bei der Validierung ist zu unterscheiden zwischen neuen Verfahren, für die keine genormten Prüfanweisungen (Standard- bzw. Referenzmethoden) bestehen und Alternativverfahren zu bereits existierenden Referenzmethoden.

Von der Validierung nach den festgelegten Kriterien ausgenommen sind:

- Kommerzialisierte Testsysteme (Kits) sofern die Validierungsresultate die Validierungsanforderungen dieses Leitfadens erfüllen, dokumentiert und erhältlich sind. Diese sind an Stelle von eigenen Validierungsunterlagen abzulegen.
- Abgeänderte genormte Prüfanweisungen, die keine substantiellen Änderungen aufweisen (Änderung ist für das Resultat nicht relevant).
- Publizierte Methoden, wenn die in der Fachliteratur publizierten Spezifikationen die unten angeführten Validierungsanforderungen erfüllen. Die Publikation ist an Stelle von eigenen Validierungsunterlagen abzulegen.

Validiert werden müssen somit:

- Nicht genormte mikrobiologische Prüfverfahren.
- Abgeänderte genormte mikrobiologische Prüfverfahren, sofern die Änderungen substantiell, d.h. relevant für die Beurteilung des Resultates sind.
- Mikrobiologische Prüfverfahren, die nicht in einem Ringversuch und/oder «proficiency test» validiert wurden. Die Teilnahme an einem «proficiency test» gilt nicht «a priori» als Validierung.

Validierungskriterien

In der Regel hat die Validierung einer Methode mindestens nachfolgende Kriterien zu umfassen (Tabelle 1). Falls auf die Bearbeitung einzelner Punkte (z.B. Wiederholbarkeit und Vergleichbarkeit) verzichtet wird, ist dies schriftlich zu begründen und festzuhalten (z.B. durch Verweis auf internationale Erfahrungswerte aus proficiency tests).

Tabelle 1
Validierungskriterien

	qualitative Methode		*quantitative Methode*	
Kriterium	*Alternativ*	*Neu*	*Alternativ*	*Neu*
Anwendungsbereich	✓	✓	✓	✓
Spezifität	✓	✓	✓	✓
Sensitivität	✓	✓	✓	✓
(Relative) Richtigkeit	✓	✓	✓	✓
Wiederholbarkeit (Präzision)	✓	✓	✓	✓
Vergleichbarkeit (Robustheit)	✓	✓	✓	✓
Nachweisgrenze	✓	✓		
Bestimmungsgrenze			✓	✓
Linearität			✓	✓
Falsch-Positiv-Rate	✓	✓		
Falsch-Negativ-Rate	✓	✓		
Statistische Übereinstimmung	✓		✓	

Vorgehen

Allgemeines

Alternativmethoden werden grundsätzlich mittels Methodenvergleich validiert. Gemäss Hygieneverordnung (HyV) Artikel 4 (6) sind andere Prüfverfahren mit Messgrössen analog zu Kapitel 56 SLMB für amtliche Untersuchungen zulässig, wenn sie nachweislich zu gleichen Beurteilungen führen wie die Referenzmethoden.

Bei neuen Methoden gründet die Validierung auf internen Laborversuchen mit definiertem Referenzmaterial (künstlich kontaminiert).

Daraus ergeben sich folgende Auswerteschemata (Tabelle 2 und Tabelle 3):

Tabelle 2
Auswerteschema für qualitative Nachweisverfahren

		zu validierende Methode		
		+	*–*	*Summe*
Referenzmethode oder reelle Kontamination	+	a	b	a+b
	–	c	d	c+d
	Summe	a+c	b+d	a+b+c+d=n

a = Anzahl der bei beiden Methoden positiven Analysenergebnisse
b = Anzahl der falsch negativen Analysenergebnisse
c = Anzahl der falsch positiven Analysenergebnisse
d = Anzahl der bei beiden Methoden negativen Analysenergebnisse

Tabelle 3
Auswerteschema für quantitative Nachweisverfahren

	Abhängige Wertepaare (Resultate)		
Probe	*Referenzmethode resp. reelle Kontamination (R)*	*zu validierende Methode (A)*	*Differenz (Methode A-R)*
1	x_{R1}	x_{A1}	d_1
i	x_{Ri}	x_{Ai}	d_i
n	x_{Rn}	x_{An}	dn
Mittelwerte	$\bar{x}_R$	$\bar{x}_A$	$\bar{d} \pm s_d$

x_{Ri} = i-ter Messwert mit Referenzmethode
x_{Ai} = i-ter Messwert mit Alternativmethode

Die Validierungsparameter im einzelnen

Anwendungsbereich

In Abhängigkeit des vorgesehenen Anwendungsbereiches sind die Untersuchungen an einer bzw. mehreren Produkte- bzw. Lebensmittelkategorien (Matrices) durchzuführen. Dabei sind insbesondere Aspekte der bei der Produktion verwendeten Lebensmitteltechnologien (z.B rohes, fermentiertes oder gekochtes Lebensmittel) zu berücksichtigen.

Mögliche Produkte- bzw. Lebensmittelkategorien sind in den Anhängen der AOAC-Richtlinie (4) und der ISO-Norm 16140 (5) aufgeführt. Bei der Auswahl möglicher Produkte- bzw. Lebensmittelkategorien für die Validierung eines mikrobiologischen Prüfverfahrens sind die Kenntnisse über Prävalenz der nachzuweisenden Mikroorganismen in den einzelnen Produkte- bzw. Lebensmittelkategorien sowie deren Relevanz für die menschliche Gesundheit zu berücksichtigen.

Ist die Methode nur für den Nachweis von Mikroorganismen in einem Produkt (z.B. Trinkwasser) bestimmt, ist diese Matrix bei den Untersuchungen einzusetzen. Handelt es sich um eine horizontale Methode (z.B. Nachweis in allen Lebensmitteln) sind die Untersuchungen jeweils an mindestens vier verschiedenen Produkte- bzw. Lebensmittelkategorien durchzuführen.

Für einen Methodenvergleich sind pro jeweilige Produkte- bzw. Lebensmittelkategorie mindestens 20 natürlich kontaminierte und mindestens 20 natürlich nicht kontaminierte Proben sowohl mit der zu validierenden Alternativ- als auch mit der Referenzmethode zu untersuchen. Ist es nicht möglich, eine genügende Anzahl natürlich kontaminierter Proben aufzutreiben, ist eine künstliche Kontamination von Proben erlaubt. Dabei sind die Proben nach der Kontamination einer physikalischen Behandlung zu unterziehen, die repräsentativ für die natürlichen Bedingungen ist. Das Vorgehen der künstlichen Kontamination ist genau zu beschreiben (siehe auch Annex C der prEN ISO/FDIS 16140:2000(E), (5)).

Bei neuen Methoden sind pro Produkte- bzw. Lebensmittelkategorie mindestens 20 mit verschiedenen Stämmen (wo möglich fünf) des Zielorganismus künstlich kontaminierte und mindestens 20 nicht mit dem Zielorganismus, jedoch mit anderen Spezies kontaminierte Proben zu untersuchen. Dabei sollte die Konzentration zehnmal über der Nachweisgrenze (qualitative Methoden, siehe Nachweisgrenze) respektive der Bestimmungsgrenze (quantitative Methoden, siehe Bestimmungsgrenze) liegen. Die künstlich kontaminierten Proben sollen das Spektrum der (natürlicherweise) vorkommenden Eigenflora abdecken.

