

Genauigkeit (Richtigkeit und Präzision) von Messverfahren und Messergebnissen

Teil 6: Anwendung von Genauigkeitswerten in der Praxis
(ISO 5725-6:1994 einschließlich Technisches Korrigendum 1:2001)

DIN

ISO 5725-6

ICS 17.020

Accuracy (trueness and precision) of measurement methods and results — Part 6: Use in practice of accuracy values
(ISO 5725-6:1994 including Technical Corrigendum 1:2001)

Exactitude (justesse et fidélité) des résultats et méthodes de mesure — Partie 6: Utilisation dans la pratique des valeurs d'exactitude (ISO 5725-6:1994, Rectificatif Technique 1:2001 inclus)

Die Internationale Norm ISO 5725-6:1994-12-15, "Accuracy (trueness and precision) of measurement methods and results — Part 6: Use in practice of accuracy values", einschließlich Technisches Korrigendum 1:2001, ist unverändert in diese Deutsche Norm übernommen worden.

Nationales Vorwort

Die Internationale Norm ISO 5725 wurde vom Technischen Komitee 69 „Anwendung statistischer Methoden“ der Internationalen Organisation für Normung (ISO) erarbeitet. Die 2. Ausgabe 1986 wurde in der nachfolgenden 3. Ausgabe 1994 in sechs Teile unterteilt. In diesen sind die Teilaspekte zum Gesamtkomplex „Genauigkeit (Richtigkeit und Präzision) von Messverfahren und Messergebnissen“ einzeln dargestellt. Die deutschen Übersetzungen sind vom Normenausschuss Qualitätsmanagement, Statistik und Zertifizierungsgrundlagen im DIN erstellt worden. Das Vorwort in jedem der sechs Teile gibt Auskunft über die jeweils behandelten Teilkomplexe.

Die Normen der Reihe DIN ISO 5725 befassen sich mit der Ermittlung und Anwendung von Präzisions- und Richtigkeitsmaßen eines vereinheitlichten Messverfahrens mittels Ringversuchen. Die Genauigkeit eines Messverfahrens wird dadurch beschrieben, dass sowohl ein Richtigkeits- als auch ein Präzisionsmaß gemeinsam angegeben werden. Nach DIN 55350-13 ist Präzision die qualitative Bezeichnung für das Ausmaß der gegenseitigen Annäherung voneinander unabhängiger Ermittlungsergebnisse bei mehrfacher Anwendung eines vereinheitlichten Messverfahrens unter vorgegebenen Bedingungen. Nach der gleichen Norm ist Richtigkeit die qualitative Bezeichnung für das Ausmaß der Annäherung des Erwartungswertes des Ermittlungsergebnisses an den Bezugswert, wobei dieser je nach Festlegung oder Vereinbarung der wahre oder der richtige Wert sein kann. Der Erwartungswert ist dabei als gedanklich feststellbarer Wert das mittlere Ergebnis bei unablässig wiederholter Anwendung des vereinheitlichten Messverfahrens.

Die Normen der Reihe DIN ISO 5725 sind nur auf kontinuierliche Merkmale anwendbar (siehe 1.2).

Als Maß für die Richtigkeit wird die Differenz zwischen einem Schätzwert für den Erwartungswert und dem gewählten Bezugswert benutzt. Die Differenz zwischen Erwartungswert und Bezugswert nennt man auch „bias“. Im Deutschen wird dieses Maß für die Richtigkeit als „systematische Abweichungskomponente“ bezeichnet, um sie von einzelnen systematischen Abweichungen zu unterscheiden.

Fortsetzung Seite 2 bis 91

Normenausschuss Qualitätsmanagement, Statistik und Zertifizierungsgrundlagen (NQSZ)
im DIN Deutsches Institut für Normung e.V.

Maße für die Präzision eines vereinheitlichten Messverfahrens sind im Allgemeinen die Standardabweichungen der unter Wiederholbedingungen, Zwischenbedingungen oder Vergleichbedingungen gewonnenen Messwerte. Aus dem Präzisionsmaß Standardabweichung kann man auch die früher häufig benutzten speziellen Präzisionsmaße Wiederholgrenze r und Vergleichgrenze R als „kritische Differenzbeträge“ berechnen. Diese sind im Vergleich mit den Standardabweichungen um einen angebbaren Faktor größere Beträge, unter denen oder höchstens gleich denen der Betrag der Differenz zweier Einzelergebnisse (gemessen unter den betreffenden Bedingungen) mit einer vorgegebenen Wahrscheinlichkeit P erwartet werden kann. Für diese Wahrscheinlichkeit wird bei Abwesenheit anderer Festlegungen oder Vereinbarungen $P = 95\%$ eingesetzt. Der genannte Faktor hängt von der gewählten oder vorgegebenen Wahrscheinlichkeit ab und hat bei $P = 95\%$ den Zahlenwert 2,8¹⁾. Dem Anwender der Norm (beispielsweise einem Normenausschuss) ist es nach wie vor prinzipiell freigestellt, ob er als Präzisionsmaße eines vereinheitlichten Messverfahrens

- Wiederholstandardabweichung, Zwischenstandardabweichung und Vergleichstandardabweichung oder
- Wiederholgrenze, Zwischengrenze und Vergleichgrenze

angibt. Empfohlen wird, nur die im vorausgehenden ersten Anstrich genannten Größen zu verwenden.

Nicht jeder für ein betrachtetes vereinheitlichtes Messverfahren durchgeführte Ringversuch führt zu den gleichen Werten der Richtigkeitsmaße bzw. der Präzisionsmaße. Diese Ergebnisse sind nämlich Zufallsergebnisse. Überdies können die Anzahl der beteiligten Labors und die Anzahl der Ermittlungen im Labor unterschiedlich sein. Die ermittelten Werte der Richtigkeitsmaße bzw. der Präzisionsmaße sind also Schätzwerte. Für einen Industrie- oder Wirtschaftssektor oder für ein Fachgebiet können für sie jedoch verbindliche Grenzwerte eingeführt werden, falls es erforderlich ist. Das kann beispielsweise aus juristischen Gründen nötig sein.

In den Normen der Reihe DIN ISO 5725 werden außer den eingangs für Präzision und Richtigkeit genannten die folgenden Begriffe benutzt:

- Ringversuch = Versuch unter Beteiligung mehrerer Labors, z. B. zur Ermittlung der Präzision oder der Richtigkeit eines Messverfahrens nach dieser Normenreihe (siehe auch DIN ISO 5725-1:1997-11, 3.22).

ANMERKUNG Ein Ringversuch kann unterschiedliche Zwecke haben. Beispielsweise kann er nur auf Präzisionsmaße oder nur auf Richtigkeitsmaße oder auf beide ausgerichtet sein. Unterbegriffe zum Ringversuch für dessen unterschiedliche Ziele sind nicht erforderlich.

- Labor = Stelle, die Prüfungen oder nur Ermittlungen ausführt.

ANMERKUNG Gelegentlich wird dieser Begriff auch im Sinn des Untersuchungsorts verwendet.

- Vereinheitlichtes Messverfahren: Siehe DIN ISO 5725-1:1997-11, 4.1.

ANMERKUNG Bisher wurde es „festgelegtes Messverfahren“ genannt.

Im Übrigen wird auf die Normen über Begriffe zu Qualitätsmanagement und Statistik der Normenreihen DIN 55350 und DIN 13303 hingewiesen.

Die vorliegende Normenreihe DIN ISO 5725 ist eine Rahmenregelung für zahlreiche Fachgebiete. Sie bedarf bei der Anwendung gegebenenfalls ergänzender Festlegungen, insbesondere zur Anzahl der am Ringversuch beteiligten Labors und zur Anzahl der Merkmalsniveaus und der Wiederholmessungen. Für beides enthält die Normenreihe Empfehlungen. Branchenspezifisch kann eine feste Anzahl oder Mindestanzahl von Labors, Merkmalsniveaus und von Wiederholmessungen pro Merkmalsniveau vorgegeben werden.

Die vorliegende Norm behandelt unter Anwendung der in DIN ISO 5725-1 dargelegten, für alle Normen der Reihe DIN ISO 5725 geltenden allgemeinen Grundlagen und Begriffe, die Art und Weise der Anwendung von Genauigkeitswerten in unterschiedlichen Situationen der Praxis. Dabei wird unterstellt, dass die Genauigkeitswerte nach den grundlegenden Methoden für die Ermittlung der beiden Präzisionsmaße eines vereinheitlichten Messverfahrens gewonnen wurden, bei denen dasselbe vereinheitlichte Messverfahren an identischen Versuchsobjekten benutzt wurde, wobei die Untersuchungen entweder im selben Labor durch denselben Bearbeiter mit derselben Geräteausrüstung, oder in verschiedenen Labors durch verschiedene Bearbeiter mit verschiedenen Geräteausrüstungen durchgeführt wurden.

¹⁾ Dieser Faktor ergibt sich aus dem auf 2,0 gerundeten Quantil $u_{97,5\%} = 1,96$ sowie aus $\sqrt{n}$ mit $n = 2$. Aus deren Multiplikation folgt, für Zwecke der Schätzung hinreichend genau, der Faktor 2,8.

Insbesondere wird ein vereinheitlichtes Verfahren zur Errechnung der Grenzen für die Wiederholpräzision, die Vergleichpräzision und die Zwischenpräzision sowie für andere Prüfmerkmale vorgestellt. Sie werden durch ausführlich dargelegte Praxisbeispiele veranschaulicht. Außerdem wird die Annehmbarkeit der Ergebnisse eines Ringversuchs behandelt. Weiter wird beschrieben, wie man prüft, inwieweit ein betrachtetes Labor in der Lage ist, ein vereinheitlichtes Messverfahren in zufriedenstellender Weise anzuwenden, und inwieweit während einer betrachteten Zeitspanne die Ergebnisse eines Labors als stabil angesehen werden können.

Ein letztes Kapitel befasst sich mit Verfahren, die einen systematischen Vergleich alternativer Messverfahren ermöglichen. Das ist unter anderem für die Vereinheitlichung neuer Messverfahren von Bedeutung. Dieser Vergleich wird mittels eines gemeinsamen Versuchsprogramms von Labors durchgeführt, der in vielen Elementen einem Ringversuch gleichkommt. Auch hierzu werden erläuternde Beispiele angegeben.

Die Norm DIN ISO 5725 wurde in DIN ISO 5725-1 bis DIN ISO 5725-6 aufgeteilt und erweitert. Diese Normen enthalten jetzt außer der bisher behandelten Ermittlung von Präzisionsmaßen auch

- die Ermittlung von Richtigkeitsmaßen;
- den Vergleich von Labors;
- die Verwendung von Genauigkeitsmaßen in der Praxis;
- alternative Verfahren für die Ermittlung von Präzisionsmaßen.

