

Positionspapier der Fachgruppe Arzneimittelkontrolle/Pharmazeutische Analytik der Deutschen Pharmazeutischen Gesellschaft zum Thema Festlegung von Akzeptanzkriterien

1. Präambel

In diesem Positionspapier wird dargestellt, wie aus der analytischen Variabilität auf sinnvolle Akzeptanzkriterien bei der Gehaltsbestimmung von Wirkstoffen und Fertigprodukten sowie bei Reinheitsuntersuchungen geschlossen werden kann.

Das vorliegende Papier wurde zunächst in einem kleineren Kreis von Experten diskutiert. Neben den Vorsitzenden der Fachgruppe, Prof. Dr. H. Wätzig, TU Braunschweig, und Dr. J. Ermer, Aventis AG, waren an dieser Diskussion die Herren Doktoren Manfred Fischer, Lilly Pharma, Hamburg, Bernd Renger, Baxter AG, Wien, und Holger van Lishaut, Abbott, Ludwigshafen beteiligt.

Der erstellte Entwurf wurde veröffentlicht [1-3] und anlässlich der Hauptjahrestagung der Fachgruppe Arzneimittelkontrolle/Pharmazeutische Analytik der DPhG vom 8.-9.10.2002 in Berlin nach intensiver Diskussion mit einigen Veränderungen beschlossen. Das Positionspapier mit den eingearbeiteten Veränderungen wurde vor der endgültigen Veröffentlichung nochmals im Internet vorgestellt, damit alle beteiligten Fachgruppenmitglieder kontrollieren konnten, ob alle Korrekturen korrekt eingearbeitet wurden.

Die Begründungen und Diskussionen zu den einzelnen Punkten werden in [4] ausführlich dargestellt.

2. Akzeptanzkriterien des Gehaltes für Wirkstoffe

Die Fachgruppe empfiehlt, bei der Festlegung von Akzeptanzkriterien für Wirkstoffe im wesentlichen das Konzept von DAAS und MILLER [5] zu übernehmen. Die Fachgruppe schlägt einige Konkretisierungen vor und stellt klar, dass Zielstandardabweichungen in vielen Fällen heute noch nicht ausreichend bekannt sind.

Die Berechnung der Akzeptanzkriterien erfolgt nach den Gleichungen 1a und 1b [6]:

$$(1a) \quad LSL = 100\% - \%MGV - 3 \cdot TSD$$

$$(1b) \quad USL = 100\% + 3 \cdot TSD$$

%MGV: maximale prozentuale Gesamtverunreinigungen laut Spezifikation, die nicht bei der gewählten Bestimmungsmethode miterfasst werden

TSD: Zielstandardabweichung (target standard deviation) aus Ringversuchen [5]. Als Näherung kann die Gesamt-Wiederholpräzision (pooled repeatability) aus mehreren Serien verwendet werden.

Daas und Miller verwenden die englischen Begriffe lower bzw. upper specification limit (LSL und USL). Im deutschen werden die Begriffe Akzeptanzkriterium und Spezifikationsgrenze weitgehend synonym verwendet, im englischen hat sich der Begriff acceptance criterium durchgesetzt, der auch von der ICH verwendet wird (vgl. [10]). Wir verwenden im folgenden den Begriff Akzeptanzkriterium.

Das Konzept der Zielstandardabweichung für eine Klasse von Methoden zur Abschätzung der wahren Wiederholstandardabweichung σ_r wird grundsätzlich begrüßt. Wenn die Zielstandardabweichungen für eine Klasse von Methoden ähnlich ist, gestattet dies eine schnelle Ermittlung von sinnvollen Akzeptanzkriterien.

Diese TSD's können aber nur nach umfangreicher experimenteller Absicherung als zuverlässige Abschätzungen von σ_r betrachtet werden. Es ist heute noch zu früh, TSD's für alle Klassen von Methoden endgültig festzulegen.

Als Alternative kann auch die Gl. 2 verwendet werden (vgl. 3.2.), die Basisgrenze PL entspricht dann %MGV. Der analytisch erforderliche Bereich wird an Stelle der allgemeinen Abschätzung mit dem Dreifachen der TSD aus dem konkreten Prognosebereich des Prüfverfahrens berechnet.

