

イ法は、MDCK 上皮細胞の遊走、BaF3 細胞増殖 ( $^3\text{H}$  チミジンの取り込み)、同細胞における c-Met のチロシンリン酸化、MvILu 細胞の増殖 (クリスタルバイオレット法、 $^3\text{H}$  チミジンの取り込み)、初代培養ラット肝細胞の増殖 ( $^3\text{H}$  チミジンの取り込み)、乳腺上皮細胞 B5/589 細胞の増殖 ( $^3\text{H}$  チミジンの取り込み) であった。これらの方法を用いて各標準品候補組換え HGF の含量を各施設が所有する濃度既知の自家組換え HGF 標準品と比較することにより調べた。その結果、ほとんどの場合お互いに平行な用量反応直線が濃度を対数に変換した場合において得られた。しかしながらある施設において初代培養ラット肝細胞の増殖アッセイ ( $^3\text{H}$  チミジンの取り込み) で自家標準品が作用を示さなかった。また、ある施設において乳腺上皮細胞 B5/589 細胞の増殖アッセイ ( $^3\text{H}$  チミジンの取り込み) 及び BaF3 細胞における c-Met のチロシンリン酸化アッセイで標準品候補組換え HGF が平行な用量反応性を示さなかった。そこでこれら 2 施設におけるバイオアッセイは不適当と判断され、力価の検定からは除かれ、その他のバイオアッセイにおいて平行な用量反応直線が得られた施設における検討結果がまとめられた。作成した I アンプル当たりの標準品候補組換え HGF 充填量は CHO 細胞由来で  $5\mu\text{g}$ 、sf 21 細胞由来で  $4\mu\text{g}$ 、マウスミエローマ細胞由来で  $5\mu\text{g}$  であったが、全体の検討結果から CHO 細胞由来で  $4.03\mu\text{g}$  (変動係数 233%)、sf 21 細胞由来で  $1.70\mu\text{g}$  (変動係数 199%)、マウスミエローマ細胞由来で  $4.89\mu\text{g}$  (変動係数 267%) と含量が算出された。なお各施設における値の再現性は良好であった。

各施設の自家標準品は 1 本鎖あるいは 2 本鎖の組換え HGF であるが、1 本鎖及び 2 本鎖標準品候補組換え HGF の含量及び変動係数をそれぞれ 1 本鎖及び 2 本鎖自家標準品組換え HGF の値を元に算出しなおしても、その値は先ほどの全体の自家標準品を元にした値とさほど変化はみられなかった。また、算出された各標準品候補の含量をお互いの相対比で取り直すと変動係数は 44% 以下に低下した。標準品候補組換え HGF の中で充填量に比べ特に測定結果から算出された含量が低かったものは sf 21 細胞由来のもので、その値は約 40% であった。先の予備的検討から標準品候補 1 本鎖組換え HGF の生物活性を有する 2 本鎖への変換は FCS 濃度に依存することが明らかになっていることから、用いられたバイオアッセイでは FCS 濃度が低いためこの変換が充分ではなく含量が低く算出された可能性が考えられる。変動係数が全体的に極めて高い理由であるが、各施設における自家標準品と標準品候補の比活性がお互い異なっていることが主な原因と考えられる。その違いは由来及び調製法の違い、保存時における失活の可能性が考えられる。実際、充填量に比べ約 4 倍高い含量が得られたある施設では自家標準品の  $4^{\circ}\text{C}$  における長期保存による失活の可能性を指摘している。

標準品候補組換え HGF の免疫学的活性については市販あるいは各施設で確立された ELISA 系を用い、先と同様に自家組換え HGF 標準品を基にした測定がなされた。その結果全体の検討結果から、CHO 細胞由来で  $3.86\mu\text{g}$  (変動係数 51%)、sf 21 細胞由来で  $2.11\mu\text{g}$  (変動係数 147%)、マウスミエローマ細胞由来で  $3.71\mu\text{g}$  (変動係数

71%) と含量が算出された。なお各施設における値の再現性は良好であった。各施設で測定に用いられた抗体の特異性については明らかではないが、全体的に2本鎖 HGF の値は1本鎖 HGF に比べ2倍高い値が得られた。これは恐らく1本鎖に比べ2本鎖 HGF のほうが抗体により認識しやすいことによると思われる。1本鎖組換え HGF は昆虫細胞である sf 21 細胞で産生されるが、本細胞においては哺乳類細胞でみられるガラクトースとシアル酸を含む複合型糖鎖は形成されないことから、HGF の抗体による認識は形成される糖鎖の違いにより異なる可能性が考えられる。また、1本鎖と2本鎖では形成される HGF の立体構造も異なる可能性も考えられることから、HGF の抗体による認識は立体構造の違いにより異なる可能性も考えられる。又、自家標準品として2本鎖組換え HGF における値を基に2本鎖標準品候補組換え HGF の値を算出しなおすと、CHO 細胞由来で  $3.81\mu\text{g}$  (変動係数 60%)、マウスミエローマ細胞由来で  $2.97\mu\text{g}$  (変動係数 54%) であった。同様に、自家標準品として1本鎖組換え HGF における値を基に1本鎖標準品候補組換え HGF の値を算出しなおすと、sf 21 細胞由来で  $4.86\mu\text{g}$  (変動係数 5%) であった。以上のように同様な型の組換え HGF の観点で全体の結果を算出しなおすと標準品候補である CHO 及び sf 21 細胞由来組換え HGF の含量はそれぞれの充填量4及び5  $\mu\text{g}$  とよく一致し変動係数も比較的良かった。

以上の検討結果から、WHO 生物学的物質標準化の専門家の会議において CHO 細胞由来の HGF を組換え HGF 1 次国際標準

品とし、1アンプル当たり生物学的な力価として 4000IU、免疫学的な値として  $4\mu\text{g}$  と設定された。また、同様に1本鎖の HGF に関しては sf 21 細胞由来の組換え HGF を1次国際標準品とし、1アンプル当たり 2000IU 及び  $4\mu\text{g}$  と設定された。