Nationaler Anhang NA (informativ)

Literaturhinweise

DIN 13303-1:1982-05, *Stochastik — Wahrscheinlichkeitstheorie, Gemeinsame Grundbegriffe der mathematischen und der beschreibenden Statistik — Begriffe und Zeichen.*

DIN 13303-2:1982-11, *Stochastik — Mathematische Statistik — Begriffe und Zeichen.*

DIN 55350-12:1989-03, *Begriffe der Qualitätssicherung und Statistik — Merkmalsbezogene Begriffe.*

DIN 55350-13:1987-07, *Begriffe der Qualitätssicherung und Statistik — Begriffe zur Genauigkeit von Ermittlungsverfahren und Ermittlungsergebnissen.*

DIN 55350-14:1985-12, *Begriffe der Qualitätssicherung und Statistik — Begriffe der Probenahme.*

DIN 55350-15:1986-02, *Begriffe der Qualitätssicherung und Statistik — Begriffe zu Mustern.*

DIN 55350-17:1988-08, *Begriffe der Qualitätssicherung und Statistik — Begriffe der Qualitätsprüfungsarten.*

DIN 55350-21:1982-05, *Begriffe der Qualitätssicherung und Statistik — Begriffe der Statistik — Zufallsgrößen und Wahrscheinlichkeitsverteilungen.*

DIN 55350-22:1987-02, *Begriffe der Qualitätssicherung und Statistik — Begriffe der Statistik — Spezielle Wahrscheinlichkeitsverteilungen.*

DIN 55350-23:1983-04, *Begriffe der Qualitätssicherung und Statistik — Begriffe der Statistik — Beschreibende Statistik.*

DIN 55350-24:1982-11, *Begriffe der Qualitätssicherung und Statistik — Begriffe der Statistik — Schließende Statistik.*

DIN EN ISO 9000:2000-12, *Qualitätsmanagementsysteme – Grundlagen und Begriffe.*

Deutsche Übersetzung

Genauigkeit (Richtigkeit und Präzision) von Messverfahren und Messergebnissen —

Teil 6: Anwendung von Genauigkeitswerten in der Praxis

Inhalt

	Seite
1 Anwendungsbereich	10
2 Normative Verweisungen	10
3 Begriffe	12
4 Ermittlung von Grenzen	12
4.1 Grenzen für Wiederholpräzision und Vergleichpräzision	12
4.2 Auf mehr als zwei Werten beruhende Vergleiche	14
5 Verfahren zur Prüfung der Annehmbarkeit von Ermittlungsergebnissen und zur Ermittlung des anzugebenden Schlussergebnisses	16
5.1 Allgemeines	16
5.2 Verfahren zur Prüfung der Annehmbarkeit von Ermittlungsergebnissen, gewonnen unter Wiederholbedingungen	16
5.3 Verfahren zur Prüfung der Annehmbarkeit von Ermittlungsergebnissen, gewonnen unter Wiederholbedingungen wie auch unter Vergleichbedingungen	30
6 Verfahren zur Prüfung der Stabilität von Ermittlungsergebnissen eines Labors	34
6.1 Zusammenhänge	34
6.2 Verfahren zur Stabilitätsprüfung	34
7 Anwendung von Wiederhol- und Vergleichstandardabweichungen bei der Beurteilung von Labors	58
7.1 Beurteilungsverfahren	58
7.2 Beurteilung der Anwendung eines Messverfahrens durch ein vorher nicht beurteiltes Labor	60
7.3 Fortgesetzte Beurteilung von vorher zugelassenen Labors	66
8 Vergleich alternativer Messverfahren	72
8.1 Herkunft alternativer Messverfahren	72
8.2 Zweck des Vergleichs von Messverfahren	72
8.3 Verfahren B als Anwärter für ein alternatives vereinheitlichtes Messverfahren („Vereinheitlichungsversuch“ nicht festgelegt)	74
8.4 Ringversuch	74
8.5 Messverfahren B ist ein Anwärter für ein Routine-Messverfahren	84
Anhang A (normativ)	
Formelzeichen und Abkürzungen in ISO 5725	88

Accuracy (trueness and precision) of measurement methods and results —

Part 6:

Use in practice of accuracy values

Contents	Page
1 Scope	11
2 Normative references	11
3 Definitions	13
4 Determination of limits	13
4.1 Repeatability and reproducibility limits	13
4.2 Comparisons based on more than two values	15
5 Methods for checking the acceptability of test results and determining the final quoted result	17
5.1 General	17
5.2 Methods for checking the acceptability of test results obtained under repeatability conditions	17
5.3 Methods for checking the acceptability of test results obtained under both repeatability and reproducibility conditions	31
6 Method for checking the stability of test results within a laboratory	35
6.1 Background	35
6.2 Methods for checking stability	35
7 Use of repeatability and reproducibility standard deviations in assessing laboratories	59
7.1 Assessment method	59
7.2 Evaluation of the use of a measurement method by a laboratory not previously assessed	61
7.3 Continued assessment of previously approved laboratories	67
8 Comparison of alternative measurement methods	73
8.1 Origin of alternative measurement methods	73
8.2 Purpose of comparing measurement methods	73
8.3 Method B is a candidate for an alternative standard method ("Standardization experiment" not defined)	75
8.4 Accuracy experiment	75
8.5 Method B is a candidate for a routine method	85
Annex	
A Symbols and abbreviations used in ISO 5725	89

Vorwort

Die ISO (Internationale Organisation für Normung) ist die weltweite Vereinigung nationaler Normungsinstitute (ISO-Mitglieds Körperschaften). Die Erarbeitung Internationaler Normen obliegt den Technischen Komitees der ISO. Jede Mitglieds Körperschaft, die sich für ein Thema interessiert, für das ein Technisches Komitee eingesetzt wurde, ist berechtigt, in diesem Komitee mitzuarbeiten. Internationale (staatliche und nichtstaatliche) Organisationen, die mit der ISO in Verbindung stehen, sind an den Arbeiten ebenfalls beteiligt. Die ISO arbeitet in allen Angelegenheiten der elektrotechnischen Normung eng mit der Internationalen Elektrotechnischen Kommission (IEC) zusammen.

Die von den Technischen Komitees verabschiedeten internationalen Norm-Entwürfe werden den Mitglieds Körperschaften zur Abstimmung vorgelegt. Die Veröffentlichung als Internationale Norm erfordert Zustimmung von mindestens 75 % der abstimmenden Mitglieds Körperschaften.

Die Internationale Norm ISO 5725-6 wurde durch das Technische Komitee ISO/TC 69, *Anwendung statistischer Methoden*, Unterkomitee SC 6, *Messmethoden und Messergebnisse*, erarbeitet.

ISO 5725 besteht unter dem Haupttitel „*Genauigkeit (Richtigkeit und Präzision) von Messverfahren und Messergebnissen*“ aus den folgenden Teilen:

- Teil 1: Allgemeine Grundlagen und Begriffe
- Teil 2: Grundlegende Methode für die Ermittlung der Wiederhol- und Vergleichpräzision eines vereinheitlichten Messverfahrens
- Teil 3: Präzisionsmaße eines vereinheitlichten Messverfahrens unter Zwischenbedingungen
- Teil 4: Grundlegende Methoden für die Ermittlung der Richtigkeit eines vereinheitlichten Messverfahrens
- Teil 5: Alternative Methoden für die Ermittlung der Präzision eines vereinheitlichten Messverfahrens
- Teil 6: Anwendung von Genauigkeitswerten in der Praxis

Teil 1 bis Teil 6 von ISO 5725 ersetzen gemeinsam die Norm ISO 5725:1986 und machen diese ungültig, welche erweitert wurde, um Richtigkeit (ergänzend zu Präzision) sowie Zwischenbedingungen (ergänzend zu Wiederhol- und Vergleichbedingungen) abzudecken.

Der Anhang A ist normativer Bestandteil dieses Teils von ISO 5725.

Foreword

ISO (the International Organization for Standardization) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established has the right to be represented on that committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work. ISO collaborates closely with the International Electrotechnical Commission (IEC) on all matters of electrotechnical standardization.

Draft International Standards adopted by the technical committees are circulated to the member bodies for voting. Publication as an International Standard requires approval by at least 75 % of the member bodies casting a vote.

International Standard ISO 5725-6 was prepared by Technical Committee ISO/TC 69, *Applications of statistical methods*, Subcommittee SC 6, *Measurement methods and results*.

ISO 5725 consists of the following parts, under the general title *Accuracy (trueness and precision) of measurement methods and results*:

- *Part 1: General principles and definitions*
- *Part 2: Basic method for the determination of repeatability and reproducibility of a standard measurement method*
- *Part 3: Intermediate measures of the precision of a standard measurement method*
- *Part 4: Basic methods for the determination of the trueness of a standard measurement method*
- *Part 5: Alternative methods for the determination of the precision of a standard measurement method*
- *Part 6: Use in practice of accuracy values*

Parts 1 to 6 of ISO 5725 together cancel and replace ISO 5725:1986, which has been extended to cover trueness (in addition to precision) and intermediate precision conditions (in addition to repeatability and reproducibility conditions).

Annex A forms an integral part of this part of ISO 5725.

Einleitung

0.1 ISO 5725 benutzt die beiden Ausdrücke „Richtigkeit“ und „Präzision“, um die Genauigkeit eines Messverfahrens zu beschreiben. „Richtigkeit“ bezieht sich auf das Ausmaß der Annäherung zwischen dem arithmetischen Mittelwert einer großen Anzahl von Ermittlungsergebnissen und dem wahren oder dem richtigen Wert. „Präzision“ bezieht sich auf das Ausmaß der gegenseitigen Annäherung von Ermittlungsergebnissen.

0.2 Die Notwendigkeit zur Betrachtung der Präzision ergibt sich daraus, dass an vermutlich identischem Material unter vermutlich identischen Umständen ausgeführte Ermittlungen im Allgemeinen nicht zu identischen Ergebnissen führen. Dies ist unvermeidbaren zufälligen Abweichungen zuzuschreiben, die jedem Messverfahren eigen sind. Die Faktoren, welche das Ergebnis einer Messung beeinflussen, können nämlich nicht alle vollständig beherrscht werden. Bei der praktischen Verwertung von Messwerten muss diese Variabilität in Rechnung gestellt werden. Beispielsweise kann der Unterschied zwischen einem Ermittlungsergebnis und einem vorgegebenen Wert im Rahmen unvermeidbarer zufälliger Abweichungen liegen, so dass eine tatsächliche Abweichung von einem solchen vorgegebenen Wert nicht festgestellt werden kann. Ebenso kann man beim Vergleich von Ermittlungsergebnissen an zwei Materiallosen nicht auf eine prinzipielle Unterschiedlichkeit der Beschaffenheit schließen, sofern der Unterschied zwischen ihnen der dem Messverfahren eigenen Variabilität zugeordnet werden kann.

0.3 Die Teile 1 bis 5 von ISO 5725 behandeln im Hinblick auf ein vereinheitlichtes Messverfahren die Zusammenhänge und vorgegebenen Verfahren für die Beurteilung der Präzision (was die Wiederholstandardabweichung und die Vergleichstandardabweichung betrifft) sowie der Richtigkeit (was die unterschiedlichen Bestandteile der systematischen Abweichungskomponente betrifft) aus den Messergebnissen. Solche Beurteilungen wären wohl sinnlos, wenn es keine praktischen Anwendungen gäbe, für welche die Schätzergebnisse eingesetzt werden könnten.

0.4 Unter der Voraussetzung, dass die Genauigkeit eines (vereinheitlichten) Messverfahrens festgestellt worden ist, wendet dieser Teil von ISO 5725 diese Kenntnis für die praktische Anwendung in einer Weise an, welche die Vereinfachung der geschäftlichen Abwicklung sowie die Überwachung und Verbesserung der Arbeitsabläufe von Labors bezweckt.

Introduction

0.1 ISO 5725 uses two terms “trueness” and “precision” to describe the accuracy of a measurement method. “Trueness” refers to the closeness of agreement between the arithmetic mean of a large number of test results and the true or accepted reference value. “Precision” refers to the closeness of agreement between test results.

0.2 The need to consider “precision” arises because tests performed on presumably identical materials in presumably identical circumstances do not, in general, yield identical results. This is attributed to unavoidable random errors inherent in every measurement procedure; the factors that influence the outcome of a measurement cannot all be completely controlled. In the practical interpretation of measurement data, this variability has to be taken into account. For instance, the difference between a test result and some specified value may be within the scope of unavoidable random errors, in which case a real deviation from such a specified value has not been established. Similarly, comparing test results from two batches of material will not indicate a fundamental quality difference if the difference between them can be attributed to the inherent variation in the measurement procedure.

0.3 Parts 1 to 5 of ISO 5725 discuss the background to, and given methods for, the assessment of the precision (in terms of the repeatability standard deviation and the reproducibility standard deviation) and the trueness (in terms of the various components of bias) of measurements produced by a standard measurement method. Such assessment would, however, be pointless if there were no practical uses to which the results could be put.

0.4 Given that the accuracy of a measurement method has been established, this part of ISO 5725 applies that knowledge in practical situations in such a way as to facilitate commercial transactions and to monitor and improve the operational performance of laboratories.

