

Statistisches Protokoll / *Statistical Protocol*

Das Statistische Protokoll beschreibt sämtliche statistische Berechnungen, welche zur Auswertung von Datensätzen bzw. zur Bewertung der analytischen Leistung der Laboratorien bei der muva kempten verwendet werden.

The statistical protocol describes all statistical calculations which are used for the evaluation of datasets and for the assessment of the laboratory's performance.

INHALTSVERZEICHNIS / CONTENTS

	Seite Page	
1. Begriffe und Definitionen	1	1. <i>Terms and Definitions</i>
2. Angewendete Statistik	6	2. <i>Used Statistics</i>
3. Auswertung der sensorischen Ringversuche	9	3. <i>Evaluation of the proficiency testing schemes in sensory</i>
4. Homogenität des Prüfmaterials	10	4. <i>Homogeneity of the testing material</i>
5. Bewertung der analytischen Leistung	13	5. <i>Assessment of the analytical performance</i>
6. Verzeichnis der verwendeten Standards / Normen und Literatur	16	6. <i>Register of used standards / norms and literature</i>

1. BEGRIFFE UND DEFINITIONEN / *TERMS AND DEFINITIONS*

Begriffe	Abk. /abbr.	Formel / formula	Terms
Laboreinzelwert: Einzelwert einer Doppelbestimmung	x_{lab}		Laboratory single value: <i>Single value of a double determination</i>
Labormittelwert [6, 7]: Mittelwert einer Doppelbestimmung	x	$x = \frac{1}{n} \sum_{i=1}^n x_{lab_i}$	Laboratory mean value [6, 7]: <i>Mean of a double determination</i>
Anzahl der Werte eines Datensatzes	n		Number of values of a dataset

Begriffe	Abk./abbr.	Formel / formula	Terms
<p>Mittelwert eines Datensatzes [6, 7]:</p> <p>Mittelwert sämtlicher Labormittelwerte (vor bzw. nach Eliminierung der Ausreißer)</p>	\bar{x}	$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$	<p>Mean of a dataset [6, 7]:</p> <p>Mean of all laboratory mean values (before or after elimination of outliers)</p>
<p>Spannweite [4, 14]:</p> <p>Differenz des größten und des kleinsten Wertes eines Datensatzes</p>	w	$w = x_{\max} - x_{\min}$	<p>Range [4, 14]:</p> <p>Difference of the major and the minor value of the dataset</p>
<p>Median [10]:</p> <p>Ist der Zentralwert, d.h.:</p> <p>falls n ungerade ist</p> <p>falls n gerade ist</p>	M	$M = med\{y_i\}$ $M = y_{\left(\frac{n+1}{2}\right)}$ $M = \frac{1}{2} \left(y_{\frac{n}{2}} + y_{\left(\frac{n}{2}+1\right)} \right)$	<p>Median [10]:</p> <p>Is the central value, that is:</p> <p>if n is odd</p> <p>if n is even</p>
<p>Bester Schätzwert für den wahren Wert (= zugewiesener Wert) [3, 6, 7, 8]:</p> <p>Mittelwert der Teilnehmer ohne Ausreißer, die eine zur Analyse geeignete Methode eingesetzt haben. Zur Festlegung des besten Schätzwertes wird die unten beschriebene Statistik verwendet.</p> <p>Beim <u>Referenzmaterial</u> entspricht der zugewiesene Wert dem Referenzwert.</p>	x_{ref}		<p>Best estimate of the true value (= assigned value) [3, 6, 7, 8]:</p> <p>Mean value of all participants, using an adequate method for analysis. The statistical methods described below are used for determination of the best estimate</p> <p>For <u>reference material</u> the assigned value is the reference value.</p>
<p>Bester Schätzwert für den wahren Wert (= zugewiesener Wert) in relativer Konzentration [3, 6, 7, 8, 15]:</p> <p>z.B.: 1 mg/kg = 0,000001 kg/kg</p>	$x_{\text{ref, rel}}$		<p>Best estimate of the true value (= assigned value) in relative concentration [3, 6, 7, 8]:</p> <p>e.g: 1 mg/kg = 0,000001 kg/kg</p>

Begriffe	Abk. /abbr.	Formel / formula	Terms
<p>Standardabweichung [6, 7]: Drückt die Streuung der Datensätze um den Mittelwert aus.</p>	s	$s = \sqrt{\frac{1}{n-1} \cdot \sum_{i=1}^n (x_i - \bar{x})^2}$	<p>Standard deviation [6, 7]: <i>Expresses the deviation around the mean</i></p>
<p>Standardabweichung des besten Schätzwertes Drückt die Streuung der Werte um den besten Schätzwert aus.</p>	S _{ref}		<p>Standard deviation of the best estimate <i>Expresses the deviation of the values around the best estimate.</i></p>
<p>Vergleichsstandardabweichung [4, 14]: In Ringversuchen zur Validierung von Methoden unter Vergleichsbedingungen ermittelte Standardabweichung. Sie ist unter dem Begriff Präzisionsdaten in national und international standardisierten Methoden angegeben. Sie dient als Zielstandardabweichung zur Berechnung der z-Werte.</p>	s(R)		<p>Reproducibility standard deviation [4, 14]: <i>It is determined under reproducible conditions in collaborative studies for validation of methods. It is mentioned as precision data in national and international standardized methods. It is used as target standard deviation for the calculation of the z-scores.</i></p>
<p>Wiederholstandardabweichung: In Ringversuchen zur Validierung von Methoden unter Wiederholbedingungen ermittelte Standardabweichung. Sie ist unter dem Begriff Präzisionsdaten in national und international standardisierten Methoden angegeben.</p>	s(r)		<p>Repeatability standard deviation: <i>It is determined under repeatable conditions in collaborative studies for validation of methods. It is mentioned as precision data in national and international standardized methods.</i></p>
<p>Relative Standardabweichung nach Horwitz [15]: Berrechnet in Abhängigkeit von der relativen Konzentration des Analyten (bester Schätzwert) X_{ref,rel}-</p>	s(H) _{rel}	$s(H)_{rel} = 2^{(1-0,5 \cdot \log(x_{ref,rel}))}$	<p>Relative standard deviation acc. to Horwitz [15]: <i>Calculated on the basis of the relative concentration of the analyte (best estimate) X_{ref,rel}.</i></p>

