

# Statistische Auswertung von Ringversuchen

... mehr als nur „erfolgreich“ oder „nicht erfolgreich“

---

Niedersächsisches Landesgesundheitsamt  
Bereich: Wasserhygiene  
- Standort Aurich -  
Dr. Sven Gebhardt

*16. Jahrestagung Trinkwasserringversuche NRW - Niedersachsen,  
Osnabrück, 04. März 2015*



Roesebeckstr. 4-6  
30449 Hannover  
Tel. 0511/4505-0  
Fax 0511/4505-140



Lüchtenburger Weg 24  
26603 Aurich  
Tel. 04941/9171-0  
Fax 04941/9171-10

# Normen

- DIN 38402-42 (2005-09)  
*Ringversuche zur Verfahrensvalidierung, Auswertung (A42)*
- DIN 38402-45 (2014-06)  
*Ringversuche zur Eignungsprüfung von Laboratorien (A45)*
- DIN ISO 13528 (2009-01), derzeit in Überarbeitung  
*Statistical methods for use in proficiency testing by interlaboratory comparisons*

# Ablauf Ringversuchsauswertung

1. Sichtung der Einzelergebnisse



2. Festlegung des zugewiesenen Wertes (Sollwert)



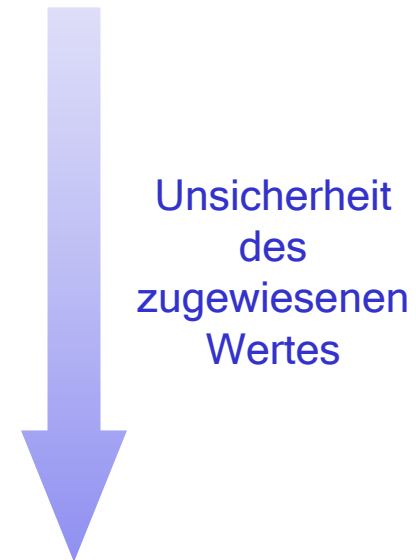
3. Ermittlung / Festlegung der Vergleichs- bzw. Soll-Standardabweichung



4. Leistungsbewertung (z.B. anhand von z-scores)

# Festlegung des zugewiesenen Wertes (Sollwert)

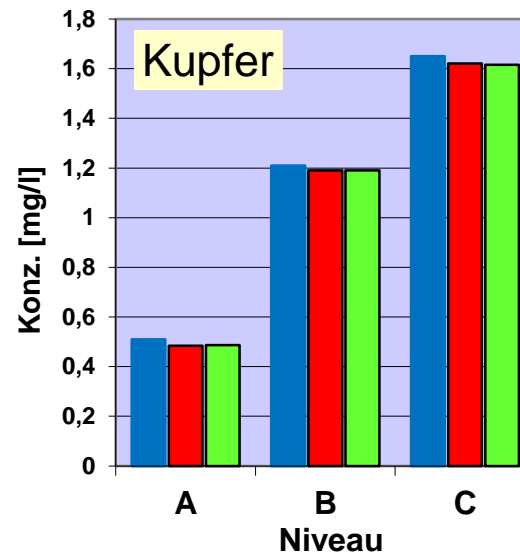
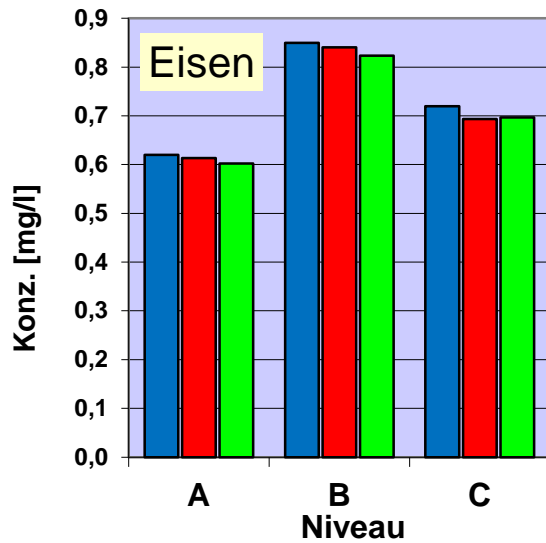
- a) Bekannter Wert
- b) Zertifizierter Referenzwert
- c) Referenzwert
- d) Konsenswert von Expertenlaboratorien
- e) Konsenswert aus Teilnehmerergebnissen