Spezifität

Definition: Die Spezifität ist der Grad, zu welchem eine Methode beeinflusst wird von anderen Komponenten der Probe. Eine spezifische Methode wird von keiner anderen Komponente beeinflusst (2).

Bei mikrobiologischen Prüfverfahren wird das, was unter dem Begriff «Spezifität» zu verstehen ist, meistens durch die Referenzmethode bestimmt. So werden zum Beispiel «Enterobacteriaceen» durch das zu verwendende Selektivmedium definiert (SLMB Kap. 56 E2 (7)), was zu einer nicht unerwünschten Normierung führt. Dies muss bei der Validierung von alternativen Methoden wie PCR-basierenden Nachweisverfahren berücksichtigt werden. Es stellt sich in diesem Kontext jedoch auch die Frage, was die Spezifität von pathogenen Keimen bestimmt: ist dies der Metabolismus, sind dies Pathogenitätsfaktoren (wie z.B. Staphylococcus-Toxine) oder sind dies Gene für Pathogenitätsfaktoren (wie z.B. die Gene für Shiga-Toxin)?

AOAC schlägt vor, dass das «Erstlabor» (primary laboratory) reine Bakterienkulturen von inklusiven und exklusiven Keimen untersucht (reine Kulturen in nichtselektiven Medien, keine gespikten Proben). Inklusive Keime sind Analyten der Testmethode und müssen detektiert werden, während exklusive Keime nicht detektiert werden dürfen. Bei quantitativen Methoden sollen mindestens fünf «analyte strains» und 10–20 «nonanalyte strains» getestet werden, wovon mindestens fünf nahe verwandte Vertreter derselben Gattung sein sollen. Bei qualitativen Methoden sind mindestens die Anforderungen für quantitative Methoden einzuhalten.

Für mikrobielle Prüfverfahren wird die Spezifität als $d/(c+d) \cdot 100$ berechnet und gibt an, wieviel % aller sicher negativen Proben als negativ erkannt werden (vgl. Tabelle 2).

Bei alternativen Prüfverfahren bezieht sich die Spezifität auf die Anzahl der mit der Referenzmethode als nicht kontaminiert erkannten Proben, bei neuen Prüfverfahren auf die Anzahl aller nicht mit dem Zielorganismus kontaminierten Proben.

Sensitivität

Definition: Bei quantitativen Methoden ist die Sensitivität die Fähigkeit eines Prüfverfahrens, innerhalb einer gegebenen Matrix leichte Änderungen in der Anzahl Mikroorganismen nachzuweisen. Bei qualitativen Methoden wird unter Sensitivität die Fähigkeit der zu validierenden Methode verstanden, den Zielorganismus nachzuweisen, wenn er von der Referenzmethode nachgewiesen wird bzw. in der Probe enthalten ist (2, 3).

Für mikrobielle Prüfverfahren wird die Sensitivität als $a/(a+b) \cdot 100$ berechnet und gibt an, wieviel % aller sicher positiven Proben als positiv erkannt werden (vgl. Tabelle 2).

Bei alternativen Prüfverfahren bezieht sich die Spezifität auf die Anzahl der mit der Referenzmethode als kontaminiert erkannten Proben, bei neuen Prüfverfahren auf die Anzahl aller mit dem Zielorganismus kontaminierten Proben. Bei quantitativen Methoden wird die Sensitivität durch eine Regressionsanalyse (siehe Linearität) zusätzlich überprüft.

Richtigkeit/Relative Richtigkeit (Effizienz)

Allgemeine Definition: Die Richtigkeit ist ein Mass für die Abweichung vom wahren bzw. richtigen Wert. In der Mikrobiologie wird als Mass für die Richtigkeit der Grad der Übereinstimmung der Resultate heranzgezogen, welche mit der zu validierenden Methode und der reellen Kontamination an gleichen Proben erhalten werden. Sinngemäss gibt die relative Richtigkeit den Grad der Übereinstimmung der Resultate an, welche mit der zu validierenden Methode und der Referenzmethode an gleichen Proben erhalten werden.

Für qualitative alternative Prüfverfahren wird die relative Richtigkeit als $(a+d)/n \cdot 100$ berechnet und gibt die Häufigkeit an, mit welcher die beiden Methoden des Methodenvergleichs die gleichen Resultate ergeben. Für qualitative neue Prüfverfahren wird die relative Richtigkeit ebenfalls als $(a+d)/n \cdot 100$ berechnet und gibt die Häufigkeit an, mit welcher die neue, zu validierende Methode die reellen Kontaminationen der Proben nachweist.

Bei quantitativen alternativen Prüfverfahren wird der Mittelwert ($\bar{d}$) aller Differenzen zwischen den Resultaten mit der Referenzmethode ($\bar{x}_R$) und mit der Alternativmethode ($\bar{x}_A$) erhaltenen Resultate (vgl. Tabelle 3) berechnet. Dieser Wert darf nicht signifikant von null verschieden sein. Analog wird für quantitative neue Prüfverfahren der Mittelwert ($\bar{d}$) aller Differenzen zwischen den reellen Kontaminationswerten ($\bar{x}_R$) und mit der neuen Methode ($\bar{x}_A$) erhaltenen Resultate (vgl. Tabelle 3) berechnet. Dieser Wert darf ebenfalls nicht signifikant von null verschieden sein.

Wiederholbarkeit (Präzision)

Definition: Vergleich von Resultaten wiederholter Messungen derselben Probe unter gleichen Bedingungen (gleiche Personen, Laboratorien, Geräte, Reagenzien, Umgebungsbedingungen). Die Wiederholbarkeit (repeatability) wird mit r («klein r») angegeben.

Zur Abschätzung der Wiederholbarkeit ist pro Produkte- bzw. Lebensmittelkategorie (Matrix) mindestens eine Probe mit einem Gehalt des Zielkeimes nahe der Nachweisgrenze (qualitative Methoden) resp. nahe der Bestimmungsgrenze (quantitative Methoden) künstlich zu kontaminieren und fünf Mal unter gleichen Bedingungen mit der zu validierenden Methode zu untersuchen.

Bei qualitativen Prüfverfahren beträgt $r = \frac{x}{n}$

x = Anzahl Messungen mit übereinstimmenden Resultaten unter Wiederholbarkeitsbedingungen

n = Anzahl Messungen

Für quantitative Prüfverfahren gilt $r = 2{,}8 \cdot s_r$, s_r = Standardabweichung der unter Wiederholbarkeitsbedingungen erzielten Messresultate. Der Faktor 2,8 setzt sich zusammen aus: $2 \cdot \sqrt{2}$, 2 kommt von der Normalverteilung (95 % Konfidenzintervall); Wurzel 2 beruht darauf, dass sich r respektive R (siehe Vergleichbarkeit) auf die Differenz von zwei einzelnen Prüfreihen beziehen (SLMB, Kap. 60 (8)). Wenn

die absolute Differenz zwischen zwei Resultaten kleiner r ist, dann können die Resultate als nicht verschieden betrachtet werden.

