

DIN ISO 5725-2



ICS 03.120.30; 17.020

Ersatz für
DIN ISO 5725-2:2002-12 und
DIN 38402-42:2005-09

**Genauigkeit (Richtigkeit und Präzision) von Messverfahren und
Messergebnissen –
Teil 2: Grundlegende Methode für die Ermittlung der Wiederhol- und
Vergleichpräzision eines vereinheitlichten Messverfahrens
(ISO 5725-2:2019);
Text Deutsch und Englisch**

Accuracy (trueness and precision) of measurement methods and results –
Part 2: Basic method for the determination of repeatability and reproducibility of a
standard measurement method (ISO 5725-2:2019);
Text in German and English

Exactitude (justesse et fidélité) des résultats et méthodes de mesure –
Partie 2: Méthode de base pour la détermination de la répétabilité et de la reproductibilité
d'une méthode de mesure normalisée (ISO 5725-2:2019);
Texte en allemand et anglais

Gesamtumfang 157 Seiten

DIN-Normenausschuss Qualitätsmanagement, Statistik und Zertifizierungsgrundlagen (NQSZ)
DIN-Normenausschuss Wasserwesen (NAW)

Inhalt

	Seite
Nationales Vorwort	6
Nationaler Anhang NA (informativ) Literaturhinweise	9
Vorwort	10
Einleitung	11
1 Anwendungsbereich	12
2 Normative Verweisungen	12
3 Begriffe	13
4 Symbole	13
5 Schätzwerte der Parameter im Grundmodell	15
6 Anforderungen an ein Präzisionsexperiment	16
6.1 Layout des Experiments	16
6.2 Rekrutierung der Labors	17
6.3 Vorbereitung des Materials	17
7 An einem Präzisionsversuch beteiligtes Personal	18
7.1 Versuchsgremium	18
7.2 Statistische Funktionen	19
7.3 Führungsaufgaben	19
7.4 Laborverantwortliche	20
7.5 Bearbeiter	21
8 Statistische Analyse des Präzisionsversuchs	21
8.1 Vorbemerkungen	21
8.2 Tabellarische Zusammenstellung der Ergebnisse sowie angewendete Darstellungsweise	22
8.2.1 Zellen	22
8.2.2 Überflüssige Daten	22
8.2.3 Fehlende Daten	22
8.2.4 Ausreißer	22
8.2.5 Ausreißerlabors	23
8.2.6 Fehlerhafte Daten	23
8.2.7 Prüfergebnisse ausgewogener einheitlicher Experimente	23
8.2.8 Zusammenstellung der Daten und Zwischenwerte	23
8.2.9 Ursprüngliche Ermittlungsergebnisse	23
8.2.10 Arithmetische Zellmittelwerte in einer Zelle (Formblatt B in Bild 2)	24
8.2.11 Maße für die Streuung innerhalb einer Zelle (Formblatt C in Bild 2)	24
8.2.12 Korrigierte oder als Ausreißer abgelehnte Daten	24
8.3 Ergebnisprüfung bezüglich Vereinbarkeit und Ausreißer	24
8.3.1 Ansätze für die Untersuchung von Daten	24
8.3.2 Grafische Vereinbarkeitsprüfungen	25
8.3.3 Numerische Methode der Ausreißerprüfung	28
8.3.4 Cochran-Test	28
8.3.5 Grubbs-Test	30
8.3.6 Wiederholte Prüfung für Mittelwert- und Datenpunktausreißer	32
8.3.7 Alternative Ausreißer-Prüf- und Testverfahren	32
8.4 Berechnung des Gesamtmittelwerts und der Varianzen	33
8.4.1 Auswertungsverfahren	33
8.4.2 Grunddaten	33
8.4.3 Nichtleere Zellen	33
8.4.4 Berechnung des Gesamtmittelwerts $\hat{m}$	33
8.4.5 Berechnung der Varianzen	33
8.4.6 Alternative Berechnungsverfahren für Varianzen	34
8.4.7 Abhängigkeit der Varianzen von m	35

8.5	Feststellen einer funktionalen Beziehung zwischen Präzisionswerten s und dem arithmetischen Gesamtmittelwert m	35
8.5.1	Wahl der funktionalen Beziehung	35
8.5.2	Anpassungsbeziehungen I und II	36
8.5.3	Anpassung der Beziehung III	37
8.5.4	Anpassung der Beziehung IV	38
8.6	Statistische Analyse als schrittweises Verfahren	40
8.7	Bericht an das Versuchsgremium und vom Versuchsgremium zu treffende Entscheidungen	42
8.7.1	Bericht durch die statistische Fachkraft	42
8.7.2	Vom Versuchsgremium zu treffende Entscheidungen	45
8.7.3	Abschlussbericht	46
9	Statistische Tabellen	46
Anhang A (informativ)	Anzahl der für eine Präzisionsschätzung erforderlichen Labors	51
Anhang B (informativ)	Berechnungsvarianten für Varianzkomponenten	54
B.1	Berechnung auf der Basis einer ANOVA-Tabelle der einfachen Varianzanalyse	54
B.2	Berechnung des Restricted Maximum Likelihood-Verfahrens (REML)	54
Anhang C (informativ)	Beispiele der statistischen Analyse von Präzisionsexperimenten	56
C.1	Beispiel 1: Ermittlung des Schwefelgehalts von Kohle (mehrere Merkmalsniveaus ohne fehlende Daten oder Ausreißerdaten)	56
C.1.1	Hintergrund	56
C.1.2	Ursprüngliche Daten	56
C.1.3	Berechnung der Zellenmittelwerte, $\bar{y}_{ij}$	57
C.1.4	Berechnung der Standardabweichungen, s_{ij}	57
C.1.5	Vereinbarkeits- und Ausreißerprüfung	57
C.1.6	Berechnung von $\hat{m}_j$, s_{Rj} und s_{Rj}	63
C.1.7	Abhängigkeit der Präzision von m	64
C.1.8	Schlussfolgerungen	64
C.1.9	Alternative Berechnung	65
C.2	Beispiel 2: Erweichungspunkt von Pech (mehrere Merkmalsniveaus mit fehlenden Daten)	65
C.2.1	Hintergrund	65
C.2.2	Ursprüngliche Daten	66
C.2.3	Arithmetische Mittelwerte einer Zelle	66
C.2.4	Absolute Differenzen innerhalb der Zellen	69
C.2.5	Untersuchung auf Vereinbarkeit und Ausreißer	70
C.2.6	Berechnung von $\hat{m}_j$, s_{Rj} und s_{Rj}	70
C.2.7	Abhängigkeit der Präzision von m	71
C.2.8	Zusammenfassung	71
C.2.9	Alternative Berechnung	71
C.3	Beispiel 3: Thermometrische Titration von Kreosotöl (mehrere Merkmalsniveaus mit Ausreißerwerten)	72
C.3.1	Hintergrund	72
C.3.2	Ursprüngliche Daten	72
C.3.3	Arithmetische Mittelwerte einer Zelle	73
C.3.4	Absolute Differenzen innerhalb der Zellen	74
C.3.5	Vereinbarkeits- und Ausreißerprüfung	74
C.3.6	Berechnung von $\hat{m}_j$, s_{Rj} und s_{Rj}	77
C.3.7	Abhängigkeit der Präzision von m	78
C.3.8	Alternative Berechnung	78
Anhang D (informativ)	Berechnung kritischer Werte und Indikatoren	80
D.1	Berechnung kritischer Werte für den Cochran-Test	80
D.2	Berechnung kritischer Werte für den Grubbs-Test	80
D.2.1	Einzelner Ausreißer	80
D.2.2	Zwei Ausreißer	81
D.3	Berechnung der Indikatoren für Mandels Vereinbarkeits-Prüfgrößen h und k	81
D.3.1	Mandels Vereinbarkeits-Prüfgröße h	81

D.3.2 Mandels Vereinbarkeits-Prüfgröße k	82
Literaturhinweise	83

Bilder

Bild 1 — Anwendungs-Fragebogen für einen Ringversuch	18
Bild 2 — Empfohlene Formblätter für die Zusammenstellung von Ergebnissen zur Analyse	26
Bild 3 — Flussdiagramm der Hauptschritte der statistischen Analyse	45
Bild A.1 — Betrag, um den s_r erwartungsgemäß von dem wahren Wert innerhalb eines Wahrscheinlichkeitsniveaus von 95 % abweichen wird	52
Bild A.2 — Betrag, um den s_R erwartungsgemäß von dem wahren Wert innerhalb einer Wahrscheinlichkeit von 95 % abweichen wird	53
Bild C.1 — Schwefelgehalt von Kohle, Probe 1	59
Bild C.2 — Schwefelgehalt von Kohle, Probe 2	60
Bild C.3 — Schwefelgehalt von Kohle, Probe 3	61
Bild C.4 — Schwefelgehalt von Kohle, Probe 4	62
Bild C.5 — Erweichungstemperatur von Pech — Zellenmittelwerte	68
Bild C.6 — Erweichungstemperatur von Pech — Absolute Differenzen innerhalb der Zellen	69
Bild C.7 — Titration von Kreosotöl — Mandels Vereinbarkeits-Prüfgröße h zwischen den Labors, gruppiert für die Labors dargestellt	75
Bild C.8 — Titration von Kreosotöl — Mandels Vereinbarkeits-Prüfgröße k innerhalb der Labors, gruppiert für die Labors dargestellt	76
Bild C.9 — Grafische Darstellung von s_{rj} und s_{Rj} gegen $\hat{m}_j$, nach den Werten aus Tabelle C.18, welche die nach 8.5 an diese Daten angepassten funktionalen Beziehungen I bis IV abbildet	79

Tabellen

Tabelle 1 — Beziehung I, $s = bm$	39
Tabelle 2 — Beziehung II, $s = a + bm$	39
Tabelle 3 — Beziehung III, $s^2 = a_v^2 + (b_v m)^2$	40
Tabelle 4 — Beziehung IV, $\lg s = c + d \lg m$	40
Tabelle 5 — Kritische Werte für den Cochran-Test	46
Tabelle 6 — Kritische Werte für den Grubbs-Test	47
Tabelle 7 — Indikatoren für die Mandels Vereinbarkeits-Prüfgrößen h und k auf dem Signifikanzniveau von 1 %	49
Tabelle 8 — Indikatoren für die Mandels Vereinbarkeits-Prüfgrößen h und k auf dem Signifikanzniveau von 5 %	50
Tabelle A.1 — Werte der Schätzwertunsicherheit der Wiederhol- und Vergleichstandardabweichungen	52
Tabelle B.1 — Layout einer Einweg-ANOVA-Tabelle	54
Tabelle C.1 — Ursprungsdaten — Schwefelgehalt von Kohle (Massenanteil, %)	57
Tabelle C.2 — Zellenmittelwerte — Schwefelgehalt von Kohle (Massenanteil, %)	58
Tabelle C.3 — Standardabweichungen — Schwefelgehalt von Kohle (Massenanteil, %)	58
Tabelle C.4 — Anwendung des Grubbs-Tests auf Zellenmittelwerte	62
Tabelle C.5 — Berechnete Werte von $\hat{m}_j$, s_{rj} und s_{Rj} für den Schwefelgehalt von Kohle	64
Tabelle C.6 — REML-Schätzwerte von $\hat{m}_j$, s_{rj} und s_{Rj} für den Schwefelgehalt von Kohle	65
Tabelle C.7 — Ursprüngliche Daten — Erweichungstemperatur von Pech (°C)	66
Tabelle C.8 — Zellenmittelwerte — Erweichungstemperatur von Pech (°C)	67
Tabelle C.9 — Absolute Differenzen innerhalb der Zellen — Erweichungstemperatur von Pech (°C)	68
Tabelle C.10 — Werte der Cochran-Prüfgröße, C	69
Tabelle C.11 — Anwendung des Grubbs-Tests auf Zellenmittelwerte	70

Tabelle C.12 — Berechnete Werte von $\hat{m}_j$, s_{rj} und s_{Rj} für die Erweichungstemperatur von Pech .	71
Tabelle C.13 — REML-Schätzwerte von $\hat{m}_j$, s_{rj} und s_{Rj} für die Erweichungstemperatur von Pech .	72
Tabelle C.14 — Ursprüngliche Daten — Thermometrische Titration von Kreosotöl	73
Tabelle C.15 — Zellenmittelwerte — Thermometrische Titration von Teeröl	73
Tabelle C.16 — Zellenbereiche — Thermometrische Titration von Kreosotöl	73
Tabelle C.17 — Anwendung des Grubbs-Tests auf Zellenmittelwerte	76
Tabelle C.18 — Berechnete Werte von $\hat{m}_j$, s_{rj} und s_{Rj} für die thermometrische Titration von Kreosotöl	77
Tabelle C.19 — REML-Schätzwerte von $\hat{m}_j$, s_{rj} und s_{Rj} für die thermometrische Titration von Teeröl	78
Tabelle D.1 — Koeffizienten für die kritischen Werte des Grubbs-Tests	81

Nationales Vorwort

Dieses Dokument enthält die deutsche Übersetzung der Internationalen Norm ISO 5725-2:2019, die vom Technischen Komitee ISO/TC 69 „Applications of statistical methods“ erarbeitet wurde, dessen Sekretariat von ANSI (Vereinigte Staaten von Amerika) gehalten wird.

Das zuständige deutsche Normungsgremium ist der Arbeitsausschuss NA 147-00-02 AA „Angewandte Statistik“ im DIN-Normenausschuss Qualitätsmanagement, Statistik und Zertifizierungsgrundlagen (NQSZ).

Das zuständige Gremium des Mitträgers ist der Unterausschuss NA 119-01-03-06 UA „Qualitätssicherung“ des Arbeitsausschusses NA 119-01-03 AA „Wasseruntersuchung“ im DIN-Normenausschuss Wasserwesen (NAW).

Für die in diesem Dokument zitierten Dokumente wird im Folgenden auf die entsprechenden deutschen Dokumente hingewiesen:

ISO 3534-1	siehe	DIN ISO 3534-1
ISO 3534-2	siehe	DIN ISO 3534-2
ISO 5725-1	siehe	DIN ISO 5725-1
ISO 5725-3	siehe	DIN ISO 5725-3
ISO 5725-5	siehe	DIN ISO 5725-5
ISO 5725-6	siehe	DIN ISO 5725-6

Aktuelle Informationen zu diesem Dokument können über die Internetseiten von DIN (www.din.de) durch eine Suche nach der Dokumentennummer aufgerufen werden.

Bezeichnung des Verfahrens

Genauigkeit (Richtigkeit und Präzision) von Messverfahren und Messergebnissen — Teil 2: Grundlegende Methode für die Ermittlung der Wiederhol- und Vergleichpräzision eines vereinheitlichten Messverfahrens:

Verfahren DIN ISO 5725-2 — A 42

Es ist erforderlich, bei den Untersuchungen nach dieser Norm Fachleute oder Facheinrichtungen einzuschalten und bestehende Sicherheitsvorschriften zu beachten.

Bei Anwendung der Norm ist im Einzelfall je nach Aufgabenstellung zu prüfen, ob und inwieweit die Festlegung von zusätzlichen Randbedingungen erforderlich ist.

Die vorliegende Norm enthält das vom DIN-Normenausschuss Wasserwesen (NAW) und von der Wasserchemischen Gesellschaft — Fachgruppe in der Gesellschaft Deutscher Chemiker (GDCh) — gemeinsam erarbeitete Deutsche Einheitsverfahren zur Wasser-, Abwasser- und Schlammuntersuchung:

Genauigkeit (Richtigkeit und Präzision) von Messverfahren und Messergebnissen — Teil 2: Grundlegende Methode für die Ermittlung der Wiederhol- und Vergleichpräzision eines vereinheitlichten Messverfahrens (A 42).

Die als DIN-Normen veröffentlichten Deutschen Einheitsverfahren sind bei der Beuth Verlag GmbH einzeln oder zusammengefasst erhältlich. Außerdem werden die genormten Deutschen Einheitsverfahren in der Loseblattsammlung „Deutsche Einheitsverfahren zur Wasser-, Abwasser- und Schlammuntersuchung“ gemeinsam von der Beuth Verlag GmbH und der Wiley-VCH Verlag GmbH & Co. KGaA publiziert.

Normen oder Norm-Entwürfe mit dem Gruppentitel „*Deutsche Einheitsverfahren zur Wasser-, Abwasser- und Schlammuntersuchung*“ sind in folgende Gebiete (Haupttitel) aufgeteilt:

Allgemeine Angaben (Gruppe A)

Sensorische Verfahren (Gruppe B)

Physikalische und physikalisch-chemische Kenngrößen (Gruppe C)

Anionen (Gruppe D)

Kationen (Gruppe E)

Gemeinsam erfassbare Stoffgruppen (Gruppe F)

Gasförmige Bestandteile (Gruppe G)

Summarische Wirkungs- und Stoffkenngrößen (Gruppe H)

Mikrobiologische Verfahren (Gruppe K)

Testverfahren mit Wasserorganismen (Gruppe L)

Biologisch-ökologische Gewässeruntersuchung (Gruppe M)

Einzelkomponenten (Gruppe P)

Schlamm und Sedimente (Gruppe S)

Suborganismische Testverfahren (Gruppe T)

Über die bisher erschienenen Teile dieser Normen gibt die Geschäftsstelle des Normenausschusses Wasserwesen (NAW) im DIN Deutsches Institut für Normung e.V., Telefon 030 2601-2448, oder die Beuth Verlag GmbH, 10772 Berlin, Auskunft.

Änderungen

Gegenüber DIN ISO 5725-2:2002-12 und DIN 38402-42:2005-09 wurden folgende Änderungen vorgenommen:

Gegenüber DIN ISO 5725-2:2002-12:

- a) es ist zulässig, alternative genaue Untersuchungen und Prüfungen zur Erkennung von Ausreißern anzuwenden, solange die Leistung vergleichbar ist;
- b) es ist zulässig, verfügbare moderne statistische Verfahren zur Berechnung der relevanten Präzisions- und Richtigkeitseigenschaften anzuwenden;
- c) eine Anleitung zur Anzahl der erforderlichen Labors für eine Präzisionsstudie wurde eingefügt;
- d) Informationen zur Berechnung kritischer Werte wurden eingefügt.

Gegenüber DIN 38402-42:2005-09:

- e) die Normnummer und der Normtitel wurden geändert;
- f) der Inhalt von ISO 5725-2:2019 wurde unverändert übernommen. DIN 38402-42:2005-09 war inhaltlich bereits weitgehend an ISO 5725-2 angelehnt;
- g) weitere Änderungen siehe bei Änderungen gegenüber DIN ISO 5725-2:2002-12.