Begriffe	Abk. /abbr.	Formel / formula	Terms
<p>Standardabweichung nach Horwitz [15]:</p> <p>Standardabweichung berechnet aus der relativen Standardabweichung und dem besten Schätzwert in der üblichen Konzentration.</p> <p>Sie dient in einigen Ringversuchen als Zielstandardabweichung zur Berechnung der z-Werte.</p>	s(H)	$s(H) = \frac{s(H)_{rel}}{100} \cdot x_{ref}$	<p>Standard deviation acc. to Horwitz [15]:</p> <p>Standard deviation calculated on the basis of the relative standard deviation acc. To Horwitz and the best estimate (in original concentration).</p> <p>It is used as target standard deviation for the calculation of the z-scores in some PT schemes.</p>
<p>Standardabweichung innerhalb den Proben (= analytische Unsicherheit aus dem Homogenitätstest) [4, 14]:</p> <p>Drückt die analytische Streuung der Datensätze um den zugehörigen Mittelwert innerhalb der Proben im Homogenitätstest aus.</p>	S _{wb}		<p>Standard deviation within the bottles (= analytical uncertainty deriving from the homogeneity test) [4, 14]:</p> <p>Expresses the analytical deviation around the mean within the samples of the homogeneity test.</p>
<p>Standardabweichung zwischen den Proben (= Materialunsicherheit aus dem Homogenitätstest) [4, 14]:</p> <p>Drückt die Streuung der Datensätze einer repräsentativen Anzahl von Stichproben um den zugehörigen Mittelwert aus dem Homogenitätstest aus.</p>	S _{bb}		<p>Standard deviation between the bottles (=uncertainty of the material deriving from the homogeneity test) [4, 14]:</p> <p>Expresses the deviation around the mean of a representative number of samples of the homogeneity test.</p>
<p>Standardabweichung der Mittelwerte [4]:</p> <p>Berechnet aus den zwei Teilstichproben des Homogenitätstest nach DIN ISO 13528</p>	S _{xt}		<p>Standard deviation of the means [4]:</p> <p>Calculated of the two sub samples in the homogeneity test of DIN ISO 13528</p>
<p>Homogenitätskriterium [4, 14]:</p> <p>Dient zur Bewertung der Homogenität der Prüfmaterialien</p>	S _{erl}		<p>Criterion for homogeneity [4, 14]:</p> <p>It is used for the assessment of the homogeneity of the testing materials</p>
<p>Robuste Standardabweichung [5]:</p> <p>Ermittelt aus der Q-Methode</p>	S _R		<p>Robust standard deviation [5]:</p> <p>Q-method</p>

Begriffe	Abk. /abbr.	Formel / formula	Terms
<p>Unsicherheit eines Datensatzes:</p> <p>Unsicherheitsintervall (95% Wahrscheinlichkeit) um den Mittelwert eines Datensatzes</p>	MU		<p>Uncertainty of a dataset:</p> <p>Uncertainty interval (95 % probability) around the mean value of a dataset</p>
<p>Erweiterte Materialunsicherheit:</p> <p>Entspricht dem im Homogenitätstest ermittelten Streubereich (95% Wahrscheinlichkeit)</p>	U _{mat}	$U_{mat} = s_{bb} \cdot t$	<p>Extended uncertainty of the testing material:</p> <p>Corresponds to the dispersion range from the homogeneity test (95% probability)</p>
<p>Kritische Differenz:</p> <p>Die kritische Differenz ist unter dem Begriff Präzisionsdaten in national und international standardisierten Methoden angegeben.</p>	CD		<p>Critical difference</p> <p>The critical difference is mentioned as precision data in national and international standardized methods.</p>
<p>Studentfaktor [6,7]:</p> <p>Erweiterungsfaktor zur Bestimmung der erweiterten Unsicherheit (95% Wahrscheinlichkeit) für kleine Datensätze. Er ist von der Anzahl der Datensätze abhängig.</p>	t		<p>Student factor [6,7]:</p> <p>Coverage factor for calculating the extended uncertainty (95% probability) for small datasets. It depends on the number of datasets.</p>
<p>Streubereich [6,7]:</p> <p>Unsicherheitsintervall (95% Wahrscheinlichkeit) um einen Mittelwert</p>	T	$T = s \cdot t$	<p>Range of variation [6,7]:</p> <p>Uncertainty interval (95% probability) around the mean value</p>
<p>Vertrauensbereich [6,7]:</p> <p>Unsicherheitsintervall (95% Wahrscheinlichkeit) um den wahren Wert (Maß für die Zuverlässigkeit des Mittelwertes)</p>	$\Delta \bar{x}$	$\Delta \bar{x} = \frac{s \cdot t}{\sqrt{n}}$	<p>Confidential interval [6,7]:</p> <p>Uncertainty interval (95% probability) around the true value (dimension for the „reliability“ of the mean)</p>
<p>Prüfgröße David-Test zur Normalverteilung [13]:</p> <p>Gegebener tabellarischer Vergleichswert</p>	PG		<p>Test statistic for the check for normal distribution (David-Test) [13]:</p> <p>Given tabulated standard value</p>