2. 疾患モデル動物を用いた組換え HGF の各種疾患の治療効果と治療効果に関連した細胞レベルにおける作用

肝臓において劇症肝炎、急性肝炎、肝硬変、肝繊維症、脂肪肝などにおける HGF の治療薬としての有用性の可能性が各種肝疾患モデル動物を用いた検討から明らかにされている。これら治療効果は HGF の以下のような作用によることが明らかになっている。劇症肝炎、急性肝炎の HGF による治療効果については抗アポトーシス、肝細胞の増殖促進、抗肝炎作用、胆管の形態形成誘導、タンパク質(血液凝固・線溶系関連タンパク質、アルブミン等)の合成促進作用によることが示されている。肝硬変、肝繊維症の HGF による治療効果については細胞外マトリックスの融解、抗アポトーシス、肝細胞の増殖促進、 $\text{TGF}\cdot\beta$  の発現抑制、タンパク質(血液凝固・線溶系関連タンパク質、アルブミン等)の合成促進作用によることが示されている。脂肪肝の HGF による治療効果については脂質の分泌促進によることが示されている。肝移植、部分肝切除、肝虚血、肝癌などの外科的手術に関する HGF による治療効果については抗アポトーシス、肝細胞の増殖促進、タンパク質(血液凝固・線溶系関連タンパク質、アルブミン等)の合成促進、肝癌細胞の増殖抑制作用によることが示されている。

これら治療効果と関連した細胞レベルに

おける HGF の作用については以下のような知見が得られている。初代培養ラット肝細胞において HGF は 1~10ng/ml で濃度依存的に DNA 合成促進作用を促進し、最大で 10 倍以上の促進がみられることが知られている。HGF は初代培養ラット肝細胞においてアルブミン合成及び分泌を同様な濃度依存性で約 1.5 倍促進することが知られている。初代培養ラット肝細胞において HGF はグルコース 6 リン酸脱水素酵素の活性を同様な濃度依存性で約 2 倍上昇させることが知られている。初代培養マウス肝細胞において HGF はインターフェロン $\gamma$  による LDH の遊離を 0.1~10 $\mu$ g/ml で濃度依存的に抑制し、最大で約 50%の抑制がみられることが知られている。肝癌細胞の一つである Hep G2 細胞において HGF は 50ng/ml で約 64 倍アポトーシスを誘導することが TUNEL 法を用いた測定により明らかになっている。また、同細胞において DNA 合成を約 30%抑制することが知られている。なお、これら作用の濃度依存性については不明である。形質転換したラット肝上皮細胞 AFLB8 及び C4T において HGF は増殖を 10ng/ml でそれぞれ約 50%及び 70%抑制することが知られている。また、AFLB8 において HGF は 10ng/ml でアポトーシスを約 5 倍促進することが知られている。なお、これらの作用の濃度依存性については不明である。

腎臓においては急性腎不全及び慢性腎不全における HGF の治療薬としての有用性の可能性について、急性及び慢性腎不全モデル動物において HGF は腎臓機能低下による血中尿酸窒素及び血清クレアチンの上昇を HGF は強く抑制することから明らか

になっている。これら治療効果は HGF による尿細管上皮の増殖促進、尿細管再構築、抗アポトーシス作用、コラゲナーゼの活性化、TGF $\cdot$  $\beta$  の発現抑制などによることが知られている。

これら治療効果と関連した細胞レベルにおける HGF の作用については以下のような知見が得られている。腎臓近位細管細胞の一つである PT-1 細胞において HGF は 0.01~100nM で濃度依存的に増殖を促進し、最大で約 6 倍の促進がみられることが知られている。この結果は培地に 10%FCS 添加した場合であるが、同じ細胞の無血清培養において HGF は 1nM で増殖、遊走、幹腔形成を約 2 倍、微繊毛の形成を約 1.5 倍促進するという報告もある。なお、これらの作用の濃度依存性については不明である。牛尿細管上皮 (MDCK) 細胞のコラゲンゲルにおける培養において HGF は 2~40ng/ml で濃度依存的に全体に占める幹腔構造を形成したコロニーの割合を増加し、最大で約 80%の増加がみられることが知られている。腎糸球体細胞においてシクロスポリン A 処理によるアポトーシスの誘導に対して HGF は 50ng/ml 及び 100ng/ml でそれぞれ約 50%及び 100%阻害することが知られている。なおアポトーシスの測定はヒストンに結合した DNA 断片の細胞質への遊離を元にした市販の ELISA キットを用いて行われている。腎臓細管上皮細胞の一つである mIMCD-3 細胞において 90%ATP 除去により接着抑制に対して HGF は 40ng/ml でコントロールレベルまで回復することが知られている。ヒト近位尿細管細胞の一つである HKC で HGF は 20ng/ml で血清除去による生存細胞数の低下を約 2

倍抑制することが知られている。なお、これら作用の濃度依存性については不明である。ヒト腎臓動脈内皮細胞において腎臓障害への関与が知られているエンドセリン-1の遊離を HGF は 0.001~10nM で濃度依存的に抑制し、最大で約 75%の抑制がみられることが知られている。

心筋梗塞、慢性動脈閉塞症などの血管機能障害における HGF の治療薬としての有用性の可能性が各種血管機能障害モデル動物を用いた検討結果から明らかにされている。これらの治療効果は血管内皮細胞の増殖、遊走、管腔形成、MT1-MMP 産生、VEGF 産生、MMP-2 活性化 Flk-1 (VEGF のレセプター) 発現などの促進及び抗アポトーシス作用によることが明らかになっている。

ヒトさい帯静脈由来血管内皮細胞 (HUVEC) において HGF は 1~10ng/ml の範囲で濃度依存的に増殖を促進し、最大で約 2 倍促進させることが知られている。また、同細胞で HGF は 1~50ng/ml の範囲で濃度依存的に遊走を促進し、最大で約 1.2 倍促進させることが知られている。HUVEC における HGF の遊走及び増殖促進作用については作用の程度は若干異なるものの多数報告されている。また、ヒト大静脈由来血管内皮細胞において HGF は 1~100ng/ml の範囲で濃度依存的に DNA 合成を促進し、最大で約 10 倍促進させることが知られている。ヒト大静脈由来内皮細胞において HGF は 1~100ng/ml の範囲で濃度依存的に血清除去による細胞数の低下を抑制し、最大で約 30%抑制することが示されている。同細胞において HGF は 100ng/ml で血清除去によるアポトーシス及びネクローシスの誘導を約 40%抑制する

ことが知られている。また、同様な実験系で HGF は 10ng/ml でアポトーシスの誘導を約 80%抑制するという報告もある。ヒト大静脈由来内皮細胞におけるアポトーシス及びネクローシスの誘導は高濃度の D-グルコースによっても誘導されるが、HGF は 100ng/ml でその誘導をほぼ完全に抑制することが示されている。なおこれら抑制作用の濃度依存性については不明である。正常ヒト内皮細胞の一つである HMVEC 及び HCAEC において HGF は 2~10ng/ml の範囲で MT1-MMP タンパク質を 6~9 倍増加させることが知られている。同様な細胞において HGF は 2ng/ml で MMP-2 の活性化を約 3~4 倍促進することが知られている。なお、その濃度依存性については不明である。HUVEC において HGF は 1~50ng/ml の範囲で濃度依存的に遊走を促進し、最大で約 2.5 倍促進することが示されている。ヒト表皮細胞 A431 及びヒト角化細胞 HaCaT において HGF は 1~100ng/ml の範囲で濃度依存的に培養上清における VEGF の蓄積を促進し、最大で約 4 倍促進することが知られている。同様な作用がヒト初代角化細胞でも知られているがその最大促進効果は約 2.5 倍である。また、HGF は 20ng/ml で HUVEC における VEGF 受容体である flk-1 の発現を約 2 倍増加させることが FACS を用いた解析から明らかになっている。なおその濃度依存性については不明である。