Frühere Ausgaben

DIN ISO 5725: 1981-11, 1988-04
DIN 38402-42: 1984-05, 2005-09
DIN ISO 5725-2: 2002-12

Nationaler Anhang NA (informativ)

Literaturhinweise

DIN ISO 3534-1, *Statistik — Begriffe und Formelzeichen — Teil 1: Wahrscheinlichkeit und allgemeine statistische Begriffe*

DIN ISO 3534-2, *Statistik — Begriffe und Formelzeichen — Teil 2: Angewandte Statistik*

DIN ISO 5725-1, *Genauigkeit (Richtigkeit und Präzision) von Meßverfahren und Meßergebnissen — Teil 1: Allgemeine Grundlagen und Begriffe*

DIN ISO 5725-3, *Genauigkeit (Richtigkeit und Präzision) von Messverfahren und Messergebnissen — Teil 3: Präzisionsmaße eines vereinheitlichten Messverfahrens unter Zwischenbedingungen*

DIN ISO 5725-5, *Genauigkeit (Richtigkeit und Präzision) von Messverfahren und Messergebnissen — Teil 5: Alternative Methoden für die Ermittlung der Präzision eines vereinheitlichten Messverfahrens*

DIN ISO 5725-6, *Genauigkeit (Richtigkeit und Präzision) von Messverfahren und Messergebnissen — Teil 6: Anwendung von Genauigkeitswerten in der Praxis*

Vorwort

ISO (die Internationale Organisation für Normung) ist eine weltweite Vereinigung nationaler Normungsinstitute (ISO-Mitgliedsorganisationen). Die Erstellung von Internationalen Normen wird üblicherweise von Technischen Komitees von ISO durchgeführt. Jede Mitgliedsorganisation, die Interesse an einem Thema hat, für welches ein Technisches Komitee gegründet wurde, hat das Recht, in diesem Komitee vertreten zu sein. Internationale staatliche und nichtstaatliche Organisationen, die in engem Kontakt mit ISO stehen, nehmen ebenfalls an der Arbeit teil. ISO arbeitet bei allen elektrotechnischen Normungsthemen eng mit der Internationalen Elektrotechnischen Kommission (IEC) zusammen.

Die Verfahren, die bei der Entwicklung dieses Dokuments angewendet wurden und die für die weitere Pflege vorgesehen sind, werden in den ISO/IEC-Direktiven, Teil 1 beschrieben. Im Besonderen sollten die für die verschiedenen ISO-Dokumentenarten notwendigen Annahmekriterien beachtet werden. Dieses Dokument wurde in Übereinstimmung mit den Gestaltungsregeln der ISO/IEC-Direktiven, Teil 2 erarbeitet (siehe www.iso.org/directives).

Es wird auf die Möglichkeit hingewiesen, dass einige Elemente dieses Dokuments Patentrechte berühren können. ISO und IEC sind nicht dafür verantwortlich, einige oder alle diesbezüglichen Patentrechte zu identifizieren. Details zu allen während der Entwicklung des Dokuments identifizierten Patentrechten finden sich in der Einleitung und/oder in der ISO-Liste der erhaltenen Patenterklärungen (siehe www.iso.org/patents).

Jeder in diesem Dokument verwendete Handelsname dient nur zur Unterrichtung der Anwender und bedeutet keine Anerkennung.

Für eine Erläuterung des freiwilligen Charakters von Normen, der Bedeutung ISO-spezifischer Begriffe und Ausdrücke in Bezug auf Konformitätsbewertungen sowie Informationen darüber, wie ISO die Grundsätze der Welthandelsorganisation (WTO, en: World Trade Organization) hinsichtlich technischer Handelshemmnisse (TBT, en: Technical Barriers to Trade) berücksichtigt, siehe: www.iso.org/iso/foreword.html.

Dieses Dokument wurde vom Technischen Komitee ISO/TC 69, *Applications of statistical methods*, Unterkomitee SC 6, *Measurement methods and results* erarbeitet.

Diese zweite Ausgabe ersetzt die erste Ausgabe (ISO 5725-2:1994), die technisch überarbeitet wurde. Sie enthält auch die Technische Berichtigung ISO 5725-2:1994/Cor 1:2002.

Die wesentlichen Änderungen im Vergleich zur Vorgängerausgabe sind folgende:

- es ist zulässig, alternative genaue Untersuchungen und Prüfungen zur Erkennung von Ausreißern anzuwenden, solange die Leistung vergleichbar ist;
- es ist zulässig, verfügbare moderne statistische Verfahren zur Berechnung der relevanten Präzisions- und Richtigkeitseigenschaften anzuwenden;
- eine Anleitung zur Anzahl der erforderlichen Labors für eine Präzisionsstudie wurde eingefügt;
- Informationen zur Berechnung kritischer Werte wurden eingefügt.

Eine Auflistung aller Teile der Normenreihe ISO 5725 kann auf der ISO-Internetseite abgerufen werden.

Rückmeldungen oder Fragen zu diesem Dokument sollten an das jeweilige nationale Normungsinstitut des Anwenders gerichtet werden. Eine vollständige Auflistung dieser Institute ist unter www.iso.org/members.html zu finden.

Einleitung

ISO 5725 verwendet die Begriffe „Richtigkeit“ und „Präzision“ zur Beschreibung der Genauigkeit eines Messverfahrens. „Richtigkeit“ bezieht sich auf den Grad der Übereinstimmung zwischen dem arithmetischen Mittelwert einer großen Anzahl von Prüfergebnissen und dem wahren oder akzeptierten Referenzwert. „Präzision“ bezieht sich auf den Grad der Übereinstimmung zwischen Prüfergebnissen.

Allgemeine Erwägungen zu diesen Größenwerten sind in ISO 5725-1 enthalten und werden daher nicht in diesem Dokument wiederholt. ISO 5725-1 sollte in Verbindung mit allen anderen Teilen von ISO 5725 einschließlich des vorliegenden Teils gelesen werden, weil dieser Teil die zugrundeliegenden Definitionen und allgemeinen Grundsätze enthält.

Dieses Dokument befasst sich ausschließlich mit der Schätzung der Wiederholstandardabweichung und der Vergleichstandardabweichung auf der Grundlage eines laborübergreifenden Konzepts, bei dem jedes Labor eine Reihe unabhängiger Messungen derselben Probe unter Wiederholbedingungen durchführt. Es bestehen andere Konzepte (wie etwa geschachtelte, faktorielle oder Split-Level-Experimente), die für die Schätzung der Präzision angewendet werden können. Diese Konzepte werden nicht in diesem Dokument behandelt, sondern sind vielmehr Gegenstand anderer Teile von ISO 5725. Darüber hinaus betrachtet dieses Dokument keine anderen Präzisionsmesswerte zwischen den beiden Hauptmesswerten. Diese werden in ISO 5725-3 behandelt.

Die aus einem zur Schätzung der Präzision durchgeführten Experiment erhaltenen Daten werden unter bestimmten Umständen auch zur Schätzung der Richtigkeit verwendet und können zur Bewertung der Messunsicherheit genutzt werden. Die Schätzung der Richtigkeit wird in diesem Dokument nicht behandelt. Alle Aspekte der Schätzung der Richtigkeit werden in ISO 5725-4 behandelt. Die Bewertung der Messunsicherheit unter Verwendung laborübergreifender Schätzungen der Richtigkeit und Präzision ist Gegenstand von ISO 21748.

Anhang C enthält praktische Beispiele zur Schätzung der Präzision von Messverfahren durch Experimente. Ausgearbeitete Beispiele werden angeführt, um ausgewogene einheitliche Sätze von Prüfergebnissen zu zeigen, auch wenn in einem Beispiel eine variable Anzahl von Replikaten pro Zelle angegeben wurde (unausgewogenes Konzept) und in einem anderen einige Daten fehlten. Dies liegt daran, dass sich ein als ausgewogen konzipiertes Experiment als unausgewogen erweisen kann. Fastausreißer und Ausreißer werden ebenfalls berücksichtigt.

1 Anwendungsbereich

1.1 Dieses Dokument

- erläutert ausführlich die allgemeinen Grundsätze für die Gestaltung von Experimenten zur numerischen Schätzung der Präzision von Messverfahren mittels Ringversuch;
- bietet eine ausführliche praktische Beschreibung des grundlegenden Verfahrens der routinemäßigen Anwendung in der Schätzung der Präzision von Messverfahren;
- bietet eine Anleitung für alle mit der Gestaltung und Durchführung der Prüfungen zur Schätzung der Präzision oder Analyse ihrer Ergebnisse befassten Personen.

ANMERKUNG Abwandlungen dieses grundlegenden Verfahrens für bestimmte Zwecke werden in anderen Teilen von ISO 5725 behandelt.

1.2 Dieses Dokument behandelt ausschließlich Messverfahren, die Messungen auf einer kontinuierlichen Skala erzielen und einen einzigen Wert als Prüfergebnis ergeben, auch wenn dieser einzelne Wert das Ergebnis der Berechnung aus einer Reihe von Beobachtungen sein kann.

1.3 Dieses Dokument geht davon aus, dass bei der Gestaltung und Durchführung des Präzisionsexperiments alle in ISO 5725-1 festgelegten Grundsätze beachtet werden. Das grundlegende Verfahren verwendet die gleiche Anzahl an Prüfergebnissen in jedem Labor, wobei jedes Labor die gleichen Niveaus der Prüfproben analysiert, d. h. in einem ausgewogenen Experiment auf einheitlicher Ebene. Das grundlegende Verfahren gilt für Verfahren, die genormt sind und regelmäßig in einer Reihe von Labors angewendet werden.

1.4 Das statistische Modell nach ISO 5725-1:1994, Abschnitt 5, wird als geeignete Grundlage für die Deutung und Auswertung der Prüfergebnisse angenommen, deren Verteilung näherungsweise eine Normalverteilung ist.

1.5 Das in diesem Dokument beschriebene grundlegende Verfahren schätzt (üblicherweise) die Präzision eines Messverfahrens:

- a) wenn die Wiederhol- und Vergleichstandardabweichungen nach ISO 5725-1 bestimmt werden müssen;
- b) wenn die zu verwendenden Materialien homogen sind oder die Auswirkungen der Inhomogenität in die Präzisionswerte einbezogen werden können und
- c) wenn die Anwendung eines ausgewogenen Versuchsplans auf einheitlichem Merkmalsniveau annehmbar ist.

1.6 Der gleiche Ansatz kann angewendet werden, um eine vorläufige Schätzung der Präzision für Messverfahren vorzunehmen, die noch nicht genormt sind oder nicht regelmäßig angewendet werden.

2 Normative Verweisungen

Die folgenden Dokumente werden im Text in solcher Weise in Bezug genommen, dass einige Teile davon oder ihr gesamter Inhalt Anforderungen des vorliegenden Dokuments darstellen. Bei datierten Verweisungen gilt nur die in Bezug genommene Ausgabe. Bei undatierten Verweisungen gilt die letzte Ausgabe des in Bezug genommenen Dokuments (einschließlich aller Änderungen).

ISO 3534-1, *Statistics — Vocabulary and symbols — Part 1: Probability and general statistical terms*

ISO 3534-2, *Statistics — Vocabulary and symbols — Part 2: Applied statistics*

ISO 3534-3, *Statistics — Vocabulary and symbols — Part 3: Design of experiments*

ISO 5725-1, *Accuracy (trueness and precision) of measurement methods and results — Part 1: General principles and definitions*

3 Begriffe

Für die Anwendung dieses Dokuments gelten die Begriffe nach ISO 3534-1, ISO 3534-2, ISO 3534-3 und ISO 5725-1.

ISO und IEC stellen terminologische Datenbanken für die Verwendung in der Normung unter den folgenden Adressen bereit:

— ISO Online Browsing Platform: verfügbar unter <https://www.iso.org/obp>

— IEC Electropedia: verfügbar unter <http://www.electropedia.org/>

4 Symbole

α	Wahrscheinlichkeit in Verbindung mit einem kritischen Wert einer Prüfgröße, auch bezeichnet als Signifikanzniveau
a	Achsenabschnitt in der Beziehung $s = a + bm$
a_v	Achsenabschnittparameter in der Beziehung $s_f^2 = a_v^2 + (b_v m)^2$
A	Faktor zur Berechnung der Unsicherheit einer Schätzung
b	Steigung in der Beziehung $s = a + bm$
b_v	Steigungsparameter in der Beziehung $s_f^2 = a_v^2 + (b_v m)^2$
B	Laboranteil des Bias unter Wiederholbedingungen
c	Achsenabschnitt in der Beziehung $\lg s = c + d \lg m$
C, C', C''	Prüfgrößen
$C_{\text{crit}}, C'_{\text{crit}}, C''_{\text{crit}}$	Kritische Werte für statistische Prüfungen
d	Steigung in der Beziehung $\lg s = c + d \lg m$
e	Komponente in einem Prüfergebnis, die für die zufällige Messabweichung in jedem Prüfergebnis steht
G	Grubbs-Prüfgröße
h	Vereinbarkeits-Prüfgröße zwischen den Labors nach Mandel
k	Vereinbarkeits-Prüfgröße innerhalb eines Labors nach Mandel
$L(\theta)$	log-Wahrscheinlichkeit für Varianzkomponenten θ
m	Gesamtmittelwert des untersuchten Merkmals; Niveau
$\hat{m}$	Schätzwert des Gesamtmittelwerts des untersuchten Merkmals
$\mathbf{M}$	In der REML-Schätzung verwendete Umwandlungsmatrix

N	Anzahl der Iterationen
n	Anzahl der in einem Labor auf einem Niveau (d. h. pro Zelle) erzielten Prüfergebnisse
n_j	Gesamtzahl der auf Niveau j im Ringversuch erzielten Prüfergebnisse
p	Anzahl der am Ringversuch teilnehmenden Labors
P	Wahrscheinlichkeit
q	Niveauanzahl der Prüfeigenschaft im Ringversuch
r	Wiederholgrenze
R	Vergleichgrenze
s	Schätzwert der Standardabweichung
$\hat{s}$	Vorhergesagte Standardabweichung
T	Gesamtwert oder Summe eines Ausdrucks
t	Anzahl der Prüfobjekte oder Gruppen
$\mathbf{V}(\boldsymbol{\theta})$	in der REML-Schätzung angewendete Kovarianz-Matrix
W	Zur Berechnung einer gewichteten Regression angewendeter Gewichtungsfaktor
w	Zur Berechnung eines gewichteten Mittelwerts verwendeter Gewichtungsfaktor
x	Für den Grubbs-Test verwendete Größe
$\mathbf{X}$	Gestaltungsmatrix für REML-Schätzungen
y	Prüfergebnis
$\bar{y}$	Gesamtmittelwert der Prüfergebnisse
$\mathbf{Y}$	Vektor aller Beobachtungen auf einem Niveau j
$\boldsymbol{\theta}$	Vektor der in der REML-Schätzung verwendeten Varianzkomponenten
μ	Wahrer Wert oder akzeptierter Referenzwert einer Prüfeigenschaft
σ	Wahrer Wert einer Standardabweichung

Indizes

i	Bezeichner eines bestimmten Labors, Index für die Summierung (Anhang A)
j	Bezeichner eines bestimmten Niveaus, Index für die Summierung (Anhang A)
k	Bezeichner für ein bestimmtes Prüfergebnis in einem Labor i im Niveau j
L	Zwischen Labors (laborübergreifend)
P	Wahrscheinlichkeit

r	Wiederhol-
R	Vergleich-
REML	Schätzwert aus einer Berechnung der eingeschränkten maximalen Wahrscheinlichkeit
v	In der Berechnung einer Beziehung zwischen dem Mittelwert und der kombinierten Varianz verwendete Ausdrücke (siehe 8.5.1.3, Beziehung III)
W	Innerhalb des Labors (laborintern)
1, 2, 3, ...	Prüfergebnisse werden in der Reihenfolge, in der sie erhalten werden, nummeriert; in anderen Fällen (Labors) werden beliebige Bezeichner verwendet
(1), (2), (3), ...	Für Prüfergebnisse, (1), (2) ... wird die Bezeichnung der Statistik 1., 2. ... usw. Ordnung verwendet, d. h. 1., 2.... usw. Wert nummeriert in aufsteigender Größenreihenfolge

5 Schätzwerte der Parameter im Grundmodell

5.1 Die in diesem Dokument beschriebenen Verfahren basieren auf dem statistischen Modell nach ISO 5725-1:1994, Abschnitt 5, und der Ausarbeitung in ISO 5725-1:1994, 1.2. Diese Verfahren basieren insbesondere auf ISO 5725-1:1994, Abschnitt 5, Gleichungen (2) bis (6).

Das Modell ist

$$y = m + B + e$$

Dabei ist für das jeweilige geprüfte Material

m der Gesamtmittelwert (Erwartungswert);

B die systematische Abweichungskomponente des Labors unter Wiederholbedingungen;

e die zufällige Abweichung, die bei jeder Messung unter Wiederholbedingungen vorkommt.

ANMERKUNG Der Laboranteil des Bias B repräsentiert die Abweichung eines Labormittelwerts von dem Gesamtmittelwert m .

5.2 ISO 5725-1:1994, Gleichungen (2) bis (6), sind mittels der wahren Standardabweichungen der betrachteten Grundgesamtheiten ausgedrückt. In der Praxis sind die wirklichen Werte dieser (wahren) Standardabweichungen unbekannt. Deshalb müssen Schätzwerte für die Präzisionswerte aufgrund einer vergleichsweise kleinen Stichprobe aller in Frage kommenden Labors ermittelt werden, und innerhalb dieser Labors aufgrund einer kleinen Stichprobe aller in Frage kommenden Ermittlungsergebnisse.

5.3 In der statistischen Praxis, wo der wahre Wert einer Standardabweichung, σ , unbekannt ist und durch einen Stichprobenschätzwert ersetzt wird, wird demzufolge das Formelzeichen σ durch das Formelzeichen s ersetzt, um deutlich zu machen, dass dies ein Schätzwert ist. Das muss in ISO 5725-1:1994, Gleichungen (2) bis (6), geschehen, woraus folgt:

- s_L^2 ist der Schätzwert für die „Varianz zwischen den Labors“;
- s_W^2 ist der Schätzwert für die „Varianz innerhalb eines Labors“;
- s_F^2 ist der arithmetische Mittelwert von s_W^2 und zugleich Schätzwert für die Wiederholvarianz; dieser arithmetische Mittelwert ist über alle Labors zu erstrecken, die sich am Ringversuch beteiligen und nach dem Ausscheiden von Ausreißern verbleiben;

— s_R^2 ist der Schätzwert für die Vergleichsvarianz nach der Gleichung

$$s_R^2 = s_L^2 + s_r^2 \quad (1)$$

6 Anforderungen an ein Präzisionsexperiment

6.1 Layout des Experiments

6.1.1 Bei der für die grundlegende Methode angewendeten Gestaltung werden Stichproben aus q Materialien, die q unterschiedliche Merkmalsniveaus für den Versuch repräsentieren, an p Labors gesandt, deren jedes sich für jedes der q Merkmalsniveaus genau n wiederholte Ermittlungsergebnisse unter Wiederholbedingungen verschafft. Diese Art von Versuch wird „Ausgewogener Versuch auf einheitlichem Merkmalsniveau“ genannt.