2. ANGEWENDETE STATISTIK / USED STATISTICS

<p>Prüfung auf Normalverteilung – Vortest nach David [13]:</p> <p>Voraussetzung für die weiter angewendete Ringversuchsstatistik (Mittelwert, Standardabweichung, Streubereich, Vertrauensbereich) ist die Prüfung der Verteilung der Daten auf Normalverteilung. Sie erfolgt in mehreren Schritten:</p> <ol style="list-style-type: none"> 1. Berechnung der Spannweite w des Datensatzes $w = (x_{max} - x_{min})$	<p>Check for normal distribution – Test David [13]:</p> <p>Condition for all further statistical calculations of the proficiency testing (mean, standard deviation, dispersion range) is the check of the data for a normal distribution. It is performed by the following steps:</p> <ol style="list-style-type: none"> 1. Calculation of the range w of the data $w = (x_{max} - x_{min})$
<ol style="list-style-type: none"> 2. Berechnung der Prüfgröße $PG = w / s$ <ol style="list-style-type: none"> 3. x_{max}: größter Wert des Datensatzes x_{min}: kleinster Wert des Datensatzes s: Standardabweichung des Datensatzes Suchen der Werte für die kritischen Grenzen in der Tabelle von Pearson und Stephens in dem Signifikanzniveau von 95 % 4. Liegt die Prüfgröße PG innerhalb dieser kritischen Grenzen ist die Verteilung normal. 	<ol style="list-style-type: none"> 2. Calculation of the test statistic $PG = w / s$ <ol style="list-style-type: none"> 3. x_{max}: major value of the dataset x_{min}: minor value of the dataset s: standard deviation of the dataset Looking for the values of the critical limits in the table of Pearson and Stephens with the significant level of 95 % 4. Is the test statistic PG between the two critical limits, the distribution is normal

<p>Ausreißer Test (Grubbs) [4, 6, 12]:</p> <p>Für die Ermittlung von Ausreißern innerhalb eines Datensatzes, der sich aus Ergebnissen von mehreren Laboratorien zusammensetzt, wird der Ausreißer Test nach Grubbs herangezogen [6]. Die Ausreißer werden wie folgt bestimmt.</p> <ol style="list-style-type: none"> 1. Berechnung der Puffergröße: 	<p>Outlier test (Grubbs) [4, 6, 12]:</p> <p>For detection of outliers in a dataset, obtaining results of different laboratories, the outlier test according to Grubbs is used [6]. The outliers are determined as follows.</p> <ol style="list-style-type: none"> 1. Calculation of the difference of the buffer size:
$\text{Puffergröße / buffersize} = \frac{x - \bar{x}}{s}$	<p>x = Labormittelwert / laboratory mean</p> <p>\bar{x} = Gesamtmittelwert / total mean</p> <p>s = Standardabweichung / standard deviation</p>
<ol style="list-style-type: none"> 2. Vergleich der Puffergröße mit den Grubbs-Werten (Abhängig von der Anzahl der Teilnehmer) <p><u>Ausreißer liegt vor</u>, wenn die Puffergröße größer als der Grubbs-Wert ist.</p> <p><u>Kein Ausreißer</u> liegt vor, wenn die Puffergröße kleiner als der Grubbs-Wert ist.</p>	<ol style="list-style-type: none"> 2. Comparing the buffer size with the Grubbs-value <p><u>An outlier is detected</u>, if the buffer size is bigger than the Grubbs-value.</p> <p><u>No outlier is detected</u>, if the buffer size is smaller than the Grubbs-value.</p>



Bester Schätzwert (Referenzwert) [3, 6, 7, 8]:

Der beste Schätzwert (= zugewiesener Wert) für den wahren Wert wird mit Hilfe der sensiblen Statistik wie folgt festgelegt:

1. Für chemische Ringversuche dienen als Ausgangsdatensätze die Labormittelwerte, die mit Hilfe einer standardisierten Methode ermittelt wurden.

Wurde für einen Parameter eine definierende Methode (Codex Typ 1) international festgelegt, so werden zur Festlegung des besten Schätzwertes nur Werte berücksichtigt, die nach dieser Methode ermittelt wurden.

Best estimate (reference value) [3, 6, 7, 8]:

The best estimate (= assigned value) for the true value is determined by the use of the sensitive statistics as follows:

1. *For the chemical proficiency testing schemes, the initial dataset is based on the mean values of the laboratories, which used standardized methods for analysis.*

For parameters where a defining method (Codex type 1) has been internationally established, only results obtained using this method were considered.

Bei eindeutig chemisch definierten Parametern werden alle unabhängigen Methoden insoweit berücksichtigt, als die Vertrauensbereiche der Mittelwerte nach den einzelnen Methoden überlappen.

Für mikrobiologische Ringversuche die Labormittelwerte zur Basis 10 logarithmiert. Alle weiteren Berechnungen von statistischen Daten und Bewertungskriterien erfolgen auf Basis dieser logarithmierten Werte.

2. Prüfung des Datensatzes auf Normalverteilung (David-Test) und Vergleich der Ergebnisse mit denen der robusten Statistik (Mittelwert, Standardabweichung, Chi²-Wert).
3. Bei normaler Verteilung Berechnung des Mittelwertes des Datensatzes nach Ausschluss der eliminierten Laboratorien (Grubbs-Test).
4. Bei nicht normaler Verteilung wird der Median herangezogen.