Idealerweise gilt für die Standardabweichung der Anzahl ausgezählter Kolonien auf einer Platte die Poisson-Verteilung: $\sigma = \sqrt{n}$. In der Praxis sind doppelt so grosse Standardabweichungen jedoch durchaus akzektabel (9). Laboruntersuchungen zeigten des weiteren auf, dass mit 50 % bis 70 % der grösste Anteil an der Gesamtvarianz eines quantitativen mikrobiologischen Nachweisverfahrens auf die Komponente «Probe» fällt (Stichwort: Probeninhomogenität), während die methodenbedingte Varianz bei geübtem Laborpersonal nur 4 % bis 10 % und die aufgrund der Poisson-Verteilung unvermeidliche statistische Plattierungsvarianz ca. 25 % ausmacht (9). Die Wiederholbarkeit ist in der Regel kleiner als die im nachfolgenden aufgeführte Vergleichbarkeit. Anstelle der experimentellen Bestimmung der Wiederholbarkeit kann deshalb bei der Validierung von quantitativen mikrobiologischen Prüfverfahren die bei internationalen mikrobiologischen proficiency tests für die Berechnung des z-score üblicherweise verwendete Standardabweichung von 0,25 $\log_{10}$ herangezogen werden (siehe auch Abb. 1).

Vergleichbarkeit (Robustheit)

Definition: Vergleich von Resultaten wiederholter Messungen derselben Probe unter wechselnden Bedingungen (verschiedene Personen, Orte, Geräte, Reagenzien, Umgebungsbedingungen). Die Standardabweichung der Vergleichbarkeit ist normalerweise 2–3 x grösser als die für die Wiederholbarkeit (quantitative Methoden). Die Vergleichbarkeit (reproducibility) wird mit R («gross R») angegeben.

Zur Abschätzung der Vergleichbarkeit ist pro Produkte- bzw. Lebensmittelkategorie (Matrix) mindestens eine Probe mit einem Gehalt des Zielkeimes nahe der Nachweisgrenze resp. nahe der Bestimmungsgrenze künstlich zu kontaminieren und durch mindestens zwei Personen je 5-mal unter wechselnden Bedingungen mit der zu validierenden Methode zu untersuchen. Es ist dabei wegen der oben erwähnten Probeninhomogenität wichtig, diese Bestimmungen mit der Stammlösung durchzuführen. Für das Austesten der Wiederholbarkeit eignen sich auch «proficiency tests», nach entsprechender Absprache der teilnehmenden Laboratorien.

Bei qualitativen Prüfverfahren beträgt $R = \frac{y}{n}$

y = Anzahl Messungen mit übereinstimmenden Resultaten unter Vergleichbarkeitsbedingungen

n = Anzahl Messungen

Für quantitative Prüfverfahren gilt $R = 2{,}8 \cdot s_R$; s_R = Standardabweichung der unter Vergleichbarkeitsbedingungen erzielten Messresultate. (Zu Faktor 2,8 siehe Wiederholbarkeit.) Wenn die absolute Differenz zwischen zwei Resultaten kleiner R ist, dann können die Resultate als nicht verschieden betrachtet werden. Die Vergleichbarkeit wird bei internationalen proficiency tests auf eine $\log_{10}$-Einheit ange-

setzt. Dies bedeutet, dass 95 % aller Resultate erwartungsgemäss in das Intervall $\log(x) \pm 0{,}5 \log_{10}$ fallen (z-score = 2; siehe auch Abb. 1).

Dabei steht x für den robusten Mittelwert oder für den Median der von den Teilnehmern gemeldeten Resultate oder für den Median des Referenzlabors.

Nachweisgrenze (limit of detection)

Definition: Die Nachweisgrenze ist die kleinste Konzentration an Mikroorganismen, die mit genügender statistischer Sicherheit nachgewiesen werden kann.

Zur Abschätzung der Nachweisgrenze sind pro Produkte- bzw. Lebensmittelkategorie Verdünnungsreihen über mindestens drei verschiedene Konzentrationen (tief: 1 bis 10 KBE, mittel: 10 bis 100 KBE und hoch: grösser als 100 KBE) mit vier verschiedenen Stämmen des Zielorganismus anzusetzen, die Matrices künstlich zu kontaminieren und mit der zu validierenden Methode zu untersuchen. Eine Negativkontrolle ist zusätzlich mitzuführen.

Die Resultate der Untersuchungen der künstlich kontaminierten Proben werden mit der reellen Kontamination verglichen. Die Nachweisgrenze ist der tiefste Wert, bei dem alle Resultate positiv ausfallen. Die Nachweisgrenze kann auch als Bereich angegeben werden, welcher zwischen den beiden Kontaminationsniveaus liegt, bei denen in weniger als 50 % bzw. in mehr als 50 % der Proben der Zielorganismus nachgewiesen werden konnte.

AOAC gibt folgende Werte an: Nachweisgrenze 1 KBE/25 g (qualitativ) und 10–100 KBE/g bei quantitativen Plattierverfahren (4). Gemäss prEN ISO/FDIS 16140:2000(E) wird die Nachweisgrenze bei quantitativen Prüfverfahren als $3{,}3 \cdot s_b$ berechnet; dabei ist s_b die Streuung des y-Achsenabschnitts der Regressionsgerade $y = b + mx$ (5, siehe auch Statistische Übereinstimmung).

Bestimmungsgrenze (limit of quantification)

Definition: Die Bestimmungsgrenze ist die kleinste Konzentration an Mikroorganismen, die noch mit akzeptabler Richtigkeit und Präzision bestimmt werden kann.

Zur Ermittlung der Bestimmungsgrenze sind pro Produkte- bzw. Lebensmittelkategorie Verdünnungsreihen über mindestens drei verschiedene Konzentrationen – wo möglich im Bereich der gesetzlich festgesetzten Toleranz- und Grenzwerte – mit vier verschiedenen Stämmen des Zielorganismus anzusetzen, die Matrices künstlich zu kontaminieren und mit der zu validierenden Methode zu untersuchen. Eine Negativkontrolle ist zusätzlich mitzuführen. Die Resultate der Untersuchungen der künstlich kontaminierten Proben werden mit der reellen Kontamination verglichen. Die Bestimmungsgrenze hängt sehr stark von der Präzision des Nachweisverfahrens ab. Im überarbeiteten Kapitel 56 des Schweizerischen Lebensmittelbuches (7) werden für eine statistisch gesicherte numerische Auswertung jedes beschriebenen mikrobiologischen Nachweisverfahrens Angaben zur Gesamtbelegungsdichte und zur spezifischen Belegungsdichte gemacht. Eine numerische

Resultatsangabe erfolgt in der Regel nur bei Auszählung von mindestens 10 Kolonien (Ausnahme unverdünntes Untersuchungsgut wie Trinkwasser oder Milch.)