#### D. 結論

医薬品審査の国際調和を目指した研究の一環として、現在肝疾患、腎不全、血管機能障害に対する治療薬としての応用が期

待されている肝細胞増殖因子 (HGF) について生物学的及び免疫化学的国際標準品の設定、さらには各種疾患の治療薬として用

国際標準品に関しては最終的に1本鎖及び2本鎖組換え HGF の生物学的及び免疫学的標準品が設定されたが、その設定にあたって数々の問題点が明らかになった。生物学的標準品の設定にあたってはアッセイ系における1本鎖と2本鎖の活性の違い、特に対照として用いた各施設における自家標準品の比活性の違いによる思われる充填量と測定結果のずれ及び測定結果のばらつきが大きさが問題となった。また、免疫学的標準品に関しては恐らく産生細胞における糖鎖の違いによる抗体の反応性の違いが問題となった。このように、いろいろな面で複雑なタンパク性生理活性物質の標準化にあたってはどのような対照品を基準とするかが重要であることが明らかになった。

HGF の各種疾患の治療効果に関連した細胞レベルの生物学的活性については多数の生物学的活性が明らかになった。今後、各生物活性について再現性、濃度依存性、反応性、操作性などの種々の観点からその規格試験法への応用について検討する必要がある。

以上 HGF を例にとり示した多機能性バイオテクノロジー応用医薬品に関する具体的な標準品の設定及びバイオアッセイに関する知見は共通性が高いことから、今後他の多機能性バイオテクノロジー応用医薬品候補について標準品及び適切なバイオアッセイを設定する上で有用と考えられる。

#### E. 参考資料

[1] Rafferty B, Maile P, Rigsby P,

いる場合、各治療効果に関連した品質を保証するうえにおいて適切な細胞を用いた生物学的活性測定法について調査研究した。

Gaines DasRE, Robinson CJ, J  
Immuno Methods 258:1-11 (2001)

[2] HGF の分子医学 松本邦夫、森下竜一 編 メディカルレビュー社 (1998)

[3] 別所一彦、中村敏一 実験医学 20:1124-1132 実験医学 (2002)

#### F. 研究発表

[1] . Niimi S, Oshizawa T, Yamaguchi T, Harashima M, Seki T, Ariga T, Kawanishi T, Hayakawa T, Specific expression of annexin III in rat small-hepatocytes Biochem. Biophys. Res. Commun., 2003:300 770-774

[2] . Niimi S, Hyuga M, Kazama H, Inagawa M, Seki T, Ariga T, Kobayashi T, Hayakawa T, Activins A, AB and B inhibit hepatocyte growth factor synthesis by MRC-5 human lung fibroblasts Biol. Pharm. Bull., 2002:25 1405-1408

[3] Kobayashi T, Niimi S, Fukuoka M, Hayakawa T, Regulation of inhibin  $\beta$  chains and follistatin mRNA levels during rat hepatocyte growth induced by the peroxisome proliferator di-n-butyl phthalate. Biol. Pharm. Bull., 2002:25 1214-1219

[4] 原島 瑞、新見 伸吾、長岡 陽子、関 泰一郎、有賀 豊彦、早川 堯

夫 初代培養肝細胞においてプロテ  
アソーム特異的阻害剤であるラクタ  
シスチンはグルココルチコイド依存  
的なチロシンアミノトランスフェラ  
ーゼの誘導を阻害する 第 75 回 日

本生化学会大会 (2002 年 10 月)

## 安定性試験における品質確保基準に関する研究

分担研究者 吉岡澄江 国立医薬品食品衛生研究所 薬品部第二室室長

安定性試験において複数のロットあるいは、複数の含量違いなどから得られたデータがある場合には、それらをまとめて一括して解析してよいかどうかを評価しなければならない。そのための方法としては、以前からFDAのガイダンスの中に記載されていた分解曲線の傾きと切片の差を共分散分析法で検定する方法がある。また、リテストないし有効期間の推定値のレンジ、すなわち範囲に基づいて有効期間の差を直接評価する方法もある。さらに、統計的取扱いをより簡単に行う評価法も考案されている。本研究では、これらの各方法の差異や特徴を比較検討し、安定性試験の目的が適切に達成されるような安定性試験データ評価法の確立を目指し、国際調和のための基礎的検討を行うことを目的とする。

### 協力研究者

阿曾幸男 国立医薬品食品衛生研究所  
主任研究官

### A. 研究目的

日米欧3極内において新有効成分含有医薬品の原薬及び製剤の承認申請を行うときに必要な安定性試験成績を得るための標準的な試験方法は、「新原薬及び新製剤の安定性試験」に関するICHガイドラインに記載されている。しかし、それに従って得られた安定性データをどのように利用してリテスト期間又は有効期間を提示したらよいかについては統一した方法が確立されていない。また、どの程度まで長期保存条件での実測期間を超えた原薬のリテスト期間又は製剤の有効期間を提示することができるかに関しての考え方も各極間で異なっている。したがって、医薬品承認申請のグローバル化のためには、安定性試験データの評価法に関する国際調和が必要になっている。

安定性試験は、3ロット以上の原薬又は製剤について行う試験に基づいて、将来にわたって同じ状況で製造、包装される全てのロットに適用できるリテスト期間又は有効期間、さらに貯蔵方法のラベル表示を設定するために行われる試験であり、試験で得られたデータの評価が試験の最も重要な部分といえる。本研究は、安定性試験の目的が適切に達成されるような安定性試験データ評価法の確立を目指し、国際調和のための基礎的検討を行うことを目的とする。

### B. 研究方法

「新原薬及び新製剤の安定性試験」に関するICH

ガイドラインでは、リテスト期間又は有効期間を推定するために定量的な安定性データを解析するための方法として回帰分析が認められており、全ロットを一括して評価できるか否かに関する統計的検定を有意水準0.25で実施することが推奨されているが、その詳細は記載されていない。