For chemically defined parameters all results obtained using independent methods were considered as far as confidence intervals of means for a particular method overlap with others.

For microbiological proficiency testing schemes, the initial dataset of the mean values of the laboratories are calculated to common logarithms (base 10). Further calculation of statistical values and performance criteria are based on the logarithm of these values.

2. *Check of the normal distribution of the dataset (test David) and comparison of the results with the results of the robust statistics (mean value, standard deviation, Chi²-value).*
3. *If the distribution is normal, calculation of the mean value of the dataset after elimination of the detected outliers (Grubbs-test).*
4. *If the distribution is not normal, the median is determined as best estimate.*

Unsicherheit des Referenzwertes [3, 4, 6, 7]:

- Bei chemischen Ringversuchen:
Festlegung der Unsicherheit (95% Wahrscheinlichkeit) auf Basis des Vertrauensbereiches des Referenzwertes. Verwendung des Streubereiches aus dem Homogenitätstest, sofern dieser größer ist als der Vertrauensbereich.
- Bei mikrobiologischen Ringversuchen:
Die Unsicherheit der Referenzwerte wird als kombinierte Standardabweichung aus den Standardabweichungen des Homogenitätstests ($\log s_{wb}$ und $\log s_{bb}$) und des Ringversuchs ($\log s_{ref}$) berechnet [4]:

Uncertainty of the reference value [3, 4, 6, 7]:

- For chemical proficiency tests:
Determination of the uncertainty based on the confidence interval (95% probability) of the reference value. Use of the range of variation deriving from the homogeneity test, if the latter one is bigger than the confidence interval.
- For microbiological proficiency tests:
Uncertainty of the reference value is calculated as combined standard deviation deriving from the standard deviations of the homogeneity ($\log s_{wb}$ and $\log s_{bb}$) and the proficiency test ($\log s_{ref}$) [4]:

$$MU = \sqrt{(s_{bb})^2 + (s_{wb})^2 + (s_{ref})^2}$$

Robuster Mittelwert [5]:

Hampelschätzer:

Robust mean [5]:

Hampel estimator:

$$\sum_{j=1}^J \Psi \left(\frac{y_j - m}{s_R} \right) = 0$$

$$\Psi(x) = \begin{cases} 0 & x \leq -4,5 \\ -4,5 - x & -4,5 < x \leq -3 \\ -1,5 & -3 < x \leq -1,5 \\ x & -1,5 < x \leq 1,5 \\ 1,5 & 1,5 < x \leq 3 \\ 4,5 - x & 3 < x \leq 4,5 \\ 0 & x > 4,5 \end{cases}$$

Robuste Standardabweichung [5]:

Q-Methode:

$$s_R = \frac{G_1^{-1}(q)}{\sqrt{2\Phi^{-1}(0,5 + 0,5q)}}$$

Robust standard deviation [5]:

Q-method:

- q = Quantilparameter
- Φ = Verteilungsfunktion der Standardnormalverteilung
- $G_1(x_i)$ = Funktion zur Stetigkeitskorrektur


3. AUSWERTUNG DER SENSORISCHEN RINGVERSUCHE / EVALUATION OF THE PROFICIENCY TESTING SCHEMES IN SENSORY

<p>Page Test [11]:</p> <p>Der Test basiert auf Rangsummen, die in die natürliche Reihenfolge gebracht werden können.</p> $L = R_1 + 2R_2 + 3R_3 + \dots + kR_k$ <p>Ist der berechnete L-Wert gleich oder größer dem tabellierten kritischen Wert, folgt daraus dass die Rangreihenfolge, wie sie von den Prüfern wahrgenommen wird, mit der vorgegebenen Reihenfolge übereinstimmt.</p>	<p>Page test [11]:</p> <p><i>The test is based of the rank data, which should be arranged in the correct order.</i></p> $L = R_1 + 2R_2 + 3R_3 + \dots + kR_k$ <p><i>If the calculated L-value is equal or higher as the critical value, is the conclusion, that the rank data, identified of the panellist, agrees with the given order.</i></p>
--	--

<p>Fehlersumme:</p> <p>Die Fehlersumme wird aus der Reihenfolge und den Abständen der 4 Prüfproben berechnet. Die Abstände der 4 Prüfproben werden aus dem Homogenitätstest entnommen.</p> <p>Die ermittelte Rangreihenfolge wird anschließend mit der tatsächlichen verglichen und für jede Fehlplatzierung die Fehlerpunkte aus den Abständen ermittelt. Die Summe der Fehlerpunkte ergibt die Fehlersumme.</p>	<p>Error sum:</p> <p><i>The error sum is calculated out of the sequence and the intervals of the four ranking test samples. The intervals are taken from the homogeneity test.</i></p> <p><i>The ranking test of each panellist is compared with the real sequence and the error points are calculated out of the intervals (table) and the sum makes the error sum.</i></p>
--	---