Gemäss AOAC liegt die Bestimmungsgrenze bei 2000–3000 KBE/g bei quantitativen Plattierverfahren. Gemäss prEN ISO/FDIS 16140:2000(E) wird für die Bestimmungsgrenze ein Variationskoeffizient kleiner 10% vorgeschlagen. Dies bedeutet, dass im Idealfall mindestens 100 Kolonien pro Platte ausgezählt werden müssten (vgl. mit Wiederholbarkeit). Die Bestimmungsgrenze kommt dadurch ebenfalls in den von der AOAC angegebenen Bereich (grösser 1000 KBE/g bei verdünntem Untersuchungsgut) zu liegen. Andere wiederum empfehlen, nur Platten mit mindestens fünf Kolonien auszuwerten (8). Es erscheint uns wichtig, in diesem Zusammenhang darauf hinzuweisen, dass gemäss SLMB Kapitel 56 D «in einem gesetzlich festgelegten Höchstwert für Mikroorganismen die verfahrensbedingte Streuung bereits mitberücksichtigt ist, so dass jedes über einem Höchstwert liegende Resultat immer eine Überschreitung desselben bedeutet» (7).

Linearität

Definition: Fähigkeit der Methode, in einer gegebenen Matrix Resultate zu geben, die proportional zur Anzahl Mikroorganismen in der Probe sind.

Die Linearität soll im Bereich der gesetzlich festgesetzten Toleranz- und Grenzwerte mit fünf verschiedenen Konzentrationen ausgetestet werden.

Die Resultate der Untersuchungen der künstlich kontaminierten Proben (KBE/g bzw. ml) und die Werte der reellen Kontamination (KBE/g bzw. ml) werden graphisch dargestellt und einer linearen Regressionsanalyse unterzogen. Die Steigung der Regressionsgeraden darf nicht signifikant von eins verschieden sein. Der y-Achsenabschnitt sollte nicht signifikant von null verschieden sein (siehe auch Statistische Übereinstimmung).

Falsch-Positiv-Rate

Definition: Anzahl Resultate mit positiver Abweichung: Anzahl falsch positiver durch Anzahl negativer Proben beim Referenzverfahren bzw. durch Anzahl nicht mit dem Zielorganismus künstlich kontaminierter Proben.

Die Falsch-Positiv-Rate wird für qualitative Prüfverfahren als $c/c+d \cdot 100$ (vgl. Tabelle 2) berechnet und gibt an, wieviel % der Proben mit dem alternativen bzw. neuen Prüfverfahren als falsch positive Befunde gewertet wurden. Falsch positive Befunde müssen unbedingt als solche durch weitere Charakterisierung der Keime bestätigt werden.

Falsch-Negativ-Rate

Definition: Anzahl Resultate mit negativer Abweichung: Anzahl falsch negativer durch Anzahl positiver Proben beim Referenzverfahren bzw. durch Anzahl mit dem Zielorganismus künstlich kontaminierter Proben.

Die Falsch-Negativ-Rate wird für qualitative Prüfverfahren als $b/a+b \cdot 100$ (vgl. Tabelle 2) berechnet und gibt an, wieviel % der Proben mit dem alternativen bzw. neuen Prüfverfahren als falsch negative Befunde gewertet wurden.

Statistische Übereinstimmung

Für den Methodenvergleich von qualitativen Methoden bieten sich zur statistischen Auswertung sogenannte «nichtparametrische» Testverfahren wie der Vierfeldertest (z.B.: χ^2-Test nach McNemar) oder die Bestimmung des Konkordanzindex Kappa (10). Da für die Anwendung des χ^2-Test nach McNemar die Summe der abweichenden Resultate (b+c, vgl. Tabelle 2) grösser als acht sein muss (10), was insbesondere beim Nachweis von pathogenen Bakterien wie *Listeria monocytogenes* (siehe auch Beispiel «Validierung eines alternativen qualitativen Nachweisverfahrens») nur durch eine hohe Anzahl Untersuchungsproben und/oder durch klare Unterschiede in der Sensitivität der beiden Testverfahren erfüllt werden kann, ziehen wir die Berechnung des Konkordanzindex Kappa für die «in house»-Validierung von mikrobiologischen Prüfverfahren als praktikable Alternative dem χ^2-Test nach McNemar vor.

Der Konkordanzindex Kappa zeigt das Mass der Übereinstimmung zweier Prüfverfahren bezüglich eines Analysenparameters an und wird wie folgt berechnet:

$\text{Kappa} = 2\ (ad-bc)/((a+c)(c+d)+(a+b)(b+d))$ (Bedeutung von a, b, c, d siehe Tabelle 2).

Die Bewertung erfolgt als Übereinstimmungsstufen je nach Wert von Kappa (Tabelle 4).

Tabelle 4
Bewertung des Konkordanzindex Kappa (nach Sachs, 1997 (10))

Kappa	*Übereinstimmung*
<0,10	keine
0,10–0,40	schwache
0,41–0,60	deutliche
0,61–0,80	starke
0,81–1,00	fast vollständige

Die Übereinstimmung zweier Prüfverfahren ist als genügend zu bezeichnen, wenn Kappa grösser gleich 0,81 ist.

Methodenvergleiche von quantitativen Methoden werden mittels parametrischer Testverfahren statistisch ausgewertet. Dabei handelt es sich um Tests für den

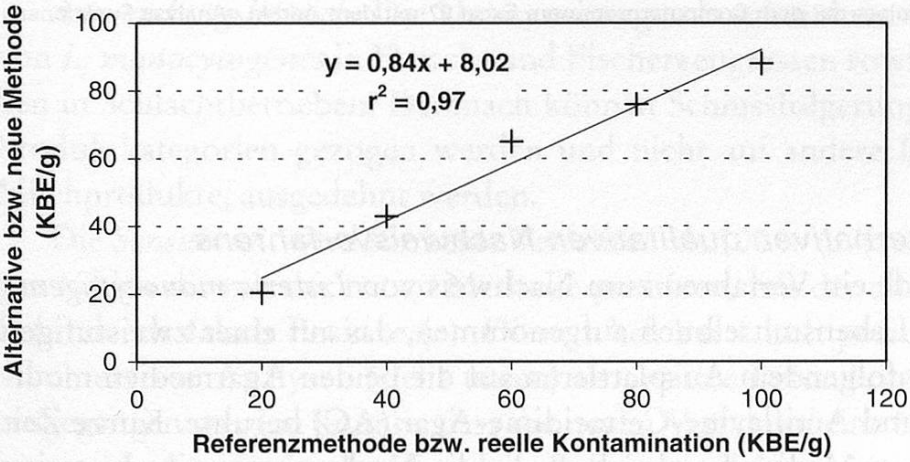
Vergleich zweier abhängiger Stichproben bzw. für den Vergleich gepaarter Beobachtungen (z.B.: z-Test, t-Test). Akzeptanzkriterium ist Signifikanz mit Irrtumswahrscheinlichkeit $p = 5\,\%$.

Wenn Keimzahlen grösser als 100 KBE/g gemessen werden, sind vor der statistischen Auswertung die Keimzahlen zu logarithmieren. Fällt die statistische Auswertung für die Keimzahlen besser aus als die für die logarithmierten Keimzahlen, ist erstere zu nehmen. In der Regel fällt die Auswertung der logarithmierten Keimzahlen besser aus bei Vorliegen hoher Keimzahlen, schlechter bei Vorliegen vieler Keimzahlen nahe Null.