3. Akzeptanzkriterien für den Gehalt im Fertigarzneimittel

Für diese Fragestellung existiert bisher kein einheitliches Konzept. Die Fachgruppe stellt eine eigene Ausarbeitung vor.

3.1. Regelakzeptanzkriterien: 95 - 105%

Für Gehaltsbestimmungen aus Formulierungen bei der Beantragung von europäischen Zulassungen gilt die Grundregel, dass der ermittelte Gehalt ohne angemessene Begründung den Bereich zwischen 95 und 105% des deklarierten Gehalts nicht überschreiten darf. Werden diese Grenzen eingehalten, können sie bei der Zulassung eines Arzneimittels angegeben werden und müssen nicht extra begründet werden.

Es besteht Konsens darüber, dass dies in der Regel eine vernünftige Vorgehensweise ist und dass diese Grenzen in den meisten Fällen eingehalten werden können. Ebenso besteht Übereinstimmung, dass es begründete Einzelfälle gibt, in denen Arzneimittel mit weiter gesteckten Grenzen zugelassen werden müssen. Mögliche Ursachen, warum durch erhöhte Variabilität die Grenzen 95% bis 105% nicht eingehalten werden können, sind:

- Unvermeidbar hohe herstellungsbedingte Chargenvariabilität
- Sehr geringe Analytkonzentrationen
- Matrixeffekte
- Unvermeidbar hohe Variabilität des Prüfverfahrens

3.2. Akzeptanzkriterien in Abhängigkeit von der analytischen Variabilität

Mit Hilfe einer einfachen Berechnung kann ermittelt werden, welche Akzeptanzkriterien bei einer jeweils vorgegebenen produktionsbedingten und analytisch bedingten Variabilität resultieren:

Die Akzeptanzgrenzen AL müssen sowohl die Herstellungs- als auch die analytische Variabilität berücksichtigen. Letztere kann als Prognosebereich des Mittelwertes beschrieben werden:

$$(2) \quad AL = 100\% \pm PL \pm \frac{t_{df,95\%} * \hat{\sigma}_R}{\sqrt{n_{assay}}}$$

PL: sich aus der Herstellung ergebende Basisgrenzen

$\hat{\sigma}_R$: Vergleichsstandardabweichung

Die **Ringversuchs-Standardabweichung** (Interlabor-Standardabweichung, Vergleichsstandardabweichung i.e.S, reproducibility) wird mit identischer Methode und identischem Testmaterial in unterschiedlichen Labors ermittelt. Bei diesem Präzisionsmaß werden zusätzlich zur zufälligen Streuung der Analysenergebnisse (einschließlich des Beitrags des Referenzstandards) besonders auch systematische Veränderungen durch äußere Einflußgrößen (z.B. Temperatur, Raumfeuchte, Reagentienqualität, Ausbildungsstand des Laborpersonals ...) berücksichtigt.

Über Monate und Jahre nähert sich die Mehrtages-Standardabweichung (intermediate precision) immer stärker der Vergleichs-Standardabweichung an, da auch innerhalb eines Labors über längere Zeiträume Personal und Instrumentierung wechseln. Wir verwenden daher in diesem Positionspapier den Begriff **Vergleichsstandardabweichung** ($\hat{\sigma}_R$, RSD_R(%)) als übergeordneten Begriff, der die Ringversuchs-Standardabweichung und die langfristige Intralabor-Standardabweichung umfasst.

n_{assay} : in der Routineanalytik geplante Anzahl von Mehrfachmessungen nach unabhängiger Probenaufarbeitung (z.B. von verschiedenen Einwaagen). Durch die unabhängige Aufarbeitung soll der gesamte analytische Fehler abgebildet werden.

t_{df} : t-Faktor für die Anzahl der Freiheitsgrade bei der Ermittlung der Standardabweichung

$\hat{\sigma}_R / \sqrt{n_{assay}}$: Standardfehler unter Vergleichsbedingungen; „Vergleichsstandardfehler“