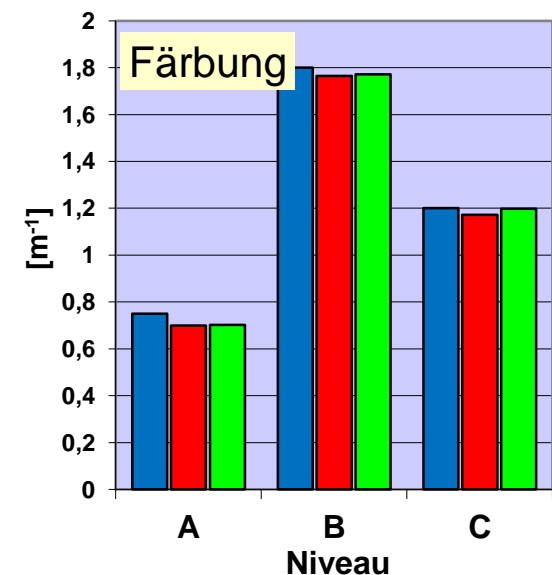
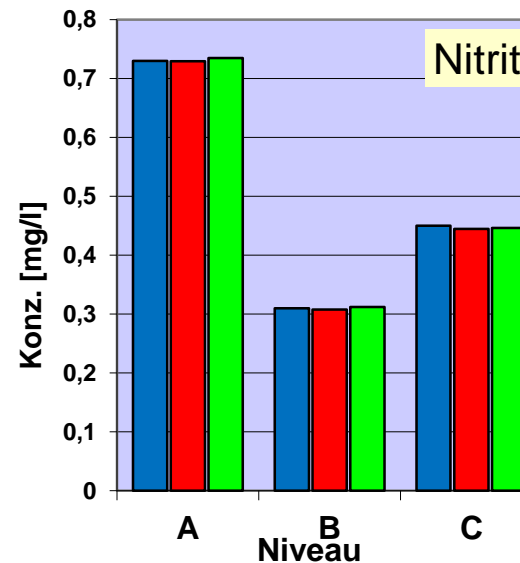
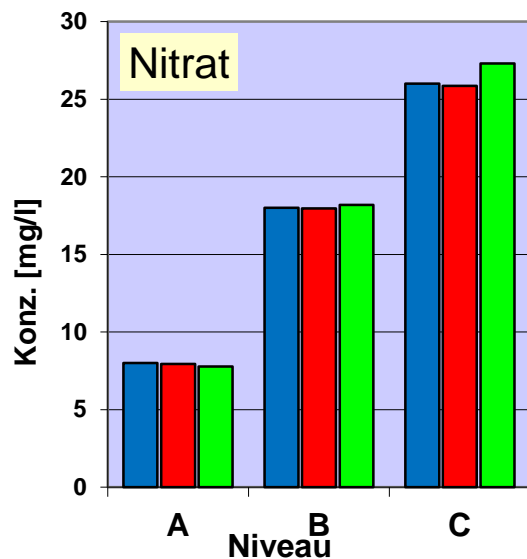
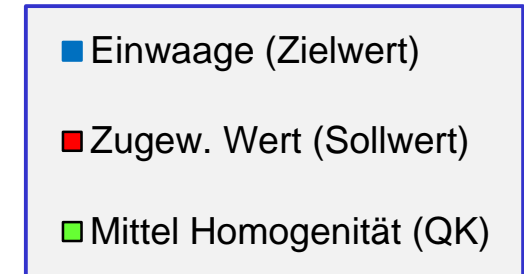
Die verschiedenen Möglichkeiten unterscheiden sich in Ihren Eigenschaften (z. B. Unsicherheit, Rückführbarkeit, Aufwand, Kosten, Machbarkeit)

⇒ „Fit for purpose“

# Einwaage vs. Zugewiesener Wert (Konsenswert)

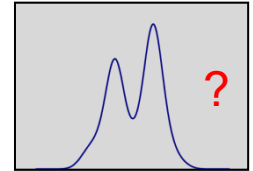


Ergebnisse aus RV 1-2014



# Voraussetzung für die Verwendung eines Konsenswertes

- Es besteht ein Konsens zwischen den Teilnehmerergebnissen
- Ermittlung durch ein robustes statistisches Verfahren
- Genügend hohe Anzahl an Teilnehmerergebnissen (mind. 12)
- Konsenswertes sollte durch Messungen bestätigt werden



⇒ Frage nach Ziel des Ringversuchs?  
Bewertung der Richtigkeit oder der Vergleichbarkeit von Ringversuchsergebnissen

# Konsenswert mittels robuster Statistik

→ DIN 38402-45: HAMPEL-Schätzer

Dabei haben Laborergebnisse, die vom Mittelwert mehr als das 4,5-fache der Vergleichsstandardabweichung abweichen keinerlei Einfluss auf das Ergebnis.

