

# Auswertung von Ringversuchsdaten: Horwitz vs. DIN ISO 13528

Michael Herbold<sup>1</sup>, Georg Schmitt<sup>2</sup>, Gisela Skopp<sup>2</sup>

<sup>1</sup>ARVECON GmbH, Kleinfeldweg 52, D-69190 Walldorf

<sup>2</sup>Institut für Rechtsmedizin und Verkehrsmedizin, Universitätsklinikum Heidelberg, Voßstraße 2, D-69115 Heidelberg

## Abstract

Since 1995, the Society of Toxicological and Forensic Chemistry (GTFCh) has been providing round robin tests for the analysis of illegal drugs in serum. Since that time, results have been evaluated using the participant mean as a reference value and the Horwitz standard deviation to determine its permissible variability. That implies that the permissible variability depends on the participants mean, only, whereas the range of variation of all single values is not considered. Using the DIN ISO 13528 instead provides a mean to cover the range of variation of all results when estimating the permissible variability. Both data analysis methods will be compared by appropriate examples simultaneously demonstrating the pros and cons.

## 1. Einleitung

Seit dem Start des Ringversuchsprogramms der GTFCh im Jahr 1995 wird für fast alle quantitativen Ringversuche die Horwitz-Standardabweichung als Kriterium für die Festlegung der Bewertungsgrenzen angewandt. Die Festlegung der Grenzwerte ist dabei nur von der Konzentration des Zielwertes (Mittelwert der Ergebnisse der Teilnehmer) abhängig. Die Schwankungsbreite der Ergebnisse wird nicht berücksichtigt. Diesbezüglich sollte untersucht werden, inwieweit dieses Verfahren noch angemessene Leistungsanforderungen an die Teilnehmer stellt und ob das Verfahren Algorithmus A nach DIN ISO 13528 [1] eine Alternative darstellt bzw. dieses bei der Beurteilung des bisherigen Verfahrens hilfreich sein kann.

## 2. Methoden

### 2.1. Das IUPAC-Verfahren

Das derzeit praktizierte Auswerteverfahren für die Ringversuche der GTFCh basiert auf den Vorgaben der IUPAC im International Harmonized Protocol for Proficiency Testing of Analytical Laboratories [2]. Die Auswertung der Teilnehmerergebnisse ist wie folgt:

1. Durchführung eines Ausreisser-Tests nach Grubbs 99% und 95% [3]
2. Berechnung des Zielwertes ZW aus dem ausreisserfreien Teilnehmermittelwert
3. Berechnung der Standardabweichung nach Horwitz auf Basis des Zielwertes ZW  
 $VK = 2 \cdot (1 - 0,5 \log ZW)$   $VK = \text{Variationskoeffizient} / ZW = \text{Konz. des Analyten (kg/L)}$
4. Berechnung der Bestehensgrenzen:  $ZW \pm 2 \cdot SD_{\text{Horwitz}}$  bzw.  $\pm 2 \cdot 1/2 \cdot SD_{\text{Horwitz}}$

### 2.2. Das Verfahren nach DIN ISO 13528 (Algorithmus A)

Das Verfahren Algorithmus A in der DIN ISO 13528 (auch in ISO 5275 Teil 5) orientiert sich am Median der Teilnehmerergebnisse und bezieht die Schwankungsbreite der Ergebnisse in der Bildung der Bewertungsgrenzen ein. Die Verfahrensweis ist wie folgt:

1. Berechnung der Anfangswerte:  $X^* = \text{Median } X_i$  ( $i = 1, 2, \dots, N$ ) und  $S^* = 1,483 \cdot \text{Median } |X_i - X^*|$
2. Berechnung der Grenzwerte:  $X_{UG} = X^* - S^*$ ;  $X_{OG} = X^* + S^*$  (UG=untere, OG=obere Grenze)
3. Falls  $X_i < X_{UG}$ , dann  $X_i = X_{UG}$  bzw. falls  $X_i > X_{OG}$ , dann  $X_i = X_{OG}$
4. Berechnung  $x^*$  und  $s^*$ :  $x^* = \sum x_i^*/N$ ;  $s^* = 1,134 \cdot \text{Wurzel}[(\sum(x_i^*-x^*)^2)/(N-1)]$