Ergibt die zu validierende Methode die gleichen Resultate wie die Referenzmethode oder entspricht sie den reellen Kontaminationen, beträgt der Mittelwert der Differenzen der beiden Messresultate ($\bar{d}$) null (vgl. Tabelle 3). Der berechnete Mittelwert dieser Differenzen wird überprüft durch Berechnung des 95 % Vertrauensintervalls und dessen Vergleich mit dem theoretischen Wert Null:

Wenn $|\bar{d}| < \frac{t_{krit} \cdot s_d}{\sqrt{n}}$, dann besteht kein signifikanter Unterschied zwischen den beiden Messreihen. Dabei bezeichnet t_{krit} den kritischen Wert aus Student's Test Tabelle für df_{n-1} (zweiseitig; 95 % Vertrauensintervall; $t_{dfn-1;\,0,05;\,zweiseitig}$), s_d die Streuung der gemessenen Differenzen und n die Anzahl Wertepaare.

Zur Bestimmung der Korrelation zweier Methoden sind die durch beide Untersuchungsmethoden ermittelten Keimzahlen einer linearen Regressionsanalyse zu unterziehen (Abb. 2). Die graphische Darstellung der Mittelwerte der durch die Referenz- und Alternativmethode gewonnenen Resultate von jeder Probe, wobei die x-Achse für die Referenz- und die y-Achse für die Alternativmethode zu verwenden ist, wird auf Ausreisser überprüft. Neben der graphischen Überprüfung bietet sich als Ausreissertest der Dixon-Test an (8, 10). Wenn Ausreisser gefunden



x	y
20	20,4
40	43
60	65
80	76
100	88

Abbildung 2 **Regressionsanalyse (hypothetisches Beispiel)**

werden, ist die entsprechende Probe, sofern möglich, wiederholt zu untersuchen, ansonsten ist das Resultat zu verwerfen. Erbringt auch die Wiederholung ein Ausreisser-Resultat, ist dieses ebenso zu verwerfen.

Der Regressionskoeffizient (r) beträgt bei vollständiger Korrelation 1, bei nicht vorhandener Korrelation 0. Ist die beobachtete Korrelation signifikant von Null verschieden, dann ist sie nicht zufälliger Natur.

Beide Methoden sind äquivalent, wenn die berechnete Regressiongleichung nicht signifikant von der theoretischen Gerade «y=x» abweicht. Die Steigung der Regressionsgeraden beträgt bei vollständiger Übereinstimmung eins. Die berechnete Steigung der Regressionsgeraden (siehe Abb. 2) wird überprüft durch Berechnung des 95 % Vertrauensintervalls für die Geradensteigung und dessen Vergleich mit dem theoretischen Wert eins:

Wenn $|m-1|<t_{krit}\cdot s_m$, dann ist die Steigung m der Regressionsgerade statistisch nicht von eins verschieden. Dabei bezeichnet t_{krit} den kritischen Wert aus Student's Test Tabelle für df_{n-2} (zweiseitig; 95 % Vertrauensintervall; $t_{dfn-2;\ 0,05;\ zweiseitig}$), s_m die Streuung der Steigung der Regressionsgeraden und n die Anzahl Wertepaare.

Falls die Steigung m signifikant von 1 verschieden ist, sind die beiden Methoden nicht äquivalent. Die statistische Auswertung des Beispiels aus Abbildung 2 findet sich in Tabelle 5.

Tabelle 5
Statistische Auswertung der Regressionsgerade aus Abbildung 2

	Koeffizienten	*Streuung*	*t-Wert*[a]	*P-Wert*[b]	*Untere 95 %*[c]	*Obere 95 %*[d]
y-Achsenabschnitt (b)	8,02	5,47	1,47	0,24	–9,4	25,4
Geradensteigung m	0,84	0,08	10,2	0,002	0,58	1,10

[a] = der t-Wert wird aus dem Wert des Koeffizienten geteilt durch die Streuung berechnet
[b] = der P-Wert gibt an, mit welcher Wahrscheinlichkeit der Koeffizient von Null verschieden ist
[c] = der Wert «Untere 95%» wird aus dem Wert des Koeffizienten minus t_{krit} mal die Streuung berechnet
[d] = der Wert «Obere 95%» wird aus dem Wert des Koeffizienten plus t_{krit} mal die Streuung berechnet
Die statistische Auswertung erfolgte mit dem Computerprogramm Excel 97 mit dem Add-In «Analyse Funktionen»

Praktische Beispiele

Validierung eines alternativen qualitativen Nachweisverfahrens

Im Jahre 1988 wurde ein Verfahren zum Nachweis von *Listeria monocytogenes* in das Schweizerische Lebensmittelbuch aufgenommen, das auf einer zweistufigen Anreicherung mit nachfolgendem Ausplattieren auf die beiden Agarmedien modified McBride (mMB) und Acriflavine-Ceftacidime-Agar (AC) beruhte. Kurze Zeit später wurden zwei neue Medien beschrieben, die die Nachweisrate von Listerien deutlich steigern sollen. Es handelt sich dabei um den Oxford- und den Palcam-Agar. Um das neue Verfahren (Oxford und Palcam) mit der Referenzmethode

(mMB und AC) zu vergleichen, wurde ein Methodenvergleich anhand natürlich kontaminierter Proben durchgeführt. Dies war im vorliegenden Beispiel möglich, da die Prävalenz über 10 % betrug. Eine Validierung ist nur möglich, wenn die gleiche Einwaage für die Alternativmethode wie für die Referenzmethode verwendet werden kann oder wenn die Voraussetzung erfüllt ist, dass eine homogene Verteilung des Keimes im Probengut vorliegt. In der Regel ist jedoch von einer inhomogenen Verteilung der Zielkeime auszugehen!

Insgesamt wurden 362 Proben auf das Vorkommen von *L. monocytogenes* untersucht. Dabei handelte es sich um 211 Fleischerzeugnisse (meist Rohwurstwaren, Rohpökelwaren und Kochwursterzeugnisse), 100 Fischerzeugnisse (geräucherte Fische, Gravad Lachs usw.) und 51 Proben aus Umgebungsuntersuchungen in Schlachtbetrieben. In 68 Proben konnten Listerien gefunden werden, wovon 18 mal *L. monocytogenes* (Tabelle 6).

Tabelle 6
Nachweishäufigkeiten von *Listeria spp.* (A) und *L. monocytogenes* (B) im Methodenvergleich

A		*zu validierende Methode*		
		+	*–*	*Summe*
Referenz-	+	52	0	52
methode	–	16	294	310
	Summe	68	294	362

B		*zu validierende Methode*		
		+	*–*	*Summe*
Referenz-	+	14	0	14
methode	–	4	344	348
	Summe	18	344	362

Anwendungsbereich: Die Untersuchungen beschränkten sich auf den Nachweis von *L. monocytogenes* in Fleisch- und Fischerzeugnissen sowie auf Umgebungsproben in Schlachtbetrieben. Demnach können Schlussfolgerungen auch nur für diese Produktkategorien gezogen werden und nicht auf andere Lebensmittel, wie z.B. Milchprodukte, ausgedehnt werden.