# Unsicherheit des zugewiesenen Wertes

## Bedingung:

$$u(x) \leq 0,3 \delta_{PT}$$

$u(x)$  ... Standardunsicherheit des zugewiesenen Wertes  
 $\delta_{PT}$  ... Standardabweichung für die Eignungsbeurteilung

## Berechnung für Konsenswert:

$$u(x) = 1,25 \times \frac{s_R}{\sqrt{n}}$$

$u(x)$  ... Standardunsicherheit des zugewiesenen Wertes  
 $s_R$  ... Vergleichsstandardabweichung der Ergebnisse  
 $n$  ... Anzahl der Teilnehmerergebnisse

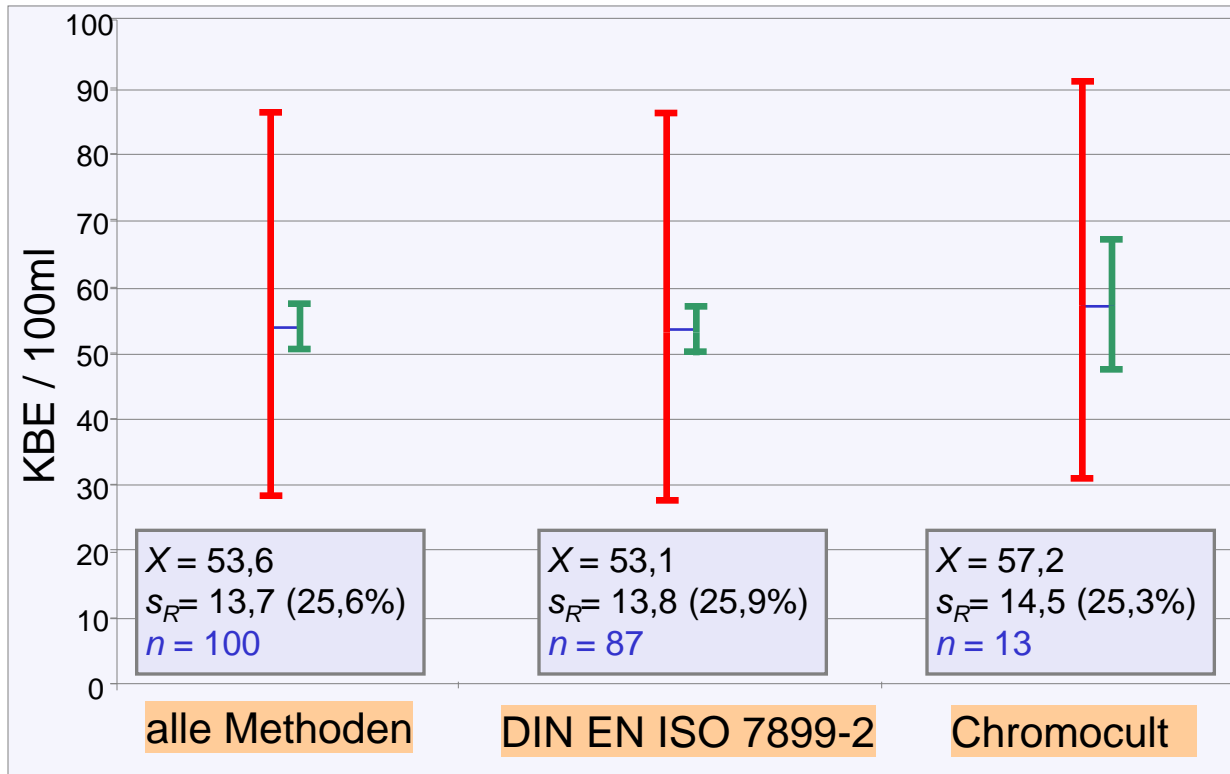
(DIN ISO 13528)

Multiplikation von  $u(x)$  mit dem Faktor  $k=2$  ergibt die erweiterte Unsicherheit  $U(x)$ .