6.1.2 Die Ausführung dieser Messungen muss systematisch geplant werden. Dazu müssen folgende Anweisungen gegeben werden:

- a) Jedes vorausgehende Prüfen von Messgeräten muss so durchgeführt werden, wie es in dem vereinheitlichten Messverfahren festgelegt ist.
- b) Jede Gruppe von n Messungen, die ein Merkmalsniveau betrifft, muss unter Wiederholbedingungen ausgeführt werden, d. h. in einer kurzen Zeitspanne, durch denselben Bearbeiter, und ohne irgendwelche zwischengeschaltete erneute Kalibrierung des Messgeräts; es sei denn, dies ist ein Bestandteil der Ausführung einer Messung.
- c) Es ist unbedingt erforderlich, dass jeder der n Versuche unter Wiederholbedingungen unabhängig voneinander ausgeführt wird, so als ob es n Versuche an unterschiedlichen Materialien wären. In der Regel wird allerdings der Bearbeiter oder die Bearbeiterin wissen, dass er (sie) identisches Material prüft. Deshalb sollte in den Anleitungen besonders der Gesichtspunkt hervorgehoben werden, dass es der ganze Zweck des Versuchs ist, zu ermitteln, welche Unterschiede in den Ergebnissen während der Versuchsdurchführung vorkommen können. Wenn trotz dieser Hinweise zu befürchten ist, dass vorausgehende Ermittlungsergebnisse nachfolgende Ermittlungsergebnisse und damit die Wiederholvarianz beeinflussen könnten, sollte überlegt werden, ob auf jedem der q Merkmalsniveaus n verschiedene Proben verwendet werden sollten, die in einer Weise codiert sind, dass der Bearbeiter oder die Bearbeiterin nicht weiß, was die Proben für ein gegebenes Merkmalsniveau sind. Ein solches Verfahren kann allerdings dennoch Probleme bezüglich der Sicherstellung von tatsächlich zutreffenden Wiederholbedingungen zwischen den Proben verursachen. Das ist nur dann möglich, wenn die Messungen in einer Art und Weise ausgeführt werden, dass alle q mal n Messungen innerhalb einer kurzen Zeitspanne ausgeführt werden können.
- d) Es ist nicht unbedingt nötig, dass jede der q Gruppen von je n Messungen alle genau in einer kurzen Zeitspanne ausgeführt werden. Unterschiedliche Gruppen von Messungen dürfen an verschiedenen Tagen ausgeführt werden.
- e) Messungen aller q Merkmalsniveaus müssen durch ein und denselben Bearbeiter ausgeführt werden, und zusätzlich müssen die n Messungen auf einem gegebenen Merkmalsniveau durchgehend mit demselben Messgerät ausgeführt werden.
- f) Sofern im Verlauf der Messungen ein Bearbeiter oder eine Bearbeiterin ausfallen sollte, darf ein anderer Bearbeiter oder Bearbeiterin die Messungen zu Ende führen, vorausgesetzt, dass der Wechsel nicht innerhalb ein und derselben Gruppe von n Messungen auf einem Merkmalsniveau vorkommt, sondern nur zwischen zwei der q Gruppen. Über einen solchen Wechsel muss mit den Ergebnissen berichtet werden.
- g) Für die vollständige Durchführung aller Messungen muss eine Höchstzeitspanne vorgegeben werden. Das kann nötig sein, um die Zeitspanne zu begrenzen, die zwischen dem Tag des Empfangs der Proben und dem Tag des Abschlusses der Messungen verstreichen darf.

h) Alle Proben müssen mit dem Namen des Versuchs und einer Kennzeichnung der Probe eindeutig etikettiert sein.

6.1.3 In 6.1.2 und an anderer Stelle in diesem Dokument ist der Bearbeiter erwähnt. Für einige Messungen kann es in Wirklichkeit ein Team von Bearbeitern sein, von denen jeder einen speziellen Teil des Verfahrens ausführt. In einem solchen Fall muss das Team als „der Bearbeiter“ betrachtet werden, und irgendeine Änderung innerhalb des Teams muss als Bereitstellung eines anderen „Bearbeiters“ betrachtet werden.

6.1.4 In der Geschäftspraxis können Ergebnisse ziemlich grob gerundet werden. In einem Präzisionsversuch hingegen müssen Ergebnisse mindestens mit einer Stelle mehr als in dem vereinheitlichten Messverfahren verlangt aufgeschrieben werden. Wenn das vereinheitlichte Messverfahren die Anzahl Stellen nicht festlegt, darf die Rundung nicht gröber sein als die Hälfte des Schätzwertes der Wiederholstandardabweichung. Sollte die Präzision vom Merkmalsniveau m abhängen, können für unterschiedliche Merkmalsniveaus unterschiedliche Rundungsgrade nötig sein.

6.2 Rekrutierung der Labors

6.2.1 Die allgemeinen Prinzipien bezüglich der Auswahl der Labors, die an einem Ringversuch teilnehmen sollen, sind in ISO 5725-1 angegeben. Eine Orientierungshilfe zur Anzahl der Labors ist in Anhang A angegeben. Bei der Anwerbung der Mitarbeit der erforderlichen Anzahl von Labors müssen deren Verantwortlichkeiten zweifelsfrei festgelegt werden. Ein Beispiel für einen geeigneten Anwerbungs-Fragebogen zeigt Bild 1.

6.2.2 Für den Zweck dieses Dokumentes wird ein „Labor“ als eine Kombination aus Bearbeiter, dem Messgerät und der Prüfstelle betrachtet. Eine einzelne Prüfstelle (oder ein Labor im herkömmlichen Sinn) kann deshalb mehrere „Labors“ verkörpern, sofern es mehrere Bearbeiter mit voneinander unabhängigen Messgeräten bereitstellen kann, mit denen die Arbeit unabhängig voneinander durchgeführt wird.

6.3 Vorbereitung des Materials

6.3.1 Eine Besprechung der Gesichtspunkte, die betrachtet werden müssen, wenn Material für die Anwendung in einem Präzisionsversuch ausgewählt wird, ist in ISO 5725-1 zu finden.

6.3.2 Bei der Entscheidung über die Mengen des bereitzustellenden Materials müssen versehentliche Mengenverluste oder Irrtümer bei der Beschaffung einiger Prüfergebnisse berücksichtigt werden, was den Einsatz von zusätzlichem Material nötig machen kann. Die Menge des vorbereiteten Materials muss für den Versuch ausreichend sein und einen angemessenen Reserve-Vorrat erlauben.

6.3.3 Es sollte überlegt werden, ob es für einige Labors wünschenswert ist, zur Eingewöhnung in das (vereinheitlichte) Messverfahren noch vor Erarbeitung des offiziellen Ermittlungsergebnisses einige vorläufige Ermittlungsergebnisse zu gewinnen, und ob gegebenenfalls für diesen Zweck zusätzliches Material (nicht aber Proben für den Präzisionsversuch) bereitgestellt werden sollte.

6.3.4 Wenn ein Material homogenisiert werden muss, dann muss das in der für dieses Material zweckmäßigsten Art und Weise geschehen. Wenn das zu prüfende Material nicht homogen ist, ist es wichtig, die Proben so vorzubereiten, wie es in dem (vereinheitlichten) Messverfahren festgelegt ist, vorzugsweise beginnend mit einem Los handelsüblichen Materials für jedes Merkmalsniveau. Im Fall von instabilem Material müssen spezielle Anleitungen für Lagerung und Behandlung festgelegt werden.

ANMERKUNG ISO Guide 35 enthält Informationen zur Bewertung der Homogenität und Stabilität von Referenzmaterialien.

6.3.5 Für die Proben zu jedem Merkmalsniveau müssen n getrennte Behälter für jedes Labor verwendet werden, sofern irgendeine Gefahr der Beeinträchtigung des Materials besteht, sobald die Behälter geöffnet worden sind (z. B. durch Oxydation, durch Verlust flüchtiger Komponenten oder bei hygroskopischem Material). Im Fall von instabilem Material müssen spezielle Anweisungen für Lagerung und Behandlung festgelegt werden. Vorsichtsmaßnahmen können nötig werden, um sicherzustellen, dass Proben bis zum Zeitpunkt der Durchführung der Messungen unverändert bleiben. Besteht das zu prüfende Material aus einem Gemisch von

Pulvern unterschiedlicher relativer Dichte oder unterschiedlicher Korngröße, ist Vorsicht geboten im Hinblick auf Entmischung infolge von Erschütterungen, beispielsweise während eines Transports. Wenn eine Reaktion mit der Atmosphäre erwartet werden kann, können die Proben in Ampullen versiegelt werden, die entweder evakuiert oder mit einem Inertgas gefüllt sind. Für verderbliches Material, wie etwa Lebensmittel- oder Blutproben, kann es notwendig werden, sie in tiefgefrorenem Zustand mit ins Einzelne gehenden Anweisungen für das Verfahren des Auftauens an die beteiligten Labors zu senden.

Fragebogen für einen Ringversuch	
Titel des vereinheitlichten Messverfahrens (siehe beigegefügte Kopie):	
1. Unser Labor hat die Absicht, sich an einem Präzisionsversuch für dieses vereinheitlichte Messverfahren zu beteiligen JA ☐ NEIN ☐ (Zutreffendes ankreuzen)	
2. Als Teilnehmer wissen wir, dass a) alle erforderlichen Geräte, Chemikalien und andere durch das Verfahren festgelegten Anforderungen in unserem Labor verfügbar sein müssen, sobald das Programm beginnt; b) die festgelegten Anforderungen bezüglich „Zeit“, wie das Datum des Beginns, das Anfordern von Prüfproben sowie das Datum für die Beendigung des Programms, strikt erfüllt werden müssen; c) wir uns an das vereinheitlichte Messverfahren genau zu halten haben; d) Proben genau nach den Anleitungen behandelt werden müssen; e) ein qualifizierter Bearbeiter die Messungen ausführen muss.	
Wir haben das vereinheitlichte Messverfahren studiert sowie eine faire Abschätzung unserer Fähigkeiten und Möglichkeiten vorgenommen. Wir glauben, dass wir angemessen für die gemeinsame Prüfung dieses Messverfahrens vorbereitet sein werden.	
3. <u>Kommentare</u>	(Unterschrift) (Unternehmen oder Labor)

Bild 1 — Anwendungs-Fragebogen für einen Ringversuch

7 An einem Präzisionsversuch beteiligtes Personal

ANMERKUNG Es ist nicht zu erwarten, dass die Arbeitsmethoden in verschiedenen Labors identisch sind. Deshalb soll der Inhalt dieses Abschnitts lediglich als Leitfaden dienen, der in zweckmäßiger Weise modifiziert und auf den einzelnen Fall ausgerichtet werden muss.

7.1 Versuchsgremium

7.1.1 Das Präzisionsexperiment sollte von einem Versuchsgremium beaufsichtigt werden, das mit dem (vereinheitlichten) Messverfahren und seiner Anwendung vertraut ist.

7.1.2 Die Aufgaben des Versuchsgremiums sind:

- a) den Präzisionsversuch zu planen und (seinen Ablauf) zu koordinieren;
- b) über die Anzahl der Labors, der Merkmalsniveaus und der auszuführenden Messungen sowie über die Anzahl der zu verlangenden signifikanten Zahlenstellen zu entscheiden;
- c) jemanden für die statistischen Aufgaben zu benennen (siehe 7.2);
- d) jemanden für die Führungsaufgaben zu benennen (siehe 7.3);
- e) die zusätzlich zum vereinheitlichten Messverfahren für die Laborverantwortlichen zu erstellenden Anweisungen zu überlegen;
- f) zu entscheiden, ob einigen Bearbeitern erlaubt werden kann, einige wenige inoffizielle Messungen auszuführen, um nach einer langen Zeitspanne Erfahrung wiederzugewinnen (solche Messungen dürfen niemals an offiziellen Ringversuchs-Proben ausgeführt werden);
- g) den Bericht über die statistische Auswertung zum Zweck der Vervollständigung der Auswertung der Ermittlungsergebnisse zu diskutieren;
- h) Schlusswerte für die Wiederholstandardabweichung und die Vergleichstandardabweichung festzulegen;
- i) zu entscheiden, ob weitere Maßnahmen zur Verbesserung der Vereinheitlichung für das Messverfahren oder im Hinblick auf solche Labors verlangt sind, deren Ergebnisse als Ausreißer abgelehnt wurden.

7.2 Statistische Funktionen

Mindestens ein Mitglied des Versuchsgremiums sollte Erfahrung in statistischer Versuchsplanung und in der Auswertung von Versuchen haben. Seine/ihre Aufgaben sind:

- a) sein/ihr Spezialwissen bei der Versuchsplanung einzubringen;
- b) die Daten auszuwerten;
- c) einen Bericht zur Vorlage beim Versuchsgremium zu schreiben, und zwar nach den in 8.7 enthaltenen Anweisungen.

7.3 Führungsaufgaben

7.3.1 Das Organisieren des Versuchs sollte einem einzelnen Labor anvertraut werden. Ein Mitglied des Mitarbeiterstabs dieses Labors sollte die volle Verantwortung dafür übernehmen; er bzw. sie wird „leitende Führungskraft“ genannt und ist bestellt vom Versuchsgremium.

7.3.2 Die Aufgaben der leitenden Führungskraft sind:

- a) die Mitarbeit der benötigten Anzahl von Labors anzuwerben und sicherzustellen, dass (dort) die betreffenden Laborverantwortlichen bestellt sind;
- b) die Vorbereitung der Materialien und Proben zu organisieren und zu überwachen, ebenso die Versendung der Proben; für jedes Merkmalsniveau sollte eine angemessene Menge Material als Reservevorrat zur Seite gelegt werden;
- c) Anleitungen zur Abdeckung aller Gesichtspunkte nach 6.1.2 a) bis h) zu entwerfen und diese an die Laborverantwortlichen zu verteilen, frühzeitig genug für diese und im Voraus, um Kommentare zu geben und Fragen zu stellen und um sicherzustellen, dass die ausgewählten Bearbeiter solche sind, die üblicherweise bei ihrer Routinearbeit solche Messungen ausführen;

- d) geeignete Formblätter zu entwerfen zur Benutzung als Arbeitsbericht für die Bearbeiter und zur Benutzung für die Laborverantwortlichen zum Bericht über die Ermittlungsergebnisse mit der verlangten Anzahl von Dezimalstellen (oder signifikanten Stellen, falls erforderlich). Solche Formblätter können den Namen des Bearbeiters, das Empfangs- und Messdatum der Proben, das benutzte Messgerät und andere relevante Informationen enthalten;
- e) irgendwelche Fragen von Labors im Hinblick auf die Ausführung der Messungen zu behandeln;
- f) darauf zu achten, dass ein Gesamtzeitplan eingehalten wird;
- g) die Formblätter mit den Daten einzusammeln und diese an die statistische Fachkraft weiterzugeben.

ANMERKUNG Die in 7.3.2 d) genannten Formblätter können elektronisch sein, wie etwa eine Tabellenkalkulation, die ausreichend gegen versehentliche Manipulation geschützt ist.

7.4 Laborverantwortliche

7.4.1 Ein Mitglied des Mitarbeiterstabes in jedem der teilnehmenden Labors sollte für das Organisieren der eigentlichen Durchführung der Messungen in Übereinstimmung mit den von der leitenden Führungskraft erhaltenen Anweisungen sowie für das Berichten über die Ermittlungsergebnisse verantwortlich gemacht werden.

7.4.2 Die Aufgaben des Laborverantwortlichen sind:

- a) sicherzustellen, dass die ausgewählten Bearbeiter solche sind, die üblicherweise bei ihrer Routinearbeit solche Messungen auszuführen haben;
- b) das Aushändigen der Proben an den (die) Bearbeiter in Übereinstimmung mit den Anweisungen der leitenden Führungskraft (und nötigenfalls das Bereitstellen von Material für Versuche zur Einarbeitung);
- c) das Überwachen der Ausführung der Messungen (der Laborverantwortliche darf nicht an der Ausführung der Messungen teilnehmen);
- d) sicherzustellen, dass die Bearbeiter die geforderte Anzahl von Messungen ausführen;
- e) die Befolgung des vorgegebenen Terminplans für die Ausführung der Messungen sicherzustellen;
- f) die Ermittlungsergebnisse zu sammeln, aufgezeichnet entsprechend der vereinbarten Anzahl von Dezimalstellen (oder signifikanten Stellen), eingeschlossen irgendwelche in Erfahrung gebrachte Unregelmäßigkeiten und Schwierigkeiten sowie Kommentare durch die Bearbeiter.

7.4.3 Der Verantwortliche jedes Labors sollte einen vollständigen Bericht schreiben, der die folgende Information enthalten sollte:

- a) die Ermittlungsergebnisse, lesbar eingetragen durch ihren Erzeuger in die bereitgestellten Formblätter, nicht übertragen oder maschinengeschrieben (als Alternative können Computer- oder Prüfmaschinenaussagen annehmbar sein);
- b) die original beobachteten Werte oder Ablesungen (sofern vorhanden), aus denen die Ermittlungsergebnisse abgeleitet sind, lesbar durch den Bearbeiter in die bereitgestellten Formblätter eingetragen, nicht übertragen oder maschinengeschrieben;
- c) Kommentare durch die Bearbeiter über die Vereinheitlichung des Messverfahrens;
- d) Information über Unregelmäßigkeiten oder Störungen, die während der Messungen vorgekommen sein können, eingeschlossen irgendwelche Änderungen von Bearbeitern, und zwar zusammen mit einer Darle-

gung, welche Messungen durch welche Bearbeiter ausgeführt wurden, und schließlich Gründe für irgendwelche fehlenden Ergebnisse;

- e) der Zeitpunkt (die Zeitpunkte), zu dem (zu denen) die Proben in Empfang genommen wurden;
- f) der Zeitpunkt (die Zeitpunkte), zu dem (zu denen) jede Probe gemessen wurde;
- g) Information über die Geräteausrüstung, sofern wesentlich;
- h) irgendeine andere sachdienliche Information.

ANMERKUNG Die in 7.4.3 a) und b) genannten Datenausgaben und Formblätter können elektronisch sein, wie etwa eine Tabellenkalkulation, die ausreichend gegen versehentliche Manipulation geschützt ist.

7.5 Bearbeiter

7.5.1 In jedem Labor müssen die Messungen durch nur einen Mitarbeiter ausgeführt werden, der unter jenen als repräsentativ ausgewählt wurde, die in ihrer üblichen Arbeit geeignet sind, die Messungen auszuführen.

7.5.2 Weil das Ziel des Versuchs die Ermittlung der Präzision ist, wie sie aufgrund der allgemeinen Grundgesamtheit von Bearbeitern gewonnen werden kann, die mit dem vereinheitlichten Messverfahren arbeiten, sollten den Bearbeitern im Allgemeinen keine ergänzenden Erläuterungen zu den Anforderungen an das Messverfahren gegeben werden. Allenfalls sollte gegenüber den Bearbeitern zum Ausdruck gebracht werden, dass es Zweck der Übung ist, das Ausmaß herauszufinden, um das sich in der Praxis Ergebnisse unterscheiden können, so dass für sie eine kleinere Versuchung besteht, Ergebnisse auszusondern oder nachzuarbeiten, von denen sie das Gefühl haben, sie seien widersprüchlich.

7.5.3 Obwohl die Bearbeiter üblicherweise keine ergänzenden weiteren Erläuterungen zum vereinheitlichten Messverfahren erhalten sollten, sollten sie ermutigt werden, die Vereinheitlichungsregeln zu kommentieren, insbesondere aber darzulegen, ob die in diesen enthaltenen Anweisungen ausreichend eindeutig und klar sind.

