

Planung und statistische Auswertung von Ringversuchen zur Validierung von quantitativen Untersuchungsverfahren

Inhalt

	Seite		Seite
Kapitel 1 Einleitung	1	1.1 Eliminierung ungültiger Ergebnisse und Plausibilitätsprüfung der Daten	12
Kapitel 2 Planung von Methodvalidierungsstudien	3	1.2 Ausreißerprüfung und -eliminierung.	12
1. Vorüberlegungen.	3	1.3 Kerndichteschätzung, Test auf Normalverteilung	13
1.1 Liegt eine Verfahrensvorschrift vor?	3	1.4 Schlussfolgerungen.	13
1.2 Gibt es für die geplante MVS spezielle Vorgaben?	3	2. Orthogonales MVS-Design mit variierenden Laborbedingungen	13
1.3 Welche Matrices sind relevant?	3	2.1 Eliminierung ungültiger Ergebnisse und Plausibilitätsprüfung der Daten	13
1.4 Welche Konzentrations- bzw. Gehaltsbereiche sind relevant?	3	2.2 Ausreißerprüfung und -eliminierung.	13
1.5 Welche Probenmaterialien sind für die MVS geeignet?	3	2.3 Kerndichteschätzung, Test auf Normalverteilung	14
1.6 Stehen ausreichend Ressourcen zur Durchführung der MVS zur Verfügung?	4	2.4 Ermittlung des Mittelwertes und der Varianzkomponenten.	15
1.7 Welche Probemengen sollen an die Labore verschickt werden?	4	2.5 Überprüfung des Modells	15
1.8 Sollen im Rahmen der MVS zusätzliche Proben mit Konzentrationen bzw. Gehalten, die den Teilnehmenden bekannt sind, an die Labore gesandt werden?	4	3. Validierung nicht urheberrechtlich geschützter Verfahren (faktoriell) (ohne Bezug auf ein Referenzverfahren)	15
1.9 Welche Anweisungen sollen die Labore erhalten?	4	4. Bestimmung der Verfahrenskenndaten	15
1.10 Welche Informationen sollen in welcher Form von den Laboren übermittelt werden?	4	4.1 Klassisches MVS-Design oder geschachteltes MVS-Design mit variierenden Laborbedingungen	15
2. Durchführung der Methodvalidierungsstudie	4	4.1.1 Auswertung nach DIN ISO 5725-2 bzw. ISO 5725-3	15
2.1 Design der MVS	4	4.1.2 Auswertung nach DIN 38402-45 / DIN ISO 5725-5.	15
2.1.1 Quantitative Verfahren.	4	4.1.3 Verfahrenskenndaten.	15
2.1.2 MVS-Design in der Mikrobiologie.	8	4.2 Orthogonales MVS-Design mit variierenden Laborbedingungen	15
2.2 Auswahl der Labore	8	4.2.1 Auswertung nach ISO 5725-3	15
2.3 Materialien und Proben	8	4.2.2 Verfahrenskenndaten.	16
2.3.1 Auswahl.	8	Kapitel 4 Bewertung der Ergebnisse der MVS	17
2.3.2 Homogenitätstests.	10	1. Bewertung der ermittelten Kenndaten	17
2.3.3 Stabilität	10	2. Festlegung des Anwendungsbereichs	17
3. Durchführung der MVS	10	2.1 Matrix	17
3.1 Probenversand.	10	2.2 Arbeitsbereich.	18
3.2 Unterlagen für Teilnehmende der MVS	10	Kapitel 5 Dokumentation der MVS	19
3.3 Ergebnisübermittlung	10	Literaturverzeichnis	19
Kapitel 3 Auswertung von Methodvalidierungsstudien für quantitative Untersuchungsverfahren.	12		
1. Klassisches MVS-Design und geschachteltes MVS-Design mit variierenden Laborbedingungen.	12		

Kapitel 1 Einleitung

Die Feststellung analytisch quantitativ zu bestimmender (und/oder auch qualitativ nachzuweisender) Merkmale bei der Untersuchung von Lebensmitteln, Tabakerzeugnissen, kosmetischen Mitteln, Bedarfsgegenständen, Futtermitteln und Proben, die im Rahmen der gentechnikrechtlichen Über-

wachungstätigkeit der Bundesländer durchgeführt werden, erfolgt häufig unter Anwendung sehr unterschiedlicher Verfahren. Die Folge ist, dass die Ergebnisse dieser Untersuchungen teilweise erheblich voneinander abweichen und dadurch sowohl eine wissenschaftliche als auch eine rechtliche Beurteilung der erhaltenen Daten erschwert wird.

Gemäß § 64 Lebensmittel-, Bedarfsgegenstände- und Futtermittelgesetzbuch (LFGB) [1], § 38 Tabakerzeugnisgesetz (TabakerzG) [2] sowie § 28b Gentechnikgesetz (GenTG) [3] hat das Bundesamt für Verbraucherschutz und Lebensmittelsicherheit (BVL) die Aufgabe, eine Amtliche Sammlung von Verfahren zur Probenahme und Untersuchung (ASU) zu veröffentlichen. Mit dieser Standardisierung von Untersuchungsverfahren wird die Grundlage für eine statistisch abgesicherte Beurteilung analytischer Daten geschaffen.

Zuverlässige Aussagen über die Eignung und Leistungsfähigkeit von Untersuchungsverfahren werden erst nach Durchführung von Ringversuchen unter Beteiligung mehrerer Labore und deren Auswertung unter Anwendung statistischer Methoden erhalten. Daher wird vor der Veröffentlichung eines bereits entwickelten Untersuchungsverfahrens in der Amtlichen Sammlung eine Methodvalidierungsstudie (MVS) durch die Arbeitsgruppen (AG) des BVL durchgeführt.

Grundlage für die Planung und statistische Auswertung von MVS zur Validierung von quantitativen Untersuchungsverfahren sind beispielsweise die Normen der Reihe ISO 5725 [4] [5] [6] [7] und DIN EN ISO 16140 [8] [9]. Die in EU-Rechtsvor-

schriften für die Standardisierung allgemeiner bzw. spezieller Untersuchungsverfahren beschriebenen Kriterien sind zu beachten und gegebenenfalls dazu getroffene Festlegungen zu erfüllen. Darüber hinaus können weitere international anerkannte Standards einbezogen werden.

Die Planung von Ringversuchen zur Validierung von qualitativen Untersuchungsverfahren sollte prinzipiell ebenfalls auf Grundlage von Normen, u. a. der Reihe ISO 5725, erfolgen. Für die statistische Auswertung solcher Ringversuche gibt es bisher jedoch keine allgemeingültigen, einheitlichen Vorgaben aus der Normung oder aus Rechtsvorschriften. Etablierte Ansätze für die Auswertung sind daher momentan noch auf bestimmte Arbeitsbereiche (z. B. die DIN EN ISO 16140 Reihe für Mikrobiologie) begrenzt.

Die Planung, Durchführung und statistische Auswertung einer MVS richtet sich maßgeblich nach dem zu prüfenden Untersuchungsverfahren. Dabei können die dem Untersuchungsverfahren zugrundeliegenden Fragestellungen, Analyseprinzipien und Ziele sehr vielfältig sein (Abbildung 1). Diese Konditionen sind im Vorfeld der MVS explizit zu definieren, um deren Effizienz zu gewährleisten.

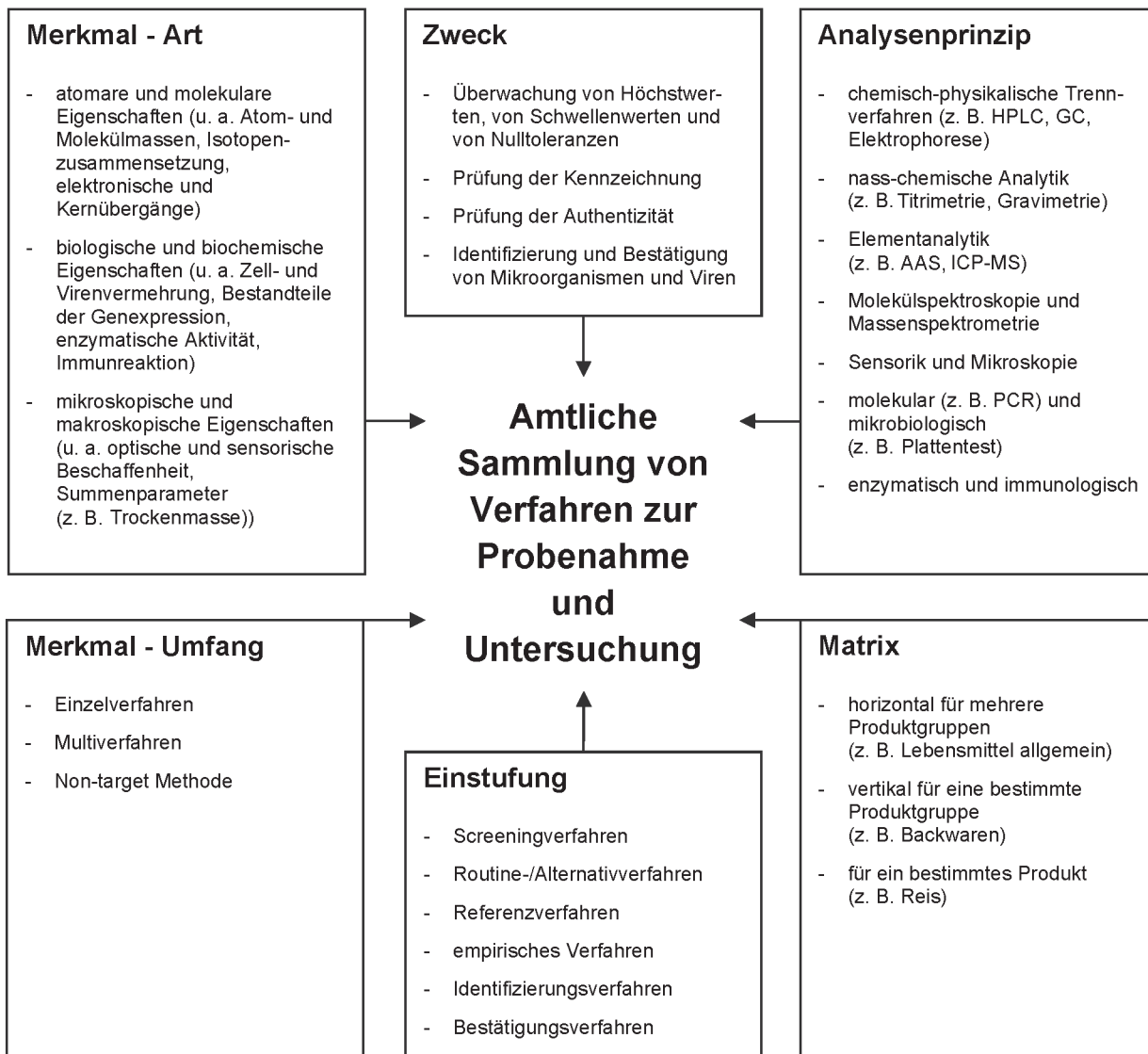


Abbildung 1: Übersicht des Spektrums der Amtlichen Sammlung von Verfahren zur Probenahme und Untersuchung

Kapitel 2 Planung von Methodenvalidierungsstudien

Die Planung von MVS erfolgt unter Einbeziehung von Sachverständigen innerhalb der AGs, die mit dem Untersuchungsverfahren, seiner Anwendung und der statistischen Auswertung von Ringversuchen sowie vorliegender rechtlicher Anforderungen bzw. Empfehlungen zur Methodenvalidierung in dem jeweiligen Bereich vertraut sind. Die jeweilige AG benennt in Abstimmung mit der Referentin/ dem Referenten eine oder mehrere Versuchsleitende, die für die Organisation der MVS zuständig sind. Die Versuchsleitenden haben die Aufgabe, die MVS so zu organisieren, wie es die AG zuvor beschlossen hat. Sie organisieren die Vorbereitung der Proben, geben die entsprechenden Anweisungen weiter und sind während der MVS die fachlichen Ansprechpersonen für die Geschäftsstelle und die teilnehmenden Labore. Die Versuchsleitenden werden von der Geschäftsstelle bei ihrer Arbeit unterstützt.

1. Vorüberlegungen

Zur effizienten Planung der MVS muss eine eindeutige und umfassende Fragestellung formuliert sein, aus der klar hervorgeht, was das zu validierende Verfahren in der Praxis leisten soll und was das Ziel der MVS ist. Bei der Formulierung der Fragestellung ist es zwingend erforderlich, die Anforderungen der jeweils gültigen Rechtsvorschriften zu berücksichtigen (z. B. vorgegebene Leistungskriterien zur Nachweis- und Bestimmungsgrenze, Vergleich- und Wiederholpräzision usw.).

Weiterhin sollten insbesondere die im Folgenden aufgeführten Aspekte berücksichtigt werden.

1.1 Liegt eine Verfahrensvorschrift vor?

Voraussetzung für die Planung von MVS ist das Vorliegen einer innerhalb der AG abgestimmten Verfahrensvorschrift. Falls für den betreffenden Bereich Mindestleistungskriterien festgelegt wurden (z. B. Verordnung (VO) (EU) 2023/2782 [10]), sollte das Verfahren diese vollumfänglich erfüllen. Vor der Durchführung der MVS sollten mit der Problematik vertraute Sachverständige der AG die Verfahren auf ihre störungsfreie Anwendung überprüft und gegebenenfalls optimiert haben.

Zur endgültigen Erstellung der standardisierten Verfahrensvorschrift kann es notwendig sein, Vorversuche mit einer statistisch geeigneten Anzahl von Laboren durchzuführen. Hierdurch sollen mögliche Störeinflüsse erfasst und die Handhabbarkeit des Verfahrens getestet werden.