すなわち、単一のロットから得られたデータに基づいてリテストないし有効期間を決める際には、通常、定量的な項目が時間に対して直線的な関係にあると仮定し、平均値の95%両側信頼限界を求め、片側下限が規格の下限と交差する点をリテスト期間ないし有効期間として設定することができるが、複数のロットあるいは複数の含量違いの製剤などから得られたデータがある場合には、それらをまとめて一括して解析してよいかどうかを評価しなければならない。そのための方法としては、以前からFDAのガイダンスの中に記載されていた分解曲線の傾きと切片の差を共分散分析法で検定する方法がある。また、リテストないし有効期間の推定値のレンジ、すなわち範囲に基づいて有効期間の差を直接評価する方法もある。さらに、統計的取扱いをより簡単に行う評価法も考案されている。本研究では、これらの各方法の特徴や差異を比較検討する。

### C. 研究結果

#### 1) 回帰曲線の傾きおよび切片の差の共分散分析に基づく一括評価の検定

含量や容器サイズあるいは容れ目などの要因が単一の原薬および製剤について、異なるロットに由来するデータを一括し、すべての原薬あるいは製剤に適用できる単一の有効期間を推定できるか否かを判

断する方法として、共分散分析による方法がある。

すなわち、異なるロット由来のデータを一括する前に、予備的な統計的検定を実施して、時間を共変量とみなす共分散分析によってロット間の回帰直線の傾きおよび縦軸切片の差を検定し、異なるロットの回帰直線の傾きと時間ゼロ時の縦軸切片が共通であるかどうかを判断する。正式な安定性試験では検体サイズが比較的小さく、検出力が低いことが予想されるので、それを補償するために、有意水準は0.25として各検定を行う。

傾きが等しいとする仮説が検定で棄却された場合、すなわち、ロット間に傾きの有意差が存在する場合には、全ロットのデータを一括することは不適切とみなされる。この場合、個々のロットのリテスト期間又は有効期間は、個々の縦軸切片及び個々の傾き、並びにすべてのデータから計算した平均二乗誤差を用いて推定することができる。全ロットのリテスト期間又は有効期間として、個々のロットの推定値のうち最も短いものを選ぶ。

縦軸切片が等しいとする仮説が検定で棄却されるが、傾きが等しいことを棄却できない場合、すなわち、ロット間で縦軸切片には有意差があるが、傾きには有意差がない場合には、データを一括して共通の傾きを求めることができる。この場合、個々のロットに対するリテスト期間又は有効期間は、共通の傾き及び個々の縦軸切片を用いて推定し、個々のロットの推定値のうち最も短いものを全ロットのリテスト期間又は有効期間として選ぶ。

傾きが等しいとする仮説および縦軸切片が等しいとする仮説が有意水準0.25で棄却されない場合、すなわち、ロット間で傾きおよび縦軸切片に有意差がない場合には、全ロットのデータを一括したデータから単一のリテスト期間/有効期間を推定し、それを全ロットに適用できる。ロットを一括することによってデータ量が増加するにしたがって回帰直線の信頼限界幅が狭くなるので、通常、一括したデータから推定したリテスト期間/有効期間は個々のロットから求めたものよりも長くなる。この一括評価に関する検定は、まず傾きの項を検定し、次に縦軸切片の項を検定して、項の数を最小にしたモデルにしたがってリテスト期間/有効期間を設定する。

複数の要因がある製剤の場合には、異なる要因のデータを一括し、すべての製剤に適用できる単一の有効期間を推定できるか否かを共分散分析によって判断する。すなわち、共分散分析によって、要因及び要因の組み合わせ間の回帰直線の傾きと縦軸切片の差を検定する。

全ての主効果及び交互作用の縦軸切片及び傾きの項、及び測定ランダム誤差を反映する項が含まれた統計的モデルを使う。高次の交互作用が極めて小さいと考えられる場合には、モデルにその項を含める必要はない。

通常、まず傾きの項を検定し、次に、縦軸切片の項を検定する。また交互作用は主効果よりも前に検定する。すなわち、高次交互作用の傾きの項から検定を開始し、次に高次交互作用の縦軸切片の項を検定し、さらに主効果の傾きの項、そして主効果の縦軸切片の項という順番で検定する。統計的に有意であることが明らかな項のみからなるモデルを利用して、有効期間を推定する。

検定に用いる有意水準としては、ロットが関わる項には0.25、ロットが関与しない項には0.05を用いる。異なる要因の組み合わせから求めたデータを一括評価できることが示された場合には、一括したデータを用いて有効期間を推定できる。

一括評価に関する検定によって特定の要因又は要因の組み合わせから得たデータを一括すべきではないと示された場合には、有意差があると判断された要因及び要因の組み合わせのそれぞれについて、別々の有効期間を推定するか、または、最も短い有効期間の推定値を単一の有効期間とすることができる。

## 2) 推定値のレンジに基づく一括評価の検定

上記の共分散分析による一括評価の検定のように回帰曲線の傾きと切片の変動を検定するのではなく、個々のロットや要因で得られる有効期間の推定値についてその同等性を推定値のレンジに基づいて検定し、異なるロットや要因に由来するデータを一括し、すべての製剤に適用できる単一の有効期間を推定できるか否かを判断することもできる。この方法では、例えば含量違いの製剤について推定した有効期間にある一定の割合の差が認められたとき、有意な差があるとみなす。

## 3) 各要因のリテスト期間/有効期間の推定値と設定値との比較に基づく評価（一括評価に関する検定を行わない方法）

含量や容器サイズあるいは容れ目などの要因が単一の原薬および製剤について、異なるロットに由来するデータを一括できるかどうかを判断するための一括評価に関する検定を行わずに有効期間を決める方法として、ロットの中に、提示したリテスト期間又は有効期間より短いリテスト期間又は有効期間を



持つものがあるか否かを評価する方法がある。

まず、個々のロットのリテスト期間又は有効期間を、個々の縦軸切片、個々の傾き及びすべてのデータから計算した平均二乗誤差を用いて推定する。各ロットのリテスト期間又は有効期間の推定値が提示したものより長い場合には、提示したリテスト期間又は有効期間は適切であると考えられる。

しかし、1つ、又は複数のリテスト期間又は有効期間の推定値が提示した期間より短い場合は、一括評価に関する検定を実施して、ロットを一括できるか否かを判断できる。

また、この方法は、一括評価に関する検定のプロセスで、使うこともできる。ロットの回帰直線が共通の傾きを持ち、共通の傾きと個々の縦軸切片に基づいて推定した有効期間が提示した有効期間よりも長い場合には、一括評価が可能であるか否かについて縦軸切片の検定を継続する必要はない。