Die *Sensitivität* wird für beide Verfahren als 100 % (a/(a+b)) berechnet. Hingegen zählen die mit der neuen Methode positiven, mit der alten Methode negativen Befunde als falsch Positive (c). Diese scheinbar falsch positiven Befunde wurden durch externe Analysen (siehe unten) als positive Befunde bestätigt, so dass davon ausgegangen werden konnte, dass die neuen Verfahren eine höhere Sensitivität aufweisen als die entsprechenden alten Referenzverfahren. Um zu valablen Aussagen bezüglich der Validierungsparameter zu gelangen, wurde die Tabelle 6 transformiert, indem die neuen Verfahren als Standard gesetzt wurden (Tabelle 7).

Tabelle 7
Nachweishäufigkeiten von *Listeria spp.* (A) und *L. monocytogenes* (B) im Methodenvergleich

A		*altes Verfahren*		
		+	*–*	*Summe*
neues Verfahren	+	52	16	68
	–	0	294	294
	Summe	52	310	362

B		*altes Verfahren*		
		+	*–*	*Summe*
neues Verfahren	+	14	4	18
	–	0	344	344
	Summe	14	348	362

Damit ergeben sich für die alten Verfahren Sensitivitätswerte von 76,5 % für *Listerien spp.* resp. von 77,8 % für *Listeria monocytogenes*. Ausschlaggebend für die Umkehrung der Tabelle war die Bestätigung der vermeintlich falsch positiven Befunde gemäss Tabelle 6.

Die *Spezifität* wird mit d/(c+d) als jeweils 100 % für die alten Verfahren berechnet. Dies drückt aus, dass die alten Verfahren keine falsch positiven Befunde ergaben. Die Spezifität ist in der Regel jedoch schwierig zu bestimmen, da in den natürlich kontaminierten Proben meistens eine inhomogene Verteilung vorliegt. Hingegen ist die Selektivität ein wichtiges Kriterium. Auf dem mMB, dem AC-Agar und dem Oxford-Agar wurde in der Mehrzahl der Fälle starkes Wachstum der Begleitflora festgestellt (70 % und mehr). Beim Palcam-Medium waren es aber nur 15,6 % (11).

Die *relative Richtigkeit* beträgt 95,6 % für Listerien und 98,9 % für *L. monocytogenes*. Wir müssen uns aber hier im Klaren sein, dass dies relativ zu sehen ist, da es sich bei den im Vergleich zur alten Referenzmethode zusätzlichen positiven Befunden um bestätigte Listeriennachweise handelt. Der Begriff Richtigkeit steht somit immer in Relation zum definierten Standard und gibt im Wesentlichen nur an, bei wievielen Messungen die alten und neuen Verfahren zum gleichen Ergebnis führten.

Die theoretische *Nachweisgrenze* beträgt ein *L. monocytogenes*-Bakterium pro eingewogene Menge. Die praktische Nachweisgrenze hängt von verschiedenen Faktoren ab und kann in der Praxis kaum bestimmt werden. Dies ist darauf zurückzuführen, dass lebende Materie gemessen werden soll, die sich dynamisch verhält (Vermehrung, Absterben, Schädigung, Inaktivierung). Eine wissenschaftliche Studie zeigte jedoch auf, dass die theoretische recht gut mit der praktischen Nachweisgrenze übereinstimmt (12). Dabei wurden Proben von künstlich kontaminierten Rohwürsten während verschiedenen Phasen der Herstellung und Lagerung ent-

nommen, bis unter die theoretische Nachweisgrenze dezimal verdünnt und mittels MPN-Ansatz auf Listerien untersucht.

Die *Falsch-Positiv-Rate* beträgt für die neuen Verfahren gemäss Tabelle 6 5,2 % für *Listeria spp.* resp. 1,5 % für *Listeria monocytogenes*. Diese Zahlen sind jedoch irreführend. Vielmehr interessiert uns, wie gross die falsch positive Rate des gesamten Verfahrens inklusive biochemischer Differenzierung ist. Über einen Zeitraum von sechs Jahren sandten wir insgesamt 413 von uns als *L. monocytogenes* erkannte Stämme an das «Centre national de référence pour les listérias, CHUV, 1011 Lausanne» zur Serotypisierung ein. 411 Befunde wurden bestätigt, zwei Isolate (aus dem Jahr 1992) wurden als *L. seeligeri* mit atypischem biochemischen Verhalten (Rhamnose positiv) erkannt. Anteil falsch positiver Befunde: 2/413 = 0,5 %. Die Falsch-Positiv-Rate beträgt gemäss Tabelle 7 für die alten Verfahren 0 %.

Die *Falsch-Negativ-Rate* beträgt für die alten Verfahren 23,5 % resp. 22,2 %. Diese hohe Raten führen noch einmal vor Augen, dass die neuen Nachweisverfahren sensitiver sind als die alten Referenzverfahren.

Statistische Übereinstimmung: Der Konkordanzindex Kappa beträgt 0,84 für Listerien und 0,87 für *L. monocytogenes*. Somit besteht eine fast vollständige Übereinstimmung der alten und neuen Nachweisverfahren.

Beispiel einer kleinen Änderung eines genormten quantitativen Prüfverfahren

Mit der Revision der Schweizerischen Hygieneverordnung vom 11. Juli 2000 ging die Einführung neuer mikrobiologischer Prüfverfahren im SLMB Kapitel 56 (7) einher. Für den quantitativen Nachweis von *Escherichia coli* wird in SLMB Kapitel 56 E3 (7) als Medium Tryptone Bile Glucuronid Agar (TBX) vorgeschrieben. Gemäs Herstellerangaben sollte TBX-Medium nach der Einwaage autoklaviert und im Wasserbad auf 47° C abgekühlt werden bis zum Gebrauch. Es stellte sich – nicht zuletzt aus Kostengründen – die Frage, ob angebrochenes, nicht benötigtes TBX-Medium nach dem Gebrauch in der Flasche abgekühlt und bis zu 10 Tagen bei Raumtemperatur aufbewahrt werden kann. Für den erneuten Gebrauch wurde das Medium in einem Mikrowellenherd aufgekocht und ebenfalls bei 47° C im Wasserbad bereitgehalten. Für den Vergleich wurden jeweils *E. coli* positive Proben vom Vortag verwendet. Die Resultate dieser Untersuchung sind in Tabelle 8 zusammengestellt.

Tabelle 8
Vergleich verschiedener Herstellung von TBX-Medium für den quantitativen Nachweis von *E. coli*

	KBE/g		logarithmisierte Werte	
Probenbezeichnung	*TBX alt*[a]	*TBX neu*[b]	*TBX alt*	*TBX neu*
Lachs	14400	15300	4,16	4,18
Weichkäse	24000	24000	4,38	4,38
Crèmecornet	550	430	2,74	2,63
Past. Rahm	150	220	2,18	2,34
Weichkäse	1690	1000	3,23	3,00
Ringversuch	850	890	2,93	2,95
Butter	8	11	0,90	1,04
Butter	11	17	1,04	1,23
Salat	170	160	2,23	2,20
Vitello Tonnato	3500	4900	3,54	3,69
Meerrettichschaum	140	90	2,15	1,95
Past. Rahm	100	100	2,00	2,00
Guacamole	2600	2320	3,41	3,37
Rettich	660	630	2,82	2,80

[a] = TBX alt bezeichnet bis zu 10 Tage altes TBX-Medium, welches vor dem Gebrauch wieder aufgekocht wurde
[b] = TBX neu bezeichnet frisch gemäss Herstellerangaben hergestelltes TBX-Medium (Oxoid, CM945)

Da Werte über 100 KBE/g erhalten wurden, mussten die Daten für die Regressionsanalyse logarithmiert werden. Die Regressionanalyse ist in Abbildung 3 dargestellt.