In der Verfahrensvorschrift muss in allen Details beschrieben werden, wie die Analysenprobe zu präparieren und die Prüfung durchzuführen ist.

1.2 Gibt es für die geplante MVS spezielle Vorgaben?

Falls es für den betreffenden Bereich spezielle nationale, europäische oder internationale Vorgaben bzw. Empfehlungen zur Methodenvalidierung gibt, sind diese bei der Planung und Durchführung der MVS zu berücksichtigen. Solche Vorgaben sind z. B. in folgenden, hier beispielhaft ausgewählten Bereichen in den jeweils aktuell gültigen Fassungen relevant:

- Pflanzenschutzmittel-Rückstände
 - Analytical Quality Control and Method Validation Procedures for Pesticide Residues Analysis in Food and Feed SANTE 11312/2021 [11]

- Bereich Mykotoxine
 - Durchführungsverordnung (EU) 2023/2782 der Kommission vom 14. Dezember 2023 zur Festlegung der Probenahmeverfahren und Analysemethoden für die Kontrolle des Mykotoxingehalts von Lebensmitteln und zur Aufhebung der Verordnung (EG) Nr. 401/2006 [10]
- Bereich Tierarzneimittel-Rückstände
 - Durchführungsverordnung der KOM (EU) 2021/808 über Leistungskriterien für Analysemethoden für Rückstände pharmakologisch wirksamer Stoffe in zur Lebensmittelerzeugung genutzten Tieren und über die Auswertung von Ergebnissen sowie über die für Probe [12]
- Bereich Mikrobiologie
 - Mikrobiologie von Lebensmitteln und Futtermitteln – Arbeitsvorschrift für die Validierung alternativer Verfahren (DIN EN ISO 16140 [13] [9]).

In einigen Bereichen, in denen es noch keine der o. g. Vorgaben zur Validierung gibt, werden von den für die MVS zuständigen AG Leitlinien zur Validierung erstellt (z. B. Leitlinien zur Ringversuchs-Validierung von qualitativen real-time PCR-Methoden). Diese Dokumente sind auf der Webseite des BVL hinterlegt und sollten bei der Validierung berücksichtigt werden:

https://www.bvl.bund.de/DE/Arbeitsbereiche/09_Untersuchungen/01_Aufgaben/04_Amtliche_Sammlung_Untersuchungsverfahren/untersuchungen_Sammlung_Untersuchungsverfahren_node.html [14].

1.3 Welche Matrices sind relevant?

Zu Beginn der Planungen der MVS ist festzustellen, für welche Matrices das Verfahren grundsätzlich geeignet ist und welche dieser Matrices in der Praxis auf das betreffende Merkmal hin untersucht werden. Vor diesem Hintergrund ist festzulegen, für welche Matrices oder Matrixgruppen das Verfahren validiert werden soll und wie dieser Anwendungsbereich im Rahmen der MVS möglichst effizient überprüft werden kann.

1.4 Welche Konzentrations- bzw. Gehaltsbereiche sind relevant?

Es ist zu eruieren, welche Konzentrations- bzw. Gehaltsbereiche des zu bestimmenden Merkmals in der Praxis angetroffen werden und für die rechtliche Beurteilung der jeweiligen Matrices relevant sind. Insbesondere bei qualitativen Verfahren sollten auch die analytischen Grenzen des Verfahrens (v. a. die Nachweisgrenze) einbezogen werden. Außerdem ist ggf. zu berücksichtigen, dass sich die Nachweisgrenzen deutlich von den rechtlich relevanten Konzentrations- und Gehaltsbereichen unterscheiden können. In diesem Fall sollte unter Berücksichtigung der rechtlichen Anforderungen ein Vorgehen zur Bestimmung der Nachweisgrenze festgelegt werden (z. B. durch Einbeziehung von Standards oder Kontrollproben). Auf Grundlage dieser Überlegungen ist festzulegen, für welchen Konzentrations- bzw. Gehaltsbereich das Verfahren validiert werden sollte und wie dieser Anwendungsbereich im Rahmen der MVS möglichst effizient überprüft werden kann.

1.5 Welche Probenmaterialien sind für die MVS geeignet?

Es sind Überlegungen hinsichtlich Homogenität und Stabilität von relevanten Probenmaterialien anzustellen. Diese sind bei der Probenauswahl für die MVS zu berücksichtigen. Erforderliche Homogenitäts- und Stabilitätsuntersuchungen werden vor der Durchführung der MVS vorgenommen oder können ggf. in das Ringversuchsdesign integriert werden.

1.6 Stehen ausreichend Ressourcen zur Durchführung der MVS zur Verfügung?

Neben der Frage, was das Verfahren in der Praxis leisten soll, müssen auch die zur Durchführung verfügbaren Ressourcen berücksichtigt werden. So kann es im Fall von Verfahren, die eine Vielzahl von Analyten gleichzeitig erfassen können oder die zeit-, arbeits- und/oder kostenaufwendig sind, sinnvoll sein, sich bei der Validierung vorerst auf die wichtigsten Leitanalyten bzw. relevantesten Matrices zu beschränken. Bei einer geringen Anzahl teilnehmender Labore kann auf ein orthogonales MVS-Design (faktoriell) unter variierenden Laborbedingungen zurückgegriffen werden [9] [6].

1.7 Welche Probemengen sollen an die Labore verschickt werden?

Bei der Planung einer MVS ist festzustellen, welche Probenmenge zur Durchführung der entsprechenden Untersuchungen pro Labor benötigt wird.

In diesem Zusammenhang ist auch zu klären, ob für jede Bestimmung eine Einzelprobe versandt oder aus einer Probe eine bestimmte Anzahl von Mehrfachbestimmungen durchgeführt wird. Einzelproben können z. B. bei hygroskopischem Material bzw. bei Material, das oxidiert oder flüchtige Komponenten verliert, sinnvoll sein. Einzelproben bieten sich auch dann an, wenn Homogenitäts- und Stabilitätsuntersuchungen auf Basis der Ringversuchsdaten durchgeführt werden sollen.

1.8 Sollen im Rahmen der MVS zusätzliche Proben mit Konzentrationen bzw. Gehalten, die den Teilnehmenden bekannt sind, an die Labore gesandt werden?

Um den Laboren die Möglichkeit zu geben, ihre Vorgehensweise bei der Durchführung des standardisierten Verfahrens vorab zu überprüfen, bietet es sich an, zusätzliche Proben, bei denen die Konzentrationen bzw. Gehalte bereits bekannt sind, mit zu versenden. Dies kann bei Untersuchungsverfahren, die besonders hohe Anforderungen an die Erfahrung der Bearbeitenden stellen, sinnvoll sein.

1.9 Welche Anweisungen sollen die Labore erhalten?

Es ist sicherzustellen, dass alle Labore die Untersuchungen für die MVS in der vorgesehenen Art und Weise durchführen. Dazu sollten die Labore neben dem zu standardisierenden Untersuchungsverfahren weitere Anweisungen zur Durchführung der MVS erhalten. Dazu gehören z. B.:

- Vorgaben zum Ringversuchsdesign (Wiederholanalysen, Zeitrahmen)
- Genauigkeit der Angabe der Messergebnisse
- Abgabetermin der Messergebnisse
- Anweisungen zur Behandlung der Proben (z. B. Lichtabschluss, Aufbewahrung im Kühlschrank)

1.10 Welche Informationen sollen in welcher Form von den Laboren übermittelt werden?

Zur Abfrage aller für die Auswertung der MVS notwendigen Informationen sollten innerhalb der betreffenden AG einheitliche (elektronische) Tabellen oder Formblätter bzw. Online-Eingabemasken festgelegt werden, die als Grundlage für die Übermittlung der Untersuchungsergebnisse dienen. In diesen sollten neben den Messergebnissen auch zusätzliche Informationen, die für die Auswertung der MVS relevant sind (Metadaten), abgefragt werden. Dazu gehören z. B.:

- Datum der einzelnen Analysen
- bei variablen Verfahrensabschnitten Informationen über die individuelle Vorgehensweise (z. B. Probeneinwaage, Verdünnungen, Messbedingungen, verwendeter Gerätetyp, Geräteeinstellungen)
- Nachweis- und Bestimmungsgrenzen sowie das zur Festlegung verwendete Verfahren
- ggf. Anmerkungen zu Problemen oder zu Abweichungen von der Verfahrensvorschrift

2. Durchführung der Methodvalidierungsstudie

2.1 Design der MVS

Das Design der MVS richtet sich nach dem Ziel der MVS und den für die Methodvalidierung zur Verfügung stehenden Ressourcen.

2.1.1 Quantitative Verfahren

Für quantitative Verfahren können folgende Ringversuchsdesigns zum Einsatz kommen:

- a) MVS unter Einhaltung konstanter Laborbedingungen (DIN ISO 5725-2 [5])
- b) MVS mit variierenden Laborbedingungen (ISO 5725-3 [6])
- c) MVS (Mikrobiologie) mit variierenden Laborbedingungen eines nicht urheberrechtlich geschützten Verfahrens (mit und ohne Bezug auf ein Referenzverfahren) (DIN EN ISO 16140-5 [9])

Modifikationen und die Wahl anderer MVS-Designs sind zulässig, wenn die analytische Fragestellung dies erfordert und das BVL sowie die für die betreffende MVS zuständige Arbeitsgruppe diesem Vorgehen zugestimmt haben.

2.1.1.1 Klassisches MVS-Design unter Einhaltung konstanter Laborbedingungen

Klassische MVS auf Basis konstanter Laborbedingungen sind ein weltweit anerkanntes Vorgehen zur Validierung von Analysenverfahren. Unter konstanten Laborbedingungen wird das Arbeiten unter Wiederholbedingungen nach ISO 5725-1 [4] verstanden. Dabei werden durch Anwendung desselben Verfahrens an identischen Proben im selben Labor durch dieselbe bearbeitende Person mit derselben Geräteausstattung in kurzen Zeitabständen voneinander unabhängige Untersuchungsergebnisse ermittelt.

Dabei ist darauf zu achten, dass bei der Durchführung von Untersuchungen unter Wiederholbedingungen das in der standardisierten Vorschrift beschriebene Untersuchungsverfahren im Ganzen wiederholt wird. So ist bei der Validierung eines Verfahrens, das die Aufarbeitung einer Probe für eine anschließende gerätetechnische Bestimmung beschreibt, die Aufarbeitung der Proben zu wiederholen. Die gerätetechnische Bestimmung der Messwerte erfolgt im Anschluss je Aufarbeitung.

Die Zeitabstände zwischen den Wiederholuntersuchungen sind so kurz zu halten, dass sichergestellt wird, dass die Ermittlung der Messergebnisse unter unveränderten Arbeitsbedingungen stattfindet. Gleichzeitig ist zu vermeiden, dass sich die Wiederholuntersuchungen gegenseitig beeinflussen.

Bei der Durchführung einer MVS unter konstanten Laborbedingungen ist zu berücksichtigen, dass zur Ermittlung statistisch abgesicherter Daten zur Vergleichspräzision mindestens acht ausreißerfreie Datensätze von unterschiedlichen Laboren zur Auswertung der MVS vorliegen (in Anlehnung an ISO 5725-1 [4] und AOAC Summary Statement of AOAC Recom-

mendation for Design of a Collaborative Study Minimum Criteria for Quantitative Study [15]).

2.1.1.2 MVS-Design mit variierenden Laborbedingungen

Die Validierung eines Verfahrens unter konstanten Laborbedingungen beschreibt mit der Ermittlung der Wiederholpräzision die Variabilität unter identischen und damit häufig nicht realistischen Bedingungen. Im Laboralltag werden Messungen ggf. auch an verschiedenen Tagen oder von unterschiedlichen Bearbeitenden realisiert, die dabei möglicherweise auf eine geänderte bzw. neu kalibrierte Geräteausrüstung zurückgreifen. Um dies zu berücksichtigen, bietet sich die Durchführung einer MVS mit variierenden Laborbedingungen an.

Die ISO 5725 sieht in Teil 3 [6] entsprechende alternative Designs (z. B. geschachtelt oder orthogonal) vor, die die Ermittlung der Variabilität des Untersuchungsverfahrens unter Intermediärbedingungen zum Ziel haben. Dafür werden bei der Durchführung der MVS in den einzelnen Laboren gezielt bestimmte Faktoren wie z. B. die Bearbeitenden, die Kalibrierung, die verwendeten Geräte oder das Zeitintervall zwischen den Untersuchungen einer Probe variiert. In Tabelle 1 sind Beispiele für mögliche Faktoren und entsprechende Faktorstufen aufgeführt. Zudem ermöglichen orthogonale Designs eine effizientere Schätzung der Präzisionsparameter. Dies bedeutet, dass das gleiche Maß an Zuverlässigkeit mit – im Vergleich zu Kapitel 2 Abschnitt 2.1.1.1 – weniger teilnehmenden Laboren erreicht werden kann.

Weitere Aspekte zu Vor- und Nachteilen der verschiedenen MVS-Designs unter variierenden Laborbedingungen sind in ISO 5725-3 [6] beschrieben.