7.5.4 Die Aufgaben der Bearbeiter sind:

- a) Ausführen der Messungen entsprechend dem vereinheitlichten Messverfahren;
- b) Berichten über irgendwelche erlebten Besonderheiten oder Schwierigkeiten; es ist besser, über ein Missgeschick zu berichten, als das Ermittlungsergebnis zurechtzurücken, weil ein oder zwei fehlende Ermittlungsergebnisse den Versuch nicht ungültig machen, und weil viele davon eine Unzulänglichkeit des Verfahrens anzeigen;
- c) Kommentare abgeben über die Angemessenheit der Anweisungen in den Vereinheitlichungsregeln; Bearbeiter sollten über alle Fälle berichten, in denen sie nicht in der Lage sind, ihre Anweisungen zu befolgen, zumal dies ebenfalls eine Unzulänglichkeit in den Vereinheitlichungsregeln anzeigen kann.

8 Statistische Analyse des Präzisionsversuchs

8.1 Vorbemerkungen

8.1.1 Die Auswertung der Daten, die als eine statistische Aufgabe betrachtet werden sollte, die durch eine statistische Fachkraft gelöst werden muss, umfasst drei aufeinander folgende Abschnitte:

- a) kritisches Prüfen der Daten mit dem Ziel, Ausreißer oder andere Unregelmäßigkeiten zu erkennen und zu behandeln, sowie Prüfen der Eignung des (statistischen) Modells;
- b) Berechnen vorläufiger Werte für Präzision und Mittelwerte, getrennt für jedes Merkmalsniveau;

- c) Feststellen der Schlusswerte für Präzision und Mittelwerte, eingeschlossen die Ermittlung einer Beziehung zwischen Präzision und Merkmalsniveau m , sofern die Auswertung anzeigt, dass eine solche Beziehung bestehen kann.

8.1.2 Bei der Auswertung werden als Erstes, getrennt für jedes Merkmalsniveau, Schätzwerte berechnet für

- die Wiederholvarianz, s_r^2
- die Varianz zwischen den Labors, s_L^2
- die Vergleichvarianz, s_R^2
- den arithmetischen Mittelwert, m .

ANMERKUNG Die Auswertung schließt eine systematische Anwendung statistischer Ausreißertests ein, für die von der Literatur her eine große Vielfalt verfügbar ist und für die Zwecke dieses Dokumentes angewendet werden könnte. Aus praktischen Gründen wurde nur eine begrenzte Anzahl dieser Tests (in diesem Dokument) aufgenommen. Sie sind in 8.3 erklärt.

8.2 Tabellarische Zusammenstellung der Ergebnisse sowie angewendete Darstellungsweise

8.2.1 Zellen

Jede Kombination eines Labors und eines Merkmalsniveaus wird „Zelle des Ringversuchs“ genannt. Im Idealfall bestehen die Ergebnisse eines Ringversuchs mit p Labors und q Merkmalsniveaus aus einer Tabelle mit p mal q Zellen. Jede Zelle enthält n Ergebnisse der Wiederholprüfung. Sie alle können für die Berechnung der Wiederholstandardabweichung und der Vergleichstandardabweichung eingesetzt werden. Allerdings wird in der Praxis dieser Idealfall nicht immer erreicht. Abweichungen entstehen infolge von überflüssigen Daten, fehlenden Daten und Ausreißern.

8.2.2 Überflüssige Daten

Zuweilen kommt es vor, dass ein Labor mehr als die offiziell verlangten n Ermittlungsergebnisse bereitstellt und darüber berichtet. In diesem Fall muss der Laborverantwortliche darüber berichten, warum das geschehen ist und welches die richtigen Ermittlungsergebnisse sind. Lautet die Antwort, dass alle Ergebnisse gleichermaßen gültig sind, darf aus den verfügbaren Ermittlungsergebnissen eine Zufallsauswahl vorgenommen werden, um die für die Auswertung geplante Anzahl von Ermittlungsergebnissen auszuwählen.

ANMERKUNG Die Berechnungen der Präzisionskenndaten, die in diesem Dokument erlaubt sind, können unterschiedliche Anzahlen an Prüfergebnissen in jeder Zelle berücksichtigen (siehe 8.2.7). Erhebliche Schwankungen der Anzahl der gemeldeten Beobachtungen können die Interpretation der Ausreißertests und Mandelschen Vereinbarkeitsprüfungen erheblich beeinträchtigen und die Zuverlässigkeit des Gesamtmittelwerts reduzieren.

8.2.3 Fehlende Daten

In anderen Fällen können einige Ermittlungsergebnisse fehlen, beispielsweise infolge des Verlusts einer Probe oder infolge eines Fehlers bei der Ausführung der Messung. Die in diesem Dokument empfohlene Auswertung ist dergestalt, dass vollständig leere Zellen einfach übergangen werden können, während teilweise leere Zellen in das vereinheitlichte Berechnungsverfahren einbezogen werden können.

8.2.4 Ausreißer

Diese sind Eintragungen unter den Originalermittlungsergebnissen oder in den aus diesen entwickelten Tabellen, die so sehr von vergleichbaren Eintragungen in derselben Tabelle abweichen, dass sie als unvereinbar mit den anderen Daten betrachtet werden. Erfahrung hat gelehrt, dass Ausreißer nicht immer vermieden werden können und dass sie in ähnlicher Art und Weise in Betracht gezogen werden müssen wie die Behandlung von fehlenden Daten.

8.2.5 Ausreißerlabors

Wenn einige unerklärte, stark abweichende Ermittlungsergebnisse auf unterschiedlichen Merkmalsniveaus innerhalb desselben Labors vorkommen, dann kann dieses Labor als Ausreißerlabor betrachtet werden, das eine zu große Varianz innerhalb des Labors und/oder eine zu große systematische Abweichungskomponente im Niveau seiner Ermittlungsergebnisse hat. Es kann dann vernünftig sein, einige oder alle Daten eines solchen Ausreißerlabors auszusondern.

Dieses Dokument stellt keinen statistischen Test bereit, durch den verdächtige Labors beurteilt werden können. Die primäre Entscheidung sollte der Verantwortung der statistischen Fachkraft unterliegen, jedoch muss im Hinblick auf weitere Maßnahmen über alle als Ausreißer abgelehnte Labors dem Versuchsgremium berichtet werden.

8.2.6 Fehlerhafte Daten

Offensichtlich fehlerhafte Daten sollten untersucht und korrigiert oder aussortiert werden.

ANMERKUNG Fehlerhafte Daten können offensichtlich sein, wenn beispielsweise Ergebnisse in falschen Maßeinheiten angegeben werden, die gemeldeten Werte nicht angemessen dem betrachteten Material zugewiesen werden können oder andere gemeldete Informationen auf eine erhebliche unbeabsichtigte Abweichung von dem untersuchten Normverfahren hindeuten.

8.2.7 Prüfergebnisse ausgewogener einheitlicher Experimente

Idealfall sind p Labors, bezeichnet mit dem Index i ($i = 1, 2, \dots, p$), wobei jedes Labor q Merkmalsniveaus untersucht, bezeichnet mit dem Index j ($j = 1, 2, \dots, q$), und wobei auf jedem Merkmalsniveau (jeder ij -Kombination) n Wiederholmessungen ausgeführt werden, was zu insgesamt pqn Ermittlungsergebnissen führt. Wegen fehlender (siehe 8.2.3) oder (als Ausreißer von allen anderen Ermittlungsergebnissen) abweichender (siehe 8.2.4) Ermittlungsergebnisse oder wegen Ausreißerlabors (siehe 8.2.5) oder wegen falscher Daten (siehe 8.2.6) wird dieser Idealfall nicht immer erreicht. Unter diesen Umständen sind für abweichende Anzahlen von Ermittlungsergebnissen die in 8.2.9 bis 8.2.11 vorgegebenen Bezeichnungen sowie die Verfahren nach 8.4 zugelassen.

8.2.8 Zusammenstellung der Daten und Zwischenwerte

Muster für empfehlenswerte Formblätter für die statistische Auswertung sind in Bild 2 angegeben. Formblatt A enthält zusammengestellte Einzelergebnisse, Formblatt B enthält berechnete Zellenmittelwerte und Formblatt C enthält berechnete Standardabweichungen. Aus praktischen Gründen werden sie einfach als Formblätter A, B und C (siehe Bild 2) bezeichnet.

ANMERKUNG Die Verwendung der Formblätter A, B und C für die statistische Analyse ist keine Anforderung dieses Dokuments. Die Formblätter A, B und C eignen sich jedoch für die Zusammenstellung von Zwischenwerten in der manuellen Berechnung und können einfach für die Anwendung in einem Tabellenkalkulationsprogramm angepasst werden. Alternative Layouts sind oftmals besser für die Anwendung mit Statistiksoftware geeignet.

8.2.9 Ursprüngliche Ermittlungsergebnisse

Siehe Formblatt A in Bild 2, dabei ist

n_{ij} die Anzahl von Ermittlungsergebnissen in der Zelle für das Labor i auf dem Merkmalsniveau j ;

y_{ijk} eines dieser Prüfergebnisse ($k = 1, 2, \dots, n_{ij}$);

p_j die Anzahl der Labors mit Bericht über mindestens ein Ermittlungsergebnis für das Merkmalsniveau j (nach Aussonderung von Ermittlungsergebnissen, die als Ausreißer oder als falsch gekennzeichnet wurden).

8.2.10 Arithmetische Zellmittelwerte in einer Zelle (Formblatt B in Bild 2)

Diese Werte ergeben sich aus den Ergebnissen in Formblatt A anhand der Gleichung (2) wie folgt.

$$\bar{y}_{ij} = \frac{1}{n_{ij}} \sum_{k=1}^{n_{ij}} y_{ijk} \quad (2)$$

Die arithmetischen Mittelwerte einer Zelle sollten mit mindestens einer signifikanten Stelle mehr aufgezeichnet werden als die angegebenen Ermittlungsergebnisse (Formblatt A).

8.2.11 Maße für die Streuung innerhalb einer Zelle (Formblatt C in Bild 2)

Diese Werte ergeben sich aus einzelnen Ergebnissen (Formblatt A, siehe 8.2.9) und Mittelwerten (Formblatt B, siehe 8.2.10) wie folgt. Für den allgemeinen Fall ist die Standardabweichung innerhalb einer Zelle wie in Gleichung (3) zu benutzen

$$s_{ij} = \sqrt{\frac{1}{n_{ij} - 1} \sum_{k=1}^{n_{ij}} (y_{ijk} - \bar{y}_{ij})^2} \quad (3)$$

oder nach der äquivalenten Formel

$$s_{ij} = \sqrt{\frac{1}{n_{ij} - 1} \left[\sum_{k=1}^{n_{ij}} (y_{ijk})^2 - \frac{1}{n_{ij}} \left[\sum_{k=1}^{n_{ij}} y_{ijk} \right]^2 \right]} \quad (4)$$

Bei der Anwendung dieser Formeln muss darauf geachtet werden, eine hinreichende Anzahl Stellen für die Berechnungen vorzusehen, d. h. jeder Zwischenwert muss mit mindestens doppelt so vielen signifikanten Stellen wie für die ursprünglich angegebenen Daten erforderlich berechnet werden [siehe 7.3.2 b), 7.4.3 f)].

Wird das Formblatt C verwendet, sollte die Standardabweichung mit einer signifikanten Stelle mehr als die Ergebnisse in Formblatt A dokumentiert werden. Bei Werten für n_{ij} unter 2 sollte ein Bindestrich in das Formblatt C eingefügt werden.

ANMERKUNG 1 Wenn eine Zelle ij zwei Prüfergebnisse enthält, ist die Standardabweichung innerhalb der Zelle

$$s_{ij} = |y_{ij1} - y_{ij2}| / \sqrt{2} \quad (5)$$

Zur Vereinfachung können also die Beträge der Differenzen in Formblatt C dokumentiert und für den Cochran-Test anstelle der Standardabweichung angewendet werden, sofern alle Zellen zwei Prüfergebnisse enthalten.

ANMERKUNG 2 Gleichung (4) ist für Rundungsfehler in Computerberechnungen anfällig; Gleichung (3) ist für Computerberechnungen genauer.

8.2.12 Korrigierte oder als Ausreißer abgelehnte Daten

Weil einige Daten aufgrund der in 8.3.3 und 8.3.4 erwähnten Tests korrigiert oder als Ausreißer abgelehnt werden können, können die Werte von y_{ijk} , n_{ij} und p_j , wie sie für die abschließenden Ermittlungen von Präzision und Mittelwert angewendet werden, sich von den Werten unterscheiden, wie sie in den Formblättern A, B, und C in Bild 2 aufgezeichnet sind. Aus diesem Grund muss beim Bericht über die Schlusswerte für Präzision und Richtigkeit gegebenenfalls stets erwähnt werden, welche Daten korrigiert oder ausgesondert worden sind.

8.3 Ergebnisprüfung bezüglich Vereinbarkeit und Ausreißer

8.3.1 Ansätze für die Untersuchung von Daten

Wiederhol- und Vergleichstandardabweichungen sind zu schätzen aus den für eine Anzahl spezifischer Merkmalsniveaus zusammengestellten Daten. Die Schätzwerte können sich ändern, weil einzelne Labors oder Werte

sich als unvereinbar mit allen anderen Labors oder allen anderen Werten erweisen können. Deshalb müssen Entscheidungen bezüglich (dieser Labors und) dieser Werte getroffen werden. Zwei Ansätze sind eingeführt:

- a) grafische Vereinbarkeitsprüfungen;
- b) numerische Ausreißertests.

ANMERKUNG Literaturhinweis [3] bietet weitere Erörterungen der Handhabung laborübergreifender Daten.

8.3.2 Grafische Vereinbarkeitsprüfungen

8.3.2.1 Zwei Prüfgrößen, genannt Mandels h und Mandels k , werden verwendet. Diese sind sowohl für die Laborbeurteilung hilfreich als auch für die Beschreibung der Variabilität des (vereinheitlichten) Messverfahrens.

8.3.2.2 Die Vereinbarkeits-Prüfgröße zwischen den Labors, h , ist für jedes Labor wie folgt zu berechnen: Die Zellenabweichung (arithmetischer Mittelwert einer Zelle minus Gesamtmittelwert für das betreffende Merkmalsniveau) ist durch die Standardabweichung der arithmetischen Mittelwerte einer Zelle (für dieses betrachtete Merkmalsniveau) zu dividieren:

$$h_{ij} = \frac{\bar{y}_{ij} - \bar{\bar{y}}_j}{\sqrt{\frac{1}{p_j-1} \sum_{i=1}^{p_j} (\bar{y}_{ij} - \bar{\bar{y}}_j)^2}} \quad (6)$$

wobei für $\bar{y}_{ij}$ 8.2.10 und für $\bar{\bar{y}}_j$ 8.4.4 gilt.

Die Werte h_{ij} für jede Zelle sind abhängig vom Labor aufzuzeichnen, in Gruppen für jedes Merkmalsniveau (und getrennt gruppiert für die unterschiedlichen, durch jedes Labor geprüften Merkmalsniveaus) (siehe Bild C.7).

Formblatt A – Empfohlenes Formblatt für die Zusammenstellung der ursprünglichen Daten ^a									
Prüflaboratorium	Stufe								
	1	2	...	...	<i>j</i>	...	...	<i>q</i> -1	<i>q</i>
1									
2									
⋮									
<i>i</i>	<i>y</i> _{<i>i</i>11}				<i>y</i> _{<i>i</i><i>j</i>1}				<i>y</i> _{<i>i</i><i>q</i>1}
	⋮				⋮				⋮
	<i>y</i> _{<i>i</i>1<i>k</i>}				<i>y</i> _{<i>i</i><i>j</i><i>k</i>}				<i>y</i> _{<i>i</i><i>q</i><i>k</i>}
	⋮				⋮				⋮
	<i>y</i> _{<i>i</i>1<i>n</i>}				<i>y</i> _{<i>i</i><i>j</i><i>n</i>}				<i>y</i> _{<i>i</i><i>q</i><i>n</i>}
⋮									
<i>p</i>									

Formblatt B – Empfohlenes Formblatt für die Zusammenstellung der Mittelwerte ^b									
Prüflaboratorium	Stufe								
	1	2	...	...	<i>j</i>	...	...	<i>q</i> -1	<i>q</i>
1	$\bar{y}_{11}$								$\bar{y}_{1q}$
2									
⋮									
<i>i</i>					$\bar{y}_{ij}$				
⋮									
<i>p</i>	$\bar{y}_{p1}$								$\bar{y}_{pq}$

Formblatt C – Empfohlenes Formblatt für die Zusammenstellung der Spread-Messwerte in den Zellen ^c									
Prüflaboratorium	Stufe								
	1	2	...	...	<i>j</i>	...	...	<i>q</i> -1	<i>q</i>
1	<i>s</i> ₁₁								<i>s</i> _{1<i>q</i>}
2									
⋮									
<i>i</i>					<i>s</i> _{<i>i</i><i>j</i>}				
⋮									
<i>p</i>	<i>s</i> _{<i>p</i>1}								<i>s</i> _{<i>p</i><i>q</i>}

^a Formblatt A enthält die zusammengestellten einzelnen Ergebnisse. Jede Zelle enthält alle *n* Ergebnisse für das Labor.
^b Formblatt B enthält die nach 8.2.10 berechneten Zellenmittelwerte.
^c Formblatt C enthält die nach 8.2.11 berechneten Spread-Messwerte (Standardabweichungen).

Bild 2 — Empfohlene Formblätter für die Zusammenstellung von Ergebnissen zur Analyse

8.3.2.3 Die Vereinbarkeits-Prüfgröße innerhalb des Labors, *k*, ist zu berechnen, indem zuerst die „kombinierte Standardabweichung innerhalb der Zellen“ für jedes Merkmalsniveau berechnet wird (Gleichung (7))

$$\sqrt{\frac{\sum s_{ij}^2}{p_j}}$$

(7)

Dann ist für jedes Labor und für jedes Merkmalsniveau zu berechnen:

$$k_{ij} = \frac{s_{ij}\sqrt{p_j}}{\sqrt{\sum s_{ij}^2}} \quad (8)$$

Die Werte k_{ij} für jede Zelle sind abhängig vom Labor aufzuzeichnen, in Gruppen für jedes Merkmalsniveau (und getrennt gruppiert für die unterschiedlichen, durch jedes Labor geprüften Merkmalsniveaus) (siehe Bild C.8).

8.3.2.4 Die Prüfung der grafischen Darstellungen von h und k kann anzeigen, dass spezielle Labors Ergebnismuster ausweisen, die sich markant von den anderen des Versuchs unterscheiden. Dies zeigt sich durch eine (im Vergleich mit den anderen Labors) durchgängig hohe oder niedrige Streuung innerhalb der Zelle und/oder durch extreme (hohe oder niedrige) arithmetische Mittelwerte einer Zelle über viele Merkmalsniveaus hin. Wenn das vorkommt, sollte Verbindung mit dem speziellen Labor aufgenommen werden, um zu versuchen, die Ursache für das abweichende Erscheinungsbild festzustellen. Abhängig vom betreffenden Ergebnis kann die statistische Fachkraft

- a) die Daten des betreffenden Labors vorläufig zurückhalten;
- b) das Labor bitten, die Messung nochmals auszuführen (falls realisierbar);
- c) die Daten des Labors aus dem Versuch entfernen.