<p>Profilprüfung:</p> <p>Bei der Profilprüfung werden die Intensitäten der gefragten Attribute der Prüfproben anhand von zwei Referenzproben ermittelt und entsprechend auf Skalen zwischen 1 bis 10 eingetragen.</p> <p>Statistische Auswertung:</p> <ul style="list-style-type: none"> • Ausreißertest nach Grubbs (ab 3 Panellisten) • Mittelwert und Standardabweichung eines Panels • Gesamtmittelwert und Gesamtstandardabweichung über alle Panel • Vertrauensbereich • Messunsicherheit (MU) <p>Die Messunsicherheit errechnet sich folgendermaßen:</p>	<p>Profile testing</p> <p><i>In the proficiency test profile testing, the intense of the demanded attributes are determined with the help from two references. The results are plotted on scales from 1 to 10.</i></p> <p>Statistical evaluation:</p> <ul style="list-style-type: none"> • <i>Outlier test according to Grubbs (more than 2 panellists)</i> • <i>Mean value and standard deviation for every single panel</i> • <i>Mean value and standard deviation for all panels</i> • <i>Confidence interval</i> • <i>Measurement uncertainty (MU)</i> <p><i>It is calculated as follows:</i></p>
---	--

$$MU = \sqrt{\frac{1}{n} \cdot \sum_{i=1}^n (s_i)^2}$$

<p>Verantwortlich: Anita Schott Revisions-Nr.: 6 Seite: 9 von 17</p>	<p>Revision erstellt von: Sabine Klee, 19.10.11 Revision geprüft von: Sabine Klee, 24.10.11 Revision genehmigt von: Ute Braun, 24.10.11</p>	
--	---	---

4. HOMOGENITÄT DES PRÜFMATERIALS / HOMOGENEITY OF THE TESTING MATERIAL

<p>Homogenitätstest nach BCR [9]:</p> <p>Die Homogenität zwischen den Probenpackungen wird – wenn möglich – für sämtliche Parameter bestimmt und als Streubereich (95 % statistische Sicherheit) ausgedrückt.</p>	<p>Homogeneity test according to BCR [9]:</p> <p><i>The homogeneity between sample packages is determined, if possible, for all parameters and expressed as the range of variation (95 % probability level).</i></p>
--	---

<ul style="list-style-type: none"> • Aus der Charge werden (mindestens) 10 repräsentative Stichproben entnommen und wie folgt analysiert: • 10 Untersuchungen aus 10 Proben zur der Unsicherheit zwischen den Packungen, ausgedrückt als Standardabweichung des Materials (s_{bb}) • 10 Untersuchungen aus 1 Probe zur Bestimmung der analytischen Unsicherheit ausgedrückt als Standardabweichung (s_{wb}). • Die erweiterte Materialunsicherheit (U_{mat}) errechnet wie folgt: 	<ul style="list-style-type: none"> • <i>Minimum 10 representative samples were drawn out of the whole charge. The samples were analyzed as following:</i> • <i>10 measurements out of 10 samples are performed to determine the uncertainty between the packages, expressed as standard deviation of the material (s_{bb}).</i> • <i>10 measurements out of 1 are performed to determine the analytical uncertainty, expressed as standard deviation (s_{wb}).</i> <p><i>The extended uncertainty of the material (U_{mat}) is calculated as follows:</i></p>
--	---

$$U_{mat} = [(s_{bb}) - (s_{wb})] \cdot student \ factor$$

<p>Homogenitätstest nach DIN ISO 13528 [4]:</p> <p>Die Homogenität zwischen den Probenpackungen wird – wenn möglich – für sämtliche Parameter bestimmt und als Streubereich (95 % statistische Sicherheit) ausgedrückt.</p> <ul style="list-style-type: none"> • Aus der Charge wird eine repräsentative Anzahl an Proben zufällig verteilt entnommen, aus jeder Probe werden jeweils zwei Teilstichproben hergestellt. • Die Teilproben werden unter Wiederholbedingungen analysiert. • Die erweiterte Materialunsicherheit (U_{mat}) wird aus der Standardabweichung zwischen den Proben (s_{bb}) wie folgt berechnet: 	<p>Homogeneity test according to DIN ISO 13528 [4]:</p> <p><i>If possible – the homogeneity between sample packages was determined for all parameters and expressed as the range of variation (95 % probability level).</i></p> <ul style="list-style-type: none"> • <i>A statistical representative number of random distributed samples are taken from the whole charge. Each sample is divided into two sub-samples.</i> • <i>The subsamples are analysed under repeatable conditions.</i> • <i>The extended uncertainty of the material (U_{mat}) is calculated out of the standard deviation between the bottles (s_{bb}) as follows:</i>
---	---



$$s_{bb} = \sqrt{s_{xt}^2 - \left(\frac{s_{wb}^2}{2}\right)} \quad \text{mit / with:} \quad s_{wb} = \sqrt{\frac{\sum w_t^2}{2 \cdot n}} \quad \text{und / and:} \quad w_t = |x_{t1} - x_{t2}|$$

- s_{bb} = Standardabweichung zwischen den Proben / *standard deviation between the bottles*
- s_{xt} = Standardabweichung der Mittelwerte / *standard deviation mean values*
- s_{wb} = Standardabweichung innerhalb der Proben / *standard deviation within the bottles*
- x_{t1} = Ergebnisse in Teilstichprobe 1 / *result in subsample 1*
- x_{t2} = Ergebnisse in Teilstichprobe 2 / *result in subsample 2*
- n = Anzahl der Stichproben / *number of samples*
- w_t = Spannweite zwischen den Teilproben / *between-test-portion ranges*

Und / and:

$$U_{mat} = (s_{bb}) \cdot \text{student factor}$$

Kriterien zur Bewertung der Homogenität [4, 14]:

Die Homogenität des Materials, ausgedrückt durch die im Homogenitätstest ermittelte Standardabweichung (s_{bb}), ist unter anderem für die Leistungsbeurteilung der Teilnehmer entscheidend. Die Kriterien zur Bewertung der Homogenität eines Probenmaterials sind in der Norm DIN ISO 13528 sowie im IUPAC Harmonized Protocol festgelegt.

Criteria for the assessment of the homogeneity [4, 14]:

The homogeneity of the material, expressed as the standard deviation (s_{bb}) deriving from the homogeneity test, is amongst others important for the assessment of the performance of the participants. The standard DIN ISO 13528 and the IUPAC Harmonized Protocol define the criteria for the assessment of sufficient homogeneity.