2.1.1.2.1 Geschachteltes MVS-Design mit variierenden Laborbedingungen

Welche der in ISO 5725-3 [6] verfügbaren geschachtelten Designs (balanced fully-nested design, staggered-nested design, balanced partially-nested design) sinnvoll sind und Anwendung finden sollten, wird je nach Fragestellung in den AGs ggf. mit Unterstützung statistischer Fachkräfte abgestimmt. Geschachtelte MVS unter variierenden Laborbedingungen sind so anzulegen, dass bei der Auswertung zur Ermittlung der Vergleichpräzision mindestens acht ausreißerfreie Datensätze von verschiedenen Laboren vorliegen (in Anlehnung an ISO 5725-1 [4]; die Norm sieht hier für die Anzahl einen Bereich von 8 bis 15 Laboren vor; ISO 5725-3 [6] gibt eine Anzahl von 4 bis 15 Laboren vor, wobei eine Mindestanzahl von vier Laboren in der Regel nur mit orthogonalen MVS-Designs realisiert werden kann).

Im Rahmen der Versuchsplanung sind zur Bestimmung der Wiederholpräzision mindestens zwei Untersuchungen unter konstanten Laborbedingungen durchzuführen. Zur Ermittlung der Variabilität des Verfahrens unter den vorher festgelegten Intermediärbedingungen (z. B. weitere bearbeitende Person, anderer Messtag) ist in jedem teilnehmenden Labor mindestens eine weitere Untersuchung der gleichen Probe unter Variation des entsprechenden Faktors (z. B. zweiter Messtag) notwendig.

Die Variation einzelner Faktoren ist vor der Durchführung einer solchen MVS klar zu definieren und in den Anweisungen für die teilnehmenden Labore genau zu beschreiben.

Abbildung 2 ist ein Beispiel für ein abgestuft-geschachteltes Design für Untersuchungen im Rahmen einer MVS unter Einhaltung variierender Laborbedingungen schematisch dargestellt.

Tabelle 1: Beispiele für Faktoren mit möglichem Einfluss auf das Messergebnis und dazugehörige Faktorstufen

Faktor	beispielhafte Faktorstufen
Bearbeitende	unterschiedliche Erfahrung der Bearbeitenden mit dem Untersuchungsverfahren
Gerät	unterschiedliche Chromatographiesäulen oder Detektoren
Material	unterschiedliche Materialchargen oder Hersteller
Umgebungsbedingungen	Klimatisierung des Labors
Kalibrierung	Gerätekalibrierung
Matrix	unterschiedliche Matrixvariationen aus der zu validierenden Warengruppe
Probenhandling	unterschiedliche Behandlung der Proben vor der Untersuchung
Zeit zwischen zwei Verfahrensschritten	Zeit zwischen Probenextraktion und Messung
Zeit zwischen den Untersuchungen	Untersuchung zu unterschiedlichen Zeitpunkten

2.1.1.2.2 Orthogonales MVS-Design mit variierenden Laborbedingungen

Bei der Durchführung von MVS mit konstanten oder variierenden Laborbedingungen im geschachtelten Design werden mindestens acht Labore benötigt (in Anlehnung an ISO 5725-1 [4]), um eine ausreichende Sicherheit der Präzisionsdaten sicherzustellen. Falls nach Ausreißereliminierung weniger als acht Datensätze zur Verfügung stehen, sind mit diesen MVS-Designs keine statistisch abgesicherten Präzisionsdaten zu ermitteln [4] [6].

Eine weitere Möglichkeit zur Durchführung einer MVS unter variierenden Laborbedingungen bietet das orthogonale Versuchsdesign nach ISO 5725-3 [6]. Dieses Design bietet die Möglichkeit, statistisch abgesicherte Präzisionsdaten bereits ab vier teilnehmenden Laboren zu ermitteln (die ISO 5725-3 [6] spricht von typischerweise 4 bis 15 teilnehmenden Laboren bei ausgewählten Versuchsdesigns). Bei der Planung müssen in einem orthogonalen Versuchsdesign Verfahrensschritte (z. B. Probenaufarbeitung, Extraktion, Kalibrierung, Messung) festgelegt werden, bei denen davon ausgegangen werden kann, dass sie die Präzisionsdaten beeinflussen können. Im Anschluss wird aus jedem dieser Verfahrensschritte ein Faktor ausgewählt, der bei der Durchführung der MVS variiert wird.

Ein beispielhaftes orthogonales Versuchsdesign (nach ISO 5725-3) ist in Tabelle 3 dargestellt. Die Grundlage für dieses Design bildet das „balanced partially-nested design“ (Tabelle 2), bei welchem verschiedene Faktoren, die einen Einfluss auf das Messergebnis haben, gezielt variiert werden. Durch diese systematische Erfassung der relevantesten im Labor auftretenden Effekte und deren Auswirkung auf die Vergleichspräzision, ist es möglich, bei gleichbleibender statistischer

Sicherheit der Präzisionsdaten die Anzahl der teilnehmenden Labore auf mindestens vier zu reduzieren, ohne die statistische Zuverlässigkeit der ermittelten Präzisionsdaten einzuschränken.

Im Vergleich zum „balanced partially-nested design“ (Tabelle 2) ist bei einem orthogonalen MVS-Design (Tabelle 3) der Aufwand („workload“) für die teilnehmenden Labore reduziert (hier um 50 %), ohne dass die Zuverlässigkeit der berechneten Präzisionsdaten beeinträchtigt wird. Dies geschieht durch gezieltes Weglassen bestimmter Zeilen im Versuchsplan.

Die Reliabilität der mit dem orthogonalen MVS-Design erhaltenen Präzisionsdaten hängt von der Auswahl geeigneter und relevanter Faktoren ab. Grundsätzlich könnte eine MVS auch mit Faktoren realisiert werden, die in keinerlei Beziehung zu den Messungen stehen. Jedoch ist insbesondere dann, wenn die acht Messungen des jeweiligen Labors kurz hintereinander durchgeführt werden, die Reliabilität oft schlechter als erwartet. Die allgemeine Vorgehensweise zur Bestimmung der Reliabilität anhand des relativen Standardfehlers ist in ISO 5725-3 [6] beschrieben.

2.1.1.2.3 Auswahl des Designs

Die mit der Planung der MVS betrauten Versuchsleitenden entscheiden in Abstimmung mit der jeweiligen AG und ggf. mit Unterstützung statistischer Fachkräfte, in welchem Design die MVS durchgeführt werden soll (Tabelle 4).

Bei dieser Entscheidung sind neben dem Ziel der MVS und den verfügbaren Ressourcen auch ökonomische Gesichtspunkte und Aspekte der Nachhaltigkeit (z. B. Probenversand, Anzahl teilnehmender Labore) zu berücksichtigen.

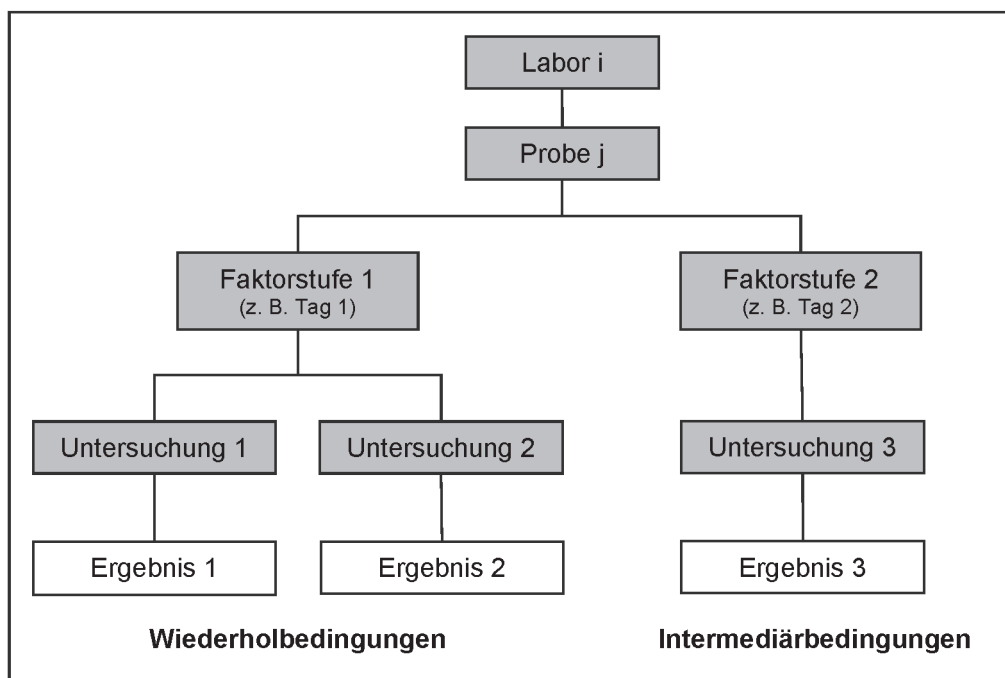


Abbildung 2: Beispieldesign für die einzelnen teilnehmenden Labore für eine MVS mit variierenden Laborbedingungen (abgestuft-geschachteltes Design)

Tabelle 2: Beispieldesign für die einzelnen teilnehmenden Labore für eine MVS mit variierenden Laborbedingungen (balanced partially-nested design)

Labor	A: Matrix a: Matrix 1 b: Matrix 2	B: Bearbeitende Person a: führt Verfahren routinemäßig durch b: hat keine routinemäßige Erfahrung	C: Lagerung des Extraktes a: sofortige Messung b: 24 h Lagerung des Extraktes bis zur Messung	D: SPE-Säule a: Hersteller 1 b: Hersteller 2	Ergebnisse ¹⁾
<i>i</i>	a	a	a	a	y_1, y_2
<i>i</i>	a	a	a	b	y_1, y_2
<i>i</i>	a	a	b	a	y_1, y_2
<i>i</i>	a	a	b	b	y_1, y_2
<i>i</i>	a	b	a	a	y_1, y_2
<i>i</i>	a	b	a	b	y_1, y_2
<i>i</i>	a	b	b	a	y_1, y_2
<i>i</i>	a	b	b	b	y_1, y_2
<i>i</i>	b	a	a	a	y_1, y_2
<i>i</i>	b	a	a	b	y_1, y_2
<i>i</i>	b	a	b	a	y_1, y_2
<i>i</i>	b	a	b	b	y_1, y_2
<i>i</i>	b	b	a	a	y_1, y_2
<i>i</i>	b	b	a	b	y_1, y_2
<i>i</i>	b	b	b	a	y_1, y_2
<i>i</i>	b	b	b	b	y_1, y_2

¹⁾ Die ausführliche Beschriftung der Ergebnisse wäre wie in folgendem Bsp.: y_{iabba1}

Tabelle 3: Beispiel für ein orthogonales Design für eine MVS mit fünf Faktoren (Labor, Matrix, bearbeitende Person, Lagerung, SPE-Säule) Design (mit 2 Replikaten)

Labor	A: Matrix a: Matrix 1 b: Matrix 2	B: Bearbeitende Person a: führt Verfahren routinemäßig durch b: keine routinemäßige Erfahrung	C: Lagerung des Extraktes a: sofortige Messung b: 24 h Lagerung des Extraktes bis zur Messung	D: SPE-Säule a: Hersteller 1 b: Hersteller 2	Ergebnisse ¹⁾
<i>i</i>	a	a	a	a	y_1, y_2
<i>i</i>	a	a	b	b	y_1, y_2
<i>i</i>	a	b	a	b	y_1, y_2
<i>i</i>	a	b	b	a	y_1, y_2
<i>i</i>	b	a	a	b	y_1, y_2
<i>i</i>	b	a	b	a	y_1, y_2
<i>i</i>	b	b	a	a	y_1, y_2
<i>i</i>	b	b	b	b	y_1, y_2

¹⁾ Die ausführliche Beschriftung der Ergebnisse wäre wie in folgendem Bsp.: y_{iabba1}

Tabelle 4: Übersicht möglicher MVS-Designs

		klassisches MVS-Design	geschachteltes MVS-Design	orthogonales MVS-Design
Laborbedingungen		konstant	Variation mehrerer Faktoren (geschachtelt)	Variation mehrerer Faktoren (orthogonal)
Mindestanzahl ausreißerfreier Datensätze		8	8	4
Information zur Verfahrenspräzision	Vergleichstandardabweichung s_R	✓	✓	✓
	Wiederholstandardabweichung s_r	✓	✓	✓
	Laborstandardabweichung s_L	✓	✓	✓
	Intermediärstandardabweichung s_I	✓	✓	✓

2.1.2 MVS-Design in der Mikrobiologie

Die Normenreihe DIN EN ISO 16140 umfasst verschiedene Möglichkeiten zur Validierung oder Verifizierung von Prüfverfahren in der Mikrobiologie der Lebensmittelkette. Sie ist für die Validierung von Verfahren zur Analyse (Nachweis oder Quantifizierung) von Mikroorganismen, insbesondere Bakterien und Pilzen, in folgenden Probenarten anwendbar:

- Erzeugnisse, die für den menschlichen Verzehr vorgesehen sind,
- Erzeugnisse, die als Futtermittel vorgesehen sind,
- Umgebungsproben im Bereich der Herstellung und Handhabung von Lebens- und Futtermitteln,
- Proben aus dem Bereich der Primärproduktion.

Für die Durchführung einer Verfahrensvalidierung im Rahmen eines Ringversuchs sind insbesondere die DIN EN ISO 16140-2 [8] und DIN EN ISO 16140-5 [9] relevant. Die Normenreihe richtet dabei ihren Fokus auf die Validierung von alternativen (urheberrechtlich geschützten (Teil 2) bzw. nicht urheberrechtlich geschützten (Teil 5)) Verfahren anhand von Referenzverfahren. Diese Art der Validierung liegt nicht im Aufgabenbereich der ASU, die ihren Schwerpunkt in der Validierung von vereinheitlichten Messverfahren (entsprechend ISO 5725) hat. Die DIN EN ISO 16140-5 [9] verwendet Arbeitsvorschriften auf der Grundlage von orthogonal faktoriellen Untersuchungen und bietet als eine mögliche Option – jedoch nur für quantitative Verfahren – die Verfahrensvalidierung ohne Bezug auf ein Referenzverfahren, welche sich aber letztendlich an dem MVS-Design der ISO 5725-3 [6] orientiert. Auch hier erfolgt, analog zum orthogonalen MVS-Design nach ISO 5725-3 [6], die Verfahrensvalidierung bestimmter Parameter.