1 Anwendungsbereich

1.1 Der Zweck dieses Teils von ISO 5725 ist es, einige Hinweise zu der Art und Weise zu geben, in welcher in unterschiedlichen Situationen der Praxis Genauigkeitswerte angewendet werden können, und zwar durch

- a) Bereitstellen eines vereinheitlichten Verfahrens zur Ermittlung der Grenze der Wiederholpräzision, der Grenze der Vergleichpräzision und anderer Grenzen, die bei der Prüfung der Ermittlungsergebnisse anzuwenden sind, die in Anwendung eines vereinheitlichten Messverfahrens gewonnen worden sind;
- b) Bereitstellen eines Verfahrens zur Beurteilung der Annehmbarkeit von Ermittlungsergebnissen, die unter Wiederholbedingungen oder Vergleichbedingungen gewonnen worden sind;
- c) Beschreiben, wie die Stabilität von Ermittlungsergebnissen in einem Labor während einer Zeitspanne zu beurteilen ist, und somit Bereitstellen eines Verfahrens zur „Qualitätsprüfung“ der Tätigkeiten in diesem Labor;
- d) Beschreiben, wie zu beurteilen ist, ob ein betrachtetes Labor geeignet ist, ein vorgegebenes vereinheitlichtes Messverfahren in zufriedenstellender Weise anzuwenden;
- e) Beschreiben, wie alternative Messverfahren zu vergleichen sind.

1.2 Dieser Teil von ISO 5725 behandelt ausschließlich Messverfahren, die Messwerte auf einer kontinuierlichen Skala liefern und als Ergebnis eine einzelne Zahl angeben, wobei diese einzelne Zahl allerdings auch das Ergebnis einer Errechnung aus einer Reihe von Beobachtungen sein kann.

1.3 Es wird angenommen, dass die Schätzwerte für Richtigkeit und Präzision für das (betrachtete vereinheitlichte) Messverfahren entsprechend den Teilen 1 bis 5 von ISO 5725 gewonnen worden sind.

1.4 Einige ergänzende Informationen im Hinblick auf das Anwendungsgebiet werden am Anfang jedes einzelnen Anwendungshinweises gegeben.

2 Normative Verweisungen

Die folgenden normativen Dokumente enthalten Festlegungen, die durch Verweisung in diesem Text Bestandteil dieses Teils von ISO 5725 sind. Zum Zeitpunkt der Veröffentlichung dieser Internationalen Norm waren die angegebenen Ausgaben gültig. Alle Normen unterliegen der Überarbeitung. Vertragspartner, deren Vereinbarungen auf diesem Teil von ISO 5725 basieren, werden gebeten, die Möglichkeit zu prüfen, ob die jeweils neuesten Ausgaben der im folgenden genannten Normen angewendet werden können. Die Mitglieder von IEC und ISO führen Verzeichnisse der gegenwärtig gültigen Internationalen Normen.

1 Scope

1.1 The purpose of this part of ISO 5725 is to give some indications of the way in which accuracy data can be used in various practical situations by:

- a) giving a standard method of calculating the repeatability limit, the reproducibility limit and other limits to be used in examining the test results obtained by a standard measurement method;
- b) providing a way of checking the acceptability of test results obtained under repeatability or reproducibility conditions;
- c) describing how to assess the stability of results within a laboratory over a period of time, and thus providing a method of "quality control" of the operations within that laboratory;
- d) describing how to assess whether a given laboratory is able to use a given standard measurement method in a satisfactory way;
- e) describing how to compare alternative measurement methods.

1.2 This part of ISO 5725 is concerned exclusively with measurement methods which yield measurements on a continuous scale and give a single numerical figure as the result, although this single figure may be the outcome of a calculation from a set of observations.

1.3 It is assumed that the estimates of trueness and precision for the method have been obtained in accordance with parts 1 to 5 of ISO 5725.

1.4 Any additional information regarding the field of application will be given at the beginning of each particular application.

2 Normative references

The following standards contain provisions which, through reference in this text, constitute provisions of this part of ISO 5725. At the time of publication, the editions indicated were valid. All standards are subject to revision, and parties to agreements based on this part of ISO 5725 are encouraged to investigate the possibility of applying the most recent editions of the standards indicated below. Members of IEC and ISO maintain registers of currently valid International Standards.

ISO 3534-1:1993, *Statistik — Begriffe und Formelzeichen — Teil 1: Wahrscheinlichkeit und allgemeine statistische Begriffe*.

ISO 5725-1:1994, *Genauigkeit (Richtigkeit und Präzision) von Messverfahren und Messergebnissen — Teil 1: Allgemeine Grundlagen und Begriffe*.

ISO 5725-2:1994, *Genauigkeit (Richtigkeit und Präzision) von Messverfahren und Messergebnissen — Teil 2: Grundlegende Methode für die Ermittlung der Wiederhol- und Vergleichpräzision eines vereinheitlichten Messverfahrens*.

ISO 5725-3:1994, *Genauigkeit (Richtigkeit und Präzision) von Messverfahren und Messergebnissen — Teil 3: Präzisionsmaße eines vereinheitlichten Messverfahrens unter Zwischenbedingungen*.

ISO 5725-4:1994, *Genauigkeit (Richtigkeit und Präzision) von Messverfahren und Messergebnissen — Teil 4: Grundlegende Methoden für die Ermittlung der Richtigkeit eines vereinheitlichten Messverfahrens*.

ISO 8258:1991, *Shewhart-Qualitätsregelkarten*.

ISO/IEC Guide 33:1989^{N1}), *Anwendung von zertifizierten Referenzmaterialien*.

ISO/IEC Guide 35:1989, *Zertifizierung von Referenzmaterialien — Allgemeine und statistische Grundlagen*.

ISO/IEC Guide 25:1990^{N2}), *Allgemeine Anforderungen an die Kompetenz von Kalibrier- und Prüflaboratorien*.

3 Begriffe

Für diesen Teil von ISO 5725 gelten die in ISO 3534-1 und ISO 5725-1 angegebenen Begriffe.

Die in (der Normenreihe) ISO 5725 angewendeten Formelzeichen sind im Anhang A angegeben.

4 Ermittlung von Grenzen

4.1 Grenzen für Wiederholpräzision und Vergleichpräzision

4.1.1 In ISO 5725-2 ist die Aufmerksamkeit auf das Schätzen der Standardabweichungen unter Wiederholbedingungen oder unter Vergleichbedingungen gerichtet. Normale Laborpraxis verlangt jedoch eine Prüfung der (des) Unterschiede(s) zwischen zwei (oder mehreren) Ermittlungsergebnissen. Für diesen Zweck ist eher ein Maß ähnlich einer kritischen Differenz erforderlich als eine Standardabweichung.

4.1.2 Wenn eine Größe auf der Summe oder der Differenz von n unabhängigen Schätzwerten beruht, deren jeder eine Standardabweichung σ hat, dann wird die resultierende Größe eine Standardabweichung von $\sigma\sqrt{n}$ haben. Die Vergleichsgrenze (R) oder die Wiederholgrenze (r) haben für Differenzen zwischen zwei Ermittlungsergebnissen demzufolge die zugehörige Standardabweichung $\sigma\sqrt{2}$. Gemäß normaler statistischer Praxis ist zur Prüfung der Differenz zwischen diesen zwei Werten die anzuwendende kritische Differenz f -mal so groß wie diese Standardabweichung, das heißt $f\sigma\sqrt{2}$. Der Wert von f (kritischer Spannweitenfaktor) hängt von dem der kritischen Differenz zuzuordnenden Wahrscheinlichkeitsniveau und von der Form der zugrunde liegenden Verteilung ab. Für die Wiederholgrenze und die Vergleichsgrenze ist das Wahrscheinlichkeitsniveau mit 95 % festgelegt, und in ISO 5725 ist durchgängig die Annahme gemacht, dass die zugrunde liegende Verteilung näherungsweise normal ist. Für eine Normalverteilung hat f auf dem Wahrscheinlichkeitsniveau 95 % den Wert 1,96. Folglich hat $f\sqrt{2}$ dann den Wert 2,77. Im Hinblick auf den Zweck dieses Teils von ISO 5725, einige einfache „Faustregeln“ zu geben, die von Nichtstatistikern bei der Prüfung von Ermittlungsergebnissen angewendet werden können, erscheint es angemessen, den gerundeten Wert 2,8 für $f\sqrt{2}$ anzuwenden.

4.1.3 Wie festgestellt wurde, führt der Prozess des Schätzens der Präzision zu Schätzwerten für wahre Standardabweichungen, wobei aber diese wahren Standardabweichungen unbekannt bleiben. Deshalb sollten diese Schätzwerte in der statistischen Praxis besser mit s bezeichnet werden als mit σ . Wenn allerdings die in ISO 5725-1 und ISO 5725-2 angegebenen Verfahren befolgt werden, werden diese Schätzwerte auf einer beträchtlichen Anzahl von Ermittlungsergebnissen beruhen. Damit geben sie die beste Information, die wir über die wahren Werte der Standardabweichungen wahrscheinlich haben können. Bei anderen nachfolgenden Anwendungen für Schätzwerte dieser Standardabweichungen, die auf einer geringeren Anzahl von Werten beruhen, wird das Formelzeichen s (Schätzwert einer Standardabweichung) angewendet. Demzufolge scheint es am besten zu sein, wenn man das Formelzeichen σ für Werte anwendet, die aus einem vollständigen Ringversuch stammen, und wenn man diese als wahre Standardabweichungen betrachtet, mit denen andere Schätzwerte (s) verglichen werden.

4.1.4 Unter Berücksichtigung von 4.1.1 bis 4.1.3 muss der Vergleich bei der Prüfung zweier einzelner Ermittlungsergebnisse, die unter Wiederholbedingungen oder unter Vergleichbedingungen gewonnen wurden, ausgeführt werden mit der Wiederholgrenze

$$r = 2,8\sigma,$$

^{N1}) Nationale Fußnote: Ersetzt durch Ausgabe 2000

^{N2}) Nationale Fußnote: Ersetzt durch ISO/IEC 17025: 1999, in Deutschland DIN EN ISO/IEC 17025: 2000-04, *Allgemeine Anforderungen an die Kompetenz von Prüf- und Kalibrierlaboratorien*.

ISO 3534-1:1993, *Statistics — Vocabulary and symbols — Part 1: Probability and general statistical terms*.

ISO 5725-1:1994, *Accuracy (trueness and precision) of measurement methods and results — Part 1: General principles and definitions*.

ISO 5725-2:1994, *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 5725-3:1994, *Accuracy (trueness and precision) of measurement methods and results — Part 3: Intermediate measures of the precision of a standard measurement method*.

ISO 5725-4:1994, *Accuracy (trueness and precision) of measurement methods and results — Part 4: Basic methods for the determination of the trueness of a standard measurement method*.

ISO 8258:1991, *Shewhart control charts*.

ISO Guide 33:1989, *Uses of certified reference materials*.

ISO Guide 35:1989, *Certification of reference materials — General and statistical principles*.

ISO/IEC Guide 25:1990, *General requirements for the competence of calibration and testing laboratories*.

3 Definitions

For the purposes of this part of ISO 5725, the definitions given in ISO 3534-1 and ISO 5725-1 apply.

The symbols used in ISO 5725 are given in annex A.

4 Determination of limits

4.1 Repeatability and reproducibility limits

4.1.1 In ISO 5725-2, attention has been focussed on estimating the standard deviations associated with operations under repeatability or reproducibility conditions. However, normal laboratory practice requires examination of the difference(s) observed between two (or more) test results, and for this purpose some measure akin to a critical difference is required, rather than a standard deviation.

4.1.2 When a quantity is based on sums or differences of n independent estimates each having a standard deviation σ , then that resultant quantity will have a standard deviation $\sigma\sqrt{n}$. The reproducibility limit (R) or repeatability limit (r) are for differences between two test results, so the associated standard deviation is $\sigma\sqrt{2}$. In normal statistical practice, for examining the difference between these two values the critical difference used is f times this standard deviation, i.e. $f\sigma\sqrt{2}$. The value of f (the critical range factor) depends on the probability level to be associated with the critical difference and on the shape of the underlying distribution. For the reproducibility and repeatability limits, the probability level is specified as 95 %, and throughout the analysis in ISO 5725 the assumption is made that the underlying distribution is approximately normal. For a normal distribution at 95 % probability level, f is 1,96 and $f\sqrt{2}$ then is 2,77. As the purpose of this part of ISO 5725 is to give some simple "rule of thumb" to be applied by non-statisticians when examining the results of tests, it seems reasonable to use a rounded value of 2,8 instead of $f\sqrt{2}$.