Nach DIN 13528 gilt die Homogenität eines Materials als bestätigt, wenn die Standardabweichung aus dem Homogenitätstest das 0,3-fache der sogenannten Zielstandardabweichung für die Berechnung der z-Werte beträgt.

According to DIN 13528, the material is sufficient homogeneous if the standard deviation of the homogeneity test is the 0.3-fold of the so called target standard deviation, which is used for the calculation of the z-scores.

Seit Einführung des z-Wertes in der Lebensmittelanalytik dient als Zielstandardabweichung zur Berechnung dieses Bewertungskriteriums die Vergleichsstandardabweichung ($s(R)$) der jeweiligen national und international anerkannten (Referenz-)methode. Sie stellt eine von den Ringversuchsergebnissen unabhängige Größe dar und hat sich als Zielvorgabe in der Praxis bestens bewährt. Die Präzision (Genauigkeit) eines Verfahrens wird durch die Präzisionsdaten (Vergleichs- und Wiederholstandardabweichung) definiert und vor allem limitiert. Aus diesem Grund können Materialien deren Homogenität mit Hilfe dieser Methoden ermittelt wird nach dem oben genannten Kriterium fälschlicherweise als nicht homogen bewertet werden.

Since the implementation of the z-score in food analysis, the reproducibility standard deviation ($s(R)$) of the national and international accredited (reference) method is used as target standard deviation for the calculation of the z-scores. This parameter is independent from the results of the proficiency testing and they delivered an optimal set target in practice over the years. The precision of a method is defined and above all limited by the precision data (reproducibility and repeatability standard deviation). For that reason the materials, which are analysed with those methods in the homogeneity test, may show a lack of sufficient homogeneity according to the above named criterion.



Das Harmonized Protocol trägt diesem Umstand Rechnung, indem ein weiteres Kriterium (c) zur Bewertung der Homogenität aufgeführt wird, das die methodisch bedingte analytische Streuung stärker berücksichtigt.

Dieses Kriterium wird zur Beurteilung der Homogenität der chemischen und mikrobiologischen Prüfmateriale herangezogen.

Berechnung des Kriteriums (c):

The harmonized protocol takes this circumstance into account by including a second criterion (c) for the assessment of the homogeneity. It considers the analytical variation of the results more intensely.

That criterion is used for the assessment of the homogeneity of the testing materials in chemistry and microbiology.

Calculation of the criterion (c):

$$c = [F_1 * (s_{erl})^2] + [F_2 * (s_{wb})^2]$$

$s_{bb}^2 < c$: Homogenität ausreichend / Homogeneity sufficient

$s_{bb}^2 > c$: Homogenität nicht ausreichend / Homogeneity not sufficient

s_{erl} = Homogenitätskriterium nach DIN 13528 (= 0,3 * Zielstandardabweichung; hier: s(R))

s_{wb} = Standardabweichung innerhalb der Packungen

s_{bb} = Standardabweichung zwischen den Packungen

F_1 / F_2 = Faktoren, im Harmonized Protocol angegeben (abhängig von n)

s_{erl} = *critterion for homogeneity accord. to DIN 13528 (= 0,3 * target standard deviation; here: s(R))*

s_{wb} = *standard deviation within the packages*

s_{bb} = *standard deviation between the packages*

F_1 / F_2 = *factors, given in the harmonized protocol (dependent from n)*

Homogenitätstest für sensorische Prüfmateriale in Anlehnung an BCR [9]:

Die Homogenität zwischen den Probenpackungen wird über ein spezielles Verfahren in Anlehnung an BCR/48/93 bestimmt. Dabei wird sowohl geprüft, ob die Konzentrationsstufen unterscheidbar sind (Prüfung auf Überlappung der Vertrauensbereiche) und ob bei Wiederholung der sensorischen Tests für die einzelnen Konzentrationsstufen reproduzierbare Ergebnisse erzielt werden.

Der Vertrauensbereich einer Konzentrationsstufe stellt das Maß der Variabilität von 12 Messungen um den Mittelwert dar.

Die Homogenität zwischen den Probenpackungen wird für alle Parameter bestimmt und als Streubereich (95 % statistische Sicherheit) ausgedrückt.

Homogeneity test for sensorical testing materials on the basis of BCR [9]:

The homogeneity between sample packages is determined with a special procedure following the principles of BCR/48/93. In the process is checked, if the concentration-levels of the four samples are discriminable or not (testing of overlapping the confidence interval) and if the repeat in the sensory tests provides reproducible results for each concentration level.

The confidence interval of each sample presents the measure of variability from twelve measurements around the mean value.

The homogeneity between sample packages is determined for all parameters and expressed as the range of variation (95 % probability level).