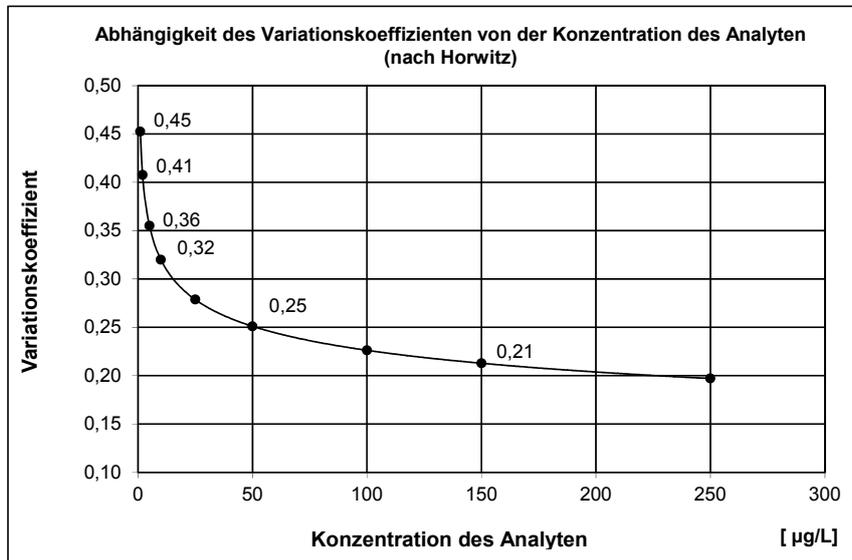


Fig. 1. Abhängigkeit des VK von der Konzentration des Analyten.

Die Schätzwerte  $x^*$  und  $s^*$  werden durch eine iterative Berechnung hergeleitet, d.h. durch mehrfaches Aktualisieren der Werte von  $x^*$  und  $s^*$  unter Verwendung der modifizierten Daten, bis Konvergenz auftritt.

### 3. Beispiele

Im ersten BTMF-Ringversuch in Jahr 1995 wurde die ganze Horwitz-Standardabweichung zur Berechnung der Bestehensgrenzen angewandt. Als die analytischen Leistungen im Laufe der Zeit besser wurden, wurde ab dem Ringversuch BTMF 3/98 die Sollwertbestimmung durch Sollwertlabors eingestellt und nur noch die halbe Standardabweichung nach Horwitz zur Berechnung der Bewertungsgrenzen angewandt. Dieses Verfahren hat sich bis heute bewährt. Die beiden Beispiele zeigen sowohl die Verbesserung der analytischen Leistungen der Teilnehmer als auch, dass die bis heute angewandte Horwitz-Standardabweichung akzeptable Bewertungsgrenzen liefert. Die Anforderungen sind heute sowohl für das Verfahren nach Algorithmus A als auch nach IUPAC ( $2 \cdot SD_{\text{Horwitz}}$ ) vergleichbar.

#### 3.1. Beispiel 1

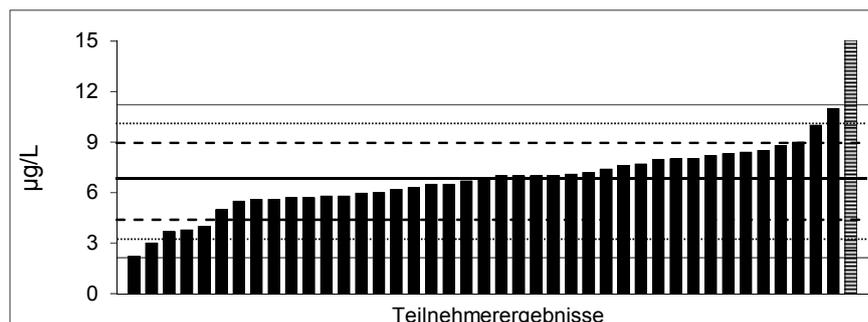


Fig. 2. Ergebnisse der THC-Bestimmung im Ringversuch BTMF 2/95.

Tab. 1. Ergebnisse THC - BTMF 2/95.

Standardabweichung	Zielwert	SD	Zielwert $\pm$ 2·SD
Horwitz	6,67	2,27	2,13 – 11,21
½·Horwitz	6,67	1,14	4,39 – 8,95
Algorithmus A	6,79	1,69	3,41 - 10,17

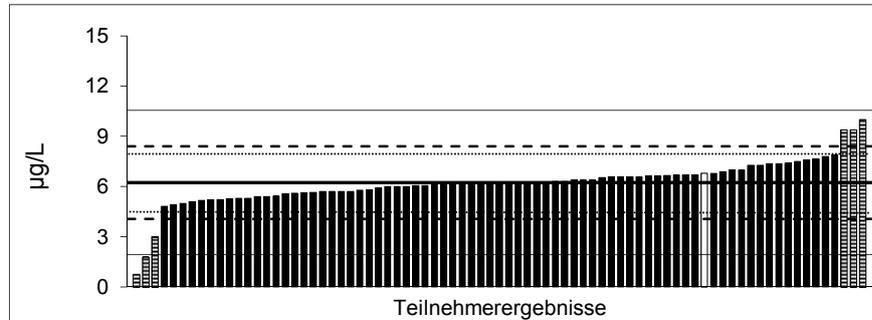


Fig. 3. Ergebnisse der THC-Bestimmung im Ringversuch BTMF 1/12.

Tab. 2. Ergebnisse THC - BTMF 1/12.