2.2 Auswahl der Labore

Der Kreis der teilnehmenden Labore wird im Rahmen der Planung der MVS durch die AG festgelegt. Hierbei ist zu berücksichtigen, dass die teilnehmenden Labore sowohl fachlich qualifiziert als auch technisch entsprechend ausgerüstet sind.

Sowohl bei MVS mit klassischem Design als auch bei MVS mit geschachteltem Design (mit variierenden Laborbedingungen) müssen zur statistischen Auswertung nach DIN ISO 5725-2 [5] von mindestens acht Laboren valide Ergebnisse vorliegen (nach Ausreißereliminierung), um die Unsicherheit der damit ermittelten Präzisionsdaten in einem akzeptablen Rahmen zu

halten. Daher empfiehlt sich die Teilnahme von mindestens zehn Laboren.

2.3 Materialien und Proben

Die mit der Planung der MVS betrauten Versuchsleitenden sind dafür verantwortlich, dass den teilnehmenden Laboren geeignete Proben zur Durchführung der MVS bereitgestellt werden. Dazu sind Überlegungen bezüglich der Probenauswahl und auch Untersuchungen zur Prüfung der Homogenität und Stabilität erforderlich, sofern hier nicht auf zertifizierte Referenzmaterialien zurückgegriffen werden kann.

2.3.1 Auswahl

Die in der MVS zu untersuchenden Proben müssen repräsentativ für die analytische Fragestellung sein und die relevanten Konzentrations- bzw. Gehaltsbereiche und Matrices abdecken. Besteht aufgrund des Umfangs der analytischen Fragestellung nicht die Möglichkeit, alle betreffenden Merkmale bzw. Analysenparameter in der MVS zu untersuchen, müssen entsprechende, aus Sicht der AG besonders relevante Leitanalyten ausgewählt werden, die den in Frage kommenden Anwendungsbereich des zu testenden Untersuchungsverfahrens bestmöglich abdecken.

Soll mit der MVS ein größerer Konzentrations- bzw. Gehaltsbereich abgedeckt werden, sind unterschiedliche Merkmalniveaus zu untersuchen, die diesen Bereich in geeigneter Weise repräsentieren. Ist auf Grund von Rechtsvorschriften eine bestimmte Konzentration bzw. ein bestimmter Gehalt relevant, sollte der zu untersuchende Bereich zwischen dem 0,5- bis 1,5-fachen der relevanten Konzentration bzw. des relevanten Gehalts liegen. Damit wird ermöglicht, dass die Messunsicherheit in statistisch abgesicherter Form (siehe hierzu das Kapitel der asymmetrischen Messunsicherheit in ISO/TS 23471 [16] ermittelt werden kann. Sehen Rechtsvorschriften z. B. eine Nulltoleranz der Analyten vor oder sind mindesteinzuhaltende Bestimmungsgrenzen gefordert, so muss die Konzentration bzw. der Gehalt des Untersuchungsmaterials so ausgewählt werden, dass eine laborübergreifende Bestimmungsgrenze ermittelt werden kann.

Die Anzahl der zu untersuchenden Proben orientiert sich neben dem Umfang des relevanten Konzentrations- bzw. Gehaltsbereichs auch an der Vielfalt der Matrices, für die das Untersuchungsverfahren im Rahmen der MVS validiert werden soll. Nach den Vorgaben der ISO 5725-1 [4] bzw. den AOAC-Kriterien [15] werden mindestens fünf Matrices empfohlen. Die Mindestzahl hängt allerdings auch von der Art des Verfahrens ab:

- Sollen im Rahmen der MVS Präzisionsdaten für eine bestimmte Matrix ermittelt werden, sollten so viele Proben dieser Matrix ausgewählt werden wie erforderlich sind, um den interessierenden Konzentrations- bzw. Gehaltsbereich abzudecken. Wenn eine funktionelle Beziehung zwischen Konzentration bzw. Gehalt einerseits und den Präzisionsdaten andererseits erhoben werden soll, sollten hierfür mindestens vier Proben untersucht werden. Ansonsten sind ggf. auch zwei bis drei Proben ausreichend.
- Ist es hingegen das Ziel der MVS, das Untersuchungsverfahren für eine bestimmte Produktgruppe (vertikales Verfahren; z. B. Backwaren) zu validieren, sollten mindestens vier Matrices aus dieser Gruppe ausgewählt werden, die die Variabilität der Matrixzusammensetzung (wie z. B. den Fettgehalt) beispielhaft abbilden.
- Falls die MVS jedoch eine große, inhomogene Gruppe an Matrices abdecken soll, um ein mehrere Produktgruppen umfassendes horizontales Verfahren (z. B. für Lebensmittel allgemein) zu validieren, sollte eine Untersuchung von mindestens acht stellvertretenden und fachlich relevanten Matrices erfolgen.

Bei der Auswahl der Proben sollte gewachsenes homogenes Material mit möglichst langer Lagerfähigkeit bevorzugt werden. Da auch die Richtigkeit (ISO 5725-3 [6], ISO 5725-4 [7]) ein wichtiges Charakteristikum des Untersuchungsverfahrens darstellt, sollten daneben auch zertifizierte Referenzmaterialien oder andere Proben mit bekannten Konzentrationen bzw. Gehalten im Rahmen der MVS untersucht werden. Steht weder gewachsenes Material noch geeignetes Referenzmaterial zur Verfügung, kann entsprechendes Untersuchungsmaterial durch Dotierung hergestellt werden. Die Richtigkeit, gemäß ISO 5725-1 [4] definiert als der Grad der Übereinstimmung zwischen dem aus einer Reihe von Prüfergebnissen erhaltenen Durchschnittswert und einem anerkannten Referenzwert, wird im Allgemeinen in Form des Bias (systematische Abweichungskomponente) ausgedrückt. Im Rahmen einer MVS wird der Bias als (relative) Differenz zwischen dem Gesamtmittelwert und der wahren Konzentration bzw. dem wahren Gehalt ausgedrückt. Bei dem wahren Wert handelt es sich idealerweise um einen Wert, der unter Verwendung von (zertifizierten) Referenzmaterialien bestimmt wurde. Wenn dieser nicht verfügbar ist, kann ein zugewiesener Wert aus einer Eignungsprüfung oder einer Kooperationsstudie herangezogen werden. Falls dies ebenfalls nicht möglich ist, so können Proben mit bekannten Konzentrationen bzw. Gehalten (idealerweise Blankmaterial) dotiert/aufgestockt werden. Richtigkeit und Wiederfindung werden häufig synonym verwendet, stellen aber nicht denselben Parameter dar, da das Maß für die Richtigkeit als Bias ausgedrückt wird. Bias und Wiederfindung korrelieren wie folgt miteinander:

$\text{Bias (\%)} = \text{Wiederfindung (\%)} - 100 \%$ [17]. Es ist auch zu beachten, dass der Begriff der Wiederfindung in manchen Arbeitsgruppen im Kontext einer Extraktion gebraucht wird, sich also auf den tatsächlich extrahierten Anteil des Analyten bezieht, während sich der Bias in der Regel auf das Endergebnis bezieht.

Um den Einfluss unterschiedlicher Verfahrensschritte zu ermitteln, können im Rahmen der MVS auch Probenextrakte oder matrixangepasste Standardlösungen untersucht werden. Die daraus resultierenden Ergebnisse dienen lediglich der Information und können die Messergebnisse von Proben, die vollständig gemäß der standardisierten Prüfvorschrift analysiert wurden, nicht ersetzen.

Bei der Beschaffung von Probenmaterial ist sicherzustellen, dass die Probenmenge für die MVS sowie die Homogenitäts- und Stabilitätstests ausreicht und eine Reserve vorrätig gehalten werden kann.

2.3.2 Homogenitätstests

In einer MVS ist es wichtig, dass alle teilnehmenden Labore gleichartiges (homogenes) Probenmaterial zur Verfügung gestellt bekommen, um die Vergleichbarkeit der Messergebnisse zu gewährleisten.

Dabei können die Homogenitätsuntersuchungen entweder vor dem Versand des Probenmaterials in einem Labor erfolgen oder ggf. alternativ im Ringversuchsdesign berücksichtigt werden.

2.3.2.1 Separate Homogenitätsuntersuchungen (vor Versand des Probenmaterials)

Von den Aliquots, die für die MVS vorbereitet wurden, werden je Probenmaterial zufällig zehn Aliquots ausgewählt und in Doppelbestimmung unter Wiederholbedingungen analysiert (Mindestanforderung gemäß ISO 13528:2022 [18], sofern keine Erfahrungen für derartiges Probenmaterial aus vorangegangenen Homogenitätstests vorliegen).

Mittels des Cochran-Tests wird überprüft, ob es statistisch signifikante Abweichungen (zum 1 % Niveau) zwischen zwei Messwerten innerhalb eines Aliquots gibt (Empfehlung der ISO 13528:2022 [18] Abschnitt B.2.1 c)). Als Ausreißer identifizierte Aliquots werden für das betroffene Merkmal bei den folgenden statistischen Berechnungen nicht berücksichtigt.

Ein Probenmaterial gilt hinsichtlich des interessierenden Merkmals als hinreichend homogen, wenn in Anlehnung an ISO 13528:2022 Abschnitt B.2.2 [18] die Standardabweichung zwischen den Proben s nicht größer als das 0,3-fache der zu erwartenden Vergleichsstandardabweichung s_R ist. Hierbei ist gleichzeitig sicherzustellen, dass die Wiederholstandardabweichung aus dem Homogenitätstest nicht mehr als 50 % der zu erwartenden Vergleichsstandardabweichung entspricht, da ansonsten eine mögliche Heterogenität des Probenmaterials unerkannt bleiben könnte.

Gibt es keine Kenntnis über die zu erwartende Vergleichsstandardabweichung, da diese im Allgemeinen nicht vorab bekannt ist, kann stattdessen die (ggf. nach Thompson modifizierte) Horwitzstandardabweichung [19] herangezogen werden.

Kann eine Heterogenität des Untersuchungsmaterials trotz aller Sorgfalt nicht vermieden werden, muss diese bei der statistischen Auswertung in geeigneter Weise berücksichtigt werden. So kann die Vergleichsstandardabweichung s_R um die Standardabweichung zwischen den Proben s_s bereinigt werden.

Weitere Vorgehensweisen zur Homogenitätsprüfung sind in der ISO 13528:2022 [18] beschrieben.

2.3.2.2 Im Ringversuchsdesign integrierte Homogenitätsuntersuchungen

Um die Homogenität des Probenmaterials direkt im Ringversuch von den teilnehmenden Laboren mit zu prüfen, kann das Ringversuchsdesign wie folgt adaptiert werden (Beschreibung zweier Minimaldesigns):

- (1) Je Labor werden an einem Untersuchungstag zwei Aliquots untersucht, von denen Aliquot 1 in Doppelbestimmung unter Wiederholbedingungen sowie Aliquot 2 in Einfachbestimmung untersucht wird. Auf Basis der Er-

gebnisse von Aliquot 1 kann die Wiederholstandardabweichung ermittelt werden und auf Basis der laborweisen Unterschiede zwischen Aliquot 1 und 2 (abzüglich der Wiederholstandardabweichung) die Standardabweichung zwischen den Proben ss. Dieses Design entspricht genau genommen einem abgestuft-geschachtelten Versuchsdesign mit dem Faktor "Aliquot" und lässt sich vor allem bei klassischen MVS mit konstanten Laborbedingungen gut umsetzen.

Die statistische Bewertung der Homogenität erfolgt analog der in Kapitel 2 Abschnitt 2.3.2.1 beschriebenen Vorgehensweise.

- (2) Wurde von vornherein ein (abgestuft-)geschachteltes Versuchsdesign gewählt, bei dem die Untersuchungen an zwei unterschiedlichen Tagen unter ansonsten konstanten Laborbedingungen durchgeführt werden, kann die Homogenität des Probenmaterials auch wie folgt untersucht werden: An Untersuchungstag 1 wird ein Aliquot in Doppelbestimmung unter Wiederholbedingungen analysiert, an Untersuchungstag 2 werden zwei Aliquots jeweils in Einfachbestimmung analysiert.

Bei hinreichend homogenem Probenmaterial ist zu erwarten, dass es zwischen der Varianz der Doppelbestimmung an Tag 1 (innerhalb eines Aliquots) und der Varianz zwischen den zwei Aliquots an Tag 2 keine Unterschiede gibt. Die statistische Überprüfung erfolgt mittels F-Test zum 5%-Signifikanzniveau (in Anlehnung an ISO 13528:2022 B.2.4 a) [18]).

2.3.3 Stabilität

Es ist sicherzustellen, dass die Proben bis zum Zeitpunkt der Durchführung der Messungen unverändert bleiben.

Hierzu sind der Zeitraum der MVS und die Anweisungen zu den Lagerbedingungen des Probenmaterials bis zur Untersuchung so festzulegen, dass unter diesen Voraussetzungen von einer ausreichenden Stabilität des Materials und der Analyten ausgegangen werden kann.

Bei bestimmten Proben (z. B. oxidationsanfällige, flüchtige Komponenten oder stark hygroskopische Proben) kann es sinnvoll sein, mehrere einzeln abgepackte Aliquots des Probenmaterials zu versenden.

