

## Methoden zur Bestimmung der Referenzdaten - Fact Sheet

### Allgemeines

Bei den von der esz AG organisierten Eignungsprüfungen stehen vier verschiedene Methoden zur Bestimmung der Referenzdaten zur Verfügung.

- Methode A:  
Messtechnische Bestimmung der Referenzdaten anhand eines Referenzlabors
- Methode B:  
Gewichteter Mittelwert nach Cox
- Methode C:  
Robuste Abwandlung des arithmetischen Mittelwertes nach DIN ISO 13528:2020, Anhang C.3
- Methode D:  
Potenz-moderierter Mittelwert nach Pommé und Keightley

Bei den Methoden B, C und D handelt es sich um statistische Methoden, die unter der Berücksichtigung von Teilnehmerdaten die Referenzdaten berechnen. Eine ausführliche Analyse dieser Methoden, sowie ein Vergleich anhand von Beispieldatensätzen und weiterführende Informationen zur Leistungsbewertung, können dem White Paper „Eignungsprüfungen zwischen Kalibrierlaboratorien - Statistische Auswertung von Vergleichsmessungen“ entnommen werden.

### Anmerkungen

- Standardmäßig kommen Methode A oder Methode D zum Einsatz.
- Die geplante Methode zur Bestimmung der Referenzdaten wird im technischen Protokoll der Eignungsprüfung dokumentiert.
- Die tatsächlich angewandte Methode zur Bestimmung der Referenzdaten wird im Bericht der Eignungsprüfung dokumentiert.
- Wenn die Methoden B, C oder D geplant sind, kann es in der tatsächlichen Auswertung, bspw. aufgrund von Inkonsistenzen im Datensatz, nötig sein, auf Methode A auszuweichen.
- Jegliche Änderungen in der Auswertung werden den Teilnehmern mitgeteilt.
- Bei Eignungsprüfungen im akkreditierten Bereich können bei Anwendung der Methoden B, C oder D für die Bestimmung der Referenzdaten nur Daten von akkreditierten Teilnehmern herangezogen werden, die eine Messunsicherheit angeben, die größer ist als deren CMC.
- Daten die als Ausreißer identifiziert werden, werden bei der Bestimmung der Referenzdaten entsprechend der gewählten Methode behandelt. Die Leistungsbewertung ist i.d.R. unabhängig davon auch für Ausreißer durchführbar.

## **Methode A**

### **Quelle**

- DIN EN ISO/IEC 17043:2023
- DIN ISO 13528:2020

### **Voraussetzungen**

- Das Referenzlabor ist ein Nationalinstitut oder ein akkreditiertes Kalibrierlabor.
- Die zu erwartende Messunsicherheit des Referenzlabors muss hinreichend klein sein.
- Um Korrelationen zu vermeiden wird nach Möglichkeit ein Referenzlabor gewählt, auf das das zu erwartende Teilnehmerfeld nicht rückgeführt ist.
- Weitere messgrößenspezifische Voraussetzungen werden bei Bedarf berücksichtigt.

### **Charakteristik**

- Es handelt sich um eine messtechnische Definition der Referenzdaten.

### **Eigenschaften des Referenzwertes**

- n.a.

### **Eigenschaften der Referenzunsicherheit**

- n.a.

### **Leistungsbewertung**

- Korrelationen mit dem Referenzlabor, bspw. aufgrund der Rückführung, können i.d.R. nicht berücksichtigt werden.

### **Anwendbarkeit und Einschränkungen**

- Die Methode ist anwendbar für alle Eignungsprüfungen, sofern ein geeignetes Referenzlabor vorliegt.
- Neben der genannten Rückführungsthematik gibt es keine weiteren Anforderungen an den Datensatz der Teilnehmer.
- Die Methode ist unabhängig von der Teilnehmerzahl und ist auch für bilaterale Eignungsprüfungen geeignet.

## Methode B

### Quelle

- „Evaluation of key comparison data“ (M. Cox, Metrologia, 2002)

### Voraussetzungen

- Es werden mindestens vier akkreditierte Teilnehmer benötigt, die im Rahmen der Eignungsprüfung eine Messunsicherheit angeben, die größer ist als ihr CMC.
- Es muss ein normalverteilter Datensatz vorliegen, d.h. der  $\chi^2$ -Test muss positiv sein. Ggf. müssen dafür Ausreißer aus dem Datensatz eliminiert werden.

### Charakteristik

- Die Referenzdaten werden auf der Basis von Teilnehmerdaten statistisch bestimmt.
- Die Methode ist unsicherheitsbasiert und zeichnet sich durch eine individuelle Gewichtung der Teilnehmerwerte aus.

### Eigenschaften des Referenzwertes

- Die Gewichtung der Teilnehmerwerte erfolgt über die dazugehörige Messunsicherheit.
- Messwerte mit kleiner Messunsicherheit können den Referenzwert maßgeblich beeinflussen.

### Eigenschaften der Referenzunsicherheit

- Die Referenzunsicherheit ist immer kleiner als die kleinste berichtete Messunsicherheit eines Teilnehmers, der in die Berechnung der Referenzdaten einfließt.
- Die Referenzunsicherheit ist umso kleiner, je größer die Datengrundlage ist, die für die Berechnung der Referenzdaten vorliegt.
- Die Streuung der Messwerte wird in der Referenzunsicherheit nicht berücksichtigt.