Diese Formel bestätigt, das die 95 - 105%-Kriterien in vielen Fällen eingehalten werden können:

Für eine Herstellvariabilität von 2.5% und eine mit 2 Serien a 6 Werten ermittelte Laborvergleichspräzision folgt ein Wert von ungefähr 1% für $\hat{\sigma}_R / \sqrt{n_{assay}}$ für das Akzeptanzkriterium 95 bis 105%.

$$AL = 100\% \pm 2.5\% \pm \frac{2.23 * \hat{\sigma}_R}{\sqrt{n_{assay}}}$$

Der Wert $t_{0,05, 10}$ beträgt 2.23 für eine zweiseitige Fragestellung. Die Anzahl der Freiheitsgrade ist die Summe der Freiheitsgrade in den zwei Einzelserien, also $2*(6-1) = 10$.

Also können die 95-105%-Grenzen mit einer relativen Vergleichsstandardabweichung von 1% mit Einfachbestimmungen eingehalten werden, oder mit einer solchen von 1.7%, wenn Dreifachbestimmungen durchgeführt werden.

Zum Zeitpunkt der Zulassung ist PL meist noch nicht exakt bekannt. Die Abschätzung, dass die Hälfte des Akzeptanzbereiches von der Herstellung beansprucht wird, dürfte jedoch bei Standardprozessen realistisch sein. Die vorgestellte Berechnung von individuellen Akzeptanzkriterien ist aber selbstverständlich auch für andere Werte als PL=2.5% gültig.

Wenn die produktions- oder die analytisch bedingte Variabilität deutlich höhere Werte als in den Beispielen diskutiert annimmt, ermöglicht Gleichung 2 eine Ermittlung sinnvoller individueller Akzeptanzkriterien. Bei einer produktionsbedingten Variabilität von 2.5% sind

weniger enge Akzeptanzkriterien erforderlich, wenn $\hat{\sigma}_R/\sqrt{n_{\text{assay}}}$ (mit einer sinnvollen Anzahl an Wiederholbestimmungen) einen anderen Wert als 1% annimmt.

Beispiel: Es ist bekannt, dass die relative Vergleichsstandardabweichung bei abnehmender Konzentration zunimmt (vgl. 3.3., [7-9]). Die im folgenden tabellierten Werte für $RSD_R(\%)$ entsprechen der bestenfalls erreichbaren Vergleichspräzision nach HORWITZ (Hälfte der im Mittel erreichten $RSD\%$, [9]). Aus den zunehmenden Werten für $RSD_R(\%)$ resultiert eine zunehmende analytische Variabilität und daher sind erweiterte Akzeptanzkriterien (LAL/UAL = lower/upper acceptance limit) erforderlich:

Tabelle 1

| c (g/kg) | $RSD_R(\%)$ | n | analytische Variabilität | LAL | UAL |
|----------|-------------|---|--------------------------|------|-------|
| 100 | 1,4 | 3 | 1,8 | 95,7 | 104,3 |
| 10 | 2 | 3 | 2,6 | 94,9 | 105,1 |
| 1 | 2,9 | 3 | 3,7 | 93,8 | 106,2 |
| 0,1 | 4 | 3 | 5,1 | 92,4 | 107,6 |

Um die Zunahme der analytischen Variabilität zu kompensieren, könnte die Datenzahl für die Routinemessungen erhöht und dadurch die 95-105%-Grenzen weiter eingehalten werden. Dies führt aber schnell zu einer sehr hohen Anzahl von Messungen:

Tabelle 2 (n angepasst)

| c (g/kg) | $RSD_R(\%)$ | n | analytische Variabilität | LAL | UAL |
|----------|-------------|----|--------------------------|------|-------|
| 100 | 1,4 | 3 | 1,8 | 95,7 | 104,3 |
| 10 | 2 | 4 | 2,2 | 95,3 | 104,7 |
| 1 | 2,9 | 7 | 2,4 | 95,1 | 104,9 |
| 0,1 | 4 | 13 | 2,5 | 95 | 105 |