5. BEWERTUNG DER ANALYTISCHEN LEISTUNG / ASSESSMENT OF THE ANALYTICAL PERFORMANCE

<p>z-Wert [3, 4, 15]:</p> <p>Der z-Wert ist ein vom vorliegenden Ringversuch weitgehend unabhängiges Bewertungskriterium.</p> <p>Zur Berechnung des z-Wertes wird die Differenz des Labormittelwertes und des Besten Schätzwertes (zugewiesenen Wertes) durch eine vom Ringversuch weitgehend unabhängige Zielstandardabweichung dividiert.</p>	<p>z-Score [3, 4, 15]:</p> <p><i>The performance criterion z-score is widely independent from the present proficiency test.</i></p> <p><i>For calculation of the z-score, the difference of the laboratory mean value and the best estimate (assigned value) is divided by a target standard deviation which is widely independent from the present proficiency test.</i></p>
--	--

<p>z-Wert auf der Grundlage von Präzisionsdaten:</p> <p>Als Zielstandardabweichung wird bei den muva-Ringversuchen in der Regel die Vergleichsstandardabweichung s(R) der national bzw. international standardisierten Methoden (Präzisionsdaten) eingesetzt.</p>	<p>z-score on the basis of precision data:</p> <p><i>Generally in muva-proficiency testing schemes the reproducibility standard deviation s(R) of the national and international standardized methods (precision data) is used as target standard deviation.</i></p>
--	---

$$z - \text{Wert} = \frac{x - x_{\text{ref}}}{s(R)}$$

<p>Bewertung:</p> <p> z < 2 analytische Leistung ist zufriedenstellend</p> <p>2 > z < 3 analytische Leistung ist fragwürdig (Warnsignal)</p> <p> z > 3 analytische Leistung ist nicht zufriedenstellend (Eingriffssignal)</p>	<p>Assessment:</p> <p> z < 2 <i>analytical performance is satisfactory</i></p> <p>2 > z < 3 <i>analytical performance is questionable, (warning signal)</i></p> <p> z > 3 <i>analytical performance is unsatisfactory, (action signal)</i></p>
--	---

<p>z-Wert auf der Grundlage der Horwitz Standardabweichung:</p> <p>Der z-Wert (Horwitz) kommt zur Anwendung wenn:</p> <ul style="list-style-type: none"> keine geeigneten Präzisionsdaten für den Parameter existieren die Parameter in Konzentrationen im Bereich mg/kg (ppm) vorliegen <p>Nicht geeignet sind Parameter die ohne Einheit angegeben werden, wie z.B. der pH-Wert.</p> <p>Vor allem in niedrigen Konzentrationsbereichen sind aufgrund der Analytik größere Streuungen der Datensätze durchaus üblich.</p>	<p>z-score on the basis of standard deviation according to Horwitz:</p> <p><i>The z-score (Horwitz) is applied, if:</i></p> <ul style="list-style-type: none"> <i>applicable precision data do not exist for the parameters</i> <i>the concentration of the parameter is in the range of mg/kg (ppm).</i> <p><i>Parameters indicated without a measurement unit are not suitable, e.g. pH value.</i></p> <p><i>Due to the analytical methods, especially in low concentration a wide variation of the results is usual.</i></p>
---	--

<p>Aus diesem Grund hat Horwitz auf der Basis von zahlreichen Ringversuchen eine Funktion entwickelt, mit deren Hilfe, in Abhängigkeit von der Konzentration, eine Zielstandardabweichung berechnet werden kann.</p> <p>Mit Hilfe des sogenannten Horrat-Wert (Horwitz-ratio) wird anschließend geprüft, ob sich die mittels der Zielstandardabweichung nach Horwitz errechneten z-Werte zur Beurteilung der Laborleistung eignen. Dies ist nach Horwitz der Fall, wenn nicht mehr als 22 % der Labore als Ausreißer gewertet werden und der Quotient aus der Standardabweichung des besten Schätzwertes und der Zielstandardabweichung kleiner 1,5, im ungünstigsten Fall kleiner 2 ist [16].</p> <p>Anschließend wird der z-Wert (Horwitz) wie bekannt ermittelt:</p>	<p><i>For that reason, on the basis of a great number of interlaboratory studies Horwitz developed a function for the evaluation of a target standard deviation in dependence of the concentration.</i></p> <p><i>Then it is checked by the help of the so called Horrat-value (Horwitz-ratio) if the z-scores, which are calculated on the basis of the target standard deviation, are suitable for the assessment of the laboratories performance.</i></p> <p><i>According to Horwitz, the z-scores are suitable if less than 22 % of the laboratories are detected as outliers and the ratio of the standard deviation of the best estimate and the target standard deviation is less than 1.5. In the most inappropriate case the ration should be less than 2 [10].</i></p> <p><i>Finally the z-score (Horwitz) is calculated as known:</i></p>
---	--

$$z - Wert = \frac{x - x_{ref}}{s(H)}$$

<p>Bewertung:</p> <p> z < 2 analytische Leistung ist zufriedenstellend</p> <p>2 > z < 3 analytische Leistung ist fragwürdig (Warnsignal)</p> <p> z > 3 analytische Leistung ist nicht zufriedenstellend (Eingriffssignal)</p>	<p>Assessment:</p> <p> z < 2 analytical performance is satisfactory</p> <p>2 > z < 3 analytical performance is questionable, (warning signal)</p> <p> z > 3 analytical performance is unsatisfactory, (action signal)</p>
--	--

<p>z'-Wert [3, 4]:</p> <p>Dieses Bewertungskriterium ist vom vorliegenden Ringversuch abhängig, da bei der Berechnung des z'-Wertes die Standardabweichung des besten Schätzwertes (s_{ref}) aus dem vorliegenden Ringversuch sowie die Standardabweichung (s_{bb}) des Materials verwendet werden.</p>	<p>z'-score [3, 4]:</p> <p><i>The z'-score depends on the present proficiency test, since the standard deviation of the best estimate (s_{ref}) and heterogeneity of the material (s_{bb}) are used for the calculation.</i></p>
--	---

$$z' - Wert = \frac{x - x_{ref}}{\sqrt{s_{ref}^2 + s_{bb}^2}}$$

<p>Bewertung:</p> <p> z < 2 analytische Leistung ist zufriedenstellend</p> <p>2 > z < 3 analytische Leistung ist fragwürdig (Warnsignal)</p> <p> z > 3 analytische Leistung ist nicht zufriedenstellend (Eingriffssignal)</p>	<p>Assessment</p> <p> z < 2 analytical performance is satisfactory</p> <p>2 > z < 3 analytical performance is questionable, (warning signal)</p> <p> z > 3 analytical performance is unsatisfactory, (action signal)</p>
--	---



CRD-Wert [6, 12]:

Bei der Berechnung des CRD-Wertes wird die Differenz des Labormittelwertes und des besten Schätzwertes durch die kritische Differenz (CD) aus den national und international anerkannten Methoden (Präzisionsdaten) dividiert. Das Bewertungskriterium ist daher wiederum vom vorliegenden Ringversuch weitgehend unabhängig.