4.1.3 As has been stated, the process of estimating precision leads to estimates of the true standard deviations while the true standard deviations remain unknown. Therefore in statistical practice they should be denoted by s rather than σ . However, if the procedures given in ISO 5725-1 and ISO 5725-2 are followed, these estimates will be based on an appreciable number of test results, and will give the best information we are likely to have of the true values of the standard deviations. In other applications that follow, for estimates of these standard deviations based on more limited data, the symbol s (estimate of a standard deviation) is used. Therefore it seems best to use the symbol σ to denote the values obtained from a full precision experiment, and treat these as true standard deviations with which other estimates (s) will be compared.

4.1.4 In view of 4.1.1 to 4.1.3, when examining two single test results obtained under repeatability or reproducibility conditions, the comparison shall be made with the repeatability limit

$$r = 2,8\sigma_r$$

or the reproducibility limit

$$R = 2,8\sigma_R$$

oder mit der Vergleichsgrenze

$$R = 2,8\sigma_R.$$

4.2 Auf mehr als zwei Werten beruhende Vergleiche

4.2.1 Zwei Gruppen von Messungen in einem Labor

Werden in einem Labor unter Wiederholbedingungen zwei Gruppen von Messungen ausgeführt, wobei die erste Gruppe mit n_1 Ermittlungsergebnissen einen arithmetischen Mittelwert von $\bar{y}_1$ und die zweite Gruppe mit n_2 Ermittlungsergebnissen einen arithmetischen Mittelwert von $\bar{y}_2$ ergibt, dann ist die Standardabweichung von $(\bar{y}_1 - \bar{y}_2)$

$$\sigma = \sqrt{\sigma_r^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}$$

und die kritische Differenz von $|\bar{y}_1 - \bar{y}_2|$ auf dem Wahrscheinlichkeitsniveau 95 %

$$CD = 2,8\sigma_r \sqrt{\frac{1}{2n_1} + \frac{1}{2n_2}}$$

ANMERKUNG 1 Sind n_1 und n_2 beide 1, reduziert sich dies erwartungsgemäß auf $r = 2,8\sigma_r$.

4.2.2 Zwei Gruppen von Messungen in zwei Labors

Werden im ersten Labor n_1 Ermittlungsergebnisse gewonnen, die einen arithmetischen Mittelwert von $\bar{y}_1$ ergeben, während das zweite Labor n_2 Ermittlungsergebnisse gewinnt, die einen arithmetischen Mittelwert von $\bar{y}_2$ ergeben, in beiden Fällen unter Wiederholbedingungen, dann ist die Standardabweichung von $(\bar{y}_1 - \bar{y}_2)$

$$\begin{aligned} \sigma &= \sqrt{\sigma_L^2 + \frac{1}{n_1}\sigma_r^2 + \sigma_L^2 + \frac{1}{n_2}\sigma_r^2} \\ &= \sqrt{2\sigma_L^2 + \sigma_r^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)} \\ &= \sqrt{2(\sigma_L^2 + \sigma_r^2) - 2\sigma_r^2 \left(1 - \frac{1}{2n_1} - \frac{1}{2n_2} \right)} \end{aligned}$$

und die kritische Differenz von $|\bar{y}_1 - \bar{y}_2|$ auf dem Wahrscheinlichkeitsniveau 95 %

$$CD = \sqrt{(2,8\sigma_R)^2 - (2,8\sigma_r)^2 \left(1 - \frac{1}{2n_1} - \frac{1}{2n_2} \right)}$$

ANMERKUNG 2 Sind n_1 und n_2 beide 1, reduziert sich dies erwartungsgemäß auf $R = 2,8\sigma_R$.

4.2.3 Vergleich mit einem Referenzwert für ein Labor

Muss für n in einem Labor unter Wiederholbedingungen gewonnene Ermittlungsergebnisse, die einen arithmetischen Mittelwert von $\bar{y}$ ergeben, ein Vergleich mit einem vorgegebenen Referenzwert μ_0 durchgeführt werden, und zwar bei fehlender spezieller Kenntnis der systematischen Abweichungskomponente des Labors, ist eine Standardabweichung für $(\bar{y} - \mu_0)$ zu benutzen von

$$\begin{aligned} \sigma &= \sqrt{\sigma_L^2 + \frac{1}{n}\sigma_r^2} \\ &= \frac{1}{\sqrt{2}} \sqrt{2(\sigma_L^2 + \sigma_r^2) - 2\sigma_r^2 \left(1 - \frac{1}{n} \right)} \\ &= \frac{1}{\sqrt{2}} \sqrt{2(\sigma_L^2 + \sigma_r^2) - 2\sigma_r^2 \left(\frac{n-1}{n} \right)} \end{aligned}$$

Die kritische Differenz für $|\bar{y} - \mu_0|$ ist

$$CD = \frac{1}{\sqrt{2}} \sqrt{(2,8\sigma_R)^2 - (2,8\sigma_r)^2 \left(\frac{n-1}{n} \right)}$$

4.2.4 Vergleich mit einem Referenzwert für mehr als ein Labor

Wenn p ($i = 1, 2, 3, \dots, p$) Labors n_i Ermittlungsergebnisse mit dem jeweiligen arithmetischen Mittelwert $\bar{y}_i$ gewonnen haben (in jedem Fall unter Wiederholbedingungen), wobei der Gesamtmittelwert von $\bar{y}$ errechnet wurde zu

$$\bar{y} = \frac{1}{p} \sum \bar{y}_i$$

und wenn dieser Gesamtmittelwert mit einem Referenzwert μ_0 verglichen werden muss, dann ist die Standardabweichung für $(\bar{y} - \mu_0)$

$$\begin{aligned} \sigma &= \sqrt{\frac{1}{p}\sigma_L^2 + \frac{1}{p^2}\sigma_r^2 \sum \frac{1}{n_i}} \\ &= \frac{1}{\sqrt{2p}} \sqrt{2(\sigma_L^2 + \sigma_r^2) - 2\sigma_r^2 + \frac{2\sigma_r^2}{p} \sum \frac{1}{n_i}} \\ &= \frac{1}{\sqrt{2p}} \sqrt{2(\sigma_L^2 + \sigma_r^2) - 2\sigma_r^2 \left(1 - \frac{1}{p} \sum \frac{1}{n_i} \right)} \end{aligned}$$

4.2 Comparisons based on more than two values

4.2.1 Two groups of measurements in one laboratory

If, in one laboratory under repeatability conditions, two groups of measurements are performed with the first group of n_1 test results giving an arithmetic mean of $\bar{y}_1$ and the second group of n_2 test results giving an arithmetic mean of $\bar{y}_2$, then the standard deviation of $(\bar{y}_1 - \bar{y}_2)$ is

$$\sigma = \sqrt{\sigma_r^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}$$

and the critical difference for $|\bar{y}_1 - \bar{y}_2|$ is

$$CD = 2,8\sigma_r \sqrt{\frac{1}{2n_1} + \frac{1}{2n_2}}$$

at the 95 % probability level.

NOTE 1 If n_1 and n_2 are both unity, this reduces to $r = 2,8\sigma_r$, as expected.

4.2.2 Two groups of measurements in two laboratories

If the first laboratory obtains n_1 test results giving an arithmetic mean of $\bar{y}_1$ while the second laboratory obtains n_2 test results giving an arithmetic mean of $\bar{y}_2$, in each case under repeatability conditions, then the standard deviation of $(\bar{y}_1 - \bar{y}_2)$ is

$$\begin{aligned} \sigma &= \sqrt{\sigma_L^2 + \frac{1}{n_1} \sigma_r^2 + \sigma_L^2 + \frac{1}{n_2} \sigma_r^2} \\ &= \sqrt{2\sigma_L^2 + \sigma_r^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)} \\ &= \sqrt{2(\sigma_L^2 + \sigma_r^2) - 2\sigma_r^2 \left(1 - \frac{1}{2n_1} - \frac{1}{2n_2} \right)} \end{aligned}$$

and the critical difference for $|\bar{y}_1 - \bar{y}_2|$ is

$$CD = \sqrt{(2,8\sigma_R)^2 - (2,8\sigma_r)^2 \left(1 - \frac{1}{2n_1} - \frac{1}{2n_2} \right)}$$

at the 95 % probability level.

NOTE 2 If n_1 and n_2 are both unity, this reduces to $R = 2,8\sigma_R$, as expected.

4.2.3 Comparison with a reference value for one laboratory

If n test results are obtained under repeatability conditions within one laboratory which give an arithmetic mean of $\bar{y}$, then the comparison with a given reference value μ_0 shall be made, in the absence of specific knowledge of the laboratory component of bias, using a standard deviation for $(\bar{y} - \mu_0)$ of

$$\begin{aligned} \sigma &= \sqrt{\sigma_L^2 + \frac{1}{n} \sigma_r^2} \\ &= \frac{1}{\sqrt{2}} \sqrt{2(\sigma_L^2 + \sigma_r^2) - 2\sigma_r^2 \left(1 - \frac{1}{n} \right)} \\ &= \frac{1}{\sqrt{2}} \sqrt{2(\sigma_L^2 + \sigma_r^2) - 2\sigma_r^2 \left(\frac{n-1}{n} \right)} \end{aligned}$$

and the critical difference for $|\bar{y} - \mu_0|$ is

$$CD = \frac{1}{\sqrt{2}} \sqrt{(2,8\sigma_R)^2 - (2,8\sigma_r)^2 \left(\frac{n-1}{n} \right)}$$

4.2.4 Comparison with a reference value for more than one laboratory

If p laboratories have obtained n_i test results giving arithmetic means of $\bar{y}_i$ (in each case under repeatability conditions) and the grand mean $\bar{\bar{y}}$ is computed by

$$\bar{\bar{y}} = \frac{1}{p} \sum \bar{y}_i$$

and this grand mean is to be compared with a reference value μ_0 , then the standard deviation for $(\bar{\bar{y}} - \mu_0)$ is

$$\begin{aligned} \sigma &= \sqrt{\frac{1}{p} \sigma_L^2 + \frac{1}{p^2} \sigma_r^2 \sum \frac{1}{n_i}} \\ &= \frac{1}{\sqrt{2p}} \sqrt{2(\sigma_L^2 + \sigma_r^2) - 2\sigma_r^2 + \frac{2\sigma_r^2}{p} \sum \frac{1}{n_i}} \end{aligned}$$

und die kritische Differenz für $\left| \bar{y} - \mu_0 \right|$ auf dem Wahrscheinlichkeitsniveau 95 % ist

$$CD = \frac{1}{\sqrt{2p}} \sqrt{(2,8\sigma_R)^2 - (2,8\sigma_r)^2 \left(1 - \frac{1}{p} \sum \frac{1}{n_i} \right)}$$

4.2.5 Angaben der Ergebnisse eines Vergleichs

Sofern der Absolutwert der Differenz die gemäß den vorausgehenden Abschnitten anzuwendende Grenze überschreitet, muss die Differenz als verdächtig betrachtet werden, und daher müssen alle Messungen, welche diese Differenz verursacht haben, ebenfalls als verdächtig eingestuft und einer weiteren Untersuchung unterzogen werden.

5 Verfahren zur Prüfung der Annehmbarkeit von Ermittlungsergebnissen und zur Ermittlung des anzugebenden Schlussergebnisses

5.1 Allgemeines

5.1.1 Das in diesem Abschnitt beschriebene Prüfverfahren sollte nur dann angewendet werden, wenn die Messung nach einem Messverfahren ausgeführt wurde, das vereinheitlicht worden ist, und dessen Standardabweichungen σ_r und σ_R bekannt sind. Wenn also die Spannweite von N Ermittlungsergebnissen die anwendbare Grenze nach Abschnitt 4 überschreitet, ist ein Ermittlungsergebnis (oder es sind zwei oder alle Ermittlungsergebnisse) als anomal zu betrachten. Es wird empfohlen, dass die Ursache für das anomale Ergebnis oder für die anomalen Ergebnisse aus technischer Sicht untersucht wird. Es kann jedoch aus handelsrechtlichen Gründen nötig sein, annehmbare Werte zu gewinnen. In solchen Fällen müssen die Ermittlungsergebnisse entsprechend den Festlegungen dieses Abschnitts behandelt werden.