### Umgang mit Ausreißern

- Anhand des  $\chi^2$ -Tests werden Ausreißer eliminiert und fließen nicht in die Bestimmung der Referenzdaten ein.

### Leistungsbewertung

- Für Teilnehmer, die in die Bestimmung der Referenzdaten einfließen wird im  $E_n$ -Wert die Korrelation mit den Referenzdaten berücksichtigt.

### Anwendbarkeit und Einschränkungen

- Benötigt werden mindestens vier, besser jedoch sieben akkreditierte Teilnehmer.
- Die Methode ist nicht anwendbar für bilaterale Vergleichsmessungen.
- Die Methode ist u.a. defizitär bei inkonsistenten Datensätzen mit einer Messwertstreuung, die größer ist als die berichteten Messunsicherheiten.

## Methode C

### Quelle

- Robuste Abwandlung des arithmetischen Mittelwertes (DIN ISO 13528:2020, Anhang C.3)

### Voraussetzungen

- Es werden mindestens vier akkreditierte Teilnehmer benötigt, die im Rahmen der Eignungsprüfung eine Messunsicherheit angeben, die größer ist als ihr CMC.
- Der Ausreißeranteil darf bei maximal 25 % liegen.

### Charakteristik

- Die Referenzdaten werden auf der Basis von Teilnehmerdaten statistisch bestimmt.
- Die Methode ist streuungsbasiert und zeichnet sich durch eine einheitliche Gewichtung der Teilnehmerwerte aus.

### Eigenschaften des Referenzwertes

- Der Referenzwert ist robust bzgl. Ausreißern und entspricht dem arithmetischen Mittelwert, falls keine Ausreißer vorliegen.
- Die Teilnehmerdaten fließen mit einer einheitlichen Gewichtung in den Referenzwert ein.
- Der Referenzwert ist unabhängig von den berichteten Messunsicherheiten der Teilnehmer.

### Eigenschaften der Referenzunsicherheit

- Die Referenzunsicherheit basiert auf der Streuung der Teilnehmerdaten und ist proportional zur empirischen Standardabweichung.
- Die Referenzunsicherheit ist umso kleiner, je größer die Datengrundlage ist, die für die Berechnung der Referenzdaten vorliegt.
- Die Messunsicherheit der Teilnehmer wird standardmäßig nicht berücksichtigt. In einer Abwandlung der Methode ist die Berücksichtigung jedoch möglich.

### Umgang mit Ausreißern

- Ausreißer werden nicht eliminiert, sondern durch den Algorithmus modifiziert.

### Leistungsbewertung

- Für Teilnehmer, die in die Bestimmung der Referenzdaten einfließen wird im  $E_n$ -Wert die Korrelation mit den Referenzdaten berücksichtigt.

### Anwendbarkeit und Einschränkungen

- Benötigt werden mindestens vier, besser jedoch sieben akkreditierte Teilnehmer.
- Die Methode ist nicht anwendbar für bilaterale Vergleichsmessungen.
- Die Methode ist u.a. defizitär bei Datensätzen mit einer Messwertstreuung, die kleiner ist als die berichteten Messunsicherheiten.

## Methode D

### Quelle

- „Determination of a reference value and its uncertainty through a power-moderated mean“  
(S. Pommé und J. Keightley, Metrologia, 2015)

### Voraussetzungen

- Es werden mindestens vier akkreditierte Teilnehmer benötigt, die im Rahmen der Eignungsprüfung eine Messunsicherheit angeben, die größer ist als ihr CMC.

### Charakteristik

- Die Referenzdaten werden auf der Basis von Teilnehmerdaten statistisch bestimmt.
- Die Methode ist unsicherheits- und streuungsbasiert und zeichnet sich durch eine vereinheitlichte individuelle Gewichtung der Teilnehmerwerte aus.
- Die Methode passt sich dem Datensatz an und kann die Methoden B und C abbilden und dazwischen interpolieren.

### Eigenschaften des Referenzwertes

- Der Referenzwert ist robust bzgl. Ausreißern und Messwerten mit kleiner Messunsicherheit.

### Eigenschaften der Referenzunsicherheit

- Die Referenzunsicherheit berücksichtigt sowohl die Streuung der Messwerte, als auch die berichteten Messunsicherheiten und moderiert je nach Datensatz zwischen den beiden Komponenten.
- Unzulänglichkeiten in der Datengrundlage, bspw. aufgrund einer geringen Teilnehmerzahl, werden berücksichtigt.

### Umgang mit Ausreißern

- Ausreißer mit  $E_n > 1,25$  werden eliminiert und fließen nicht in die Bestimmung der Referenzdaten ein.

### Leistungsbewertung

- Für Teilnehmer, die in die Bestimmung der Referenzdaten einfließen wird im  $E_n$ -Wert die Korrelation mit den Referenzdaten berücksichtigt.

### Anwendbarkeit und Einschränkungen

- Benötigt werden mindestens vier, besser jedoch sieben akkreditierte Teilnehmer.
- Die Methode ist nicht anwendbar für bilaterale Vergleichsmessungen.
- Die Methode ist anwendbar bei einer Vielzahl von Datensätzen, da Inkonsistenzen berücksichtigt werden können und dynamisch zwischen den Grenzfällen von Methode B und C interpoliert wird.