8.3.2.5 In den grafischen Darstellungen der Werte h können unterschiedliche Muster erscheinen. Alle Labors können bei den unterschiedlichen Merkmalsniveaus beides haben, positive oder negative Werte von h . Einzelne Labors können dazu neigen, entweder nur positive oder nur negative Werte h zu liefern, wobei die Anzahl der Labors mit negativen Werten etwa derjenigen entspricht, die positive Werte liefern. Keines dieser Muster ist unüblich oder erfordert Untersuchungen, obwohl das zweite dieser Muster die Vermutung nahelegen kann, dass eine gemeinsame Ursache für die systematische Abweichungskomponente dieser Labors besteht. Wenn andererseits alle Werte h für eines der Labors das eine Vorzeichen haben und die Werte h für die anderen Labors haben alle das andere Vorzeichen, dann sollte die Ursache gesucht werden. Ebenso sollte die Ursache gesucht werden, wenn die Werte h vom Merkmalsniveau des Versuchs in systematischer Weise abzuhängen scheinen. Entsprechend den in Tabelle 7 und Tabelle 8 angegebenen Indikatoren sind Linien in die grafischen Darstellungen der Werte h einzuzeichnen. Diese Indikatorlinien dienen als Anhaltspunkte bei der Prüfung der Daten auf Muster.

8.3.2.6 Wenn eines der Labors dadurch unter den k Aufzeichnungen auffällt, dass es viele große Werte hat, dann sollte die Ursache gesucht werden: Dies weist darauf hin, dass es eine schlechtere Wiederholpräzision hat als die anderen Labors. Einem Labor kann das Anlass sein, zu durchweg kleineren Werten k zu kommen, und zwar aufgrund solcher Faktoren wie übertriebene Rundung seiner Daten oder eine sehr grobe Messskala. Entsprechend den in (Tabelle 7 und Tabelle 8) angegebenen Indikatoren sind Linien in die grafischen Darstellungen der Werte k einzuzeichnen. Diese Indikatorlinien dienen als Anhaltspunkte bei der Prüfung der Daten auf Muster.

8.3.2.7 Wenn eine nach Labors geordnete Darstellung der Werte von h und k zur Vermutung Anlass gibt, dass eines der Labors einige Werte h oder k in der Nähe der Linie kritischer Werte hat, sollte die betreffende Darstellung untersucht werden, gruppiert nach Merkmalsniveaus. Häufig verwandelt sich ein Wert, der in der nach Labors geordneten Darstellung groß erschien, zu einem mit anderen Labors für dasselbe Merkmalsniveau recht gut übereinstimmenden Wert. Wenn aber dieser Wert als stark von den Werten der anderen Labors abweichend erkannt ist, dann sollte die Ursache gesucht werden.

8.3.2.8 Zusätzlich zu diesen Darstellungen von h und k können Histogramme der arithmetischen Mittelwerte und der Spannweiten der Zellen das Vorhandensein von beispielsweise zwei unterschiedlichen Grundgesamtheiten erkennen lassen. Ein solcher Fall könnte eine spezielle Behandlung verlangen, weil das allgemein für die hier beschriebene Methode zugrundeliegende Prinzip eine einzige, eingipflige Verteilung der Grundgesamtheit voraussetzt.

8.3.3 Numerische Methode der Ausreißerprüfung

8.3.3.1 Für die Behandlung von Ausreißern wird die folgende Verfahrensweise empfohlen.

- a) Die in 8.3.4 und 8.3.5 empfohlenen Tests werden angewendet, um Fastausreißer oder Ausreißer zu erkennen:
- sofern die Prüfgröße (einen Wert hat, der) kleiner oder höchstens gleich der Grenze des kritischen Bereichs für ein Signifikanzniveau von 5 % ist, wird die geprüfte Einheit als korrekt angenommen;
 - sofern die Prüfgröße (einen Wert hat, der) größer als die Grenze des kritischen Bereichs für ein Signifikanzniveau von 5 % ist, aber kleiner oder höchstens gleich der Grenze des kritischen Bereichs für ein Signifikanzniveau von 1 %, wird die geprüfte Einheit als „Fastausreißer“ bezeichnet und mit einem Einzelstern gekennzeichnet;
 - sofern die Prüfgröße (einen Wert hat, der) größer als die Grenze des kritischen Bereichs für ein Signifikanzniveau von 1 % ist, wird die geprüfte Einheit als „statistischer Ausreißer“ bezeichnet und mit einem Doppelstern gekennzeichnet.
- b) Als Nächstes ist zu untersuchen, ob die Fastausreißer und/oder die statistischen Ausreißer durch irgendein technisches Versehen erklärt werden können, beispielsweise durch
- ein Fehlverhalten bei der Ausführung der Messung,
 - einen Rechenfehler,
 - einen einfachen Schreibfehler beim Aufschreiben des Ermittlungsergebnisses oder
 - die Prüfung einer falschen Probe.

War der Fehler ein Rechenfehler oder ein Schreibfehler, sollte das verdächtige Ergebnis durch den korrigierten Wert ersetzt werden. War der Fehler durch die Prüfung einer falschen Probe verursacht, sollte das Ergebnis in seine zutreffende Zelle platziert werden. Ist diese Korrektur ausgeführt worden, sollte die Prüfung auf Fastausreißer oder Ausreißer wiederholt werden. Wenn die Erklärung des technischen Fehlers von der Art ist, dass damit die Unmöglichkeit des Ersetzens des verdächtigen Ermittlungsergebnisses bewiesen ist, dann sollte der Wert als ein „echter“ Ausreißer ausgesondert werden, der eigentlich nicht zum Versuch gehört.

- c) Wenn irgendwelche Fastausreißer und/oder Ausreißer übrig bleiben, die nicht erklärt oder als zu einem Ausreißerlabor gehörig abgelehnt wurden, werden die Fastausreißer als korrekte Einheiten beibehalten und die statistischen Ausreißer werden ausgesondert; es sei denn, die statistische Fachkraft entscheidet mit guten Gründen, dass sie beibehalten werden.
- d) Wenn die Daten einer Zelle im Formblatt B in Bild 2 entsprechend der obigen Verfahrensweise abgelehnt wurden, dann müssen die entsprechenden Daten in Formblatt C in Bild 2 (ebenfalls) als Ausreißer abgelehnt werden und umgekehrt.

8.3.3.2 Die in 8.3.4 und 8.3.5 angegebenen Tests sind zweierlei Art. Der Cochran-Test ist ein Test der Streuung innerhalb eines Labors und sollte als erster angewendet werden. Dann sollten irgendwelche erforderlichen Maßnahmen ergriffen werden, nötigenfalls unter Wiederholung von Tests. Der andere Test (Grubbs) ist vorzugsweise ein Test der Streuung zwischen den Labors und kann ebenfalls angewendet werden (wenn $n > 2$), wo der Cochran-Test den Verdacht geweckt hat, die hohe Streuung innerhalb des Labors sei nur einem einzigen unter den Ermittlungsergebnissen in der Zelle zuzuordnen.

8.3.4 Cochran-Test

8.3.4.1 Dieses Dokument geht von der Annahme aus, dass zwischen den Labors nur kleine Differenzen bezüglich der Streuungen innerhalb der Labors bestehen. Die Erfahrung zeigt aber, dass das nicht immer der

Fall ist. Deshalb wurde ein Test einbezogen, um die Gültigkeit dieser Annahme zu prüfen. Einige Tests können für diesen Zweck angewendet werden, hier aber wurde der Cochran-Test ausgewählt.

8.3.4.2 Gegeben ist eine Reihe von p Standardabweichungen s_i , alle berechnet aus derselben Anzahl (n) von Wiederholerergebnissen. Die Prüfgröße C nach Cochran ist durch Gleichung (9) gegeben:

$$C = \frac{s_{\max}^2}{\sum_{i=1}^p s_i^2} \quad (9)$$

Dabei ist $s_{\max}$ die höchste Standardabweichung in der Reihe.

- a) Sofern die Prüfgröße (einen Wert hat, der) kleiner oder höchstens gleich der Grenze des kritischen Bereichs für ein Signifikanzniveau von 5 % ist, kann die geprüfte Einheit als korrekt angenommen werden.
- b) Sofern die Prüfgröße (einen Wert hat, der) größer als die Grenze des kritischen Bereichs für ein Signifikanzniveau von 5 % ist, aber kleiner oder höchstens gleich der Grenze des kritischen Bereichs für ein Signifikanzniveau von 1 %, wird die geprüfte Einheit als „Fastausreißer“ bezeichnet und mit einem Einzelstern gekennzeichnet.
- c) Sofern die Prüfgröße (einen Wert hat, der) größer als die Grenze des kritischen Bereichs für ein Signifikanzniveau von 1 % ist, wird die geprüfte Einheit als „statistischer Ausreißer“ bezeichnet und mit einem Doppelstern gekennzeichnet.

Die Grenzen des kritischen Bereichs für den Cochran-Test sind in Tabelle 5 angegeben.

Der Cochran-Test muss auf das Formblatt C in Bild 2 angewendet werden, und zwar getrennt auf jedes Merkmalsniveau.

8.3.4.3 Die Kriterien des Cochran-Tests treffen streng nur dann zu, wenn alle Standardabweichungen aus derselben Anzahl (n) von Ermittlungsergebnissen unter denselben Messbedingungen (Wiederholbedingungen in diesem Dokument) gewonnen wurden. In konkreten Fällen kann diese Anzahl aufgrund von fehlenden oder ausgesonderten Daten unterschiedlich sein. Dieses Dokument geht aber davon aus, dass in einem richtig organisierten Ringversuch solche Unterschiede in der Anzahl von Ermittlungsergebnissen je Zelle begrenzt sind und vernachlässigt werden können. Deshalb werden die Kriterien des Cochran-Tests angewendet, wobei für n die Anzahl der in den meisten Zellen vorkommenden Ermittlungsergebnisse eingesetzt wird.

8.3.4.4 Das Kriterium des Cochran-Tests prüft lediglich den größten Wert in einer Reihe von Standardabweichungen und ist daher ein einseitiger Ausreißertest. Die Heterogenität der Varianzen kann sich sicherlich auch selbst dadurch erkennbar machen, dass einige der Standardabweichungen vergleichsweise zu klein sind. Kleine Werte der Standardabweichung dürften jedoch sehr erheblich durch den Grad der Rundung der Ursprungsdaten beeinflusst sein und sind aus diesem Grund nicht sehr verlässlich. Außerdem erscheint es unberechtigt, die Daten eines Labors deshalb abzulehnen, weil es zu einer höheren Präzision seiner Ermittlungsergebnisse fähig ist als die anderen Labors. Daher wird das Kriterium des Cochran-Tests als angemessen betrachtet.

8.3.4.5 Eine kritische Prüfung des Formblatts C in Bild 2 kann zuweilen aufzeigen, dass die Standardabweichungen eines speziellen Labors bei einzelnen oder bei den meisten Merkmalsniveaus kleiner sind als die der anderen Labors. Das kann anzeigen, dass das betreffende Labor mit einer kleineren Wiederholstandardabweichung arbeitet als die anderen Labors. Das kann, je nachdem, entweder durch eine bessere Methode und eine bessere Geräteausstattung oder durch eine veränderte oder unkorrekte Anwendung des vereinheitlichten Messverfahrens verursacht sein. Wenn so etwas vorkommt, sollte es an das Versuchsgremium berichtet werden. Dieses sollte dann entscheiden, ob dieser Punkt eine weiter detaillierte Untersuchung wert ist. (Ein Beispiel hierfür ist das Labor 2 im Versuch, der in B.1^{N1)} beschrieben ist.)

N1) Nationale Fußnote: Fehler in der Referenzfassung, anstatt Verweisung nach B.1 muss nach C.1 verwiesen werden.

8.3.4.6 Wurde die größte Standardabweichung als Ausreißer eingestuft, dann sollte der Wert weggelassen und der Cochran-Test für die verbleibenden Werte wiederholt werden. Dieser Prozess darf wiederholt werden, aber er kann zu übermäßigen Ablehnung von vermeintlichen Ausreißerdaten führen, wenn – wie es manchmal der Fall ist – die zugrunde liegende Annahme Normalverteilung nicht hinreichend angenähert ist. Die wiederholte Anwendung des Cochran-Tests ist hier ausschließlich im Hinblick auf das Fehlen eines statistischen Tests für die gemeinsame Prüfung mehrerer Ausreißer als hilfreiches Werkzeug vorgeschlagen. Der Cochran-Test ist für diesen Zweck nämlich nicht ausgelegt und große Vorsicht sollte geboten sein, wenn entsprechende Rückschlüsse geschlossen werden. Wenn zwei oder drei Labors Ergebnisse angeben, die eine hohe Standardabweichung haben, insbesondere wenn das innerhalb ein und desselben Merkmalsniveaus vorkommt, sollten Schlussfolgerungen aus dem Cochran-Test sorgfältig geprüft werden. Andererseits aber kann es, wenn einige Fastausreißer und/oder statistische Ausreißer auf unterschiedlichen Merkmalsniveaus in einem einzigen Labor gefunden wurden, ein ernster Hinweis sein, dass bei diesem Labor die Varianz innerhalb des Labors außergewöhnlich hoch ist, und dass die Gesamtheit der Daten dieses Labors als Ausreißer abgelehnt werden sollte.

8.3.5 Grubbs-Test

8.3.5.1 Einzelner Ausreißer

Gegeben ist eine Reihe von Daten $x_{(1)}, x_{(2)}, \dots, x_{(p)}$, gebildet aus $x_1, x_2, \dots, x_p$, angeordnet in aufsteigender Reihenfolge. Um zu ermitteln, ob die Beobachtung mit dem größten Wert ein Ausreißer ist, wird der Grubbs-Test angewendet. Dazu wird die Prüfgröße G_p nach Grubbs anhand von Gleichung (10) berechnet.

$$G_p = \frac{x_{(p)} - \bar{x}}{s} \quad (10)$$

wobei

$$\bar{x} = \frac{1}{p} \sum_{i=1}^p x_{(i)} \quad (11)$$

und

$$s = \sqrt{\frac{1}{p-1} \sum_{i=1}^p (x_{(i)} - \bar{x})^2} \quad (12)$$

Zur Prüfung der Signifikanz des kleinsten Beobachtungswerts ist die Prüfgröße, G_1 , anhand der Gleichung (13) zu berechnen:

$$G_1 = \frac{(\bar{x} - x_{(1)})}{s} \quad (13)$$

- Sofern die Prüfgröße (einen Wert hat, der) kleiner oder höchstens gleich der Grenze des kritischen Bereichs für ein Signifikanzniveau von 5 % ist, kann die geprüfte Einheit als korrekt angenommen werden.
- Sofern die Prüfgröße (einen Wert hat, der) größer als die Grenze des kritischen Bereichs für ein Signifikanzniveau von 5 % ist, aber kleiner oder höchstens gleich der Grenze des kritischen Bereichs für ein Signifikanzniveau von 1 %, wird die geprüfte Einheit als „Fastausreißer“ bezeichnet und mit einem Einzelstern gekennzeichnet.
- Sofern die Prüfgröße (einen Wert hat, der) größer als die Grenze des kritischen Bereichs für ein Signifikanzniveau von 1 % ist, wird die geprüfte Einheit als „statistischer Ausreißer“ bezeichnet und mit einem Doppelstern gekennzeichnet.

8.3.5.2 Zwei Ausreißer

Zur Prüfung, ob die zwei größten Beobachtungen Ausreißer sein könnten, wird die Prüfgröße, G , anhand der Gleichung (14) berechnet:

$$G = \frac{s_{p-1,p}^2}{s_0^2} \quad (14)$$

wobei

$$s_0^2 = \sum_{i=1}^p (x_{(i)} - \bar{x})^2 \quad (15)$$

und

$$s_{p-1,p}^2 = \sum_{i=1}^{p-2} (x_{(i)} - \bar{x}_{p-1,p})^2 \quad (16)$$

und

$$\bar{x}_{p-1,p} = \frac{1}{p-2} \sum_{i=1}^{p-2} x_{(i)} \quad (17)$$

Alternativ wird zur Prüfung der zwei kleinsten Beobachtungen die Prüfgröße, G , anhand der Gleichung (18) berechnet:

$$G = \frac{s_{1,2}^2}{s_0^2} \quad (18)$$

wobei

$$s_{1,2}^2 = \sum_{i=1}^{p-2} (x_{(i)} - \bar{x}_{1,2})^2 \quad (19)$$

und

$$\bar{x}_{1,2} = \frac{1}{p-2} \sum_{i=3}^p x_{(i)} \quad (20)$$

8.3.5.3 Anwendung der Grubbs-Test

Wenn ein Präzisionsversuch ausgewertet wird, kann der Grubbs-Test auf Folgendes angewendet werden:

- a) Auf die arithmetischen Mittelwerte einer Zelle (Formblatt B in Bild 2) für ein vorgegebenes Merkmalsniveau j , wobei

$$x_i = \bar{y}_{ij} \quad (21)$$

und

$$p = p_j \quad (22)$$

und wobei j ein fester Wert (für das ausgewählte Merkmalsniveau j) ist.

Mit den Daten des einen Merkmalsniveaus wird der Grubbs-Test für einen einzelnen Ausreißer auf die arithmetischen Mittelwerte einer Zelle wie in 8.3.5.1 beschrieben angewendet. Wenn ein arithmetischer Mittelwert einer Zelle durch diesen Test als Ausreißer erkannt wird, wird er ausgesondert, und der Test für den anderen extremen arithmetischen Mittelwert einer Zelle wird wiederholt (z. B. wird, wenn der größte Wert ein Ausreißer

ßer ist, dann der kleinste Wert betrachtet, wobei der größte ausgeschlossen bleibt). Der in 8.3.5.2 beschriebene Grubbs-Test für zwei Ausreißer sollte aber nicht angewendet werden. Erst dann, wenn der Grubbs-Test keinen arithmetischen Mittelwert einer Zelle als Ausreißer zeigt, wird der in 8.3.5.2 beschriebene Grubbs-Test für zwei Ausreißer angewendet.

- b) Auf ein Einzelergebnis innerhalb einer Zelle, wo der Cochran-Test gezeigt hat, dass die Standardabweichung der Zelle zweifelhaft ist.

ANMERKUNG Nach 8.3.3.1 heißt ein Wert statistischer Ausreißer, wenn die Prüfgröße größer ist als ihr kritischer Wert für das Signifikanzniveau 1 %. Wenn der Grubbs-Test erstmalig auf eine Gruppe von arithmetischen Mittelwerten einer Zelle angewendet wird, wird zum einen der kritische Wert aus Tabelle 6 verwendet, um den größten arithmetischen Mittelwert einer Zelle auf dem Signifikanzniveau von 0,5 % zu testen und auch um den kleinsten arithmetischen Mittelwert einer Zelle auf dem Signifikanzniveau von 0,5 % zu testen. Dies läuft auf einen Test des extremsten arithmetischen Mittelwerts einer Zelle auf dem Signifikanzniveau von 1 % nach 8.3.3.1 hinaus. Sofern einer der beiden extremen arithmetischen Mittelwerte einer Zelle als statistischer Ausreißer eingestuft wurde, wird anschließend der Grubbs-Test auf den anderen extremen arithmetischen Mittelwert einer Zelle angewendet. Es kann eingewendet werden, dass nun ein einseitiger Test angewendet werden sollte. Das in diesem Dokument empfohlene Verfahren ist es jedenfalls, dass ausschließlich die in Tabelle 6 aufgeführten kritischen Werte anzuwenden sind (die kritischen Werte für zweiseitige Tests auf dem Signifikanzniveau von 1 %), damit alle arithmetischen Mittelwerte einer Zelle in gleicher Weise behandelt werden. Ein analoges Argument kann angewendet werden, um die Anwendung der zweiseitigen kritischen Werte auf dem Signifikanzniveau von 5 % aus Tabelle 6 für alle Tests auf Fastausreißer zu rechtfertigen.