複数の要因がある製剤の場合には、異なる要因のデータを一括できるかどうかを判断するための一括評価に関する検定を行わずに、要因の組み合わせの中に、提示した有効期間より短い有効期間を持つものがあるか否かを判断して有効期間を決める。

共分散分析による一括評価と同様に、まず、ロットや含量などの要因で有意差があるとする統計的モデルにしたがって、各要因及び要因の組み合わせのそれぞれについて有効期間を推定する。全ての有効期間の推定値が提示した有効期間より長い場合は、提示した有効期間は適切であると考えられる。有効期間の推定値のうち提示した有効期間より短いものがある場合は、各要因について共分散分析を行っていき、その都度、新しいモデルにしたがって、各要因及び要因の組み合わせのそれぞれについて有効期間を推定する。全ての有効期間が提示するものより長い場合には、それ以上のモデルは必要でない。

#### D. 考察

ロット間あるいは要因間の差を検定し、データを一括できるかどうか判断する方法としては、回帰曲線の傾きおよび切片について共分散分析を行う方法が最も一般的であると考えられる。しかし、この方法の検出力は定量誤差の大きさによって大きく影響され、定量誤差が大きくなるほど検出力は著しく低下する。その結果、大きい定量誤差を含むデータほど、ロット間あるいは要因間の安定性の差を見逃す傾向が大きくなり、安定性の差を検定するには問題があることも分かっている。

有効期間の推定値の同等性を推定値のレンジに基

づいて検定する方法は、共分散分析法に比較して、検出力が定量誤差の大きさによって受ける影響が著しく小さい利点を持つ。しかし、有効期間の推定値の最大値と最高値から計算した単一の値に基づいて有効期間の推定値の差を検定するため、推定される有効期間の変動は考慮されないため、レンジに基づく方法の検出力は試験に用いるロットや要因の数に大きく影響されるという欠点を持つ。推定値の変動を考慮して平均有効期間の同等性を検定する方法を確立することが必要であると考えられる。

一括評価に関する検定を行わずに、各要因の有効期間を提示する値と比較し、その結果に基づいて有効期間の設定する方法は、詳細な統計解析を必要としない簡便な方法であるが、ロット間あるいは要因間に差があるのかどうかを最終的に判断することはできないため、見かけ上、適切と思われる有効期間の保証にとどまり、安定性の実態の解析は不可能である。

#### E. 結論

複数のロットあるいは複数の含量違いなどから得られたデータがある場合に、分解曲線の傾きと切片の差を共分散分析法で検定する方法、あるいはリテストないし有効期間の推定値のレンジに基づいて有効期間の同等性を評価する方法によって、異なるロットあるいは要因からのデータを一括して解析してよいかどうかを判断するか、あるいは、一括評価に関する検定を行わずに、各要因の有効期間を提示する値と比較し、その結果に基づいて有効期間を設定するか、3通りの方法について比較検討した結果、それぞれの特徴が明らかにされた。

#### F. 研究発表

##### 1. 論文発表

S.Yoshioka, Y.Aso, S.Kojima, N.F.Cappuccino: A Comparison of the ANCOVA and range-based approaches for assessing batch-to-batch variability of the stability of pharmaceutical products. *Chem. Pharm.Bull.*, 50:881-883 (2002).

吉岡澄江：安定性試験へのブラケットティング法およびマトリキシング法の適用に関するガイドライン。 *医薬品研究* 33:561-567 (2002)

##### 2. 学会発表

なし

1  
2  
3 **ICH Q1E Step-4 (Draft 3)**  
4 **EVALUATION OF STABILITY DATA**  
5  
6

7 **1. INTRODUCTION**  
8

9 **1.1 Objectives of the Guidelines**  
10

11 This guideline is intended to provide recommendations on how to use stability data  
12 generated in accordance with the principles detailed in the ICH guideline "Q1A(R)  
13 Stability Testing of New Drug Substances and Products" (hereafter referred to as the  
14 parent guideline) to propose a retest period or shelf life in a registration/marketing  
15 authorisation application. This guideline describes when and how extrapolation can  
16 be considered when proposing a retest period for a drug substance or a shelf life for  
17 a drug product that extends beyond the period covered by available data from the  
18 long-term storage condition.  
19

20 **1.2 Background**  
21

22 The guidance on the evaluation and statistical analysis of stability data provided in  
23 the parent guideline is brief in nature and limited in scope. The parent guideline  
24 states that regression analysis is an appropriate approach to analyzing quantitative  
25 stability data for retest period or shelf life estimation and recommends that a  
26 statistical test for batch poolability be performed using a level of significance of 0.25.  
27 However, the parent guideline includes few details and does not cover situations  
28 where multiple factors are involved in a full or reduced-design study. This guideline  
29 supersedes the Evaluation sections of the parent guideline.

31  
32 **1.3 Scope of the Guideline**  
33

34 This guideline addresses the evaluation of stability data that should be submitted in  
35 registration applications for new molecular entities and associated drug products  
36 and for new dosage forms. The guideline provides recommendations on establishing  
37 retest periods and shelf lives for drug substances and drug products intended for  
38 storage at or below room temperatures. It covers stability studies using single or  
39 multi-factor designs and full or reduced designs. ICH Q6A should be consulted for  
40 the setting and justification of acceptance criteria, and ICH Q1D should be  
41 referenced for full- vs. reduced-design studies.  
42

43 **2. GUIDELINES**  
44

45 **2.1 General Principles**  
46

47 The design and execution of formal stability studies should follow the principles  
48 outlined in the parent guideline. The purpose of a stability study is to establish,  
49 based on testing a minimum of three batches of the drug substance or product, a  
50 retest period or shelf life and label storage instructions applicable to all future  
51 batches manufactured and packaged under similar circumstances. The degree of  
52 variability of individual batches affects the confidence that a future production  
53 batch will remain within acceptance criteria throughout its retest period or shelf life.

54

55 Although normal manufacturing and analytical variations are to be expected, it is  
56 important that the drug product is formulated with the intent to provide 100  
57 percent of the labeled amount of the drug substance at the time of batch release. If  
58 the assay values of the batches used to support the registration application are  
59 higher than 100 percent of label claim (taking into account the manufacturing and  
60 analytical variations) at the time of batch release, the proposed shelf life granted in  
61 the application can be overestimated. On the other hand, if the assay value of a  
62 batch manufactured after approval is lower than 100 percent of label claim at the  
63 time of batch release, it may fall below the lower acceptance criterion before the end  
64 of the approved shelf life.

65

66 A systematic approach should be adopted in the presentation and evaluation of the  
67 stability information. The stability information should include, as appropriate,  
68 results from the physical, chemical, biological, and microbiological tests, including  
69 those related to particular attributes of the dosage form (for example, dissolution  
70 rate for solid oral dosage forms). The adequacy of the mass balance should be  
71 assessed. Factors that can cause an apparent lack of mass balance should be  
72 considered, including, for example, the mechanisms of degradation and the  
73 stability-indicating capability and inherent variability of the analytical procedures.