Analoge Überlegungen gelten für eine anderweitig verursachte erhöhte analytische Variabilität, z.B. durch Matrixeffekte oder durch Trends über längere Zeiträume. Auch wenn die Streuung von einzelnen Messwerten sehr hoch ist, kann durch eine hinreichend hohe Datenzahl die Streuung für den Mittelwert theoretisch beliebig klein erhalten werden. Allerdings sind für $\hat{\sigma}=10\%$ bereits $n = 100$ Messungen notwendig. Solch hohe Datenzahlen sind normalerweise nicht zu rechtfertigen, es sei denn, dass die verwendete Analytik extrem schnell und kostengünstig ist. Die Fachgruppe schlägt vor, nicht über eine Dreifachbestimmung hinauszugehen, sofern nicht Sicherheitsaspekte engere Akzeptanzkriterien erfordern, oder eine Erhöhung der Anzahl der Wiederholungsmessungen leicht möglich ist.

Diese Überlegungen zeigen, dass man sich um weniger enge, dem Spezialfall angepasste Akzeptanzkriterien bemühen muss, wenn die Wiederholstandardabweichung für Einzelbestimmungen 2% übersteigt.

3.3 Abschätzung der analytischen Variabilität?

Es gibt zur Zeit noch zu wenig Veröffentlichungen zur zuverlässigen Abschätzung von Vergleichsstandardabweichungen $\hat{\sigma}_R$ in der pharmazeutischen Analytik. Daher kann generell die analytische Variabilität nicht a priori abgeschätzt werden. Während der Methodenvalidierung wird jedoch die Mehrtages-Standardabweichung (intermediate precision) jeweils individuell ermittelt. Deren Wert kann zur Abschätzung von $\hat{\sigma}_R$ und zur Festlegung

von Akzeptanzkriterien verwendet werden, sofern die für eine künftige Anwendung des Verfahrens relevanten Einflussfaktoren im Design der Validierung berücksichtigt wurden. Die Arzneimittelkontrolle/Pharmazeutische Analytik der DPhG sieht weiteren Forschungsbedarf zur allgemeinen Abschätzung von Wiederhol- und Vergleichspräzisionen. Die Fachgruppe ist bereit, die dazu notwendige Datensammlung zu übernehmen. Für weitere Information bzw. die Zusendung einer Excel-Datenerfassungstabelle steht der Vorstand zur Verfügung (joachim.erner@aventis.com bzw. h.waetzig@tu-bs.de; s.u.) Die erhaltene Information wird in regelmäßigen Abständen zusammengestellt und, auf Wunsch sehr gern gemeinsam mit den Einsendern von Daten, publiziert.

Für erhöhte Vergleichsstandardabweichungen durch kleine Analytkonzentrationen gibt es eine umfangreiche Datensammlung durch HORWITZ, welche eine Orientierung über die Variabilität in diesen Fällen erleichtert. Horwitz stellte fest, dass die prozentuale Vergleichsstandardabweichung $RSD_R(\%)$ im wesentlichen von der Konzentration c des Analyten (in g/g) in der Ausgangssubstanz (also z.B. in einer Tablette oder einer Infusionslösung) abhängt, relativ unabhängig von der verwendeten analytischen Technik oder der Probenvorbereitung ([7-9], vgl. Tabellen 1 und 2).

Es ist offensichtlich, dass diese von Horwitz entdeckte Beziehung nur als Orientierungshilfe geeignet ist, da sie allgemein für die verschiedensten analytischen Methoden gültig ist. Konkrete Aussagen für ein bestimmtes Arzneimittel sind nicht aus der HORWITZ-Funktion ableitbar. Die HORWITZ-Gleichung ist jedoch sehr nützlich, um zu zeigen, welche Präzisionen bei einer bestimmten Konzentration nicht mehr erreichbar sind. In vielen Fällen mit geringen Wirkstoffkonzentration kann man also auch a priori davon ausgehen, dass $\hat{\sigma}_R/\sqrt{n_{\text{assay}}} \approx 1\%$ nur für eine unverhältnismäßig große Zahl von Bestimmungen (die in die Duzende gehen kann) noch möglich ist.