Bei der Ermittlung der kritischen Differenz geht man von einer zweiseitige Fragestellung aus:

CRD-value [6, 12]:

For calculation of the CRD-value, the difference of the laboratory mean value and the best estimate is divided by the critical difference (CD) deriving from the national and international standardized methods. Therefore the performance criterion is again independent from the present proficiency test.

The calculation is based on a double sided question:

$$CD = \frac{1}{\sqrt{2}} \sqrt{(2,8s(R))^2 - (2,8s(r))^2 \left(\frac{n-1}{n} \right)}$$

$$CD = \frac{1}{\sqrt{2}} \sqrt{R^2 - r^2 \left(\frac{n-1}{n} \right)}$$

Für eine Doppelbestimmung $n = 2$ / because of a double determination $n = 2$:

$$CD = 0,707 \sqrt{R^2 - \frac{r^2}{2}}$$

R = Vergleichbarkeit / *Reproducibility*

r = Wiederholbarkeit / *Repeatability*

Berechnung des CRD-Wertes:

Calculation of the CRD-value:

$$CRD = \frac{X - X_{ref}}{CD}$$

Bewertung:

CRD-Wert < 1 = Laborwert ist zufriedenstellend

CRD-Wert > 1 = Laborwert ist nicht zufriedenstellend (Eingriffssignal)

Assessment:

CRD-Wert < 1 = laboratory value is satisfying

CRD-Wert > 1 = laboratory value is not satisfying (action signal)

6. VERZEICHNIS DER VERWENDETEN STANDARDS / NORMEN UND LITERATUR / REGISTER OF USED STANDARDS / NORMS AND LITERATURE

- | | | |
|------|--|---|
| [1] | DIN EN ISO/IEC
17025:2005 | Allgemeine Anforderungen an die Kompetenz von Prüf- und Kalibrierlaboratorien / <i>General requirements for the competence of testing and calibration laboratories</i> |
| [2] | ILAC-G13: 2007 | Guidelines for the Requirements for the Competence of Providers of Proficiency Testing Schemes |
| [3] | DIN EN ISO/IEC
17043-05:2010 | Konformitätsbewertung – Allgemeine Anforderungen an Eignungsprüfungen / <i>Conformity assessment – General requirements for proficiency testing</i> |
| [4] | DIN ISO
13528:2009-1 | Statistische Verfahren für Eignungsprüfungen durch Ringversuche / <i>Statistical methods for use in proficiency testing by interlaboratory comparison</i> |
| [5] | DIN ISO
38402-45:2003-09 | Ringversuche zur externen Qualitätskontrolle von Laboratorien |
| [6] | DIN ISO
5725-2:2002-12 | Genauigkeit und (Richtigkeit und Präzision von Messverfahren und Messergebnissen, Teil 2 / <i>Accuracy (trueness and precision) of measurement methods and results – Part 2</i> |
| [7] | DIN ISO
5725-4:2000-11 | Genauigkeit und (Richtigkeit und Präzision von Messverfahren und Messergebnissen, Teil 4 / <i>Accuracy (trueness and precision) of measurement methods and results – Part 4</i> |
| [8] | IDF 135B: 1991-12 | Precision characteristics of analytical Methods – Outline of collaborative study procedure |
| [9] | BCR 48/93:
15.12.1994 | Guidelines for the production and certification of BCR reference materials |
| [10] | DIN 53804 Teil 1 | Statistische Auswertungen: Messbare (kontinuierliche) Merkmale |
| [11] | DIN ISO
8587:2010-08 | Sensorische Analyse – Prüfverfahren – Rangordnungsprüfung / <i>Sensory analysis – Methodology - Ranking</i> |
| [12] | Amtliche Sammlung von
Untersuchungsverfahren
§ 64 LFGB | Planung und statistische Auswertung von Ringversuchen zur
Methodenvalidierung |
| [13] | T.J. Simat,
TU-Dresden | Professur für Lebensmittelkunde und Bedarfsgegenstände, 2008 |
| [14] | Thomson, M., Ellison
S.L.R., Wood R. | The international harmonized protocol for the proficiency testing of analytical chemistry laboratories, 2006 |
| [15] | Horwitz, W. | Evaluation of analytical methods used for regulation of foods and drugs.
Anal Chem 54 (1): 67A-76A, 1982 |
| [16] | Horwitz, W., Britton, P.,
Chirtel, St. J. | Journal of AOAC International <u>81</u> , 1257-1265, 1998 |

VERWENDETE ABKÜRZUNGEN / USED ABBREVIATIONS:

DIN	Deutsches Institut für Normung e. V.
EN	Europäische Norm
IDF	International Dairy Federation
IEC	International Electrotechnical Commission
ILAC	International Laboratory Accreditation Cooperation
ISO	International Organization for Standardization
LFGB	Lebensmittel-, Bedarfsgegenstände- und Futtermittelgesetzbuch