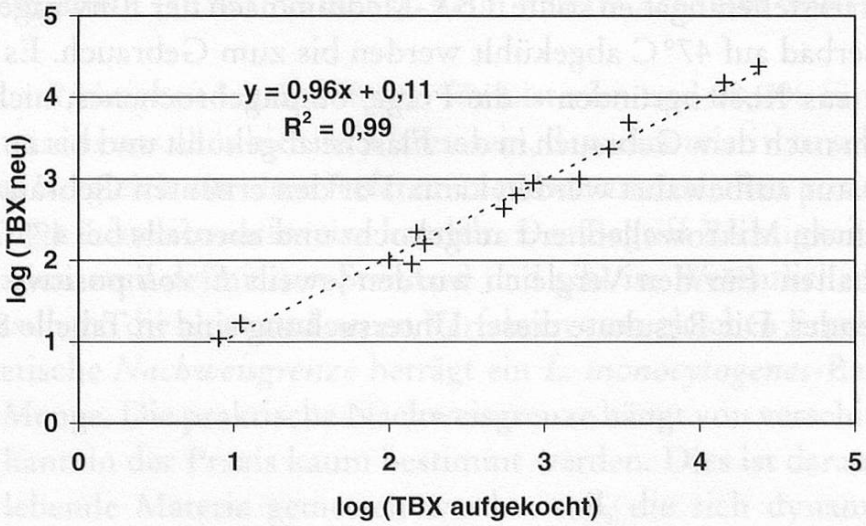


Abbildung 3 **Regressionsanalyse für Vergleich der TBX-Herstellung**

Die statistische Auswertung der Regressionskoeffizienten (siehe Tabelle 9) ergab, dass die Steigung nicht signifikant verschieden von eins und der y-Achsenabschnitt nicht signifikant verschieden von null ist. Somit ist die vorgenommene Änderung des genormten mikrobiologischen Prüfverfahrens nicht substanziell, d.h. ohne Relevanz für die Beurteilung der Resultate.

Tabelle 9
Statistische Auswertung der Regressionsgeraden aus Abbildung 3

	Koeffizienten	*Streuung*	*t-Wert*[a]	*P-Wert*[b]	*Untere 95 %*[c]	*Obere 95 %*[d]
y-Achsenabschnitt (b)	0,107	0,097	1,108	0,290	–0,104	0,318
Geradensteigung *m*	0,962	0,034	28,512	0,000	0,888	1,035

[a] = der t-Wert wird aus dem Wert des Koeffizienten geteilt durch die Streuung berechnet
[b] = der P-Wert gibt an, mit welcher Wahrscheinlichkeit der Koeffizient von Null verschieden ist
[c] = der Wert «Untere 95%» wird aus dem Wert des Koeffizienten minus t_{krit} mal die Streuung berechnet
[d] = der Wert «Obere 95%» wird aus dem Wert des Koeffizienten plus t_{krit} mal die Streuung berechnet
Die statistische Auswertung erfolgte mit dem Computerprogramm Excel 97 mit dem Add-In «Analyse Funktionen»

Validierung eines quantitativen alternativen Nachweisverfahrens mittels Vierfeldertest

Bei diesem Beispiel wurde getestet, ob der Einsatz eines Spiralplater zu den gleichen lebensmittelrechtlichen Beurteilungen führt wie die im SLMB beschriebene Gussplattentechnik. Es wurden zwei Untersuchungsreihen durchgeführt, eine mit

A

		neues Verfahren +	neues Verfahren -	Summe
altes Verfahren	+	57	3	60
	-	0	16	16
	Summe	57	19	76

Spezifität:	1,00
Sensitivität:	0,95
Richtigkeit:	0,96
Falsch Positive	0,00
Falsch Negative	0,05
Konkordanzindex Kappa:	0,89

B

		neues Verfahren +	neues Verfahren -	Summe
altes Verfahren	+	44	1	45
	-	0	21	21
	Summe	44	22	66

Spezifität:	1,00
Sensitivität:	0,98
Richtigkeit:	0,98
Falsch Positive	0,00
Falsch Negative	0,02
Konkordanzindex Kappa:	0,97

Abbildung 4 **Vergleich zweier Ausplattierverfahren**
A. vorerhitzte Lebensmittel, B. Wurstwaren
+ bedeutet Toleranzwertüberschreitung, – bedeutet keine Toleranzwertüberschreitung, altes Verfahren ist in diesem Fall die Methode nach SLMB (Gussplattentechnik), die neue Methode ist Ausplattieren mit Spiralplater

vorerhitzten Lebensmitteln und eine mit Kochpökel-, Brühwurst- und Kochwurstwaren. Für die statistische Beurteilung wurde jeweils nur unterschieden, ob in der Hygieneverordnung festgelegte Toleranzwerte überschritten wurden oder nicht. Damit reduzierte sich die Auswertung eines quantitativen Prüfverfahrens auf den Vierfeldertest und die Berechnung des Konkordanzindex Kappa (Abb. 4).

Die statistische Auswertung ergab Kappa-Werte von 0,89 und 0,97. Somit ergibt das Ausplattieren mit dem Spiralplater fast vollständige Übereinstimmung (Kappa>0,81) mit der im SLMB beschriebenen Gussplattentechnik.

Experimentelle Bestimmung der Wiederholbarkeit

Die Wiederholbarkeit des quantitativen Nachweises von Enterobacteriaceen mittels VRBD-Agar wurde wie folgt bestimmt: von fünf enterobacteriaceenpositi-

Tabelle 10
Wiederholbarkeit des Nachweises von Enterobacteriaceen

Probe	*Resultate*	*log*	*Mittelwert*	*Streuung*	*r*[a]	*Untere 95 %*[b]	*Obere 95 %*[c]
A1	2,90E+05	5,46					
A2	3,60E+05	5,56					
A3	3,70E+05	5,57					
A4	3,11E+05	5,49					
A5	3,50E+05	5,54	5,52	0,045	0,13	5,40	5,65
B1	5,30E+06	6,72					
B2	8,00E+06	6,90					
B3	6,70E+06	6,83					
B4	6,00E+06	6,78					
B5	5,80E+06	6,76	6,80	0,069	0,19	6,61	6,99
C1	8,60E+05	5,93					
C2	8,20E+05	5,91					
C3	1,02E+06	6,01					
C4	8,50E+05	5,93					
C5	6,90E+05	5,84	5,93	0,061	0,17	5,76	6,09
D1	3,40E+04	4,53					
D2	2,30E+04	4,36					
D3	8,10E+04	4,91					
D4	7,30E+04	4,86					
D5	4,30E+04	4,63	4,66	0,229	0,64	4,02	5,30
E1	1,16E+05	5,06					
E2	1,82E+05	5,26					
E3	1,35E+05	5,13					
E4	1,80E+05	5,26					
E5	1,49E+05	5,17	5,18	0,083	0,23	4,94	5,41