8.3.6 Wiederholte Prüfung für Mittelwert- und Datenpunktausreißer

8.3.6.1 Es wird oftmals festgestellt, dass eine einzige Anwendung (zum Beispiel) des Grubbs-Tests gefolgt von der Löschung der Ausreißer zu einem kleineren Datensatz führt, der bei näherer Prüfung scheinbar weitere Extremwerte im Verhältnis zu dem übrigen Datensatz enthält. Unter diesen Umständen ist die Anwendung weiterer Prüfungen für Datenpunkt- oder Labormittelwertausreißer erlaubt.

8.3.6.2 Die wiederholte Ausreißerprüfung und -ablehnung sollte für ein gegebenes Niveau eingestellt werden, wenn dies zu der Ablehnung eines übermäßigen Anteils der Daten aus diesem Niveau führt. Das Versuchsgremium muss bestimmen, ob der Anteil der abgelehnten Daten zu hoch ist.

ANMERKUNG 1 Es kann hilfreich sein, eine Obergrenze für den Anteil ausgeschlossener Daten eines gegebenen Niveaus vor der statistischen Analyse festzulegen.

ANMERKUNG 2 Laut der Empfehlung von IUPAC [15] sollte der Anteil der ausgeschlossenen Daten eines gegebenen Niveaus einer Studie den Bruchteil 2/9 nicht überschreiten. Der Bruchteil, der nach Entfernung aller Ausreißer betrachtet wird, gilt als ungültig aus technischen Gründen vor dem Ausreißertest.

8.3.7 Alternative Ausreißer-Prüf- und Testverfahren

Alternative Prüf- und statistische Testverfahren und Prüfgrößen zu den in 8.3.3 bis 8.3.5 angegebenen Verfahren und Statistiken dürfen angewendet werden, sofern:

- die gleiche Reihe von Ausreißerprüfungen und statistischen Ausreißerprüfungen auf alle Niveaus innerhalb einer gegebenen Studie angewendet wird;
- die gewählten Verfahren es ermöglichen, einzelne Datenpunktausreißer innerhalb eines Labors, Labormittelwertausreißer und Labordispersionsausreißer zu erkennen;
- jede systematische Abweichung der Varianzschätzwerte für normal verteilte Daten, die durch den Ausschluss von Ausreißern entstehen, nicht größer ist als die aus der Anwendung der Unterabschnitte 8.3.3 bis 8.3.5 entstehenden Abweichungen.

ANMERKUNG Alle Ausreißerausschlussverfahren führen zu einer gewissen systematischen Abweichung nach unten der Präzisionsschätzwerte, sofern nicht eine ausdrückliche Korrektur um die resultierende systematische Abweichung

erfolgt. Robuste statistische Verfahren, die weniger systematischen Abweichungen der Präzisionsschätzwerte für normal verteilte Daten unterliegen, sind in ISO 5725-5 beschrieben.

8.4 Berechnung des Gesamtmittelwerts und der Varianzen

8.4.1 Auswertungsverfahren

Das in diesem Dokument gewählte Auswertungsverfahren umfasst die Ausführung der Schätzung von m und der Präzision j , getrennt für jedes Merkmalsniveau. Die Ergebnisse der Berechnung werden in einer Tabelle für jeden Wert von j aufgeführt.

8.4.2 Grunddaten

Die für die Berechnung erforderlichen Grunddaten sind in den drei Tabellen in Bild 2 aufgeführt:

- Formblatt A enthält die Original-Ermittlungsergebnisse;
- Formblatt B enthält die arithmetischen Mittelwerte einer Zelle;
- Formblatt C enthält die Maße für die Streuung innerhalb einer Zelle.

8.4.3 Nichtleere Zellen

Die Anzahl nicht-leerer Zellen für die Berechnung ist, als Konsequenz der Regel in 8.3.3.1 d) für ein spezifisches Merkmalsniveau bei Formblatt B und Formblatt C üblicherweise gleich. Eine Ausnahme könnte vorkommen, wenn aufgrund fehlender Daten eine Zelle im Formblatt A nur ein einziges Ermittlungsergebnis enthält, was eine leere Zelle im Formblatt C mit sich bringen wird, nicht aber (eine leere Zelle) im Formblatt B. In diesem Fall ist es zulässig,

- a) das einzelne Ermittlungsergebnis, das zu leeren Zellen in beiden Formblättern B und C führt, auszusondern oder
- b) wenn a) als unerwünschter Informationsverlust betrachtet wird, einen Strich (fehlender Wert) in Formblatt C einzutragen.

Die Anzahl der nichtleeren Zellen kann sich für die verschiedenen Merkmalsniveaus unterschiedlich sein, daher der Index j in p_j .

8.4.4 Berechnung des Gesamtmittelwerts $\hat{m}$

Für das Merkmalsniveau j ist der (als Schätzwert für den Erwartungswert dienende) Gesamtmittelwert

$$\hat{m}_j = \bar{\bar{y}}_j = \frac{\sum_{i=1}^p n_{ij} \bar{y}_{ij}}{\sum_{i=1}^p n_{ij}} \quad (23)$$

oder mathematisch gleichwertig

$$\hat{m}_j = \bar{\bar{y}}_j = \frac{\sum_{i=1}^p \sum_{k=1}^{n_{ij}} y_{ijk}}{\sum_{i=1}^p n_{ij}} \quad (24)$$

8.4.5 Berechnung der Varianzen

Drei Varianzen werden für jedes Merkmalsniveau berechnet. Dies sind die Wiederholvarianz, die Varianz zwischen den Labors und die Vergleichvarianz.

8.4.5.1 Die Wiederholpräzisionsvarianz ist durch Gleichung (25) gegeben:

$$s_{rj}^2 = \frac{\sum_{i=1}^p (n_{ij} - 1) s_{ij}^2}{\sum_{i=1}^p (n_{ij} - 1)} \quad (25)$$

8.4.5.2 Die Varianz zwischen den Labors ist durch Gleichung (26) gegeben:

$$s_{Lj}^2 = \frac{s_{dj}^2 - s_{rj}^2}{\bar{n}_j} \quad (26)$$

wobei

$$s_{dj}^2 = \frac{1}{p-1} \sum_{i=1}^p n_{ij} (\bar{y}_{ij} - \bar{\bar{y}}_j)^2 = \frac{1}{p-1} \left[\sum_{i=1}^p n_{ij} \bar{y}_{ij}^2 - \bar{\bar{y}}_j^2 \sum_{i=1}^p n_{ij} \right] \quad (27)$$

und

$$\bar{n}_j = \frac{1}{p-1} \left[\sum_{i=1}^p n_{ij} - \frac{\sum_{i=1}^p n_{ij}^2}{\sum_{i=1}^p n_{ij}} \right] \quad (28)$$

Diese Berechnungen sind in den Beispielen in C.1 und C.3 dargestellt.

ANMERKUNG Der Ausdruck nach dem zweiten Gleichheitszeichen in Gleichung (27), der die Differenz zwischen den beiden summierten Ausdrücken enthält, ist mathematisch mit dem Ausdruck nach dem ersten Gleichheitszeichen identisch, kann aber zu erheblichen Rundungsfehlern führen, wenn es in Computerberechnungen verwendet wird. Der Ausdruck nach dem ersten Gleichheitszeichen ist deutlich weniger anfällig für Rundungsfehler.

8.4.5.3 Für den besonderen Fall, dass alle $n_{ij} = n = 2$, dürfen die einfacheren Berechnungen durch die Gleichungen (29) und (30) benutzt werden:

$$s_{rj}^2 = \frac{1}{2p} \sum_{i=1}^p (y_{ij1} - y_{ij2})^2 \quad (29)$$

und

$$s_{Lj}^2 = \frac{1}{p-1} \sum_{i=1}^p (\bar{y}_{ij} - \bar{\bar{y}}_j)^2 - \frac{s_{rj}^2}{2} \quad (30)$$

Diese (Berechnungen) sind durch das in C.2 angegebene Beispiel veranschaulicht.

8.4.5.4 Wo sich bei diesen Berechnungen zufallsbedingt ein negativer Wert für s_{Lj}^2 ergibt, sollte angenommen werden, dass dieser den Wert Null hat.

8.4.5.5 Die Vergleichsvarianz ist durch Gleichung (31) gegeben:

$$s_{Rj}^2 = s_{rj}^2 + s_{Lj}^2 \quad (31)$$

8.4.6 Alternative Berechnungsverfahren für Varianzen

8.4.6.1 Für alle Fälle, in denen Labors die gleiche Anzahl an Replikaten haben (nach Untersuchung auf Ausreißer und anschließender Löschung von Datenpunkten), dürfen die vorgenannten Varianzberechnungen durch die Anwendung der in Anhang B beschriebenen Varianzanalyse ersetzt werden.

8.4.6.2 Der Berechnungsprozess in 8.4.4 bis 8.4.5 darf durch das Restricted Maximum Likelihood-Verfahren (REML) [13], [14] mittels geeigneter Computersoftware ersetzt werden. Software, die die

REML-Varianzschätzwerte aus den Beispielen in Anhang B mit drei oder mehr signifikanten Stellen reproduziert, darf verwendet werden.

ANMERKUNG 1 Ein Verfahren zur Schätzung der Varianzkomponenten mittels REML ist in Anhang B beschrieben.

ANMERKUNG 2 Es können geringfügige Unterschiede zwischen den Ergebnissen der REML-Schätzung und an anderer Stelle in diesem Teil vorgesehener Berechnungen bestehen. Diese Unterschiede haben üblicherweise keine Konsequenzen.

8.4.6.3 Wenn ein alternatives Berechnungsverfahren für Varianzen gewählt wird, muss das gleiche Berechnungsverfahren für alle Niveaus des Präzisionsexperiments verwendet werden.

8.4.7 Abhängigkeit der Varianzen von m

Im Anschluss an die Berechnung der Präzision für jedes Niveau m sollte untersucht werden, ob die Präzision von m abhängig ist, und wenn ja, sollte die funktionale Beziehung bestimmt werden. Dies wird in 8.5 beschrieben.

8.5 Feststellen einer funktionalen Beziehung zwischen Präzisionswerten s und dem arithmetischen Gesamtmittelwert m

8.5.1 Wahl der funktionalen Beziehung

8.5.1.1 Es kann nicht immer für selbstverständlich gehalten werden, dass eine regelrechte funktionale Beziehung zwischen Präzision und m besteht. Insbesondere wird, wenn Material-Inhomogenitäten einen untrennbaren Anteil der Variabilität der Ermittlungsergebnisse ausmachen, nur dann eine funktionale Beziehung vorhanden sein, wenn diese Inhomogenität (ebenfalls) eine regelrechte Funktion des Merkmalsniveaus m ist. Bei festem Material unterschiedlicher Zusammensetzung, das aus unterschiedlichen Produktionsprozessen stammt, ist eine regelrechte funktionale Beziehung keineswegs sicher. Dieser Punkt sollte entschieden sein, bevor das folgende Verfahren angewendet wird.

8.5.1.2 Wenn keine klare funktionale Beziehung hergestellt werden kann, dürfen separate Präzisionswerte für jedes untersuchte Material angegeben werden. Wenn die verschiedenen Präzisionswerte bereichsübergreifend ähnlich sind, darf ein gemittelter Präzisionswert angegeben werden.

8.5.1.3 Die in 8.5.2 bis 8.5.4 vorgestellten Denk- und Rechenverfahren betreffen sowohl die Wiederholstandardabweichung als auch die Vergleichstandardabweichung. Sie werden hier jedoch im Interesse der Kürze nur für die Wiederholpräzision vorgelegt. Betrachtet werden nur vier Arten von Beziehungen:

I) $s_r = bm$ (eine Gerade durch den Nullpunkt)

II) $s_r = a + bm$ (eine Gerade mit einem positiven Achsenabschnitt)

III) $s_r^2 = a_v^2 + (b_v m)^2$ (eine Kombination aus konstanten und relativen Beiträgen zu einer kombinierten Standardabweichung)

IV) $\lg s_r = c + d \lg m$ (oder $s_r = Cm^d$); mit $d \leq 1$ (eine Exponentialfunktion)

Es ist zu erwarten, dass in der Mehrzahl der Fälle mindestens eine dieser drei Formeln eine zufrieden stellende Anpassung liefert. Sollte das nicht der Fall sein, dann sollte die statistische Fachkraft, welche die Auswertung ausführt, eine alternative Lösung versuchen. Um Verwirrung zu vermeiden, sollten die in diesen Gleichungen vorkommenden Konstanten a , b , c , C und d durch Indizes unterschieden werden, nämlich mit a_r , b_r , ... für die Wiederholvarianz und mit a_R , b_R , ... für die Vergleichvarianz. In diesem Abschnitt sind diese Indizes zur Vereinfachung der Bezeichnungen allerdings erneut weggelassen worden. Auch s_r wurde einfach verkürzt auf s , damit ein Index für das Merkmalsniveau j möglich ist.

8.5.1.4 Im Allgemeinfall ist $d > 0$, sodass die Beziehungen I und IV zu $s = 0$ für $m = 0$ führen. Wird das aus der Sicht des Versuchs betrachtet, kann das unannehmbar erscheinen. Beim Bericht über die Präzisionsdaten

sollte jedoch klargemacht werden, dass diese ausschließlich innerhalb der Merkmalsniveaus anwendbar sind, die durch den Ringversuch zur Ermittlung der Präzision abgedeckt sind.

8.5.1.5 Für $a = 0$ (oder $a_v = 0$) und $d = 1$ sind alle vier Beziehungen algebraisch identisch (alle können auf die Beziehung I reduziert werden), und ebenso bringen drei oder alle vier Beziehungen praktisch äquivalente Anpassungen hervor, wenn a nahe Null und/oder d nahe 1 liegt. In einem solchen Fall sollte die Beziehung I bevorzugt werden, weil sie die folgende einfache Feststellung erlaubt:

„Zwei Ermittlungsergebnisse werden als verdächtig betrachtet, wenn sie sich um mehr als $(100 b)$ % unterscheiden.“

In statistischer Terminologie bedeutet das eine Feststellung, dass der Variationskoeffizient $(100 s/m)$ für alle Merkmalsniveaus eine Konstante ist.

8.5.1.6 Wenn in einer grafischen Darstellung von s_j gegen $\hat{m}_j$, oder in einer grafischen Darstellung von $\lg s_j$ gegen $\lg \hat{m}_j$ festgestellt wird, dass die Reihe der Punkte ziemlich nahe an einer Geraden liegt, dann kann eine mit der Hand gezogene Linie eine zufrieden stellende Lösung liefern. Wenn aber aus unterschiedlichen Gründen eine numerische Anpassungsmethode bevorzugt wird, dann wird für die Beziehungen I und II das Verfahren in 8.5.2 und für die Beziehung VI das Verfahren in 8.5.4 empfohlen.

8.5.2 Anpassungsbeziehungen I und II

8.5.2.1 Statistisch gesehen ist die Anpassung an eine Gerade dadurch kompliziert, dass beide Größen, $\hat{m}_j$ und s_j , Schätzwerte und deshalb zufälligen Abweichungen unterworfen sind. Im Hinblick darauf aber, dass die Steigung b üblicherweise klein ist (von der Größenordnung 0,1 oder kleiner), haben zufällige Abweichungen von $\hat{m}$ geringen Einfluss. Die zufälligen Abweichungen bei der Schätzung von s sind vorherrschend.

8.5.2.2 Ein guter Schätzwert für die Parameter der Regressionslinie verlangt eine gewichtete Regression, weil die Standardabweichung der zufälligen Abweichungen von s proportional zum Schätzwert von s_j ($\hat{s}_j$) ist. Die Gewichtungsfaktoren müssen proportional zu $1/(\hat{s}_j)^2$ sein, wobei $\hat{s}_j$ die vorhergesagte Wiederholstandardabweichung für das Merkmalsniveau j ist. Allerdings hängt $\hat{s}_j$ von Parametern ab, die noch zu berechnen sind. Ein mathematisch korrektes Verfahren zur Ermittlung der Schätzwerte entsprechend den gewichteten kleinsten Quadraten der Residuen kann kompliziert sein. Das folgende Verfahren wird empfohlen. Es hat sich in der Praxis als zufrieden stellend erwiesen.

8.5.2.3 Mit dem Gewichtungsfaktor $W_j = 1/\hat{s}_{Nj}^2$, wobei $N = 0, 1, 2 \dots$ für aufeinander folgende Iterationen, ergeben sich die folgenden Formeln:

$$T_1 = \sum_j W_j \quad (32)$$

$$T_2 = \sum_j W_j \hat{m}_j \quad (33)$$

$$T_3 = \sum_j W_j \hat{m}_j^2 \quad (34)$$

$$T_4 = \sum_j W_j s_j \quad (35)$$

$$T_5 = \sum_j W_j \hat{m}_j s_j \quad (36)$$

Dann ist der Wert von b in der Beziehung I ($s = bm$) gegeben durch T_5/T_3 .

Für die Beziehung II ($s = a + bm$) ist:

$$a = \frac{T_3 T_4 - T_2 T_5}{T_1 T_3 - T_2^2} \quad (37)$$

und

$$b = \frac{T_1 T_5 - T_2 T_4}{T_1 T_3 - T_2^2} \quad (38)$$

8.5.2.4 Für die Beziehung I führt die algebraische Substitution für den Gewichtungsfaktor $W_j = 1/\hat{s}_j^2$ mit $\hat{s}_j^2 = b\hat{m}_j$ zum vereinfachten Ausdruck:

$$b = \frac{\sum_j (s_j/\hat{m}_j)}{q} \quad (39)$$

und eine Iteration ist unnötig.

8.5.2.5 Für Beziehung II sind die Ursprungswerte $\hat{s}_{0j}$ die Originalwerte von s , wie sie nach dem in 8.4 angegebenen Verfahren gewonnen werden. Diese werden angewendet, um zu berechnen

$$W_{0j} = 1/(s_{0j})^2 \quad (j = 1, 2, \dots, q) \quad (40)$$

sowie nach 8.5.6.2^{N2)} die Werte a_1 und b_1 zu berechnen. Das führt zu

$$\hat{s}_{1j} = a_1 + b_1 \hat{m}_j \quad (41)$$

Die Berechnungen sind dann zu wiederholen mit $W_{1j} = 1/\hat{s}_{1j}^2$, um zu

$$\hat{s}_{2j} = a_2 + b_2 \hat{m}_j \quad (42)$$

zu kommen.

Das gleiche Verfahren kann nun erneut wiederholt werden mit den aus diesen Gleichungen abgeleiteten Gewichtungsfaktoren $W_{2j} = 1/\hat{s}_{2j}^2$, was aber nur zu unwesentlichen Änderungen führt. Der Schritt von W_{0j} zu W_{1j} ist der für die Ausschaltung grober Abweichungen der Gewichtungsfaktoren wirksame, und die Gleichungen für $\hat{s}_{2j}$ sollten als das Schlussergebnis betrachtet werden.