- ⇒ Kein Rückschluss auf Messunsicherheit einer bestimmten analytischen Methode!
- ⇒ Nutzen bei der Berücksichtigung von RV-Ergebnissen als Komponente der Messunsicherheitsabschätzung  
→ s. DEV A0-4



# Beispiel: RV 1-2014, Enterokokken, Gruppe B



■ Ausschlussgrenzen

■ erw. Unsicherheit (Vertrauensbereich) des Vorgabewertes

$$u(x) = 1,25 \times \frac{s_R}{\sqrt{n}}$$

Die Zuverlässigkeit des Konsenswertes als zugewiesener Wert erhöht sich mit steigender Anzahl an Ergebnissen.

Der resultierende Vertrauensbereich erlaubt keine Verschiebung der festgelegten Ausschlussgrenzen zur Beurteilung der Teilnehmerergebnisse!

# Festlegung der Standardabweichung für die Eignungsbeurteilung

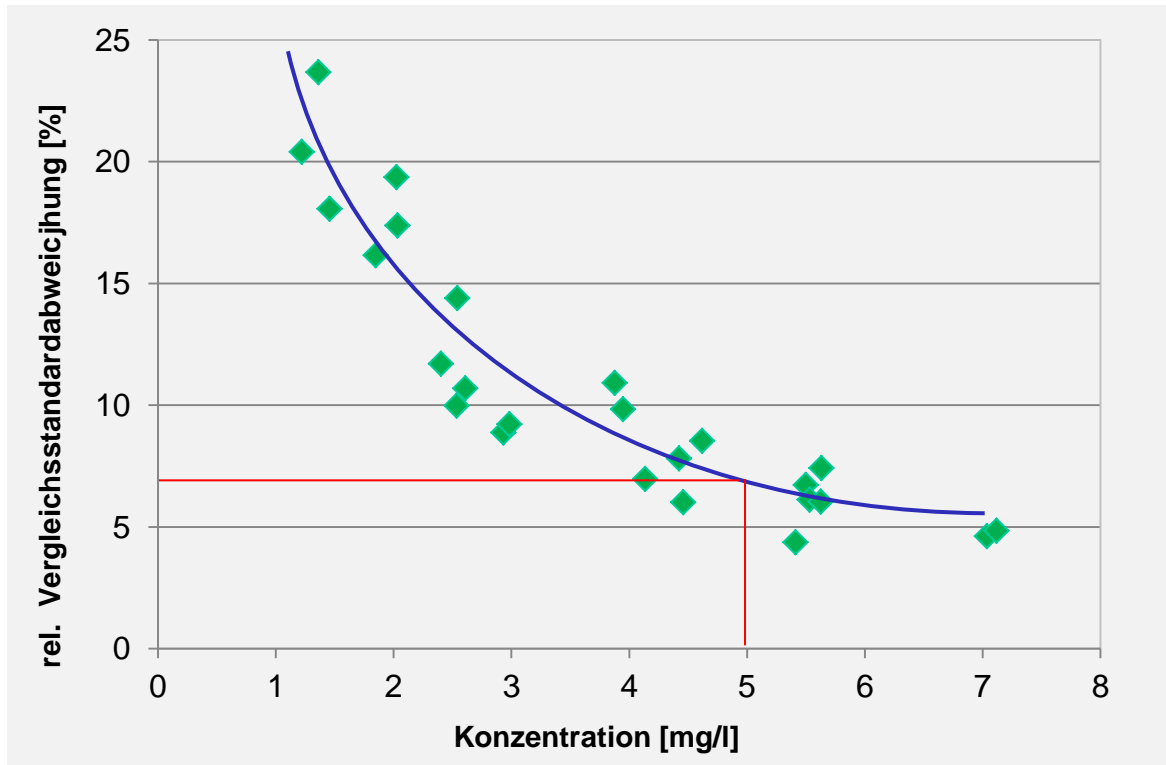
- a) Vorgabewert
- b) Wert aus Erkenntnissen
- c) Wert aus einem allgemeinen Modell
- d) Wert aus Ergebnissen von Präzisionsversuchen
- e) Statistisch ermittelter Wert aus Teilnehmerergebnissen eines Ringversuchs

## Häufig Kombination verschiedener Verfahren

- Schätzung der Vergleichsstandardabweichung  $s_R$  mit robustem statistischem Verfahren  
→ DIN 38402-45: Q-Methode
- Limitierung der Standardabweichung auf 5% bzw. 25%
- Anwendung der Varianzfunktion