Können Veränderungen des Untersuchungsmaterials trotz aller Sorgfalt nicht vermieden werden, müssen diese bei der statistischen Auswertung in geeigneter Weise berücksichtigt werden (z. B. nach ISO 13528:2022 Abschnitt B.5.3 [18]).

Im Allgemeinen können Stabilitätsuntersuchungen – falls erforderlich – von einem Labor erfolgen oder alternativ im Ringversuchsdesign berücksichtigt werden.

2.3.3.1 Separate Stabilitätsuntersuchungen (in einem Labor)

Um eine ausreichende Stabilität belegen zu können, sollte das Probenmaterial bei strikter Einhaltung der vorgeschriebenen Lagerbedingungen in regelmäßigen Intervallen von einem Labor analysiert werden. Diese Stabilitätsuntersuchungen sollten vorzugsweise von dem Labor durchgeführt werden, welches auch die Homogenitätsuntersuchungen (falls erfolgt) durchgeführt hat. Vorgehensweisen zur Stabilitätsprüfung sind z. B. in der ISO 13528:2022 und dem "International Harmonized Protocol for the Proficiency Testing of Analytical Chemistry Laboratories" beschrieben [18] [19]. Der Homogenitätstest kann als erster Punkt in der Messserie des Stabilitätstests angesehen werden.

In Anlehnung an ISO 13528:2022 [18] kann ein Material als stabil angesehen werden, wenn die absolute Differenz zwischen dem Gesamtmittelwert der Messungen der Homogenitätsstudie (bzw. dem Gesamtmittelwert, der in der Untersuchung vor dem Probenversand erzielt wurde) \bar{y}_1 und dem Gesamtmittelwert der Stabilitätsstudie \bar{y}_2 nicht größer als das 0,3-fache der zu erwartenden Vergleichstandardabweichung s_R ist ($|\bar{y}_1 - \bar{y}_2| \leq 0,3 s_R$). Falls s_R nicht bekannt ist, kann die (ggf. nach Thompson modifizierte) Horwitz-Standardabweichung herangezogen werden [19].

2.3.3.2 Im Ringversuchsdesign integrierte Stabilitätsuntersuchungen

Liegt ein (abgestuft-)geschachteltes Ringversuchsdesign vor, bei dem die Messungen an zwei unterschiedlichen Tagen (mit mindestens fünf Tagen Abstand) durchgeführt wurden, ist es möglich, Rückschlüsse von den MVS-Ergebnissen auf die Stabilität der Proben zu ziehen. Auf Basis der laborweisen Differenzen zwischen dem (Mittel-)Wert von Untersuchungstag 1 und dem (Mittel-)Wert von Untersuchungstag 2 wird für jedes Probenmaterial eine mittlere Abweichung der Messwerte zwischen den beiden Untersuchungstagen relativ zum Gesamtmittelwert gebildet. Ein negativer Wert bedeutet hierbei eine Abnahme der Konzentration bzw. des Gehalts, ein positiver Wert eine Zunahme. Von einer statistisch signifikanten Ab- bzw. Zunahme für ein Probenmaterial muss dann ausgegangen werden, wenn (gemäß t-Test) die maximal zulässige relative Abweichung überschritten ist ($= 2 \times$ Standardfehler der mittleren Abweichung; entspricht einem Signifikanzniveau von 5 %).

3. Durchführung der MVS

3.1 Probenversand

Beim Versand der Proben ist sicherzustellen, dass das Probenmaterial unbeschädigt und unter Einhaltung der im Hinblick auf dessen Stabilität erforderlichen Bedingungen bei den Teilnehmenden der MVS eintrifft.

Die Teilnehmenden sind rechtzeitig über den Probenversand und das voraussichtliche Eintreffen der Proben in ihrem Labor zu informieren.

Bei der Bezeichnung der Proben ist darauf zu achten, dass die Teilnehmenden der MVS keine Rückschlüsse auf den Gehalt bzw. die Konzentration der zu untersuchenden Analyten ziehen können. Eine Codierung der zu versendenden Proben ist hierbei sinnvoll.

3.2 Unterlagen für Teilnehmende der MVS

Allen Teilnehmenden der MVS müssen spätestens mit dem Probenversand Unterlagen bereitgestellt werden, aus denen die folgenden Informationen hervorgehen:

- Ansprechperson während der Durchführung der MVS
- Abgabetermin der Ergebnisse
- detaillierte Vorschrift zur Durchführung der MVS
- zu validierende standardisierte Prüfvorschrift des Untersuchungsverfahrens
- Anweisungen zur Behandlung und Lagerung des Probenmaterials
- Anweisungen zur Datenübermittlung

3.3 Ergebnisübermittlung

Idealerweise sollten die Daten elektronisch mittels eines geeigneten Moduls übertragen werden, damit diese ohne Übermittlungsfehler bspw. in eine Software zur statistischen Datenauswertung importiert werden können.

Es ist festzulegen, mit welcher Genauigkeit die Messergebnisse übermittelt werden sollen und wie mit Ergebnissen unterhalb der Nachweis- und Bestimmungsgrenze umzugehen ist. Bei der Durchführung einer MVS zur Ermittlung der Verfahrenspräzision ist es in der Regel sinnvoll, eine Dezimalstelle mehr anzugeben, als in der Praxis für diese Analysendaten üblich wäre. Ergebnisse unterhalb der Nachweis- und Bestimmungsgrenze sollten bevorzugt als Zahlenwert mit

dem entsprechenden Hinweis auf die Nachweis- und Bestimmungsgrenze übermittelt werden.

Neben der mit dem standardisierten Untersuchungsverfahren ermittelten Messdaten sollten im Rahmen der MVS auch weitere laborspezifische Informationen abgefragt werden, die bei Auswertung der MVS bzw. bei Bewertung des zu validierenden Untersuchungsverfahrens relevant sein könnten (vgl. Kapitel 2 Abschnitt 1.10).

Kapitel 3 Auswertung von Methodvalidierungsstudien für quantitative Untersuchungsverfahren

Die statistische Auswertung der im Rahmen der MVS erhobenen Messergebnisse erfolgt in Abhängigkeit von dem gewählten MVS-Design und nach Prüfung der jeweiligen Voraussetzungen entsprechend den in Kapitel 2 Abschnitt 2.1.1 aufgeführten Normen.

Falls die Daten einer im klassischen Design durchgeführten MVS die Voraussetzungen für die Auswertung insbesondere nach DIN ISO 5725-2 [5] nicht erfüllen, aber aus fachlicher Sicht für das jeweilige Untersuchungsverfahren plausibel erscheinen, kann im Ausnahmefall eine Auswertung mit entsprechenden robusten Verfahren z. B. gemäß DIN 38402-45 [20], ISO 13528:2022 [18] oder DIN ISO 5725-5 [21] erfolgen.

Modifikationen und die Wahl anderer statistischer Modelle zur Auswertung der MVS sind zulässig, wenn die analytische Fragestellung dies erfordert und das BVL sowie die für die betreffende MVS zuständige Arbeitsgruppe diesem Vorgehen zugestimmt haben.

1. Klassisches MVS-Design und geschachteltes MVS-Design mit variierenden Laborbedingungen

1.1 Eliminierung ungültiger Ergebnisse und Plausibilitätsprüfung der Daten

Bei der Übermittlung der Ergebnisdatensätze werden die Bemerkungen der teilnehmenden Labore auf Hinweise zu Modifikationen des Untersuchungsverfahrens geprüft. Falls solche vorgenommen wurden, ist zu bewerten, inwieweit diese Modifikationen geringfügig sind und die Durchführung prinzipiell dem Untersuchungsverfahren entspricht oder ob es sich dabei um eine klare Abweichung von der Verfahrensvorschrift handelt. Ergebnisse, die unter einer fachlichen Änderung der Verfahrensvorschrift gewonnen wurden, sind vor der weiteren Prüfung und Auswertung der Daten zu eliminieren. Genauso ist mit Ergebnissen von Laboren zu verfahren, die von gerätetechnischen Problemen berichten, die im Zeitraum der MVS aufgetreten sind und die Ergebnisse beeinflusst haben könnten.

Vor Beginn der Auswertungen werden die MVS-Daten auf auffällige Messwerte hin überprüft. Entsprechende Vorgehensweisen sind den unter Kapitel 2 Abschnitt 2.1.1 genannten Normen zu entnehmen. Bei großen und unübersichtlichen Datensätzen kann es zur Aufdeckung von auffälligen Messwerten sinnvoll sein, im ersten Schritt eine robuste Auswertung der Ergebnisdatensätze durchzuführen.

Die betreffenden Labore werden über die Feststellung von Auffälligkeiten informiert und gebeten, sowohl die angegebenen Messwerte als auch die jeweiligen Analysebedingungen genau zu prüfen.

Sollte sich bei der Prüfung herausstellen, dass ein Fehler bei der Ergebnisberechnung oder Datenübermittlung vorlag, können die übermittelten Messwerte korrigiert werden. Ist dies nicht der Fall, sind die Analysebedingungen genau zu erfassen, zu bewerten und ggf. im Kontext der Ergebnisse weiterer Proben bzw. Analyten zu betrachten, um eine sachgerechte Ausreißerprüfung gewährleisten zu können.

Im Fall, dass ein Labor deutlich abweichend von der Methodenvorschrift gearbeitet hat, werden die zugehörigen Ergebnisse bereits vorab ausgeschlossen.

1.2 Ausreißerprüfung und -eliminierung

Im Rahmen der Ausreißerprüfung wird getestet, ob zum einen die Varianzen der Labore annähernd gleich sind und ob zum anderen systematische Fehler die Mittelwerte beeinflussen.

Für den Vergleich der Varianzen wird die Mandels k-Statistik herangezogen und der Cochran-Test angewandt. Falls die MVS unter variierenden Laborbedingungen durchgeführt wurde, werden für jede Probe sowohl die Varianzen, die unter konstanten Laborbedingungen gewonnen wurden, als auch die Varianzen zwischen den variierten Laborbedingungen auf Ausreißer geprüft. Hierbei ist zu beachten, dass die Mandels k-Statistik deutlich empfindlicher auf abweichende Werte reagiert und sich besonders gut für eine grafische probenübergreifende (und je nach Untersuchungsverfahren auch merkmalsübergreifende) Beurteilung der laborinternen Streuung(en) eignet. Damit können insbesondere systematisch überhöhte Varianzen innerhalb eines Labors festgestellt werden. Der Cochran-Test ist hingegen ein klassischer Ausreißertest, bei dem das Risiko eines "falsch-positiven" Ergebnisses geringer ist als bei der Mandels k-Statistik, so dass dieser zur Identifizierung von Einzelausreißern herangezogen wird.

Für den Vergleich der Mittelwerte der Labore werden die Ergebnisse der Mandels h-Statistik und des Grubbs-Tests herangezogen. Während auch hier die Mandels h-Statistik in erster Linie als grafisches Verfahren zur Identifizierung systematischer Über- oder Unterbefunde innerhalb eines Labors dient, ist der Grubbs-Test ein klassischer Ausreißertest, der ein geringeres Risiko eines "falsch-positiven" Ergebnisses aufweist, so dass dieser zur Identifizierung von Einzelausreißern herangezogen wird.

Beim Umgang mit auffälligen Datensätzen ist zu berücksichtigen, dass die Impräzision der Ergebnisse bei übermäßiger Ausreißereliminierung stark unterschätzt wird und die ermittelten Präzisionsdaten dann kaum die Realität widerspiegeln. Die Eliminierung von potentiellen Ausreißern sollte jedoch in Betracht gezogen werden, wenn eine der folgenden Bedingungen erfüllt ist:

1. *Probenübergreifende Betrachtung je Merkmal (falls mindestens vier Proben untersucht wurden):* Ein Labor weist für ein Merkmal und alle auswertbaren Proben einen mindestens zum Signifikanzniveau von $\alpha = 5\%$ (Fastausreißer) signifikanten Wert von Mandels k oder Mandels h auf.

→ Eliminierung aller Messergebnisse für dieses Merkmal und dieses Labor aufgrund systematischer Auffälligkeiten

Insbesondere bei MVS mit sehr unterschiedlichen Matrices kann es sinnvoll sein, dieses Kriterium getrennt auf verschiedene Matrixkategorien anzuwenden (mind. vier Proben je betrachteter Kategorie; kleinere Kategorien werden i.d.R. nicht auf systematische Effekte hin untersucht).

2. *Merkmalsübergreifende Betrachtung je Probe (falls mindestens vier Merkmale untersucht wurden):* Ein Labor weist für eine Probe und alle Merkmale einen mindestens zum Signifikanzniveau von $\alpha = 5\%$ (Fastausreißer) signifikanten Wert von Mandels k oder Mandels h auf.

→ Eliminierung aller Messergebnisse für diese Probe und dieses Labor aufgrund systematischer Auffälligkeiten

Es gelten die analogen Überlegungen zur Anpassung des Kriteriums wie bei 1. (probenübergreifende Betrachtung).

3. Substanz- und probenspezifische Betrachtung:

- a) Der Cochran-Test liefert ein signifikantes Ergebnis mindestens zum 1-%-Niveau, und es besteht zugleich Grund zu der Annahme, dass das betreffende Labor Schwierigkeiten hat, die Anforderungen zur Einhaltung der Wiederholbedingungen bzw. Intermediärbedingungen zu gewährleisten.
- b) Der Grubbs-Test (Grubbs-I oder Grubbs-II) liefert ein signifikantes Ergebnis mindestens zum 1-%-Niveau, und zugleich besteht Grund zu der Annahme, dass in dem betreffenden Labor vom Messverfahren abgewichen wurde.