5.1.2 Dieser Abschnitt ist unter den Voraussetzungen abgefasst worden, dass die Ermittlungsergebnisse unter Wiederholbedingungen und unter Vergleichbedingungen gewonnen wurden, und dass das anzuwendende Wahrscheinlichkeitsniveau 95 % ist. Gelten Zwischenbedingungen (siehe ISO 5725-3), dann muss σ_r durch das anzuwendende Maß unter Zwischenbedingungen ersetzt werden.

5.1.3 In einigen Fällen, in denen die in 5.2 beschriebenen Verfahren zu dem als Schlussergebnis angegebenen Median führen, mag es besser sein, auf die Daten zu verzichten.

5.2 Verfahren zur Prüfung der Annehmbarkeit von Ermittlungsergebnissen, gewonnen unter Wiederholbedingungen

ANMERKUNG 3 Der Hinweis in 5.2.2.1 und 5.2.2.2 auf Messungen, die aufwändig oder nicht aufwändig sind, sollte nicht nur in finanzieller Hinsicht ausgelegt werden, sondern auch bezüglich der Frage, ob die Messung komplex, mühsam oder zeitaufwändig ist.

5.2.1 Einzelnes Ermittlungsergebnis

In der geschäftlichen Praxis ist es selten, nur ein einziges Ermittlungsergebnis zu gewinnen. Wenn nur ein einziges Ermittlungsergebnis gewonnen wird, ist es nicht möglich, mit Rücksicht auf den vorgegebenen Wert der Wiederholpräzision einen unmittelbaren statistischen Test auf die Annehmbarkeit dieses Ermittlungsergebnisses auszuführen. Wenn irgend ein Verdacht besteht, das Ermittlungsergebnis könnte nicht korrekt sein, sollte ein zweites Ermittlungsergebnis beschafft werden. Verfügbarkeit von zwei Ermittlungsergebnissen führt zu der häufigeren Praxis, die nachfolgend beschrieben ist.

5.2.2 Zwei Ermittlungsergebnisse

Die zwei Ermittlungsergebnisse sollten unter Wiederholbedingungen gewonnen werden. Der Betrag der Differenz zwischen den zwei Ermittlungsergebnissen sollte dann mit der Wiederholgrenze r verglichen werden, wobei $r = 2,8\sigma_r$.

5.2.2.1 Fall der nicht aufwändigen Gewinnung von Ermittlungsergebnissen

Wenn der Betrag der Differenz zwischen den zwei Ermittlungsergebnissen den Wert r nicht überschreitet, dann sind beide Ermittlungsergebnisse als annehmbar zu betrachten. Als Schlussergebnis sollte der arithmetische Mittelwert der beiden Ermittlungsergebnisse angegeben werden. Wenn der Betrag der Differenz den Wert r überschreitet, sollte das Labor zwei weitere Ermittlungsergebnisse beschaffen.

Wenn die Spannweite ($x_{\max} - x_{\min}$) der vier Ermittlungsergebnisse höchstens gleich der kritischen Spannweite auf dem Wahrscheinlichkeitsniveau 95 % für $n = 4$ ist, also höchstens gleich $CR_{0,95}(4)$, dann sollte der arithmetische Mittelwert der vier Ermittlungsergebnisse als Schlussergebnis angegeben werden. Faktoren $f(n)$ für die kritische Spannweite für $n = 2$ bis $n = 40$ sowie ausgewählte Werte von $n = 45$ bis $n = 100$ sind in Tabelle 1 angegeben. Sie sind zu benutzen zur Errechnung der kritischen Spannweite entsprechend der folgenden Gleichung:

$$CR_{0,95}(n) = f(n)\sigma_r.$$

$$= \frac{1}{\sqrt{2p}} \sqrt{2(\sigma_L^2 + \sigma_r^2) - 2\sigma_r^2 \left(1 - \frac{1}{p} \sum \frac{1}{n_i}\right)}$$

and the critical difference for $|\bar{y} - \mu_0|$ is

$$CD = \frac{1}{\sqrt{2p}} \sqrt{(2,8\sigma_R)^2 - (2,8\sigma_r)^2 \left(1 - \frac{1}{p} \sum \frac{1}{n_i}\right)}$$

at the 95 % probability level.

4.2.5 Quoting the results of a comparison

When the absolute difference exceeds the appropriate limit as given in the preceding clauses, then the difference shall be considered as suspect, and therefore all measurements that have given rise to this difference shall be considered as suspect and subject to further investigation.

5 Methods for checking the acceptability of test results and determining the final quoted result

5.1 General

5.1.1 The checking method described in this clause should be applied only to the case where the measurement was carried out according to a measurement method which has been standardized and whose standard deviations σ_r and σ_R are known. Therefore, when the range of N test results exceeds the appropriate limit as given in clause 4, it is considered that one, two or all of the N test results is or are aberrant. It is recommended that the cause of the aberrant result(s) should be investigated from the technical point of view. However, it may be necessary for commercial reasons to obtain some acceptable value, and in such cases the test results shall be treated according to the stipulations of this clause.

5.1.2 This clause has been prepared on the assumptions that the test results were obtained under repeatability and reproducibility conditions, and that the probability level to be used is 95 %. If intermediate conditions (see ISO 5725-3) were in force, then it is necessary to replace σ_r by the appropriate intermediate measure.

5.1.3 In some cases, where the procedures described in 5.2 lead to the median being quoted as the final result, it might be better to abandon the data.

5.2 Methods for checking the acceptability of test results obtained under repeatability conditions

NOTE 3 In 5.2.2.1 and 5.2.2.2, reference made to measurements being expensive or inexpensive should be interpreted not only in financial terms but also whether the measurement is complex, troublesome or time-consuming.

5.2.1 Single test result

It is not common in commercial practice to obtain only one test result. When only one test result is obtained, it is not possible to make an immediate statistical test of the acceptability of that test result with respect to the given repeatability measure. If there is any suspicion that the test result may not be correct, a second test result should be obtained. Availability of two test results leads to the more common practice which is described below.

5.2.2 Two test results

The two test results should be obtained under repeatability conditions. The absolute difference between the two test results should then be compared with the repeatability limit $r = 2,8\sigma_r$.

5.2.2.1 Case where obtaining test results is inexpensive

If the absolute difference between the two test results does not exceed r , then both test results are considered acceptable, and the final quoted result should be quoted as the arithmetic mean of the two test results. If the absolute difference does exceed r , the laboratory should obtain two further test results.

If the range ($x_{\max} - x_{\min}$) of the four test results is equal to or less than the critical range at the 95 % probability level for $n = 4$, $CR_{0,95}(4)$, the arithmetic mean of the four test results should be reported as the final quoted result. Critical range factors, $f(n)$, for $n = 2$ to $n = 40$ and selected values from $n = 45$ to $n = 100$ are given in table 1 to be used to calculate the critical range according to the following equation:

$$CR_{0,95}(n) = f(n)\sigma_r$$

Wenn die Spannweite der vier Ermittlungsergebnisse größer als die kritische Spannweite für $n = 4$ ist, dann sollte der Median der vier Ermittlungsergebnisse als Schlussergebnis angegeben werden.

Dieses Verfahren ist zusammengefasst im Flussdiagramm des Bildes 1.

5.2.2.2 Fall der aufwändigen Gewinnung von Ermittlungsergebnissen

Wenn der Betrag der Differenz zwischen den zwei Ermittlungsergebnissen den Wert r nicht überschreitet, dann sind beide Ermittlungsergebnisse als annehmbar zu betrachten. Als Schlussergebnis sollte der arithmetische Mittelwert der beiden Ermittlungsergebnisse angegeben werden. Wenn der Betrag der Differenz den Wert r überschreitet, sollte das Labor ein weiteres Ermittlungsergebnis beschaffen.

Wenn die Spannweite ($x_{\max} - x_{\min}$) der drei Ermittlungsergebnisse höchstens gleich der kritischen Spannweite (auf dem Wahrscheinlichkeitsniveau 95 %) für $n = 3$ ist, also höchstens gleich $CR_{0,95}(3)$, dann sollte der arithmetische Mittelwert der drei Ermittlungsergebnisse als Schlussergebnis angegeben werden.

Wenn die Spannweite der drei Ermittlungsergebnisse größer als die kritische Spannweite (auf dem Wahrscheinlichkeitsniveau 95 %) für $n = 3$ ist, dann sollte eine Entscheidung bezüglich einer der folgenden Fälle herbeigeführt werden.

a) Fall der Unmöglichkeit, ein viertes Ermittlungsergebnis zu gewinnen:

Das Labor sollte den Median der drei Ermittlungsergebnisse als Schlussergebnis anwenden.

Dieses Verfahren ist zusammengefasst im Flussdiagramm des Bildes 2.

b) Fall der Möglichkeit, ein viertes Ermittlungsergebnis zu gewinnen:

Das Labor sollte das vierte Ermittlungsergebnis beschaffen. Wenn die Spannweite ($x_{\max} - x_{\min}$) der vier Ermittlungsergebnisse höchstens gleich der kritischen Spannweite (auf dem Wahrscheinlichkeitsniveau 95 %) für $n = 4$ ist, also höchstens gleich $CR_{0,95}(4)$, dann sollte der arithmetische Mittelwert der vier Ermittlungsergebnisse als Schlussergebnis angegeben werden. Wenn die Spannweite der vier Ermittlungsergebnisse größer als die kritische Spannweite für $n = 4$ ist, dann sollte das Labor den Median der vier Ermittlungsergebnisse als Schlussergebnis anwenden.

Dieses Verfahren ist zusammengefasst im Flussdiagramm des Bildes 3.

Tabelle 1 — Faktoren $f(n)$ für die kritische Spannweite

n	$f(n)$	n	$f(n)$
2	2,8	25	5,2
3	3,3	26	5,2
4	3,6	27	5,2
5	3,9	28	5,3
6	4,0	29	5,3
7	4,2	30	5,3
8	4,3	31	5,3
9	4,4	32	5,3
10	4,5	33	5,4
11	4,6	34	5,4
12	4,6	35	5,4
13	4,7	36	5,4
14	4,7	37	5,4
15	4,8	38	5,5
16	4,8	39	5,5
17	4,9	40	5,5
18	4,9	45	5,6
19	5,0	50	5,6
20	5,0	60	5,8
21	5,0	70	5,9
22	5,1	80	5,9
23	5,1	90	6,0
24	5,1	100	6,1

ANMERKUNG Der Faktor $f(n)$ für die kritische Spannweite ist das 95 %-Quantil der Verteilung von $(x_{\max} - x_{\min})/\sigma$, wobei $x_{\max}$ und $x_{\min}$ die Extremwerte in einer Stichprobe des Umfangs n aus einer Normalverteilung mit der Standardabweichung σ sind.

If the range of the four test results is greater than the critical range for $n = 4$, the median of the four test results should be reported as the final quoted result.

This procedure is summarized in the flowchart given in figure 1.

5.2.2.2 Case where obtaining test results is expensive

If the absolute difference between the two test results does not exceed r , then both test results are considered acceptable, and the final quoted result should be quoted as the arithmetic mean of the two test results. If the absolute difference does exceed r , the laboratory should obtain a further test result.

If the range ($x_{\max} - x_{\min}$) of the three test results is equal to or less than the critical range for $n = 3$, $CR_{0,95}(3)$, the arithmetic mean of the three test results should be reported as the final quoted result.

If the range of the three test results is greater than the critical range for $n = 3$, a decision on one of the following two cases shall be made.

a) **Case where it is impossible to obtain a fourth test result:**

The laboratory should use the median of the three test results as the final quoted result.

This procedure is summarized in the flowchart given in figure 2.

b) **Case where it is possible to obtain a fourth test result:**

The laboratory should obtain the fourth test result. If the range ($x_{\max} - x_{\min}$) of the four test results is equal to or less than the critical range for $n = 4$,

$CR_{0,95}(4)$, the arithmetic mean of the four test results should be reported as the final quoted result. If the range of the four test results is greater than the critical range for $n = 4$, the laboratory should use the median of the four test results as the final quoted result.