# Varianzfunktion am Beispiel Oxidierbarkeit

NLGA RV-Ergebnisse 2012-2014



s. auch: [www.aqsbw.de](http://www.aqsbw.de)

# z-Score - Leistungsbeurteilung

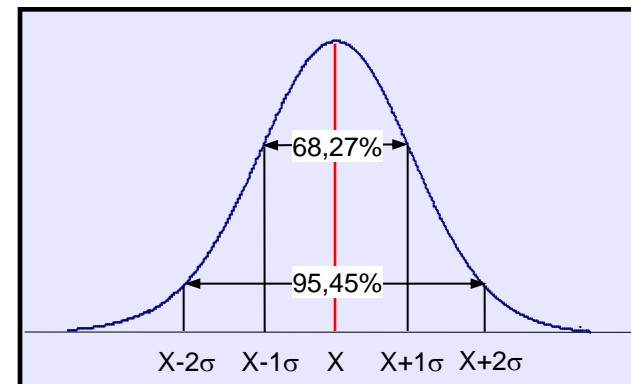
$$z - \text{score} = \frac{x - \bar{X}}{\delta_{PT}}$$

$x - \bar{X}$  ... Differenz Messwert zum Sollwert  
 $\delta_{PT}$  ... Standardabweichung für die Eignungsbeurteilung

Der z-Score wird mittels Korrekturfaktoren zum z(u)-Score modifiziert.

## Bewertung nach DIN 38402-45

- $|z(u)| \leq 2$ :            erfolgreich
- $|z(u)|$  2 bis  $< 3$ : fragwürdig
- $|z(u)| \geq 3$ :            unzureichend



Als Toleranzgrenze für eine „erfolgreiche“ Beurteilung wird  $|z(u)|=2$  vorgegeben.

# ζ-Score - Einschätzung Messunsicherheiten

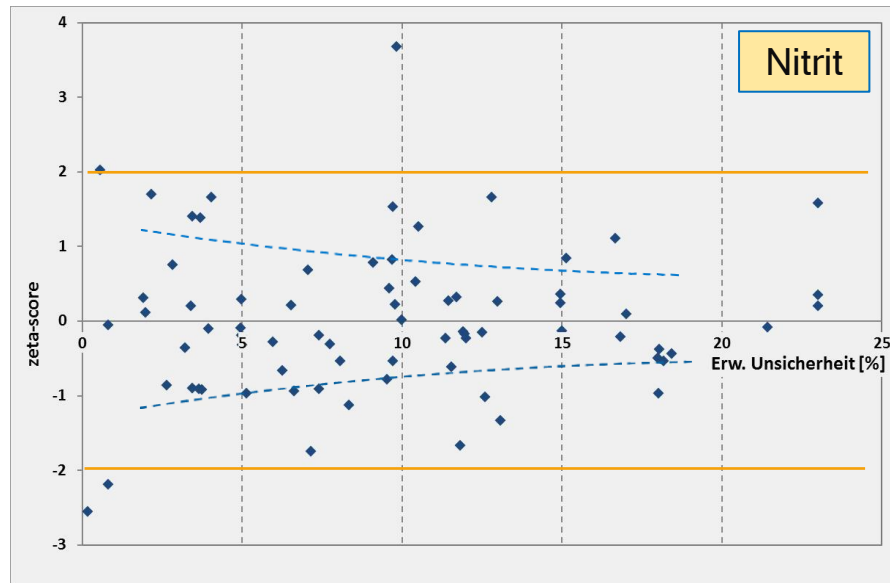
$$\zeta = \frac{x - \bar{X}}{\sqrt{u_{lab}^2 + u_{ref}^2}}$$

(DIN SO 13528)

$x - \bar{X}$  ... Differenz Messwert zum Sollwert  
 $u_{lab}$  ... Standardunsicherheit Teilnehmerergebnis  
 $u_{ref}$  ... Standardunsicherheit des Sollwertes