Bei der Ausreißereliminierung sollten nicht mehr als ~22 % (=2/9)¹⁾ der Daten von der weiteren Auswertung ausgenommen werden (DIN ISO 5725-2) [5]. Wird diese Anzahl nach einer kritischen Ursachenanalyse überschritten, ist vorab zu prüfen, ob die Ergebnisse der MVS insgesamt plausibel sind und das Untersuchungsverfahren für die betreffende Fragestellung geeignet ist.

1.3 Kerndichteschätzung, Test auf Normalverteilung

Um zu überprüfen, inwieweit die Laborergebnisse einer homogenen Verteilung folgen, welche insbesondere als normalverteilt angenommen werden kann, werden – spezifisch für jede Probe-Merkmal-Kombination – die ausreißerbereinigten Messwertverteilungen grafisch mittels eines Kerndichtediagramms auf Anomalien sowie statistisch auf Normalverteilung geprüft.

Die der Kerndichteschätzung zu Grunde liegende Bandbreite hängt von der in der MVS ermittelten Vergleichsstandardabweichung und der Anzahl der berücksichtigten Labore (nach Ausreißereliminierung) ab: $s_R / n^{0,2}$. Diese Formel wurde in Anlehnung an Silvermans „rule of thumb“ festgelegt ($1,06 \times s_R / n^{0,2}$) [22]. Letztere ist nicht unmittelbar anwendbar, weil einerseits nicht notwendigerweise von einer Normalverteilung ausgegangen werden kann, und andererseits die Messwerte jeweils eines Teilnehmers korreliert sind, d. h., die Modellvoraussetzung einer univariaten Stichprobe nicht gegeben ist. Die Erfahrung hat gezeigt, dass hingegen mit $s_R / n^{0,2}$ eine gute Anpassung der tatsächlichen Verteilung erreicht werden kann.

Anhand der Darstellung des Kerndichtediagramms lässt sich erkennen, ob eine unimodale Verteilung der Messdaten vorliegt. Falls die Messwertkurve mehr als ein Maximum (Modus) aufweist, kann dies darauf hindeuten, dass es mehrere Teilpopulationen von Laboren gibt. Sofern eine Überprüfung der Daten ergibt, dass es tatsächlich mehrere Teilpopulationen gibt, sind eine Ursachenanalyse und eine kritische Überprüfung des Untersuchungsverfahrens vorzunehmen. Beispielsweise sollte bei deutlicher Bimodalität auf eine Ermittlung der statistischen Kenngrößen zunächst verzichtet werden.

Mit dem Shapiro-Wilk-Test [23] wird zu dem spezifisch für jede Probe-Merkmal-Kombination geprüft, ob die ausreißerbereinigte Messwertverteilung als annähernd normalverteilt angesehen werden kann (Signifikanzniveau $\alpha = 1\%$). Liegt

¹⁾ Auszug aus ISO 5725-2 „NOTE 2 IUPAC recommends that the proportion of data rejected from a given level of a study should not exceed the fraction 2/9, the fraction being considered after removal of any results deemed invalid on technical grounds prior to outlier inspection.“

eine signifikante Abweichung vor, ist keine konsistente Auswertung der MVS für diese Probe-Merkmal-Kombination möglich.

1.4 Schlussfolgerungen

Könnte entsprechend Kapitel 3 Abschnitt 1.3 festgestellt werden, dass die Messwerte näherungsweise einer Normalverteilung folgen und keine auffällige Multimodalität vorliegt, wird die Auswertung der ausreißerbereinigten Daten gemäß der unter Kapitel 2 Abschnitt 2.1.1 aufgeführten Normen vorgenommen.

Wenn sich jedoch herausstellt, dass es sich bei den Ergebnissen der MVS um Daten handelt, die nicht näherungsweise einer Normalverteilung folgen, eine Multimodalität aufweisen und/oder der Datensatz eine überdurchschnittliche Anzahl von Ausreißern aufweist, ist zu prüfen, ob die Ergebnisse der MVS plausibel sind und das Untersuchungsverfahren für die betreffende Fragestellung geeignet ist. Ist dies der Fall, kann die statistische Auswertung auch mittels robuster Methoden gemäß DIN 38402-45 [20] oder DIN ISO 5725-5 [21] (in Überarbeitung) erfolgen.

In Abbildung 3 ist der Prozess der Auswertung schematisch dargestellt.

2. Orthogonales MVS-Design mit variierenden Laborbedingungen

2.1 Eliminierung ungültiger Ergebnisse und Plausibilitätsprüfung der Daten

Analog zu der Vorgehensweise bei der Auswertung von MVS mit konstanten bzw. variierenden Laborbedingungen im geschachtelten Design werden auch bei MVS, die mit einem orthogonalen Design unter variierenden Laborbedingungen durchgeführt wurden, ungültige Ergebnisse eliminiert und die Daten vor Beginn der statistischen Auswertung auf ihre Plausibilität²⁾ (visuell anhand einer geeigneten Darstellung der Rohdaten) geprüft. Im Unterschied zu MVS unter konstanten oder geschachtelt variierenden Laborbedingungen wird darauf verzichtet, Messserien bereits aufgrund erhöhter Streuung als ungültig zu eliminieren. Die Erfahrung hat gezeigt, dass durch orthogonale Designs reproduzierbare Unterschiede in den laborspezifischen Präzisionsdaten sichtbar werden. Derartige reproduzierbare Unterschiede in verschiedenen Messserien sollten nicht als Ausreißer aus den Daten entfernt werden. Das Ergebnis des MVS entspräche dann nicht mehr der Realität in den Laboren. Gleichwohl soll durch geeignete statistische Verfahren sichtbar gemacht werden, wie stark die Präzisionsunterschiede in den einzelnen Laboren sein können.

2.2 Ausreißerprüfung und -eliminierung

Zunächst wird getestet, ob einzelne Messwerte Ausreißer darstellen (nach DIN ISO 5725 2 [5]), indem messserien- und laborübergreifend der Grubbs-Test angewendet wird. Einzelwerte, die für eine Probe oder eine Dotierungsstufe auf dem Signifikanzniveau von 1 % statistisch signifikant vom jeweiligen Mittelwert abweichen, werden als Ausreißer eliminiert.

²⁾ Diese Plausibilitätsprüfung dient nicht dazu, zu überprüfen, ob eine Messserie normalverteilt ist: Dazu gibt es in der Regel nicht genügend Werte (oft nur 4 Labore). Und aufgrund faktorieller Effekte ist Normalverteilung innerhalb eines Labors auch nicht gegeben. Die visuelle Plausibilitätsprüfung dient einzig dazu, zu prüfen, ob es grobe Fehler gibt, z. B. falsche Zuordnung der Daten oder falsche Einheiten.

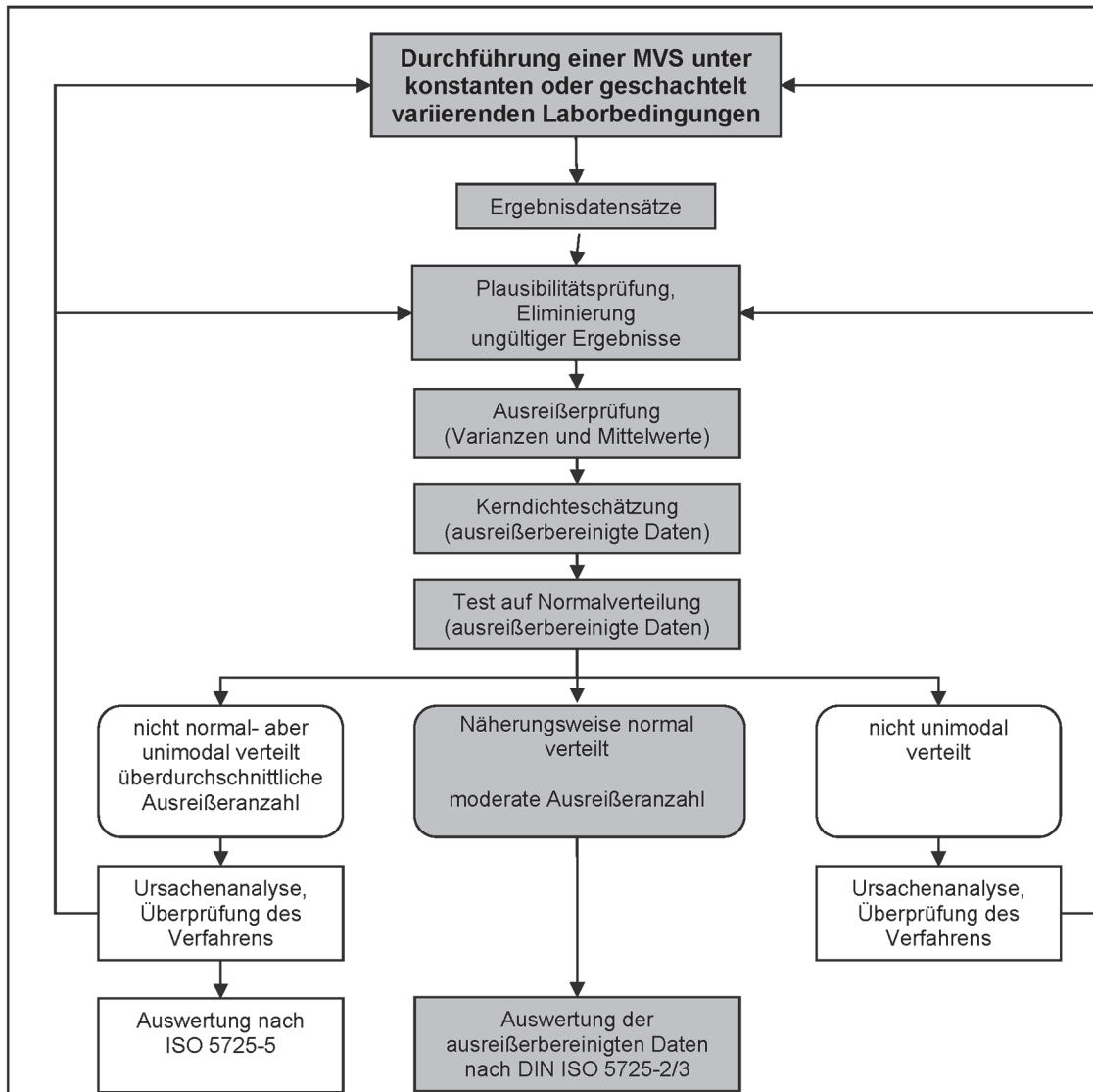


Abbildung 3: Ablauf der Auswertung für das klassische MVS-Design bzw. das geschachtelte MVS-Design mit variierenden Laborbedingungen

Bei orthogonalen MVS-Designs mit Berücksichtigung unterschiedlicher Dotierungsstufen, bei denen für jede Messserie eine lineare Regression erfolgt, werden zusätzliche Ausreißertests durchgeführt:

- Auf Basis der gewichteten messserienspezifischen Regressionsgeraden wird die *gewichtete Residualstandardabweichung* berechnet. Mit Hilfe des Cochran-Tests wird geprüft, ob eine dieser Residualstandardabweichungen zum Niveau von 1 % signifikant von den übrigen abweicht. Zur Entscheidung, ob die ganze Messserie oder ein einzelner Messwert eliminiert werden sollte, wird für die Residuen dieser Serie der Grubbs-Test durchgeführt.
- Mit Hilfe des Grubbs-Tests wird überprüft, ob statistisch signifikant *abweichende Steigungen* oder statistisch signifikant *abweichende mittlere Niveaus* (Wert der messserienspezifischen Regressionsgeraden in der Mitte des untersuchten Konzentrations- bzw. Gehaltsbereichs) identifiziert werden können. Ist dies der Fall, so werden alle Messwerte der auffälligen Serie eliminiert.

Wurden im Rahmen der MVS bei jeder Messserie Doppelbestimmungen von Proben vorgenommen, erfolgt die Ausreißerprüfung und -eliminierung der um faktorielle Effekte bereinigten Varianzen analog zu den Verfahren, die bei klassischen bzw. geschachtelten MVS-Designs Anwendung finden.

2.3 Kerndichteschätzung, Test auf Normalverteilung

Im orthogonalen MVS-Design mit variierenden Laborbedingungen wird analog Kapitel 3 Abschnitt 1.3 vorgegangen, jedoch wird für die Prüfung der Normalverteilung der Shapiro-Wilk-Test auf die standardisierten Residuen angewendet [23].

2.4 Ermittlung des Mittelwertes und der Varianzkomponenten

Für eine Probe werden der Mittelwert und die Varianzkomponenten (zufälligen Effekte), d. h.

- die Laborstandardabweichung (s_L),
- die Standardabweichung in Bezug auf die betrachteten Faktoren (für jeden Faktor gibt es eine separate Standardabweichung; $s_{(1)}, \dots, s_{(k)}$ bei k Faktoren) und
- die Wiederholstandardabweichung (s_r)

mittels des REML-Verfahrens (restricted maximum likelihood) ermittelt (ISO 5725-3 [6]). Dieses statistische Verfahren wird bei orthogonalen MVS-Designs unter Berücksichtigung unterschiedlicher Dotierungsstufen angewendet, wobei in diesem Fall sowohl der Mittelwert als auch die Varianzkomponenten in Abhängigkeit der Konzentration bzw. des Gehalts modelliert werden.