This procedure is summarized in the flowchart given in figure 3.

Table 1 — Critical range factors, $f(n)$

n	$f(n)$	n	$f(n)$
2	2,8	25	5,2
3	3,3	26	5,2
4	3,6	27	5,2
5	3,9	28	5,3
6	4,0	29	5,3
7	4,2	30	5,3
8	4,3	31	5,3
9	4,4	32	5,3
10	4,5	33	5,4
11	4,6	34	5,4
12	4,6	35	5,4
13	4,7	36	5,4
14	4,7	37	5,4
15	4,8	38	5,5
16	4,8	39	5,5
17	4,9	40	5,5
18	4,9	45	5,6
19	5,0	50	5,6
20	5,0	60	5,8
21	5,0	70	5,9
22	5,1	80	5,9
23	5,1	90	6,0
24	5,1	100	6,1

NOTE — The critical range factor $f(n)$ is the 95 % quantile of the distribution of $(x_{\max} - x_{\min})/\sigma$ where $x_{\max}$ and $x_{\min}$ are the extreme values in a sample of size n from a normal distribution with standard deviation σ .

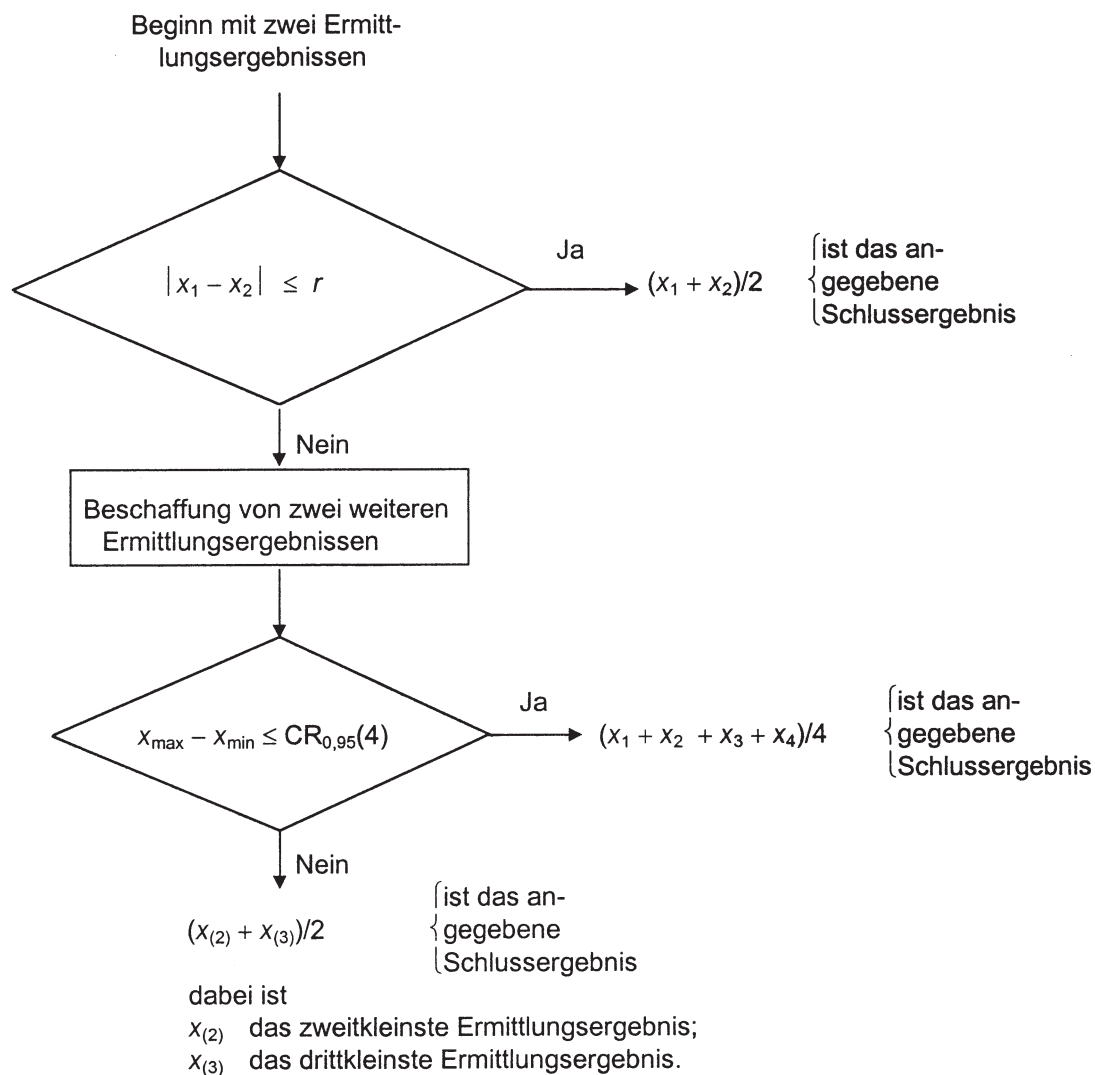
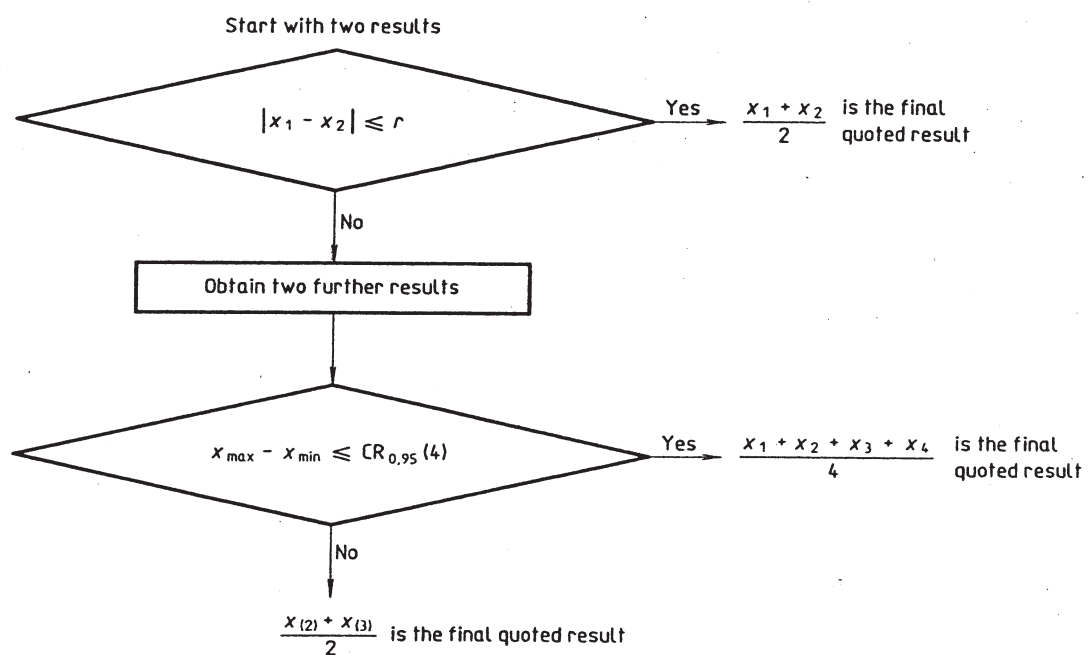


Bild 1 — Verfahren zur Prüfung der Annehmbarkeit von Ermittlungsergebnissen, gewonnen unter Wiederholbedingungen, bei anfangs zwei Ermittlungsergebnissen und bei nicht aufwändiger Gewinnung von Ermittlungsergebnissen: Fall nach 5.2.2.1



where

$x_{(2)}$ is the second smallest result

$x_{(3)}$ is the third smallest result

Figure 1 — Method for checking the acceptability of test results, obtained under repeatability conditions, when two test results are obtained to start with and obtaining test results is inexpensive: Case 5.2.2.1

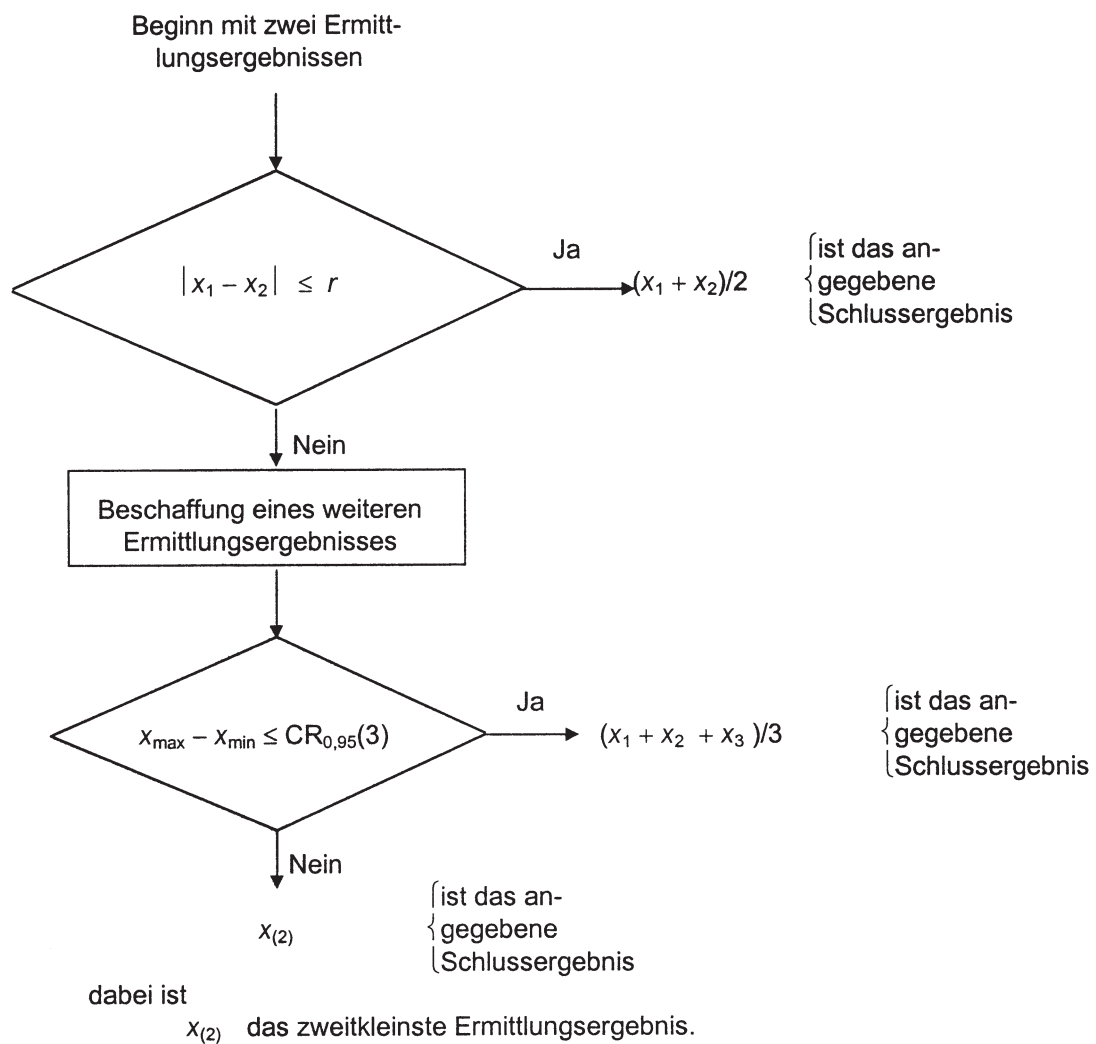
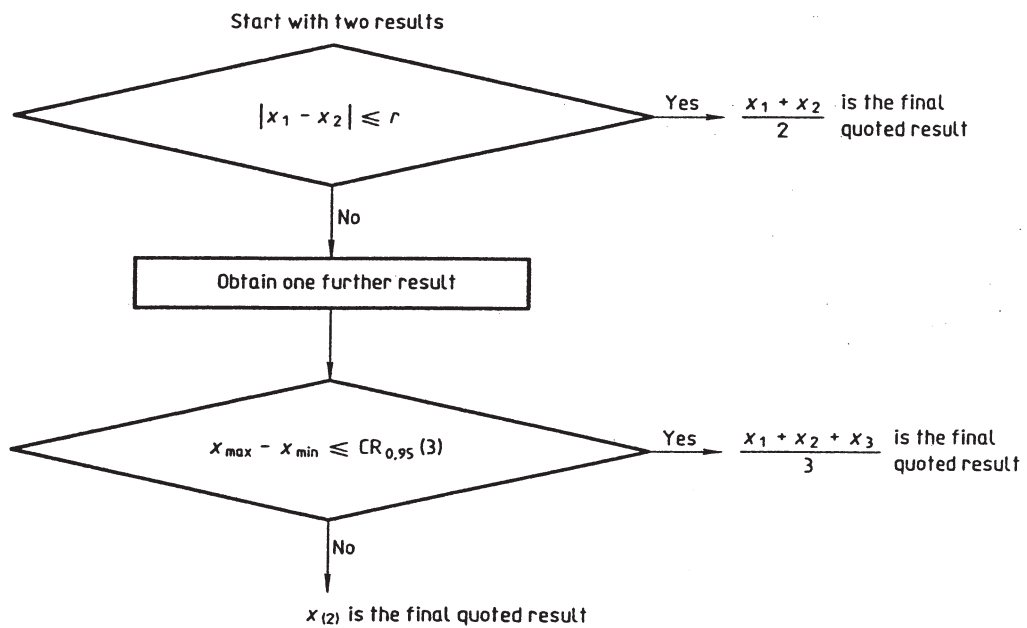