Standardabweichung	Zielwert	SD	Zielwert $\pm$ 2·SD
Horwitz	6,24	2,15	2,15 – 10,54
½·Horwitz	6,24	1,08	4,08 – 8,40
Algorithmus A	6,23	0,88	4,47 - 7,99

### 3.2. Beispiel 2

Die Bestimmung von Mirtazapin ist im Bereich des Therapeutic Drug Monitorings ebenfalls eine Routinebestimmung. In den Ringversuchen wenden 75% der Teilnehmer massenspektrometrische Bestimmungsmethoden (überwiegend LC/MS) an, was zu einer vergleichsweise hohen Präzision der Bestimmungen führt. Die bisher angewandten Bewertungsgrenzen werden über die doppelte Horwitz-Standardabweichung berechnet und stellen kein angemessenes Kriterium mehr dar. Somit konnten zwei Teilnehmer (siehe Pfeil), die statistisch gesehen Ausreisser produzierten, dennoch den Ringversuch bestehen.

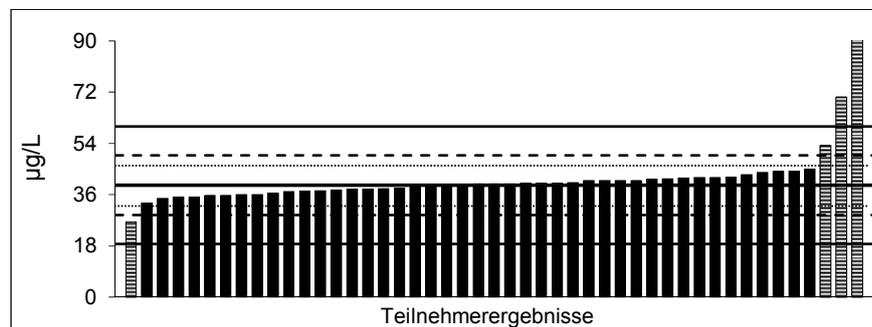


Fig. 4. Ergebnisse der Mirtazapin-Bestimmung im Ringversuch TDMD 3/12. ↑

Tab. 3. Ergebnisse Mirtazapin - TDMD 3/12.

Standardabweichung	Zielwert	SD	Zielwert $\pm$ 2·SD
Horwitz	39,3	10,3	18,7 – 59,9
½·Horwitz	39,3	5,2	28,9 – 49,7
Algorithmus A	39,6	3,6	32,4 – 46,8

### 3.3. Beispiel 3

Die Bestimmung von THC im Haar ist ebenfalls eine Routinebestimmung. Die Aufbereitung des Haarmaterials ist jedoch erheblich anspruchsvoller als die Aufbereitung von Serumproben und auch die Schwankungsbreite der Ergebnisse erheblich höher. Die mittels Algorithmus A berechneten Bewertungsgrenzen liegen daher über denen der doppelten Horwitz-Standardabweichung. Es stellt sich somit die Frage, ob die doppelte Horwitz-Standardabweichung die Mindestanforderungen in den Ringversuchen der GTFCh darstellen sollte, oder ob man bereit ist, auch weitere Bewertungsgrenzen zu akzeptieren.

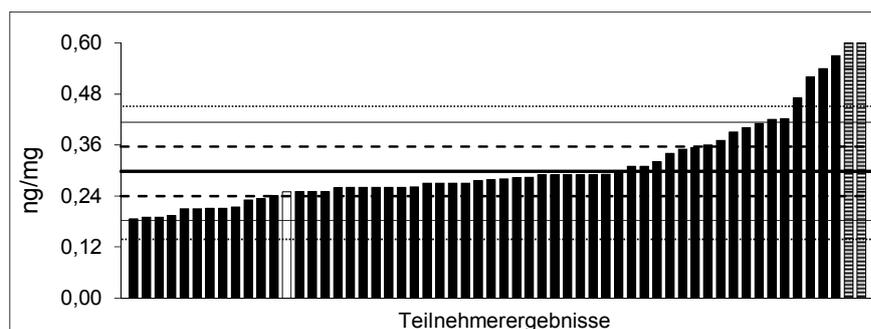


Fig. 5. Ergebnisse der THC-Bestimmung im Ringversuch DHF 3/12.

Tab. 4. Ergebnisse THC - DHF 3/12.

Standardabweichung	Zielwert	SD	Zielwert $\pm$ 2·SD
Horwitz	0,298	0,058	0,182 – 0,414
½·Horwitz	0,298	0,029	0,240 – 0,356
Algorithmus A	0,294	0,079	0,136 – 0,452

### 4. Schlussfolgerung

Das Verfahren Algorithmus A nach DIN ISO 13528 stellt ein brauchbares Verfahren zur Beurteilung der angewandten Bestehensgrenzen in den Ringversuchen der GTFCh dar. In Einzelfällen in denen das Verfahren nach Horwitz keine akzeptablen Bewertungsgrenzen liefert, sollte auch die Anwendung dieses Verfahrens berücksichtigt werden.

### 5. Literatur

- [1] DIN ISO 13538:2005, Statistische Verfahren für Eignungsprüfungen durch Ringversuche, Anhang C: 104-105
- [2] Thompson M., Wood R. International Harmonized Protocol for Proficiency Testing of Analytical Laboratories. IUPAC: Journal Of AOAC International, Vol. 76, No. 4, 1993.
- [3] Frank E. Grubbs, Procedures for Detecting Outlying Observations in Samples. In: Technometrics 1969;11:1–21.