2.5 Überprüfung des Modells

Zur Beurteilung der Güte des Modells werden die standardisierten Residuen auf Basis des Mittelwertes und der Varianzkomponenten berechnet und mittels Kerndichteschätzung visuell auf Unimodalität sowie statistisch mittels Shapiro-Wilk-Test auf Normalverteilung geprüft. Folgen die standardisierten Abweichungen nicht einer Normalverteilung, ist das zugrundeliegende Modell zu prüfen. Sofern einer der untersuchten Faktoren eine sehr starke Wirkung aufweist, muss das Untersuchungsverfahren und zugleich das Modell entsprechend angepasst werden, indem entweder der Anwendungsbereich des Untersuchungsverfahrens reduziert wird (z. B. bezüglich Probenmatrix) oder das Untersuchungsverfahren selbst angepasst wird (z. B. Lagerdauer der Extrakte reduzieren). Sofern die beobachteten Abweichungen nicht auf einen einzelnen Faktor zurückgeführt werden können, ist zu prüfen, ob Abweichungen von der Normalverteilung methodisch bedingt sein können, z. B. bei quantitativen PCR-Methoden oder aufgrund der Fundamentalvariabilität. Gleichzeitig ist zu überprüfen, ob trotz der aufgetretenen Abweichungen noch eine ausreichende Präzision erzielt werden kann.

3. Validierung nicht urheberrechtlich geschützter Verfahren (faktoriell) (ohne Bezug auf ein Referenzverfahren)

Nach DIN EN ISO 16140-5 [9] wird durch das organisierende Labor entschieden, welche Daten in die Auswertung einfließen. Bei der Validierung des nicht urheberrechtlich geschützten Verfahrens mit Bezug auf ein Referenzverfahren prüft das organisierende Labor die Rohdaten sowie weitere geforderte Angaben. Von der weiteren Auswertung werden diejenigen Daten ausgeschlossen, die möglicherweise unter ungeeigneten Bedingungen erhalten wurden bzw. wo vom Verfahren abgewichen wurde. An den ausgewählten Daten werden keine Ausreißertests vorgenommen.

Für die Verfahrensvalidierung nicht urheberrechtlich geschützter Verfahren ohne Bezug auf ein Referenzverfahren sind in der DIN EN ISO 16140-5 keine Vorgaben zum Ausschluss von Daten enthalten.

4. Bestimmung der Verfahrenskenndaten

4.1 Klassisches MVS-Design oder geschachteltes MVS-Design mit variierenden Laborbedingungen

4.1.1 Auswertung nach DIN ISO 5725-2 bzw. ISO 5725-3

Die der DIN ISO 5725-2 [5] bzw. ISO 5725-3 [6] zugrundeliegenden statistischen Modelle gehen davon aus, dass jeder Messwert die Summe der folgenden Komponenten ist:

- Gesamtmittelwert (m)
- systematische Abweichungskomponente eines Labors (B_i , $i = 1, \dots, p$ bei p Laboratorien)
- systematische Abweichungskomponente(n) in Bezug auf die variierenden Laborbedingungen (nur bei Auswertung nach ISO 5725-3 [6]) (B_{i1}, \dots, B_{ik} bei k weiteren Faktoren)
- zufällige Abweichung einer Messung unter Wiederholbedingungen (e).

4.1.2 Auswertung nach DIN 38402-45 / DIN ISO 5725-5

Die DIN 38402-45 [20] gibt ein robustes statistisches Schätzverfahren vor, das gegenüber Abweichungen von der Normalverteilung unempfindlich ist. Ausreißer und Extremwerte werden entsprechend ihrer Abweichung vom Mittelwert gewichtet und beeinflussen das Ergebnis damit nur marginal. Alternativ sind robuste Verfahren der DIN ISO 5725-5 [21] zu verwenden.

4.1.3 Verfahrenskenndaten

Grundsätzlich ist im Vorfeld der MVS zu prüfen, welche Verfahrenskenndaten aus rechtlicher Sicht für den jeweiligen analytischen Bereich gefordert sind (z. B. nach VO (EU) 2017/625 [24]). Aus den Ergebnissen einer klassischen bzw. geschachtelten MVS sind z. B. folgende Verfahrenskenndaten zu berechnen und zu dokumentieren:

- Mittelwert (m)
- absolute und relative Wiederholstandardabweichung s_r und $s_{r,rel}$
- absolute und relative Wiederholgrenze r und r_{rel}
- absolute und relative Vergleichstandardabweichung s_R und $s_{R,rel}$
- absolute und relative Vergleichgrenze R und R_{rel}
- absolute und relative Intermediärstandardabweichung s_I und $s_{I,rel}$ ³⁾
- absolute und relative Intermediärgrenze I und I_{rel} ³⁾

Zudem können je nach Fragestellung und Ziel der MVS weitere Verfahrenskenndaten zur Beurteilung des Untersuchungsverfahrens relevant sein. Diese sind ggf. aus den bei der MVS erhobenen Messdaten oder auch aus den bei Datenübermittlung von den teilnehmenden Laboren zusätzlich zur Verfügung gestellten Daten zu ermitteln. Dazu gehören zum Beispiel:

- Vertrauens- bzw. Unsicherheitsbereich des Mittelwertes
- Wiederfindung
- Richtigkeit (ISO 5725-4) [7]
- HorRat-Wert
- laborübergreifende Bestimmungsgrenze
- laborübergreifende maximale Grenze des unteren Arbeitsbereiches
- Minimum und Maximum der laborspezifischen Nachweis- und Bestimmungsgrenzen
- Minimum und Maximum der laborspezifischen Entscheidungsgrenzen ($CC\alpha$)

4.2 Orthogonales MVS-Design mit variierenden Laborbedingungen

4.2.1 Auswertung nach ISO 5725-3

MVS mit variierenden Laborbedingungen in einem faktoriell orthogonalen Versuchsdesign werden anhand des aus den MVS-Daten hervorgegangenen linearen gemischten Modells ausgewertet. Dabei werden sowohl systematische als auch zufällige Effekte berücksichtigt [6].

³⁾ Nur bei geschachtelten MVS möglich.

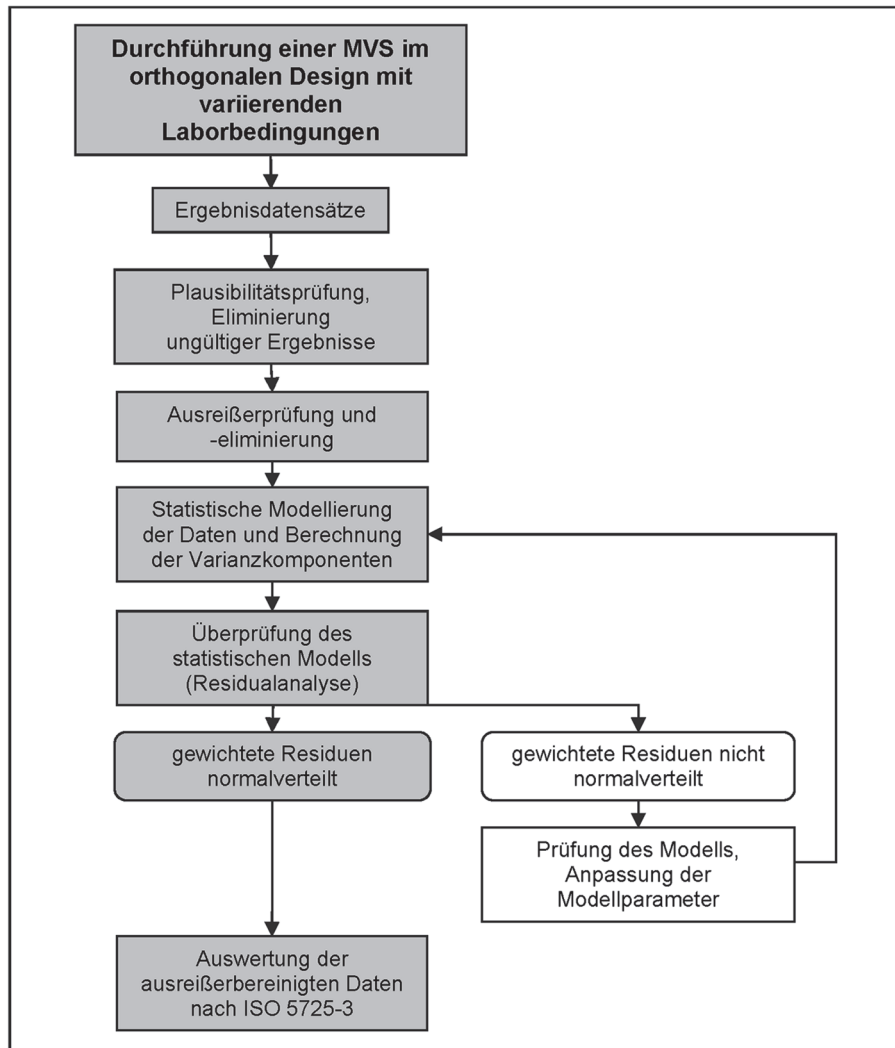


Abbildung 4: Ablauf der Auswertung im orthogonalem MVS-Design mit variierenden Laborbedingungen

4.2.2 Verfahrenskenndaten

Aus den Ergebnissen einer MVS, die in einem orthogonalem Design durchgeführt wurde, sind grundsätzlich für jede Dotierungsstufe folgende Verfahrenskenndaten zu berechnen und zu dokumentieren:

- Mittelwert (m)
- absolute und relative Wiederholstandardabweichung s_r und $s_{r,rel}$
- absolute und relative Wiederholgrenze r und r_{rel}
- absolute und relative Vergleichstandardabweichung s_R und $s_{R,rel}$
- absolute und relative Vergleichgrenze R und R_{rel}
- relativer Standardfehler der Vergleichstandardabweichung $se_{re}(s_R)$
- Anteil der Streuungskomponenten (Wiederhol- und Laborstandardabweichung sowie die Standardabweichungen der einzelnen Faktoren) an der Vergleichstandardabweichung s_R

Zusätzlich sind weitere für die Fragestellung relevante Verfahrenskenndaten aus den MVS-Ergebnissen zu ermitteln. Dazu können zum Beispiel gehören:

- relative Wiederfindung
- relative Richtigkeit
- HorRat-Wert
- modifizierter HorRat-Wert (für Konzentrationen bzw. Gehalte < 120 ppb)
- relative Inhouse-Vergleichstandardabweichung
- Minimum und Maximum der laborspezifischen Nachweis- und Bestimmungsgrenzen
 - bei orthogonalem MVS-Designs mit Berücksichtigung unterschiedlicher Dotierungsstufen: minimale und maximale Entscheidungsgrenze ($CC\alpha$)
 - minimales und maximales Erfassungsvermögen ($CC\beta$)

Darüber hinaus ist bei bekannter Konzentration bzw. bekanntem Gehalt der Probe zu ermitteln, ob eine signifikante Abweichung zwischen Gesamtmittelwert und wahren Wert besteht, und weiterhin lässt sich ein Konfidenzintervall für den absoluten bzw. relativen Bias sowie die Wiederfindung berechnen. Grundlage hierfür ist der Standardfehler für den Gesamtmittelwert aller Labore, der gemäß ISO 5725-3 [6] ermittelt werden kann.

Kapitel 4 Bewertung der Ergebnisse der MVS

1. Bewertung der ermittelten Kenndaten

Anhand der Ergebnisse der MVS ist zu beurteilen, ob das betreffende Untersuchungsverfahren als validiert betrachtet werden kann. Dabei sind insbesondere folgende Kenndaten zu berücksichtigen (**das Vorliegen gesetzlicher Festlegungen von spezifischen Leistungskriterien ist vorab zu prüfen**):

- Relative Vergleichsstandardabweichung $s_{R,rel}$
Die relative Vergleichsstandardabweichung $s_{R,rel}$ sollte grundsätzlich nicht mehr als 30 % betragen. Wenn trotz einer Überschreitung dieser Kennzahl das Verfahren als (eingeschränkt) valide betrachtet werden kann und in die ASU aufgenommen werden soll, ist dies von der zuständigen Arbeitsgruppe unter Berücksichtigung gesetzlicher Vorgaben gesondert zu begründen. Es gilt jedoch zu beachten, dass die genannten Kennzahlen nicht in allen Bereichen gültig und rechtliche Vorgaben zu prüfen sind. Beispielsweise sollte bei ausgewählten Kontaminanten die relative Vergleichsstandardabweichung $s_{R,rel}$ nicht mehr als 25 % betragen.
- Wiederholstandardabweichung s_r und $s_{r,rel}$
Sofern keine rechtlichen Vorgaben festgelegt sind, sollte die maximale Wiederholstandardabweichung nicht mehr als 1/2 bzw. 2/3 der Vergleichsstandardabweichung s_R betragen [25] [26]
- Standardfehler der Vergleichsstandardabweichung $se_{rel}(s_R)$
Anhand der Ergebnisse von orthogonalen MVS-Designs unter variierenden Laborbedingungen wird der relative Standardfehler der ermittelten Vergleichsstandardabweichung s_R ermittelt [6]. Der relative Standardfehler sollte nicht mehr als 25 % betragen. Diese Grenze kann man aus den üblichen Mindestanforderungen für MVS ablesen. Bei MVS mit acht Laboren mit Doppelbestimmung ergibt sich aus der Unsicherheitsbetrachtung der ISO 5725-1 [4] ein Standardfehler von 25 % für MVS, sofern die Wiederholstandardabweichung s_r 1/3 der Vergleichsstandardabweichung s_R entspricht (was erfahrungsgemäß ein typischer Wert ist). Bei sieben Laboren liegt der Standardfehler bereits bei 27 %, so dass ein Wert von maximal 25 % auch im Falle anderer MVS-Designs als angemessen angesehen werden kann.
- HorRat-Wert
Der HorRat-Wert ist eine etablierte Kennzahl zur Bewertung von Untersuchungsverfahren [25], welche dem Quotienten aus Vergleichsstandardabweichung s_R und Horwitzstandardabweichung (zu erwartende theoretische Standardabweichung, ggf. modifiziert nach Thompson) entspricht. Ein HorRat-Wert deutlich größer (kleiner) als 1 zeigt deshalb an, dass die im Ringversuch erreichte Vergleichsstandardabweichung s_R deutlich größer (kleiner) als die zu erwartende theoretische Standardabweichung ist. Bei Ringversuchen gelten in der Regel HorRat-Werte bis 2 als nicht auffällig. Im höheren Konzentrations- bzw. Gehaltsbereich können auch höhere HorRat-Werte vertretbar sein. Der HorRat-Wert sollte daher grundsätzlich im Zusammenhang mit der relativen Vergleichsstandardabweichung $s_{R,rel}$ beurteilt werden. Auf EU-Ebene gibt es aktuell Entwicklungen im Bereich der Gesetzgebung (z. B. im Bereich der Mykotoxine, Pflanzentoxine), die den HorRat-Wert als Leistungskriterium nicht mehr einschließen. In bestimmten Bereichen (z. B. Nickel in Lebensmitteln) findet der HorRat-Wert ebenfalls bei der Bewertung der Wiederholstandardabweichung s_r Anwendung.