Bild 2 — Verfahren zur Prüfung der Annehmbarkeit von Ermittlungsergebnissen, gewonnen unter Wiederholbedingungen, bei anfangs zwei Ermittlungsergebnissen und bei aufwändiger Gewinnung von Ermittlungsergebnissen: Fall nach 5.2.2.2 a)



where

$x_{(2)}$ is the second smallest result

Figure 2 — Method for checking the acceptability of test results, obtained under repeatability conditions, when two test results are obtained to start with and obtaining test results is expensive: Case 5.2.2.2 a)

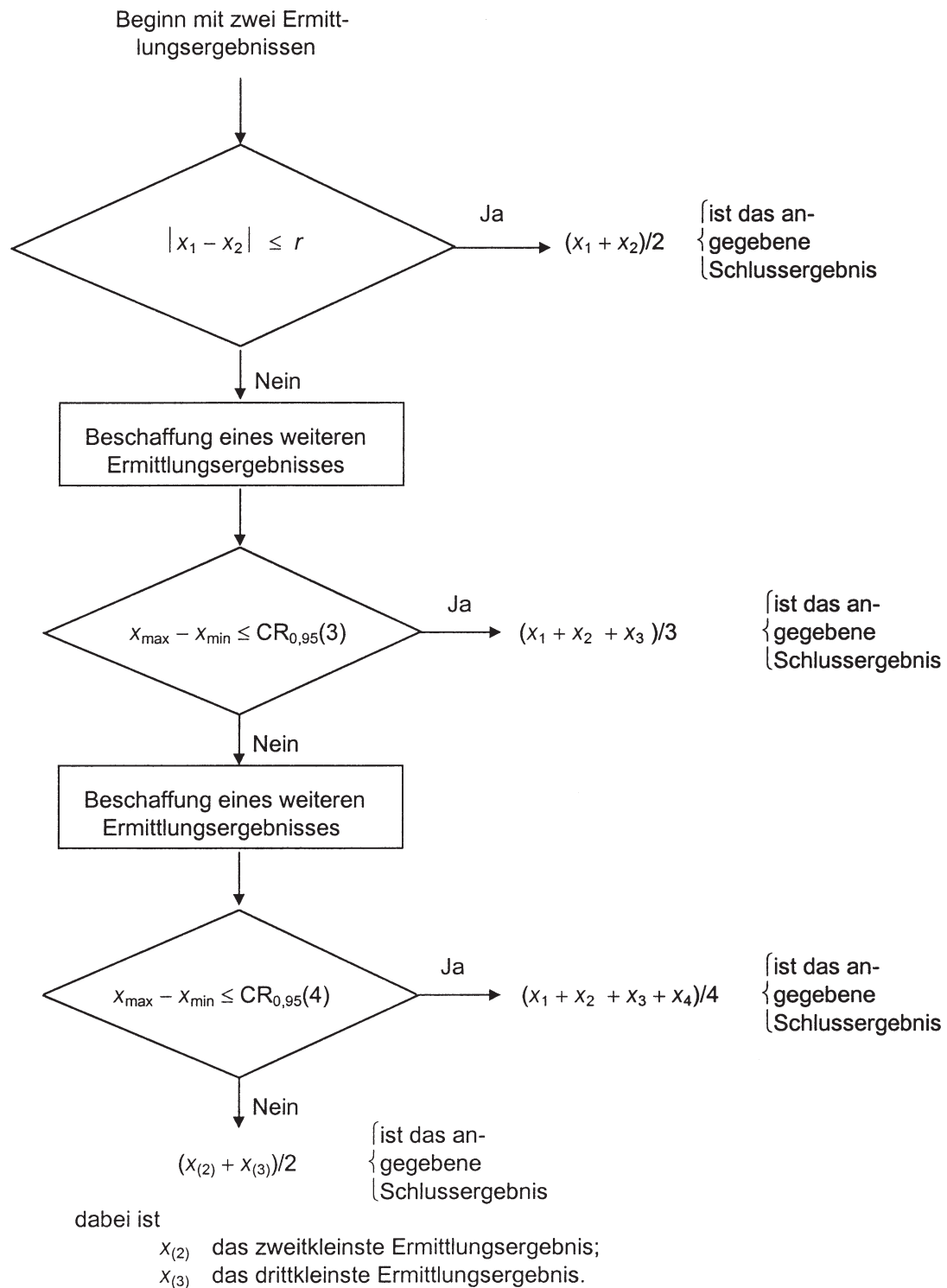
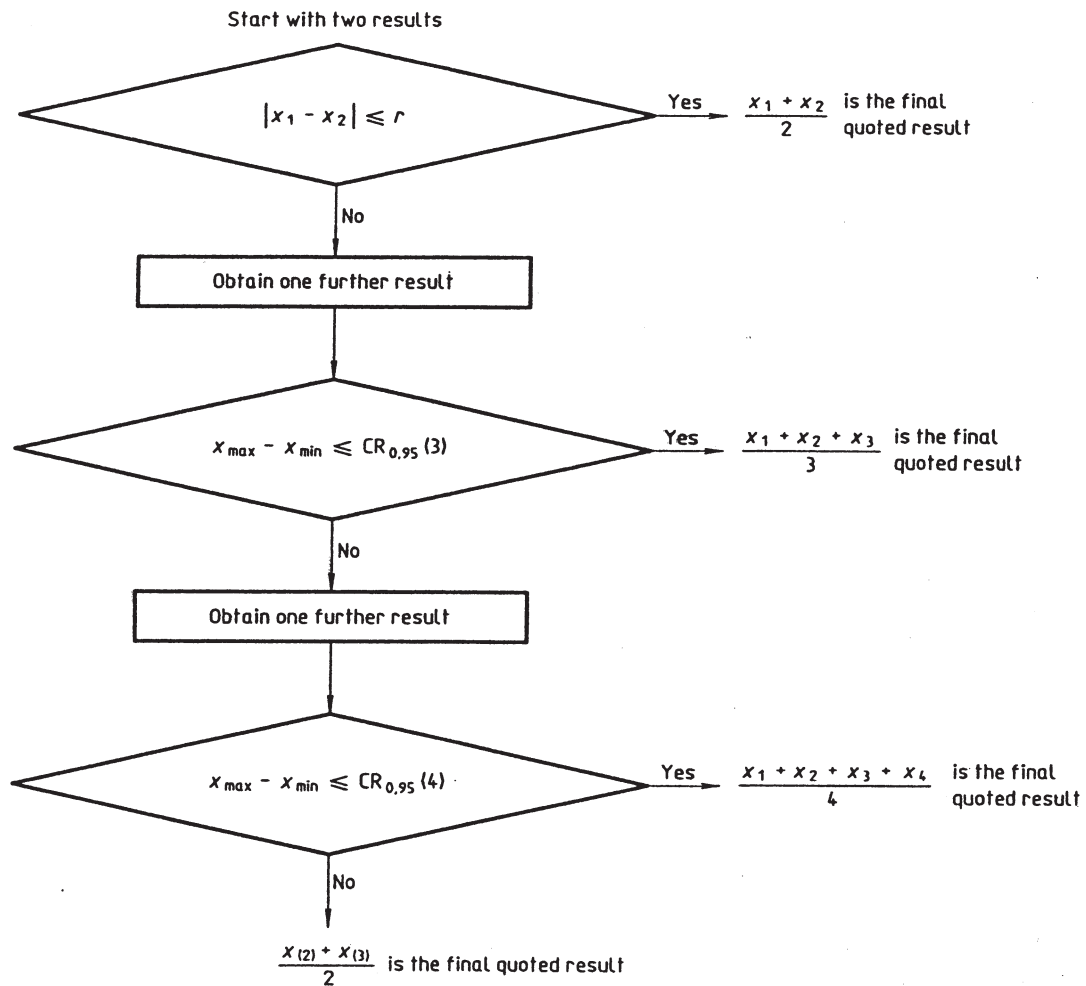


Bild 3 — Verfahren zur Prüfung der Annehmbarkeit von Ermittlungsergebnissen, gewonnen unter Wiederholbedingungen, bei anfangs zwei Ermittlungsergebnissen und bei aufwändiger Gewinnung von Ermittlungsergebnissen: Fall nach 5.2.2.2 b)



where

$x_{(2)}$ is the second smallest result

$x_{(3)}$ is the third smallest result

Figure 3 — Method for checking the acceptability of test results, obtained under repeatability conditions, when two test results are obtained to start with and obtaining test results is expensive: Case 5.2.2.2 b)

5.2.3 Beginn mit mehr als zwei Ermittlungsergebnissen

Manchmal ist es zweckmäßig, mit mehr als zwei Ermittlungsergebnissen zu beginnen. Das Verfahren zur Erzielung des anzugebenden Schlussergebnisses unter Wiederholbedingungen für die Fälle mit $n > 2$ ist ähnlich wie für den Fall mit $n = 2$.

Die Spannweite ($x_{\max} - x_{\min}$) der Ermittlungsergebnisse wird verglichen mit der kritischen Spannweite $CR_{0,95}(n)$, errechnet nach Tabelle 1 für den betreffenden Wert von n . Wenn die Spannweite die kritische Spannweite nicht überschreitet, dann wird der arithmetische Mittelwert aus allen n Ermittlungsergebnissen als Schlussergebnis angewendet.

Überschreitet die Spannweite die kritische Spannweite $CR_{0,95}(n)$, dann muss zur Erzielung des Schlussergebnisses eine Entscheidung zu einem der Fälle A, B oder C getroffen werden, wie sie in den Bildern 4 bis 6 angegeben sind.

Die Fälle A und B entsprechen den Situationen, in denen die Gewinnung von Ermittlungsergebnissen

nicht aufwändig beziehungsweise aufwändig ist. Der Fall C ist eine Alternative, die empfohlen wird, wenn die anfängliche Anzahl von Ermittlungsergebnissen fünf oder mehr und die Gewinnung jedes Ermittlungsergebnisses nicht aufwändig ist, oder wenn die anfängliche Anzahl von Ermittlungsergebnissen vier oder mehr und die Gewinnung jedes Ermittlungsergebnisses aufwändig ist.

Für nicht aufwändige Messungen besteht der Unterschied zwischen dem Fall A und dem Fall C darin, dass im Fall A n weitere Messungen verlangt sind, während der Fall C weniger als die Hälfte dieser Anzahl weiterer Messungen erfordert. Die Entscheidung wird vom Umfang n und der Leichtigkeit des Ausführens der Messungen abhängen.