74

75 The basic concepts of stability data evaluation are the same for single- versus multi-  
76 factor studies and for full- versus reduced- design studies. Data from formal  
77 stability studies and, as appropriate, supporting data should be evaluated to  
78 determine the critical quality attributes likely to influence the quality and  
79 performance of the drug substance or product. Each attribute should be assessed  
80 separately, and an overall assessment should be made of the findings for the  
81 purpose of proposing a retest period or shelf life. The retest period or shelf life  
82 proposed should not exceed that predicted for any single attribute.

83

84

85

86

87

88

89 The decision tree in Appendix A outlines the stepwise approach to stability data  
90 evaluation and when and how much extrapolation can be considered for a proposed  
91 retest period or shelf life. Appendix B provides information on how to analyze long-  
92 term stability data for appropriate quantitative test attributes from a study with a  
93 multi-factor, full or reduced design, shelf life how to use regression analysis for  
94 retest period or shelf life estimation and provides examples of statistical procedures  
95 to determine poolability of data from different batches or other factors. Additional  
96 guidance can be found in the references listed; however, the examples and  
97 references do not cover all applicable statistical approaches.

102

103 In general, certain quantitative chemical attributes (e.g., assay, degradation  
104 products, preservative content) for a drug substance or product can be assumed to  
105 follow zero-order kinetics during long-term storage<sup>1</sup>. Data for these attributes are  
106 therefore amenable to the type of statistical analysis described in Appendix B,  
107 including linear regression and poolability testing. Although the kinetics of other

108 quantitative attributes (e.g., pH, dissolution) is not known, the same statistical  
109 analysis can be applied, if appropriate. Qualitative attributes and microbiological  
110 attributes are not amenable to this kind of statistical analysis.

111  
112 The recommendations on statistical approaches in this guideline are not intended to  
113 imply that use of statistical evaluation is preferred when it can be justified to be  
114 unnecessary. However, statistical analysis can be useful in supporting the  
115 extrapolation of retest periods or shelf lives in certain situations and can be called  
116 for to verify the proposed retest periods or shelf lives in other cases.

## 118 119 **2.2 Data Presentation**

120  
121 Data for all attributes should be presented in an appropriate format (e.g., tabular,  
122 graphical, narrative), and an evaluation of such data should be included in the  
123 application. The values of quantitative attributes at all time points should be  
124 reported as measured and should not be expressed as percent of the initial values  
125 (i.e., at time-zero). If a statistical analysis is performed, the procedure used and the  
126 assumptions underlying the model should be stated and justified. A tabulated  
127 summary of the outcome of statistical analysis and/or graphical presentation of the  
128 long-term data should be included.

## 129 130 **2.3 Extrapolation**

131  
132  
133 Extrapolation is the practice of using a known data set to infer information about  
134 future data. Extrapolation to extend the retest period or shelf life beyond the period  
135 covered by available long-term data can be proposed in the application, particularly  
136 if no significant change is observed at the accelerated condition. Whether  
137 extrapolation of stability data is appropriate depends on the extent of knowledge  
138 about the change pattern, the goodness of fit of any mathematical model, and the  
139 existence of relevant supporting data. Any extrapolation should be limited so that  
140 the extrapolated retest period or shelf life would remain valid for a future batch  
141 with test results close to the in-house or release acceptance criteria at the time of  
142 batch release.

144  
145 An extrapolation of stability data assumes that the same change pattern will  
146 continue to apply beyond the period covered by available long-term data. The  
147 correctness of the assumed change pattern is crucial if extrapolation is  
148 contemplated. When estimating a regression line or curve within the available data,  
149 the data themselves provide a check on the correctness of the assumed change  
150 pattern, and statistical methods can be applied to test the goodness of fit of the data  
151 to the assumed line or curve. No such internal check is available beyond the period  
152 covered by observed data. Thus, a retest period or shelf life granted on the basis of  
153 extrapolation should always be verified by additional long-term stability data as  
154 soon as these data become available. Care should be taken to include in the  
155 protocol for commitment batches a time point that corresponds to the end of the  
156 extrapolated retest period or shelf life.

## 158 159 **2.4 Data Evaluation for Retest Period or Shelf life Estimation for Drug** 160 **Substances or Products Intended for Room Temperature Storage** 161

162 A systematic evaluation of the data from formal stability studies should be  
163 performed as illustrated in this section. A decision tree is provided in Appendix A  
164 as an aid to the guidance below. Stability data for each attribute should be assessed  
165 sequentially. For drug substances or products intended for storage at room  
166 temperature\* (see note) and studied according to the general case in the parent  
167 guideline, the assessment should begin with any significant change at the  
168 accelerated condition and, if appropriate, at the intermediate condition, and  
169 progress through the trends and variability of the data from the long-term condition.  
170 The circumstances are delineated under which extrapolation of retest period or  
171 shelf life beyond the period covered by available long-term data can be appropriate.

173  
174 **Note:** The term “room temperature” refers to the general customary environment  
175 and should not be inferred to be the storage statement for labeling.  
176

#### 177 **2.4.1 No significant change at accelerated condition**

178  
179 Where no significant change occurs at the accelerated condition, the retest period or  
180 shelf life setting would depend on the nature of the long-term and accelerated data.

##### 182 183 2.4.1.1 Long-term and accelerated data showing little or no change over time and 184 little or no variability

185  
186 Where the long-term data and accelerated data for an attribute show little or no  
187 change over time and little or no variability, it may be apparent that the drug  
188 substance or product will remain well within the acceptance criteria for that  
189 attribute during the proposed retest period or shelf life. In these circumstances a  
190 statistical analysis is normally considered unnecessary, but justification for the  
191 omission should be provided. Justification can include a discussion of the change  
192 pattern or lack of change, relevance of the accelerated data, mass balance, and/or  
193 other supporting data as described in the parent guideline.  
194

195 Extrapolation of the retest period or shelf life beyond the period covered by available  
196 long-term data can be proposed. The proposed retest period or shelf life can be up  
197 to twice, but not more than 12 months beyond, the period covered by available  
198 long-term data.