BEISPIEL 1 Tabelle 1 enthält ein Beispiel für die Anpassung der Beziehung I an die Wiederholpräzisionsdaten. Die Daten sind der Fallstudie C.3 entnommen und wurden hier nur zur Verdeutlichung des numerischen Verfahrens verwendet. Die Beziehungen werden ausführlicher in C.3 behandelt.

BEISPIEL 2 Tabelle 2 enthält ein Beispiel für die Anpassung der Beziehung II an die Wiederholpräzisionsdaten. Die Daten m_j, s_j sind dieselben wie für Tabelle 1.

8.5.3 Anpassung der Beziehung III

8.5.3.1 Für Beziehung III besteht der einfachste Ansatz darin, die beobachteten Varianzen s_j^2 als Funktion des Mittelwerts zum Quadrat anzupassen. Hieraus ergibt sich eine lineare Beziehung, in der der Schnittpunkt a_v^2 die Varianz des Nullmittelwerts ist und die Steigung b_v^2 kann als Quadratwert der relativen Standardabweichung gesehen werden, wenn m große Werte ergibt. Wie für die Beziehung II gilt auch hier: Auch wenn eine einfache lineare Regression einen nützlichen Schätzwert ergeben kann, erfordert ein guter Schätzwert eine

N2) Nationale Fußnote: Dieser Abschnitt existiert nicht. Verweisung auf 8.5.2.3 wäre richtig.

gewichtete Regression und wie für die Beziehung II hängen die optimalen Gewichtungen von den angepassten Werten ab.

8.5.3.2 Ein ungefähres Verfahren für die Anpassung von Beziehung III vergleichbar mit 8.5.2 stellt sich wie folgt dar.

- i) Bei einem Gewichtungsfaktor W_{vj} entsprechend $1/\hat{s}_{Nj}^4$, wobei $N = 0, 1, 2 \dots$ für sukzessive Iterationen, lauten die berechneten Gleichungen:

$$T_{v1} = \sum_j W_{vj} \quad (43)$$

$$T_{v2} = \sum_j W_{vj} \hat{m}_j^2 \quad (44)$$

$$T_{v3} = \sum_j W_{vj} \hat{m}_j^4 \quad (45)$$

$$T_{v4} = \sum_j W_{vj} s_j^2 \quad (46)$$

$$T_{v5} = \sum_j W_{vj} \hat{m}_j^2 s_j^2 \quad (47)$$

- ii) Dann gilt für Beziehung III, $s_r^2 = a_v^2 + (b_v m)^2$:

$$a_v^2 = \frac{T_{v3}T_{v4} - T_{v2}T_{v5}}{T_{v1}T_{v3} - T_{v2}^2} \quad (48)$$

$$b_v^2 = \frac{T_{v1}T_{v5} - T_{v2}T_{v4}}{T_{v1}T_{v3} - T_{v2}^2} \quad (49)$$

- iii) Die Ausgangswerte a_{1v}^2 und b_{1v}^2 werden anschließend genutzt, um die angepassten Werte von s_j vorherzusagen in Übereinstimmung mit

$$\hat{s}_{1j}^2 = a_{1v}^2 + b_{1v}^2 \hat{m}_j^2 \quad (50)$$

- iv) Die Berechnungen i) bis iii) werden anschließend einmal mit $W_{1vj} = 1/\hat{s}_{1j}^4$ wiederholt, um die Beziehung zu erstellen.

$$\hat{s}_{2j}^2 = a_{2v}^2 + b_{2v}^2 \hat{m}_j^2 \quad (51)$$

Die resultierenden Gleichungen $\hat{s}_{2j}$ sollten als Endergebnis betrachtet werden.

ANMERKUNG Die Gewichtungen W_{vj} basieren auf der Annahme, dass für normal verteilte Daten die Verteilung der Stichprobenvarianz eine skalierte Chi-Quadrat-Verteilung ist, deren Varianz proportional zu σ^4 ist.

BEISPIEL Tabelle 3 enthält ein Beispiel für die Anpassung der Beziehung III an die Wiederholpräzisionsdaten. Die Daten m_j, s_j sind die gleichen wie für Tabelle 1 und dem Beispiel C.3 aus Anhang C entnommen.

8.5.4 Anpassung der Beziehung IV

8.5.4.1 Der Standardfehler von $\lg s$ ist unabhängig von s , sodass eine ungewichtete Regression von $\lg s$ auf $\lg \hat{m}$ angemessen ist.

8.5.4.2 Die Berechnungsgleichungen für die Beziehung IV lauten:

$$T_1 = \sum_j \lg \hat{m}_j \quad (52)$$

$$T_2 = \sum_j (\lg \hat{m}_j)^2 \quad (53)$$

$$T_3 = \sum_j \lg s_j \quad (54)$$

$$T_4 = \sum_j (\lg \hat{m}_j) (\lg s_j) \quad (55)$$

und daher

$$c = \frac{T_2 T_3 - T_1 T_4}{q T_2 - T_1^2} \quad (56)$$

und

$$d = \frac{q T_4 - T_1 T_3}{q T_2 - T_1^2} \quad (57)$$

BEISPIEL Tabelle 4 enthält ein Beispiel für die Anpassung der Beziehung IV an die Wiederholpräzisionsdaten. Die Daten m_j , s_j sind die gleichen wie für Tabelle 1 und dem Beispiel C.3 aus Anhang C entnommen.

Tabelle 1 — Beziehung I, $s = bm$

$\hat{m}_j$	3,94	8,28	14,18	15,59	20,41
s_j	0,092	0,179	0,127	0,337	0,393
$s_j/\hat{m}_j$	0,023 4	0,021 6	0,008 9	0,021 6	0,019 3
$b = \frac{\sum_k (s_j/\hat{m}_j)}{q}$	$\frac{0,094 8}{5} = 0,019$				
$s = bm$	0,075	0,157	0,269	0,296	0,388

Tabelle 2 — Beziehung II, $s = a + bm$

W_{0j}	118	31	62	8,8	6,5
$s_1 = 0,058 + 0,009 0 m$					
$\hat{s}_{1j}$	0,093	0,132	0,185	0,197	0,240
W_{1j}	116	57	29	26	17
$s_2 = 0,030 + 0,015 4 m$					
$\hat{s}_{2j}$	0,092	0,159	0,251	0,273	0,348
W_{2j}	118	40	16	13	8
$s_3 = 0,032 + 0,015 4 m^a$					
$\hat{s}_{3j}$	0,093	0,160	0,251	0,273	0,348
ANMERKUNG Die Werte der Gewichtungsfaktoren sind nicht kritisch; zwei signifikante Stellen sind ausreichend.					
^a Die Differenz zu s_2 ist vernachlässigbar.					

Tabelle 3 — Beziehung III, $s^2 = a_v^2 + (b_v m)^2$

W_{0vj}	13 959	974	3 844	77,5	41,9
$s_1^2 = 0,088^2 + (0,008\ 5\ m)^2$					
$\hat{s}_{1j}$	0,094	0,113	0,150	0,160	0,195
W_{1vj}	12 603	6 126	1 985	1 537	687
$s_2^2 = 0,061^2 + (0,017\ 8\ m)^2$					
$\hat{s}_{2j}$	0,093	0,159	0,260	0,284	0,368
W_{2vj}	13 356	1 546	220	154	54
$s_3^2 = 0,063^2 + (0,017\ 8\ m)^{2a}$					
$\hat{s}_{3j}$	0,094	0,160	0,260	0,284	0,368
ANMERKUNG Die obigen Werte wurden mit absolut genauer Arithmetik berechnet und zu Präsentationszwecken gerundet.					
^a Die Differenz zu s_2 ist vernachlässigbar.					

Tabelle 4 — Beziehung IV, $\lg s = c + d \lg m$

$\lg \hat{m}_j$	+0,595	+0,918	+1,152	+1,193	+1,310
$\lg s_{0j}$	-1,036	-0,747	-0,896	-0,472	-0,406
$\lg s = -1,506\ 5 + 0,772 \lg m$ oder $s = 0,031\ m^{0,772}$					
s	0,089	0,158	0,239	0,257	0,316

8.6 Statistische Analyse als schrittweises Verfahren

ANMERKUNG Bild 3 zeigt das in 8.6 angegebene Verfahren als Ablaufdiagramm.

8.6.1 Es werden alle verfügbaren Ermittlungsergebnisse in einem einzigen Formblatt A in Bild 2, (siehe 8.2) gesammelt. Zu empfehlen ist, dieses Formblatt in p Zeilen zu unterteilen, die (in der Vorspalte) mit $i = 1, 2 \dots, p$ bezeichnet sind (und die p Labors repräsentieren, die Daten geliefert haben), sowie in q Spalten, die (in der Kopfzeile) mit $j = 1, 2, \dots, q$ bezeichnet sind (und die q Merkmalsniveaus in aufsteigender Folge repräsentieren).

Bei einem Versuch auf einheitlichem Merkmalsniveau brauchen die Ermittlungsergebnisse innerhalb einer Zelle des Formblatts A nicht unterschieden werden und können in beliebiger Reihenfolge eingetragen werden.

8.6.2 Das Formblatt A wird hinsichtlich offensichtlicher Unregelmäßigkeiten geprüft, dann werden etwaige offensichtlich falsche Daten geprüft (beispielsweise Daten außerhalb des Messbereichs oder Daten, die aus technischen Gründen nicht vorkommen können) und nötigenfalls ausgesondert. Darüber wird dem Versuchsgremium berichtet. Manchmal ist es unmittelbar einleuchtend, dass Ermittlungsergebnisse eines speziellen Labors oder in einer speziellen Zelle auf einem Niveau liegen, das unvereinbar mit den anderen Daten ist. Solche offensichtlich unverträglichen Daten müssen sofort ausgesondert werden, aber über diese Tatsache muss für weitere Überlegungen dem Versuchsgremium berichtet werden (siehe 8.7.1).

8.6.3 Aus dem Formblatt A, das nötigenfalls nach 8.6.2 berichtet wurde, wird das Formblatt B mit den arithmetischen Mittelwerten einer Zelle und das Formblatt C mit den Maßen für die Streuung innerhalb der Zelle berechnet.

Wenn eine Zelle im Formblatt A nur ein einziges Ermittlungsergebnis enthält, sollte eine der Optionen aus 8.4.3 gewählt werden.

8.6.4 Die grafischen Darstellungen zu den Vereinbarkeits-Prüfgrößen h und k nach Mandel werden wie in 8.3.1 beschrieben vorbereitet und auf Datenkonsistenz geprüft. Diese grafischen Darstellungen können die Eignung der Daten für weitere Auswertungen anzeigen, die Anwesenheit jeglicher Ausreißer oder von Ausreißerlabors. Allerdings werden auf dieser Stufe des Verfahrens noch keine definitiven Entscheidungen getroffen. Sie werden vielmehr bis zur Vervollständigung (der Maßnahmen) nach 8.6.5 bis 8.6.9 aufgeschoben.

8.6.5 Die Formblätter B und C (siehe Bild 2) werden Merkmalsniveau für Merkmalsniveau auf mögliche Fastausreißer und/oder statistische Ausreißer geprüft [siehe 8.3.3.1 a)]. Dazu werden die statistischen Tests nach 8.3 auf alle verdächtigen Einheiten angewendet und die Fastausreißer mit einem einzigen Stern und die statistischen Ausreißer mit einem Doppelstern markiert. Wenn keine Fastausreißer oder statistischen Ausreißer vorkommen, werden 8.6.6 bis 8.6.10 übergangen und es wird unmittelbar bei 8.6.11 fortgefahren.

8.6.6 Es wird geprüft, ob eine technische Erklärung für die Fastausreißer und/oder statistischen Ausreißer gibt oder geben kann. Eine solche Erklärung wird möglichst verifiziert. Solche zufriedenstellend erklärten Fastausreißer und/oder statistischen Ausreißer werden korrigiert oder ausgesondert, je nach Anforderung, und die Formblätter werden entsprechend korrigiert. Verbleiben nun keine Fastausreißer oder statistischen Ausreißer, die nicht erklärt wurden, werden die Schritte 8.6.7 bis 8.6.10 übergangen und es wird unmittelbar bei 8.6.11 fortgefahren.

Eine große Anzahl von Fastausreißern und/oder statistischen Ausreißern kann eine ausgesprochene Varianz-Inhomogenität oder ausgeprägte Unterschiede zwischen Labors anzeigen und kann damit Zweifel an der Eignung des (vereinheitlichten) Messverfahrens auslösen. Dies sollte an das Versuchsgremium berichtet werden.

8.6.7 Sofern die Verteilung der unerklärten Fastausreißer oder statistischen Ausreißer in Formblatt B oder Formblatt C keine Ausreißerlabors vermuten lässt (siehe 8.2.5), wird Schritt 8.6.8 übergangen, und es wird unmittelbar mit 8.6.9 fortgefahren.

8.6.8 Sofern sich Beweismaterial gegen verdächtige Ausreißerlabors als hinreichend aussagekräftig erweist, um die Ablehnung einiger oder aller Daten dieser Labors zu rechtfertigen, werden die betreffenden Daten ausgesondert und es wird an das Versuchsgremium berichtet. Die Entscheidung, einige oder alle Daten von einem speziellen Labor auszusondern, liegt in der Verantwortung der statistischen Fachkraft, welche die Auswertung ausführt. Über diese Entscheidung muss allerdings zum Zweck weiterer Überlegungen an das Versuchsgremium berichtet werden (siehe 8.7.1).

8.6.9 Sofern einige Fastausreißer und/oder statistische Ausreißer übrigbleiben, die weder erklärt noch einem Ausreißerlabor zugeordnet wurden, werden die statistischen Ausreißer ausgesondert, die Fastausreißer werden (für die weitere Auswertung) beibehalten.

8.6.10 Sofern in den vorausgehenden Schritten irgendein Eintrag im Formblatt B als Ausreißer abgelehnt worden ist, muss der entsprechende Eintrag im Formblatt C ebenfalls abgelehnt und als Ausreißer markiert werden, und umgekehrt.

8.6.11 Aus den als korrekt in den Formblättern B und C verbliebenen Eintragungen werden, getrennt für jedes Merkmalsniveau, nach dem in 8.4 angegebenen Verfahren der arithmetische Mittelwert $\bar{m}_j$ sowie die Wiederhol- und die Vergleichstandardabweichungen berechnet.

8.6.12 Wenn der Versuch nur ein einziges Merkmalsniveau anwendet oder wenn entschieden wurde, dass die Wiederhol- und die Vergleichstandardabweichungen getrennt für jedes Merkmalsniveau angegeben werden sollten (siehe 8.5.1) nicht aber als Funktionen des Merkmalsniveaus, wird 8.6.13 übergangen und es wird unmittelbar mit 8.6.14 fortgefahren.

8.6.13 Die Schritte 1) bis 6) unten werden getrennt auf s_r und s_R angewendet.

ANMERKUNG Zur Abkürzung (der Darstellung) werden die Schritte 1) bis 5) nur im Hinblick auf s_r aufgeschrieben.

- 1) s_j wird gegen $\hat{m}_j$ aufgezeichnet und anhand dieses Diagramms wird beurteilt, ob s von m abhängt oder nicht. Wird angenommen, dass s abhängig von m ist, wird Schritt 2) übergangen und es wird mit Schritt 3) fortgefahren. Wenn s als unabhängig von m angesehen wird, wird mit Schritt 2) fortgefahren. Wenn hier Zweifel bestehen sollten, ist es am besten, beide Fälle auszuarbeiten und das Versuchsgremium entscheiden zu lassen. Es gibt keinen brauchbaren statistischen Test, der für dieses Problem geeignet wäre, aber die mit dem (vereinheitlichten) Messverfahren vertrauten technischen Fachkräfte sollten über genügend Erfahrung verfügen, um eine Entscheidung zu treffen.

- 2) Gleichung (58) wird angewendet,

$$\frac{1}{q} \sum_j s_j = s_r \quad (58)$$

um den Schlusswert der Wiederholstandardabweichung zu berechnen. Die Schritte 3) bis 6) werden übergangen und es wird unmittelbar mit 8.6.14 fortgefahren.

- 3) Aus der grafischen Darstellung von s gegen m von s^2 und m^2 beurteilt, ob die Beziehungen zwischen s und m oder zwischen s^2 und m^2 durch eine gerade Linie repräsentiert werden können, und falls das so ist, ob die Beziehung I ($s = bm$), die Beziehung II ($s = a + bm$) oder die Beziehung III ($s_j^2 = a_p^2 + (b_p m)^2$) angemessen ist (siehe 8.5.1). Für die Beziehungen I und II werden der Parameter b oder die beiden Parameter a und b durch das Verfahren nach 8.5.2 bestimmt. Für die Beziehung III wird das Verfahren nach 8.5.3 verwendet. Wenn eine dieser Beziehungen als zufriedenstellend gilt, wird Schritt 4) übergangen und es wird direkt bei Schritt 5) fortgefahren. Andernfalls, wird mit Schritt 4) fortgefahren.
- 4) $\lg s_j$ wird gegen $\lg \hat{m}$ aufgezeichnet und darauf wird beurteilt, ob die Beziehung zwischen $\lg s$ und $\lg m$ gut angenähert durch eine Gerade wiedergegeben werden kann. Wenn dies als zufriedenstellend betrachtet wird, wird die Beziehung IV ($\lg s = c + d \lg m$) durch Anwendung des Verfahrens nach 8.5.4 (an diese Gerade) angepasst.
- 5) Ist in Schritt 3) oder 4) eine zufriedenstellende Beziehung festgestellt worden, dann sind die Schlusswerte von s_r (oder s_R) die aus dieser Beziehung erhaltenen (durch die Gerade von zufälligen Streuungen befreiten), geglätteten Werte für gegebene Werte von m . Schritt 6) wird übergangen und es wird mit 8.6.14 fortgefahren.
- 6) Wenn in Schritt 3) oder 4) keine zufriedenstellende Beziehung festgestellt wurde, sollte die statistische Fachkraft entscheiden, ob eine andere Beziehung zwischen s und m festgestellt werden kann, oder alternativ, ob die Daten so irregulär sind, dass die Feststellung einer funktionalen Beziehung als unmöglich betrachtet wird. Wenn die Daten zu irregulär sind, um eine funktionale Beziehung festzustellen, dürfen Präzisionswerte separat für jedes Niveau angegeben werden.

8.6.14 Ein Bericht wird angefertigt, der die Basisdaten sowie die Ergebnisse und Schlussfolgerungen der statistischen Auswertung zeigt. Dieser Bericht wird dem Versuchsgremium übergeben. Die grafischen Darstellungen nach 8.3.2 können bei der Darlegung der Vereinbarkeit oder der Variabilität der Ergebnisse nützlich sein.