Für aufwändige Messungen besteht der Unterschied zwischen dem Fall B und dem Fall C darin, dass der Fall C weitere Messungen verlangt, während im Fall B keine weiteren Messungen ausgeführt werden. Der Fall B darf nur in Betracht gezogen werden, wo die Ausführung weiterer Messungen so aufwändig ist, dass sie unerschwinglich wird.

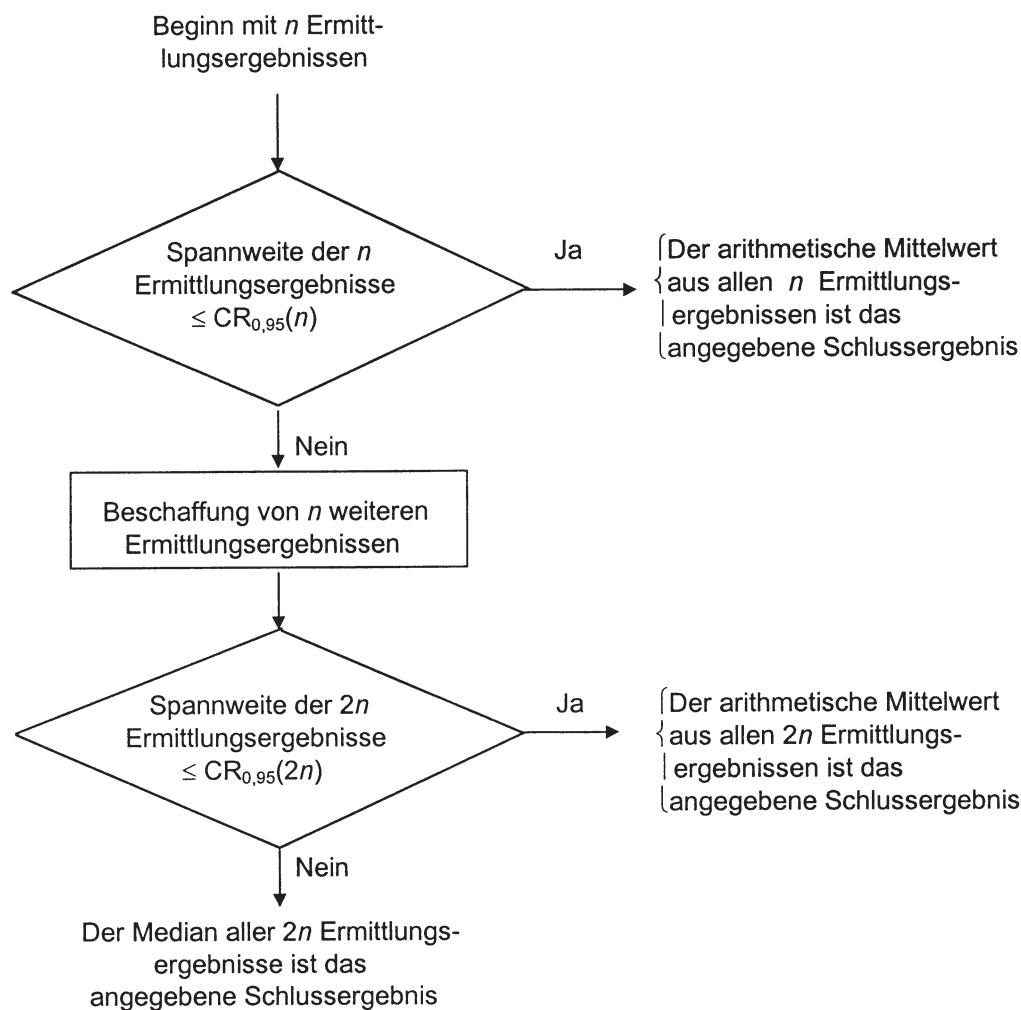


Bild 4 — Verfahren zur Prüfung der Annehmbarkeit von Ermittlungsergebnissen, gewonnen unter Wiederholbedingungen, bei anfangs n Ermittlungsergebnissen und bei nicht aufwändiger Gewinnung von Ermittlungsergebnissen: Fall A

5.2.3 More than two test results to start with

It is sometimes practical to start with more than two test results. The method for obtaining the final quoted result under repeatability conditions for the cases where $n > 2$ is similar to the case for $n = 2$.

The range ($x_{\max} - x_{\min}$) of the test results is compared with the critical range $CR_{0,95}(n)$ calculated from table 1 for the appropriate value of n . If the range does not exceed the critical range, then the arithmetic mean of all the n test results is used as the final quoted result.

If the range does exceed the critical range $CR_{0,95}(n)$, then a decision on one of the cases A, B or C given in figures 4 to 6 shall be made to obtain the final quoted result.

Cases A and B correspond to the cases where obtaining test results is inexpensive and expensive, re-

spectively. Case C is an alternative which is recommended when the starting number of test results is five or more and where obtaining each test result is inexpensive, or when the starting number of test results is four or more and where obtaining each test result is expensive.

For inexpensive measurements, the difference between case A and case C is that case A requires n further measurements, whereas case C requires less than half that number of further measurements. The decision will depend on the size of n and the ease of performing the measurements.

For expensive measurements, the difference between case B and case C is that case C requires further measurements, whereas in case B no further measurements are carried out. Case B shall only be considered where the performance of further measurements is so expensive as to be prohibitive.

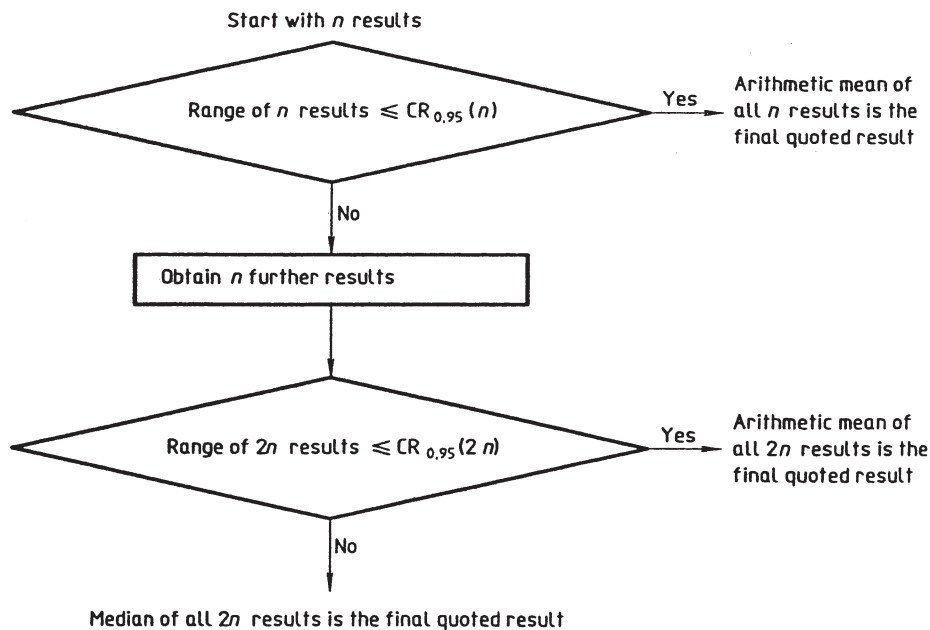


Figure 4 — Method for checking the acceptability of test results, obtained under repeatability conditions, when n test results are obtained to start with and obtaining test results is inexpensive: Case A

Beginn mit n Ermittlungsergebnissen

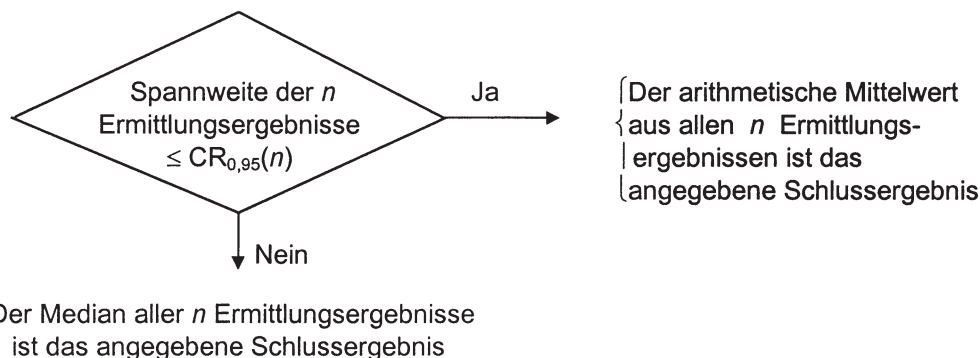
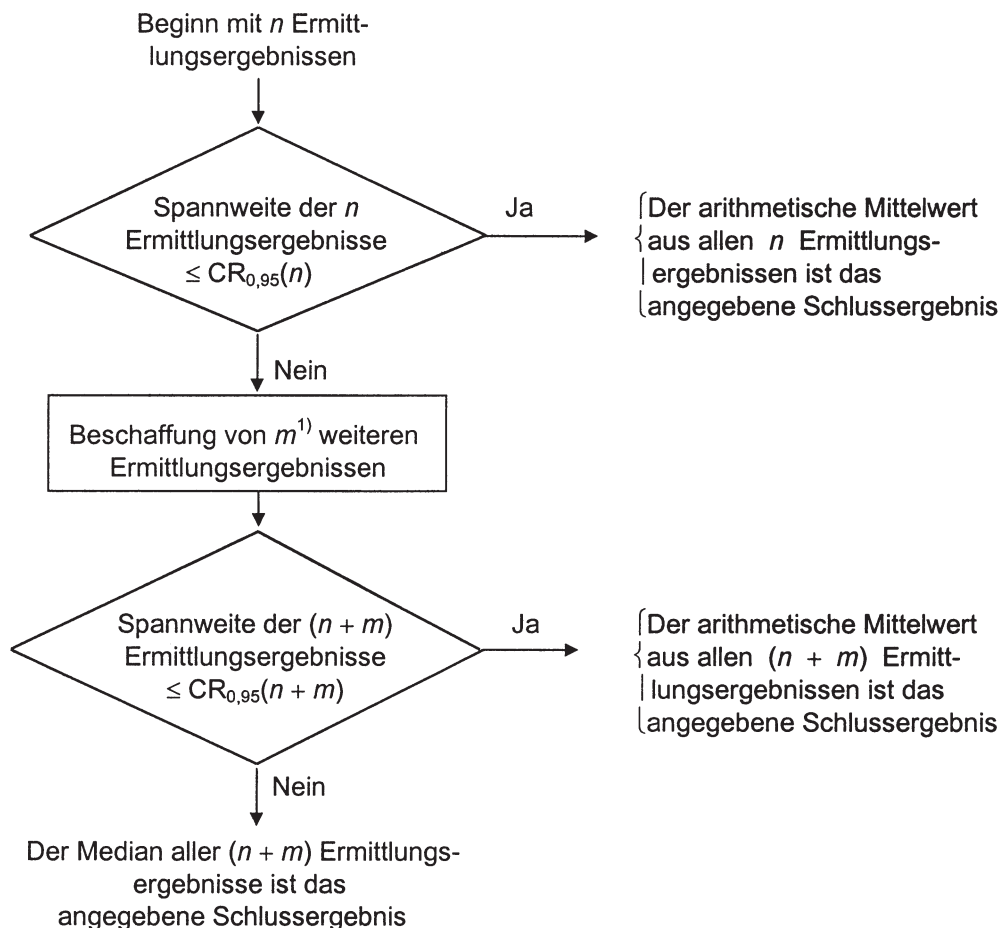


Bild 5 — Verfahren zur Prüfung der Annehmbarkeit von Ermittlungsergebnissen, gewonnen unter Wiederholbedingungen, bei anfangs n Ermittlungsergebnissen und bei aufwändiger Gewinnung von Ermittlungsergebnissen: Fall B



¹⁾ m ist als ganze Zahl zu wählen, welche die Bedingung $n/3 \leq m \leq n/2$ erfüllt

Bild 6 — Verfahren zur Prüfung der Annehmbarkeit von Ermittlungsergebnissen, gewonnen unter Wiederholbedingungen, bei anfangs $n \geq 5$ nicht aufwändig gewonnenen Ermittlungsergebnissen, oder bei $n \geq 4$ aufwändig gewonnenen Ermittlungsergebnissen: Fall C