[a] = der r-Wert wird als 2,8 · s_r berechnet, wobei s_r die Streuung der (logarithmierten) Messungen unter Wiederholbarkeitsbedingungen darstellt

[b] = der Wert «Untere 95%» wird aus dem Mittelwert minus r-Wert berechnet

[c] = der Wert «Obere 95%» wird aus dem Mittelwert plus r-Wert berechnet

ven Proben wurde je eine Einwaage gemacht und eine Stammlösung hergestellt. Ausgehend von dieser Stammlösung wurden parallel pro Probe je fünf Verdünnungsreihen angesetzt und die Platten mit VBRD-Agar gemäss den Vorschriften des SLMB inokuliert und inkubiert. Inhomogenitäten aufgrund der Probenahme und der Einwaage wurden auf diese Weise vermieden. Die Resultate dieser Untersuchung sind in Tabelle 10 zusammengestellt. Die Streuung der logarithmierten Resultate beträgt im Schnitt 0,045 bis 0,083, bei der Probe D wurde hingegen eine markant höhere Streuung gemessen. Die Berechnung des *r*-Wertes erlaubt die Abschätzung des 95 % Konfidenzintervalls. Dieses beträgt mit Ausnahme von Probe D 0,26 bis 0,46 $\log_{10}$-Einheiten. Wie einleitend erwähnt beträgt das 95 % Konfidenzintervall von «proficiency tests», welches als Mass für die Robustheit (*R*, siehe auch Vergleichbarkeit) angesehen werden kann, ca. eine $\log_{10}$-Einheit (vgl. Abb. 1). Diese Abschätzung zeigt, dass die in unserem Versuch gemessene Wiederholbarkeit (*r*) ungefähr zwei- bis viermal kleiner ist als *R* und durchaus im Rahmen dessen liegt, was erwartet werden kann.

Dank

Den Teilnehmerinnen und Teilnehmern des SGLH-Workshops vom 1. November 2000 in Spiez danken wir für die angeregten Diskussionen und für das Beisteuern mehrerer Beispiele. Insbesondere möchten wir uns bei *Renée Rieker* vom Kantonalen Laboratorium Zürich für die zur Verfügung gestellten numerischen Beispiele sowie bei *Christoph Müller* vom Kantonalen Laboratorium Aargau für die Anregung bedanken, quantitative Methoden mittels Vierfeldertest auf lebensmittelrechtliche Aspekte zu überprüfen.

Zusammenfassung

Das Vorgehen bei der In house (single laboratory)-Validierung von mikrobiologischen Prüfverfahren wird ausführlich beschrieben. Dabei muss einerseits zwischen qualitativen und quantitativen, andererseits zwischen alternativen und neuen Methoden unterschieden werden. Anhand von ausgewählten praktischen Beispielen wird auf einzelne Punkte und Probleme eingegangen. Eine grosse Vereinfachung der Validierung von quantitativen alternativen Nachweisverfahren wird durch die Anwendung des (qualitativen) Vierfeldertests und die Berechnung des Konkordanzindex Kappa erzielt. Die experimentelle Bestimmung der Präzision eines quantitativen Nachweisverfahren ergab 95 % Vertrauensintervalle in der Grössenordnung von 0,3 bis 0,5 $\log_{10}$-Einheiten. Bei internationalen mikrobiologischen proficiency tests werden üblicherweise 95 % Vertrauensintervalle von einer $\log_{10}$-Einheit für die Berechnung des z-score zugrunde gelegt.

Résumé

Ce papier décrit en détail la procédure pour la validation «in-house» (single laboratory validation) des méthodes en microbiologie. Les méthodes sont d'abord classifiées d'une part en qualitative ou quantitative, et d'autre part en alternative ou neuve. Des exemples tirés de la pratique illustrent quelques points spécifiques et les problèmes liés à la validation. Une grande simplification de la validation des méthodes quantitatives est possible par le test appliqué sur une table de contigence avec calcul de l'indice de concordance kappa. La détermination expérimentale de la précision d'une méthode quantitative a indiqué que l'intervalle de confiance 95 % est de l'ordre de 0,3 à 0,5 $\log_{10}$. Ces valeurs sont proches de la valeur d'un $\log_{10}$, utilisés pour le calcul du «z-score» pour les participants des «proficiency tests».

Summary "In House Validation (Single Laboratory Validation) of Microbiological Methods"

This paper describes the complete procedure for the in house (single laboratory) validation of microbiological methods. For this purpose the methods have to be classified on one hand in qualitative and quantitative, on the other hand in alternative and new. Some specific points and problems are illustrated and discussed on practical examples. A large simplification of the validation of quantitative methods can be achieved by applying 2x2 contingency tables and calculation of the concordance index kappa. The experimentally determination of the 95 % confidence intervall of a quantitative method yielded values in the order of 0.3 to 0.5 $\log_{10}$-units. This is close to values used in international microbiological proficiency testing where z-score calculation is often based on a 95 % confidence intervall of one $\log_{10}$-unit.

Key words

Single laboratory validation, Microbiological testing, Measuring uncertainty, Food microbiology, Proficiency testing

Literatur

1 EAL-G18 Richtlinie Accreditation for Laboratories Performing Microbiological Testing, Mai 1996.
2 MICROVAL-WG 3 Dokument Technical rules for validation criteria. Oktober 1995
3 AFNOR Richtlinie Validation AFNOR des méthodes rapides d'analyses. April 1994
4 AOAC International Qualitative and Quantitative Microbiology Guidelines for Methods Validation. JAOAC **82**, 402–415 (1999).
5 Microbiology of food and animal feeding stuffs – Protocol for the validation of alternative method, prEN ISO/FDIS 16140:2000(E).
6 Hygieneverordnung (HyV) vom 26. Juni 1995 (Stand 11.7.2000)
7 Mikrobiologie. SLMB Kapitel 56, Stand 2000
8 Statistik und Ringversuche. SLMB, Kapitel 60, Stand 1989

9 *Berg, C., Dahms, S., Hildebrandt, G., Klatwchka, S.* and *Weiss, H.:* Microbiological collaborative studies for quality control in food laboratories: Reference material and evaluation of analyst's errors. Int. J. Food Microbiol. **24**, 41–52 (1994).
10 *Sachs, L.:* Angewandte Statistik 8. Auflage. Springer Verlag, Berlin 1997.
11 *Jemmi, T., Gobat, P.F.* und *Guyer, S.:* Praktische Erfahrungen mit den Selektivmedien Oxford und Palcam zum Nachweis von *Listeria monocytogenes* aus kontaminierten Materialien. Mitt. Gebiete Lebensm. Hyg. **81**, 559–564 (1990).
12 *Trüssel, M.:* Listerien in Rohwurst- und Rohpökelwaren. Diss. med. vet. Universität Bern 1989.

Korrespondenzadresse: PD Dr. Philipp Hübner, Kantonales Laboratorium Zürich, Fehrenstrasse 15, Postfach, CH-8030 Zürich