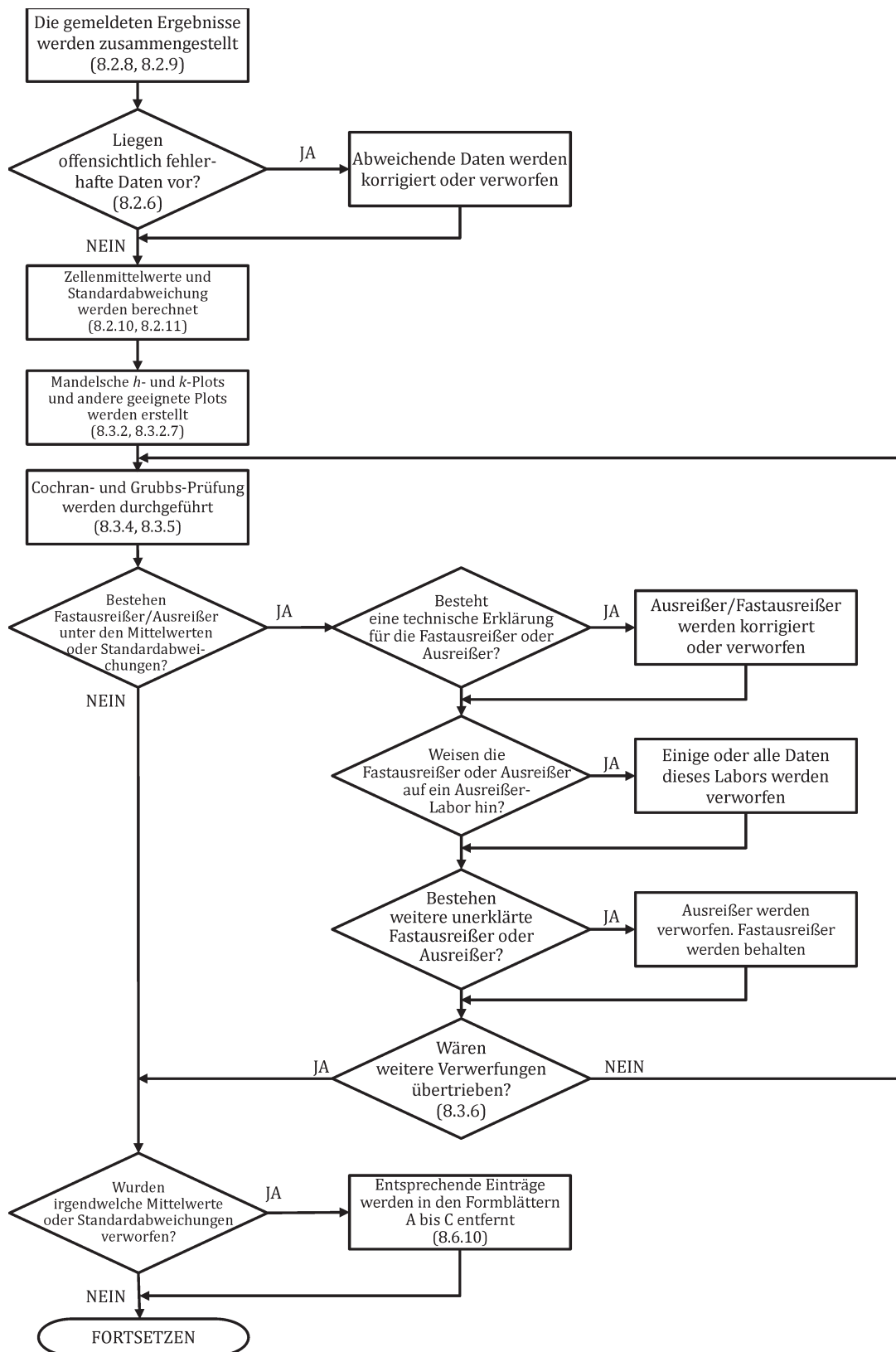
8.7 Bericht an das Versuchsgremium und vom Versuchsgremium zu treffende Entscheidungen

8.7.1 Bericht durch die statistische Fachkraft

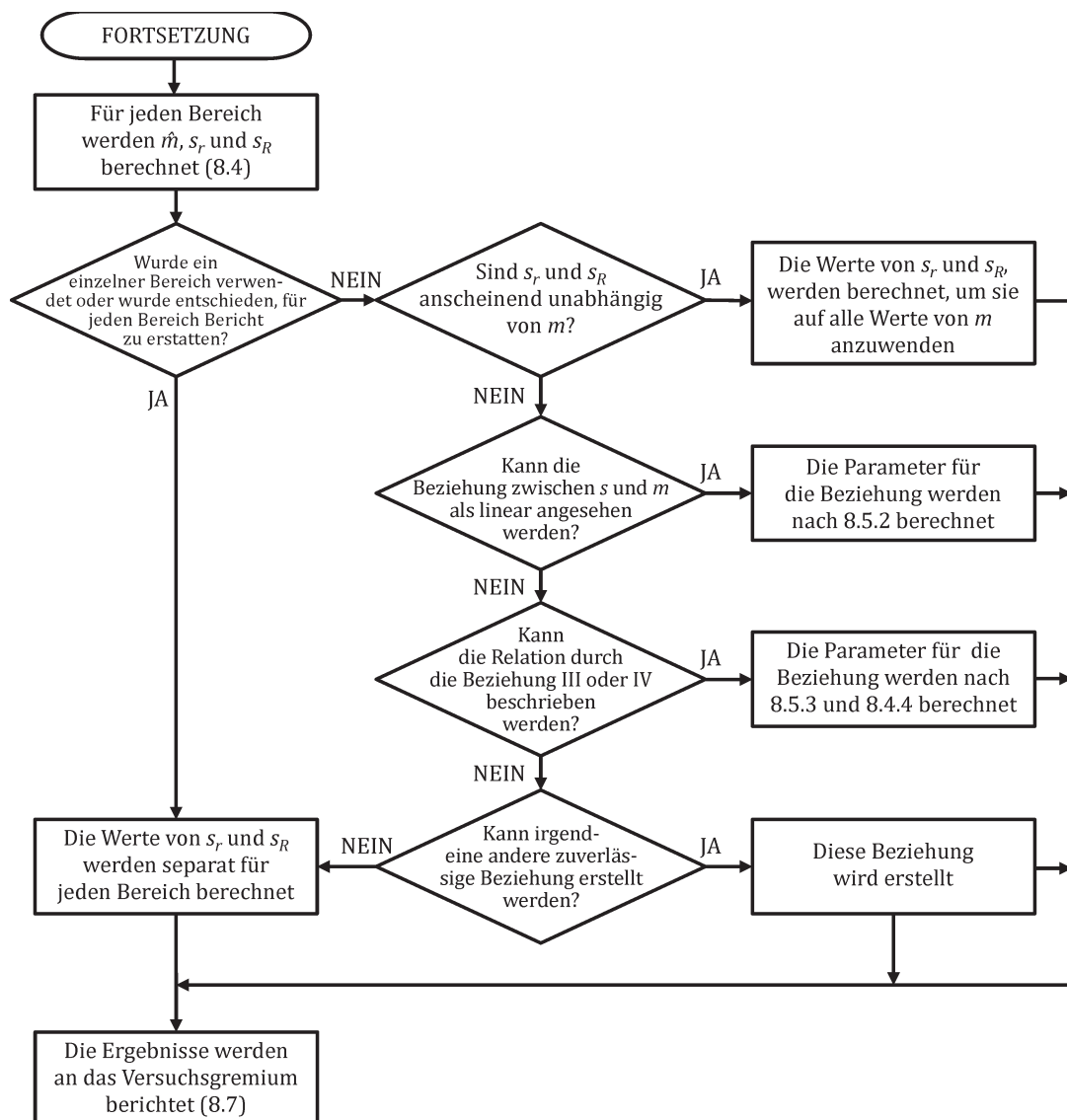
Nach Fertigstellung der statistischen Analyse sollte die statistische Fachkraft einen Bericht schreiben, der dem Versuchsgremium unterbreitet wird. In diesem Bericht sollten die folgenden Informationen gegeben werden:

- a) eine vollständige Darstellung der von den Bearbeitern und/oder Laborverantwortlichen erhaltenen Beobachtungen bezüglich des vereinheitlichten Messverfahrens;

- b) eine vollständige Aufstellung der Laboratorien, die nach 8.6.2 und 8.6.8 als Ausreißerlabors abgelehnt wurden, und die Gründe für ihre Ablehnung;
- c) eine vollständige Aufzeichnung irgendwelcher Fastausreißer und/oder statistischen Ausreißer, und ob diese erklärt und korrigiert, oder aussortiert wurden;
- d) ein Formblatt der Schlussergebnisse $\hat{m}_j$, s_r und s_R sowie eine Darstellung der Schlussfolgerungen, die nach Schritt 8.6.13 erzielt wurden, veranschaulicht durch eine der in diesen Schritten empfohlenen grafischen Darstellungen;
- e) die bei der statistischen Auswertung verwendeten Formblätter A, B und C (Bild 2), möglichst als Anlage zum Bericht.



a) Flussdiagramm der Hauptschritte der statistischen Analyse (1 von 2)



b) Flussdiagramm der Hauptschritte der statistischen Analyse (2 von 2)

Bild 3 — Flussdiagramm der Hauptschritte der statistischen Analyse

8.7.2 Vom Versuchsgremium zu treffende Entscheidungen

Das Versuchsgremium sollte anschließend diesen Bericht diskutieren und Entscheidungen zu den folgenden Fragen treffen.

- Sind die unverträglichen Ergebnisse, Fastausreißer oder Ausreißer, sofern es solche gab, etwa eine Folge von Fehlern in der Beschreibung des (vereinheitlichten) Messverfahrens?
- Welche Maßnahmen sollten im Hinblick auf die als Ausreißer abgelehnten Labors getroffen werden?
- Zeigen die Ergebnisse der Ausreißerlabors und/oder die von den Bearbeitern und Laborverantwortlichen erhaltenen Kommentare die Notwendigkeit, die Festlegungen des (vereinheitlichten) Messverfahrens zu verbessern? Wenn das so ist: Welche Verbesserungen sind nötig?
- Rechtfertigen die Ergebnisse des Präzisionsversuchs die Feststellung von Werten für die Wiederholstandardabweichung und die Vergleichstandardabweichung? Wenn das so ist: Welches sind diese Werte, in

welcher Form sollten sie bekannt gemacht werden, und was ist der Bereich, in dem die Präzisionsdaten zutreffen?

8.7.3 Abschlussbericht

Ein Bericht sollte durch die leitende Führungskraft erstellt und durch das Versuchsgremium genehmigt werden. Er sollte die Gründe für die Arbeit erläutern und wie sie organisiert wurde, den Bericht der statistischen Fachkraft einbeziehen und die abgestimmten Schlussfolgerungen herausstellen. Eine grafische Darstellung der Vereinbarkeit oder Variabilität ist oft nützlich. Der Bericht sollte an die für die Genehmigung der Arbeit Verantwortlichen und andere interessierte Stellen verteilt werden.

9 Statistische Tabellen

9.1 Kritische Werte für den Cochran-Test (siehe 8.3.4) sind in Tabelle 5 angegeben.

9.2 Kritische Werte für den Grubbs-Test (siehe 8.3.5) sind in Tabelle 6 angegeben.

Beim Grubbs-Test für eine einzelne Ausreißer-Beobachtung verursachen Ausreißer und Fastausreißer Werte, die größer als die Tabellenwerte für die kritischen Werte zu 1 % bzw. zu 5 % sind.

Beim Grubbs-Test für zwei Ausreißer-Beobachtungen (auf derselben Seite der Verteilung) verursachen Ausreißer und Fastausreißer Werte, die kleiner als die Tabellenwerte für die kritischen Werte zu 1 % bzw. zu 5 % sind.

9.3 Indikatoren für Mandels Vereinbarkeits-Prüfgrößen h und k (siehe 8.3.2) sind in Tabelle 7 und Tabelle 8 angegeben.

Tabelle 5 — Kritische Werte für den Cochran-Test

p	$n = 2$		$n = 3$		$n = 4$		$n = 5$		$n = 6$	
	1 %	5 %	1 %	5 %	1 %	5 %	1 %	5 %	1 %	5 %
2	—	—	0,995	0,975	0,979	0,939	0,959	0,906	0,937	0,877
3	0,993	0,967	0,942	0,871	0,883	0,798	0,834	0,746	0,793	0,707
4	0,968	0,906	0,864	0,768	0,781	0,684	0,721	0,629	0,676	0,590
5	0,928	0,841	0,788	0,684	0,696	0,598	0,633	0,544	0,588	0,506
6	0,883	0,781	0,722	0,616	0,626	0,532	0,564	0,480	0,520	0,445
7	0,838	0,727	0,664	0,561	0,568	0,480	0,508	0,431	0,466	0,397
8	0,794	0,680	0,615	0,516	0,521	0,438	0,463	0,391	0,423	0,360
9	0,754	0,638	0,573	0,478	0,481	0,403	0,425	0,358	0,387	0,329
10	0,718	0,602	0,536	0,445	0,447	0,373	0,393	0,331	0,357	0,303
11	0,684	0,570	0,504	0,417	0,418	0,348	0,366	0,308	0,332	0,281
12	0,653	0,541	0,475	0,392	0,392	0,326	0,343	0,288	0,310	0,262
13	0,624	0,515	0,450	0,371	0,369	0,307	0,322	0,271	0,291	0,243
14	0,599	0,492	0,427	0,352	0,349	0,291	0,304	0,255	0,274	0,232
15	0,575	0,471	0,407	0,335	0,332	0,276	0,288	0,242	0,259	0,220
16	0,553	0,452	0,388	0,319	0,316	0,262	0,274	0,230	0,246	0,208
17	0,532	0,434	0,372	0,305	0,301	0,250	0,261	0,219	0,234	0,198
18	0,514	0,418	0,356	0,293	0,288	0,240	0,249	0,209	0,223	0,189
19	0,496	0,403	0,343	0,281	0,276	0,230	0,238	0,200	0,214	0,181

Tabelle 5 (fortgesetzt)

p	$n = 2$		$n = 3$		$n = 4$		$n = 5$		$n = 6$	
	1 %	5 %	1 %	5 %	1 %	5 %	1 %	5 %	1 %	5 %
20	0,480	0,389	0,330	0,270	0,265	0,220	0,229	0,192	0,205	0,174
21	0,465	0,377	0,318	0,261	0,255	0,212	0,220	0,185	0,197	0,167
22	0,450	0,365	0,307	0,252	0,246	0,204	0,212	0,178	0,189	0,160
23	0,437	0,354	0,297	0,243	0,238	0,197	0,204	0,172	0,182	0,155
24	0,425	0,343	0,287	0,235	0,230	0,191	0,197	0,166	0,176	0,149
25	0,413	0,334	0,278	0,228	0,222	0,185	0,190	0,160	0,170	0,144
26	0,402	0,325	0,270	0,221	0,215	0,179	0,184	0,155	0,164	0,140
27	0,391	0,316	0,262	0,215	0,209	0,173	0,179	0,150	0,159	0,135
28	0,382	0,308	0,255	0,209	0,202	0,168	0,173	0,146	0,154	0,131
29	0,372	0,300	0,248	0,203	0,196	0,164	0,168	0,142	0,150	0,127
30	0,363	0,293	0,241	0,198	0,191	0,159	0,164	0,138	0,145	0,124
31	0,355	0,286	0,235	0,193	0,186	0,155	0,159	0,134	0,141	0,120
32	0,347	0,280	0,229	0,188	0,181	0,151	0,155	0,131	0,138	0,117
33	0,339	0,273	0,224	0,184	0,177	0,147	0,151	0,127	0,134	0,114
34	0,332	0,267	0,218	0,179	0,172	0,144	0,147	0,124	0,131	0,111
35	0,325	0,262	0,213	0,175	0,168	0,140	0,144	0,121	0,127	0,108
36	0,318	0,256	0,208	0,172	0,165	0,137	0,140	0,118	0,124	0,106
37	0,312	0,251	0,204	0,168	0,161	0,134	0,137	0,116	0,121	0,103
38	0,306	0,246	0,200	0,164	0,157	0,131	0,134	0,113	0,119	0,101
39	0,300	0,242	0,196	0,161	0,154	0,129	0,131	0,111	0,116	0,099
40	0,294	0,237	0,192	0,158	0,151	0,126	0,128	0,108	0,114	0,097

p = Anzahl der Labors auf einem gegebenen Merkmalsniveau.
 n = Anzahl der Ermittlungsergebnisse je Zelle (siehe 8.3.4.3).
 ANMERKUNG Anhang D bietet basierend auf Literaturhinweis [16] ein Verfahren zur Berechnung dieser kritischen Werte in einer Software.

Tabelle 6 — Kritische Werte für den Grubbs-Test

p	Ein größter oder ein kleinster Wert		Zwei größte oder zwei kleinste Werte	
	Obere 1 %	Obere 5 %	Untere 1 %	Untere 5 %
3	1,155	1,155	—	—
4	1,496	1,481	0,000 0	0,000 2
5	1,764	1,715	0,001 8	0,009 0
6	1,973	1,887	0,011 6	0,034 9
7	2,139	2,020	0,030 8	0,070 8
8	2,274	2,126	0,056 3	0,110 1
9	2,387	2,215	0,085 1	0,149 2
10	2,482	2,290	0,115 0	0,186 4

Tabelle 6 (fortgesetzt)

p	Ein größter oder ein kleinster Wert		Zwei größte oder zwei kleinste Werte	
	Obere 1 %	Obere 5 %	Untere 1 %	Untere 5 %
11	2,564	2,355	0,144 8	0,221 3
12	2,636	2,412	0,173 8	0,253 7
13	2,699	2,462	0,201 6	0,283 6
14	2,755	2,507	0,228 0	0,311 2
15	2,806	2,549	0,253 0	0,336 7
16	2,852	2,585	0,276 7	0,360 3
17	2,894	2,620	0,299 0	0,382 2
18	2,932	2,651	0,320 0	0,402 5
19	2,968	2,681	0,339 8	0,421 4
20	3,001	2,709	0,358 5	0,439 1
21	3,031	2,733	0,376 1	0,455 6
22	3,060	2,758	0,392 7	0,471 1
23	3,087	2,781	0,408 5	0,485 7
24	3,112	2,802	0,423 4	0,499 4
25	3,135	2,822	0,437 6	0,512 3
26	3,157	2,841	0,451 0	0,524 5
27	3,178	2,859	0,463 8	0,536 0
28	3,199	2,876	0,475 9	0,547 0
29	3,218	2,893	0,487 5	0,557 4
30	3,236	2,908	0,498 5	0,567 2
31	3,253	2,924	0,509 1	0,576 6
32	3,270	2,938	0,519 2	0,585 6
33	3,286	2,952	0,528 8	0,594 1
34	3,301	2,965	0,538 1	0,602 3
35	3,316	2,979	0,546 9	0,610 1
36	3,330	2,991	0,555 4	0,617 5
37	3,343	3,003	0,563 6	0,624 7
38	3,356	3,014	0,571 4	0,631 6
39	3,369	3,025	0,578 9	0,638 2
40	3,381	3,036	0,586 2	0,644 5

Nachdruck mit freundlicher Genehmigung der American Statistical Association, aus Literaturhinweis [4].

p = Anzahl der Labors auf einem gegebenen Merkmalsniveau

ANMERKUNG 1 Die in dieser Tabelle angegebenen kritischen Werte eignen sich, wenn zweiseitige Prüfungen erforderlich sind. Hierbei handelt es sich um kritische Werte, die von dem Verfahren zur Anwendung der Grubbs-Ausreißerprüfung nach 8.3.5 in diesem Dokument benötigt werden. Sie wurden aus den kritischen Werten für die entsprechenden einseitigen Prüfungen nach Literaturhinweis [5] abgeleitet.

ANMERKUNG 2 Anhang D bietet basierend auf Literaturhinweis [16] ein Verfahren zur Berechnung dieser kritischen Werte in einer Software aus anderen bekannten Verteilungen.