4. Akzeptanzkriterien für den Gehalt an Verunreinigungen: ein an ICH Q6A orientierter Vorschlag

Akzeptanzkriterien für Verunreinigungen von Arzneistoffen können auf verschiedenen Wegen sinnvoll begründet werden. Die Fachgruppe stellt eine mögliche Vorgehensweise vor, welche sich nah an der ICH-Richtlinie Q6A [10] orientiert und diese konkretisiert.

Das Akzeptanzlimit AL für eine einzelne Substanz wird nach einer einfachen Formel festgelegt:

$$(4) \quad AL = \bar{x} + 3 \cdot \hat{\sigma}$$

AL wird hier angegeben in % bezogen auf den Arzneistoff-Gehalt

Der erhaltene Wert wird auf eine Dezimale gerundet. Zur Verwendung dieses statistischen Ansatzes sollten Mittelwert und Standardabweichung aus mindestens 5 repräsentativen, möglichst aufeinanderfolgende Chargen aus den klinischen Phasen II und III bestimmt werden. Die so erhaltene Grenze muss selbstverständlich toxikologisch qualifiziert sein.

Für die Gesamtmenge an Verunreinigungen (total sum of impurities) spricht die Fachgruppe bewußt keine Empfehlungen aus. Hier müssen in jedem Fall individuelle Gegebenheiten berücksichtigt werden.

Weitere Hintergrundinformation zu diesem Positionspapier liefert [4].

Literatur:

- [1] Wätzig, H., Ermer, J., Pharmazie in unserer Zeit, 31, 501-503 (2002)
- [2] Wätzig, H., Ermer, J., Pharmazeutische Zeitung, 147, 3592 (2002)
- [3] Wätzig, H., Ermer, J., www.tu-bs.de/institute/pharmchem/dphg_pospapier.pdf
- [4] Wätzig, H., Ermer, J., Analytische Unsicherheit und Rationale Spezifikationsfindung – Schwerpunkt: Gehaltsbestimmungen, <http://www.tu-bs.de/institute/pharmchem/phchem-waetz.htm>
- [5] Daas, A. G. J., Miller, J. H. McB., Pharmeuropa 11(4), 571-577 (1999)
- [6] Miller, J. H. McB., European Directorate for the Quality of Medicines (EDQM), Strasburg, Setting of content limits in Pharm. Eur., Beitrag zum Workshop „Analytische Unsicherheit und Rationale Spezifikationsfindung – Schwerpunkt: Gehaltsbestimmungen“ der Fachgruppe Arzneimittelkontrolle/Pharmazeutische Analytik der DPhG (Frankfurt/Main, 31.01.02)
- [7] Horwitz, W., Anal. Chem. 54 (1982) 67A.
- [8] Horwitz, W., Albert, R., J. AOAC Int. 79 (1996) 589.
- [9] Albert, R., Horwitz, W. Anal. Chem. 69 (1997) 789-790.
- [10] International Conference on Harmonisation, Notes for Guidance, Q6A.
<http://www.ich.org/pdf/ICH/q6a.pdf> bzw. <http://www.fda.gov/cder/guidance/ichq6a.pdf>

Für die Fachgruppe Arzneimittelkontrolle/Pharmazeutische Analytik der Deutschen Pharmazeutischen Gesellschaft:

Prof. Dr. Hermann Wätzig
Institut für Pharmazeutische Chemie
Beethovenstr. 55
D-38106 Braunschweig
h.waetzig@tu-bs.de
<http://www.tu-bs.de/institute/pharmchem/phchem-waetz.htm>

1. Vorsitzender der Fachgruppe
Arzneimittelkontrolle/Pharmazeutische Analytik

Dr. Joachim Ermer
Global Analytical Development
QO Technical Services and Systems
Aventis Pharma AG
D-65926 Frankfurt am Main
Joachim.Ermer@aventis.com

2. Vorsitzender der Fachgruppe
Arzneimittelkontrolle/Pharmazeutische Analytik