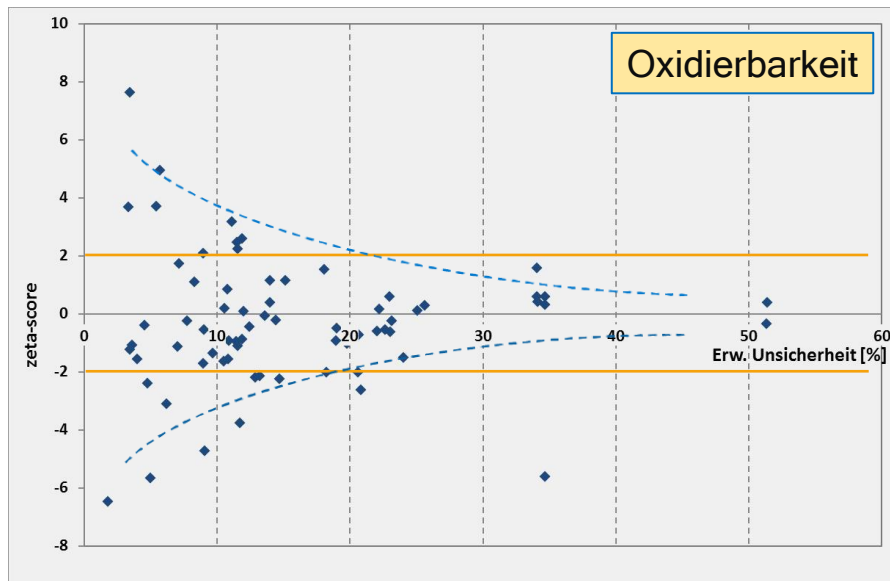
- Abweichung vom Sollwert im Verhältnis zur Unsicherheit
- Größenordnung wie der z(u)-score zu bewerten.
- Wird nicht zur Leistungsbeurteilung (Bewertung) herangezogen.

# $\zeta$ -Score (zeta) vs. Erweiterte Messunsicherheit [%]



RV-Ergebnisse NLGA, 2014

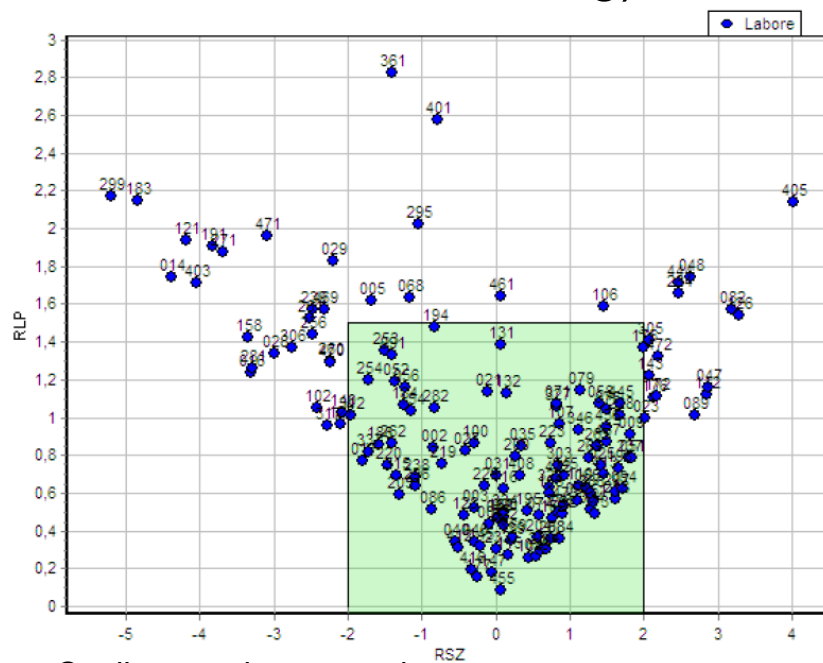
⇒ Der  $\zeta$ -score ist hervorragend geeignet, um die vom Labor angegebene Messunsicherheit in Kombination mit dem erreichten z(u)-score auf Plausibilität zu prüfen.



# Kombinationsscores - langfristige Beurteilung von RV-Ergebnissen

**RSZ** = standardisierte Summe aller z(u)-Scores (systematische Abweichung)

**RLP** = relative Leistungsfähigkeit (mittlere Abweichung im Verhältnis zur Vergleichsstandardabweichung)



Quelle: [www.lanuv.nrw.de](http://www.lanuv.nrw.de)



Vielen Dank für Ihre Aufmerksamkeit !

Kontakt:

Niedersächsisches Landesgesundheitsamtes  
Standort Aurich  
- Ringversuche -  
Tel.: 04941-9171-0  
E-Mail: [ringversuch.aurich@nlga.niedersachsen.de](mailto:ringversuch.aurich@nlga.niedersachsen.de)