- Richtigkeit/Wiederfindung

Der Mittelwert der MVS und der anerkannte Bezugswert von Referenzmaterialien in Form der systematischen Abweichungskomponente sollten sich unter Berücksichtigung des Unsicherheitsbereiches nicht signifikant voneinander unterscheiden. Die Richtigkeit kann auch anhand von dotierten Proben überprüft werden. Zusätzlich sollte entsprechend der gesetzlich vorgegebenen Leistungskriterien die Wiederfindung berücksichtigt werden. Der Zusammenhang zwischen Richtigkeit und Wiederfindung wurde oben beschrieben (Kapitel 2 Abschnitt 2.3.1).

Die für die MVS zuständige Arbeitsgruppe bewertet die Ergebnisse der MVS zudem unter dem Aspekt, inwiefern das betreffende Verfahren für die damit in der amtlichen Lebensmittelüberwachung zu bearbeitenden analytischen und rechtlichen Fragestellungen geeignet ist.

2. Festlegung des Anwendungsbereichs

Anhand der Ergebnisse der MVS formuliert die zuständige Arbeitsgruppe den in diesem Rahmen erfolgreich validierten Anwendungsbereich des Untersuchungsverfahrens. Dabei ist festzulegen, für welche Matrices und für welchen Konzentrations- bzw. Gehaltsbereich das Untersuchungsverfahren als validiert bzw. eingeschränkt validiert gelten kann.

2.1 Matrix

Zunächst bilden die ermittelten Präzisionsdaten und – falls verfügbar – die Daten zur Richtigkeit bzw. zur Wiederfindung den Zustand des in der MVS untersuchten Materials ab. Streng genommen bedeutet dies, dass selbst bei minimalen Veränderungen des Anwendungsbereiches keine Verwendung der statistischen Kenndaten mehr zulässig ist. Wurden die Untersuchungsmaterialien jedoch in der MVS sinnvoll gewählt, ist eine Extrapolation möglich. Dabei sind die folgenden Vorgehensweisen zulässig:

- Gilt für alle Labore im Ringversuch, dass der Laborbias für zwei oder mehrere Untersuchungsmaterialien jeweils (bis auf zufällige Unterschiede) der gleiche ist, kann darauf geschlossen werden, dass auch bei weiteren Materialien der gleichen Art keine signifikante Abweichung vorliegt. In diesem Falle kann – sofern keine auffälligen Unterschiede der Wiederholstandardabweichungen s_r vorliegen – gefolgert werden, dass sich auch bei weiteren Materialien die in der MVS ermittelte relative Vergleichsstandardabweichung $s_{R,rel}$ ergibt.
Im Falle zweier Materialien kann über eine lineare Regression der Labormittelwerte anhand des Korrelationskoeffizienten festgestellt werden, ob ein materialspezifischer Laboreffekt vorliegt, der eine Extrapolation verhindert. Bei der Überprüfung auf signifikante Abweichungen zwischen den Laboreffekten von mehr als zwei Materialien ist darauf zu achten, dass sich der Laboreffekt immer aus einer absoluten und einer relativen Fehlerkomponente zusammensetzt.
- Wenn materialspezifische Laboreffekte festzustellen sind, stellt sich die Frage, ob diese für alle Untersuchungsmaterialien gleichermaßen wirksam sind. Ist dies der Fall, kann ebenfalls auf weitere Materialien extrapoliert werden. Diese Möglichkeit setzt voraus, dass mindestens fünf Materialien untersucht wurden.
- Eine dritte Extrapolationsmöglichkeit ergibt sich bei Verfügbarkeit von mindestens fünf Materialien mit unterschiedlichen Konzentrationsniveaus. Sofern eine statistisch belastbare Beziehung zwischen Konzentration und

Vergleichsstandard- bzw. Wiederholstandardabweichung etabliert werden kann, und wenn keines der untersuchten Materialien von dieser Beziehung signifikant abweicht, kann ebenfalls gefolgert werden, dass die betreffende statistische Beziehung für alle Materialien des Anwendungsbereiches gültig ist.

Sollte eine in der MVS getestete Matrix sich als nicht zur Untersuchung mit der angewendeten Methode geeignet erweisen, so muss diese in der Methodenbeschreibung genannt und aus dem Anwendungsbereich ausgenommen werden.

2.2 Arbeitsbereich

Der Arbeitsbereich ist der Konzentrations- bzw. Gehaltsbereich, in dem das Untersuchungsverfahren zuverlässig angewendet werden kann. Dies muss durch die betreffende MVS belegbar sein.

Bei Untersuchungsverfahren, die auch zur Prüfung von Nulltoleranzen herangezogen werden könnten, ist aus den MVS-Daten eine laborübergreifende maximale Grenze des unteren Arbeitsbereiches abzuleiten.

Kapitel 5 Dokumentation der MVS

Die Ergebnisse der MVS werden in einem Bericht dokumentiert. Diese Dokumentation sollte mindestens folgende Angaben enthalten:

- detaillierte Beschreibung der Proben einschließlich der Angaben zur Herstellung, Homogenität und ggf. Stabilität sowie Kurzbeschreibung des angewandten Analyseverfahrens
- Liste der teilnehmenden Labore
- Erläuterungen zum gewählten statistischen Auswertemodell
- Beschreibung und Interpretation der Prozedur der Ausreißerprüfung
- ggf. graphische Darstellung der Messergebnisse wie z. B. Mandels h- und Mandels k-Statistik, Kerndichteschätzung, Boxplots der Laborergebnisse
- statistische Methodenkenndaten
- Zusammenstellung der von den teilnehmenden Laboren eingesandten Daten

Die Ergebnisse der Methodenvalidierungsstudien sind Eigentum des BVL und dürfen in wissenschaftlichen Zeitschriften nur mit Zustimmung des BVL veröffentlicht werden.

Literaturverzeichnis

- [1] Bundesministerium der Justiz, Lebensmittel- und Futtermittelgesetzbuch in der Fassung der Bekanntmachung vom 15. September 2021 (BGBl. I S. 4253), das zuletzt durch Artikel 7 des Gesetzes vom 27. September 2021 (BGBl. I S. 4530) geändert worden ist.
- [2] Bundesministerium der Justiz, Tabakerzeugnisgesetz vom 4. April 2016 (BGBl. I S. 569), das zuletzt durch Artikel 2 Absatz 5 des Gesetzes vom 20. Dezember 2022 (BGBl. I S. 2752) geändert worden ist.
- [3] Bundesministerium der Justiz, Gentechnikgesetz in der Fassung der Bekanntmachung vom 16. Dezember 1993 (BGBl. I S. 2066), das zuletzt durch Artikel 8 Absatz 7 des Gesetzes vom 27. September 2021 (BGBl. I S. 4530) geändert worden ist.
- [4] ISO 5725-1:2023-07, *Accuracy (trueness and precision) of measurement methods and results – Part 1: General principles and definitions*, (deutscher Titel: Genauigkeit (Richtigkeit und Präzision) von Messverfahren und Messergebnissen – Teil 1: Allgemeine Grundlagen und Begriffe).
- [5] DIN ISO 5725-2:2022-05, *Genauigkeit (Richtigkeit und Präzision) von Messverfahren und Messergebnissen – Teil 2: Grundlegende Methode für die Ermittlung der Wiederhol- und Vergleichpräzision eines vereinheitlichten Messverfahrens*.
- [6] ISO 5725-3:2023-06, *Accuracy (trueness and precision) of measurement methods and results – Part 3: Intermediate precision and alternative designs for collaborative studies*, (deutscher Titel: Genauigkeit (Richtigkeit und Präzision) von Messverfahren und Messergebnissen – Teil 3: Präzisionsmaße unter Zwischenbedingungen und alternative Designs für Ringversuche).
- [7] ISO 5725-4:2020-03, *Accuracy (trueness and precision) of measurement methods and results – Part 4: Basic methods for the determination of the trueness of a standard measurement method*, (deutscher Titel: Genauigkeit (Richtigkeit und Präzision) von Messverfahren und Messergebnissen – Teil 4: Grundlegende Methoden für die Ermittlung der Richtigkeit eines vereinheitlichten Messverfahrens).
- [8] DIN EN ISO 16140-2:2016-11, *Mikrobiologie der Lebensmittelkette – Verfahrensvalidierung – Teil 2: Arbeitsvorschrift für die Validierung von alternativen (urheberrechtlich geschützten) Verfahren anhand eines Referenzverfahrens*.
- [9] DIN EN ISO 16140-5:2020-12, *Mikrobiologie der Lebensmittelkette – Verfahrensvalidierung – Teil 5: Arbeitsvorschrift für eine faktorielle Ringversuch-Verfahrensvalidierung nicht urheberrechtlich geschützter Verfahren*.
- [10] Europäische Kommission, Durchführungsverordnung (EU) 2023/2782 der Kommission vom 14. Dezember 2023 zur Festlegung der Probenahmeverfahren und Analysemethoden für die Kontrolle des Mykotoxingehalts von Lebensmitteln und zur Aufhebung der Verordnung (EG) Nr. 401/2006.
- [11] Europäische Kommission, SANTE 11312/2021, *Analytical Quality Control and Method Validation Procedures for Pesticide Residues Analysis in Food and Feed*, 2021.
- [12] Europäische Kommission, Durchführungsverordnung der KOM (EU) 2021/808 über Leistungskriterien für Analysemethoden für Rückstände pharmakologisch wirksamer Stoffe in zur Lebensmittelherzeugung genutzten Tieren und über die Auswertung von Ergebnissen sowie über die für Probe, 2021.
- [13] DIN EN ISO 16140-1:2016-11, *Mikrobiologie der Lebensmittelkette – Verfahrensvalidierung – Teil 1: Terminologie*.
- [14] Bundesamt für Verbraucherschutz und Lebensmittelsicherheit, https://www.bvl.bund.de/DE/Arbeitsbereiche/09_Untersuchungen/01_Aufgaben/04_Amtliche_Sammlung_Untersuchungsverfahren/untersuchungen_SammlungUntersuchungsverfahren_node.html.
- [15] AOAC Official Methods of Analysis, Appendix D: *Guidelines for Collaborative Study Procedures To Validate Characteristics of a Method of Analysis*, p. 2, 2002.
- [16] ISO/TS 23471:2022-11 Vornorm, *Experimental designs for evaluation of uncertainty – Use of factorial designs for determining uncertainty functions*.
- [17] European Reference Laboratory – University of Wageningen, EURL-MP-guidance doc_003 (version 1.1), Guidance document on performance criteria (draft 17th September, 2021), 2021.
- [18] ISO 13528:2022-08, *Statistical methods for use in proficiency testing by interlaboratory comparison*, (deutscher Titel: Statistische Verfahren für Eignungsprüfungen durch Ringversuche).
- [19] Thompson, M. et. al., The International Harmonized Protocol for the proficiency testing of analytical chemistry laboratories (IUPAC Technical Report). *De Gruyter, Pure and Applied Chemistry*. 2006, Vol 78, No 1, pp. 145-196

- [20] DIN 38402-45:2014-06, *Deutsche Einheitsverfahren zur Wasser-, Abwasser- und Schlammuntersuchung – Allgemeine Angaben (Gruppe A) – Teil 45: Ringversuche zur Eignungsprüfung von Laboratorien (A 45)*.
- [21] DIN ISO 5725-5:2002-11, *Genauigkeit (Richtigkeit und Präzision) von Messverfahren und Messergebnissen – Teil 5: Alternative Methoden für die Ermittlung der Präzision eines vereinheitlichten Messverfahrens*.
- [22] Silverman, B.W., *Density Estimation for Statistics and Data Analysis*, London: Chapman & Hall/CRC, 1986, pp. 45; ISBN 978-0-412-24620-3
- [23] Shapiro, S.S. und Wilk, M.B., An Analysis of Variance Test for Normality (Complete Sample). *Biometrika*. 1965, Vol. 52, No 3/4, pp. 591-611
- [24] Europäisches Parlament und Rat, Verordnung (EU) 2017/625 vom 15. März 2017 über amtliche Kontrollen und andere amtliche Tätigkeiten zur Gewährleistung der Anwendung des Lebens- und Futtermittelrechts und der Vorschriften über Tiergesundheit und Tierschutz, Pflanzengesundheit und Pflanzenschutzmittel.
- [25] Albert, W., Horwitz, R., The Horwitz Ratio (HorRat) A useful index of method performance with respect to precision. *Journal of AOAC International*. 2006, 89(4), pp. 1096-1109
- [26] Uhlig, S. et. al., Reliability of Precision Data Obtained from Interlaboratory Studies. *Journal of AOAC International*. 2013, Vol. 96, No 2, pp. 466-470