##### 200 201 2.4.1.2 Long-term or accelerated data showing change over time and variability

202  
203 If the long-term or accelerated data for an attribute show change over time and/or  
204 variability within a factor or among factors, statistical analysis of the long-term data  
205 can be useful in establishing a retest period or shelf life. Where there are  
206 differences in stability observed among batches or among other factors (e.g.,  
207 container size and/or fill, strength) or factor combinations (e.g., strength-by-  
208 container size and/or fill) that preclude the combining of data, the proposed retest  
209 period or shelf life should not exceed the shortest period supported by any batch,  
210 other factor, or factor combination. Alternatively, where the differences are readily  
211 attributed to a particular factor (e.g., strength), different shelf lives can be assigned  
212 to different levels within the factor (e.g., different strengths). A discussion should  
213 be provided to address the cause for the differences and the overall significance of  
214 such a difference on the product. Extrapolation beyond the period covered by  
215 available long-term data can be proposed; however, the extent of extrapolation

216 would depend on whether long-term data for the attribute are amenable to  
217 statistical analysis.

218  
219 • *Data not amenable to statistical analysis*

220  
221 Where long-term data are not amenable to statistical analysis but relevant  
222 supporting data are provided, the retest period or shelf life can be up to one-and-a-  
223 half times, but not more than 6 months beyond, the period covered by available  
224 long-term data. Relevant supporting data include satisfactory long-term data from  
225 development batches that are (1) made with a closely related formulation to, (2)  
226 manufactured on a smaller scale than, or (3) packaged in a container closure  
227 system similar to the primary stability batches.

229  
230 • *Data amenable to statistical analysis*

231  
232 If long-term data are amenable to statistical analysis but no analysis is performed,  
233 the extent of extrapolation should be the same as when data are not amenable to  
234 statistical analysis. However, if a statistical analysis is performed, it can be  
235 appropriate to propose a retest period or shelf life of up to twice, but not more than  
236 12 months beyond, the period covered by available long-term data, when backed by  
237 statistical analysis and relevant supporting data shelf life

## 240 241 **2.4.2 Significant change at accelerated condition**

242  
243 Where significant change (see note) occurs at the accelerated condition, the retest  
244 period or shelf life setting would depend on the outcome of stability testing at the  
245 intermediate condition, as well as long-term testing.

246  
247 **Note:** The following physical changes can be expected to occur at the accelerated  
248 condition and would not be considered significant change that calls for intermediate  
249 testing if there is no other significant change (potential interaction effects should  
250 also be considered in establishing that there is no other significant change):

- 251 1) softening of a suppository that is designed to melt at 37°C, if the melting point is  
252 clearly demonstrated; and
- 253 2) failure to meet acceptance criteria for dissolution for 12 units of a gelatin  
254 capsule or gel-coated tablet if the failure can be unequivocally attributed to  
255 cross-linking.

256 However, if phase separation of semisolid dosage forms at the accelerated condition  
257 occurs, testing at the intermediate condition should be performed.

### 258 259 2.4.2.1 No significant change at intermediate condition

260  
261 If there is no significant change at the intermediate condition, extrapolation beyond  
262 the period covered by available long-term data can be proposed; however, the extent  
263 of extrapolation would depend on whether long-term data for the attribute are  
264 amenable to statistical analysis.

265  
266 • *Data not amenable to statistical analysis*

267

268 When the long-term data for an attribute are not amenable to statistical analysis,  
269 the proposed retest period or shelf life can be up to 3 months beyond the period  
270 covered by available long-term data if backed by relevant supporting data.

272

273 • *Data amenable to statistical analysis*

274

275 When the long-term data for an attribute are amenable to statistical analysis but no  
276 analysis is performed, the extent of extrapolation would be the same as when data  
277 are not amenable to statistical analysis. However, if a statistical analysis is  
278 performed, the proposed retest period or shelf life can be up to one-and-half times,  
279 but not more than 6 months beyond, the period covered by available long-term data  
280 can be proposed when backed by the statistical analysis and relevant supporting  
281 data. shelf life

284

285 2.4.2.2 Significant change at intermediate condition

286

287 Where significant change occurs at the intermediate condition, the proposed retest  
288 period or shelf life should not exceed the period covered by available long-term data.  
289 In addition, a retest period or shelf life shorter than the period covered by available  
290 long-term data could be called for.

291

292 **2.5 Data Evaluation for Retest Period or Shelf life Estimation for Drug**  
293 **Substances or Products Intended for Storage Below Room Temperature**

295

296 **2.5.1 Drug substances or products intended for storage in a refrigerator**

297

298 Data from ~~drug substances or~~ products intended to be stored in a refrigerator  
299 should be assessed according to the same principles as described in Section 2.4 for  
300 drug substances or products intended for "room temperature" storage, except  
301 where explicitly noted in the section below. The decision tree in Appendix A can be  
302 used as an aid to the guidance below.

303

304 2.5.1.1 No significant change at accelerated condition

305

306 Where no significant change occurs at the accelerated condition, extrapolation of  
307 retest period or shelf life beyond the period covered by available long-term data can  
308 be proposed. The proposed retest period or shelf life can be up to one-and-a-half  
309 times, but not more than 6 months beyond, or up to 3 months beyond the period  
310 covered by available long-term data. The extent of extrapolation depends on the  
311 accelerated stability profile, the amenability of long-term data to statistical analysis,  
312 and the performance of such analysis.

313

314 2.5.1.2 Significant change at accelerated condition

315

316 If significant change occurs between 3 and 6 months' testing at the accelerated  
317 storage condition, the proposed retest period or shelf life should be based on data  
318 available at the long-term storage condition. Extrapolation is not appropriate. In  
319 addition, a retest period or shelf life shorter than the period covered by available  
320 long-term data could be called for. If the long-term data show variability, verification  
321 of the proposed retest period or shelf life by statistical analysis can be appropriate.

322  
323 If significant change occurs within the first 3 months' testing at the accelerated  
324 storage condition, the proposed retest period or shelf life should be based on data  
325 available at the long-term storage condition. Extrapolation is not appropriate. A  
326 retest period or shelf life shorter than the period covered by available long-term data  
327 could be called for. If the long-term data show variability, verification of the  
328 proposed retest period or shelf life by statistical analysis can be appropriate. In  
329 addition, a discussion should be provided to address the effect of short-term  
330 excursions outside the label storage condition (e.g., during shipping or handling).  
331 This discussion can be supported, if appropriate, by further testing on a single  
332 batch of the drug substance or product for a period shorter than 3 months.

334  
335

### 336 **2.5.2 Drug substances or products intended for storage in a freezer**

337  
338 For drug substances and products intended for storage in a freezer, the retest  
339 period or shelf life should be based on data obtained at the long-term storage  
340 condition. In the absence of an accelerated storage condition for drug substances or  
341 products intended to be stored in a freezer, testing on a single batch at an elevated  
342 temperature (e.g.,  $5^{\circ}\text{C} \pm 3^{\circ}\text{C}$  or  $25^{\circ}\text{C} \pm 2^{\circ}\text{C}$ ) for an appropriate time period should be  
343 conducted to address the effect of short-term excursions outside the proposed label  
344 storage condition (e.g., during shipping or handling).  
345

### 346 **2.5.3 Drug substances or products intended for storage below $-20^{\circ}\text{C}$**

347  
348 For drug substances and products intended for storage below  $-20^{\circ}\text{C}$ , the retest  
349 period or shelf life should be based on data obtained at the proposed long-term  
350 storage condition. and should be assessed on a case-by-case basis.  
351

## 352 **2.6 General Statistical Approaches**

353  
354 Where applicable, an appropriate statistical method should be employed to analyze  
355 the long-term primary stability data in an original application. The purpose of this  
356 analysis is to establish, with a high degree of confidence, a retest period or shelf life  
357 during which a quantitative attribute will remain within acceptance criteria for all  
358 future batches manufactured, packaged, and stored under similar circumstances.  
359 This same method should also be applied to commitment batches to verify or extend  
360 the originally approved retest period or shelf life.  
361

362 Regression analysis is considered an appropriate approach to evaluating the  
363 stability data for a quantitative attribute and establishing a retest period or shelf life.  
364 The nature of the relationship between an attribute and time will determine whether  
365 data should be transformed for linear regression analysis. The relationship can be  
366 represented by a linear or non-linear function on an arithmetic or logarithmic scale.  
367 In some cases, a non-linear regression can better reflect the true relationship.

369  
370 An appropriate approach to retest period or shelf life estimation is to analyze a  
371 quantitative attribute (e.g., assay, degradation products) by determining the earliest  
372 time at which the 95 percent confidence limit for the mean intersects the proposed  
373 acceptance criterion.  
374



375 For an attribute known to decrease with time, the lower one-sided 95 percent  
376 confidence limit should be compared to the acceptance criterion. For an attribute  
377 known to increase with time, the upper one sided 95 percent confidence limit  
378 should be compared to the acceptance criterion. For an attribute that can either  
379 increase or decrease, or whose direction of change is not known, two-sided 95  
380 percent confidence limits should be calculated and compared to the upper and  
381 lower acceptance criteria.

382

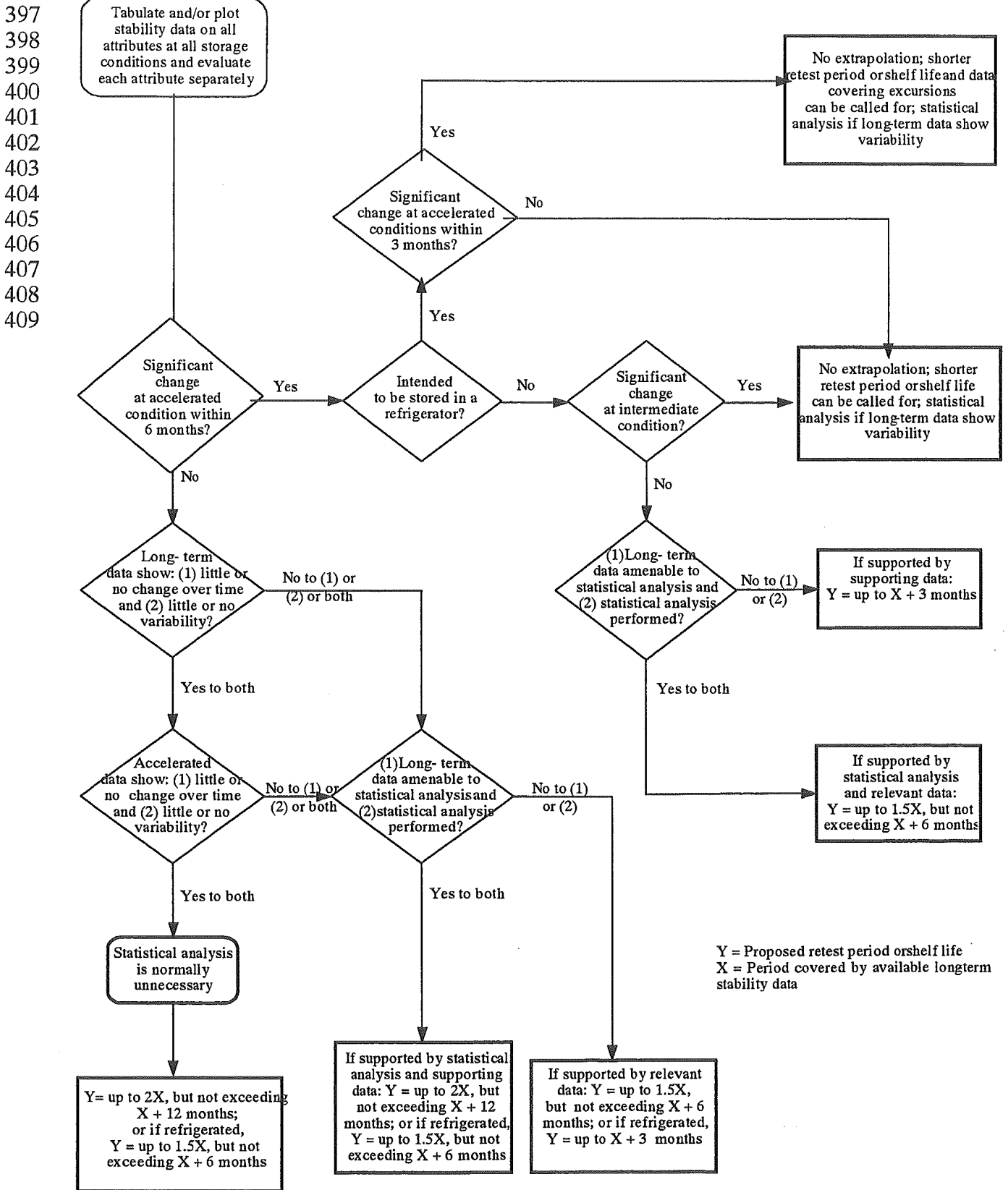
383 The statistical method used for data analysis should take into account the stability  
384 study design to provide a valid statistical inference for the estimated retest period or  
385 shelf life. The approach described above can be used to estimate the retest period or  
386 shelf life for a single batch or for multiple batches when the data are combined after  
387 an appropriate statistical test. Examples of statistical approaches to the analysis of  
388 stability data from single or multi-factor, full or reduced design studies are included  
389 in Appendix B. References to current literature sources can be found in Appendix  
390 B.6.

391

### 392 **3. APPENDICES**

393

393 **Appendix A: Decision Tree for Data Evaluation for Retest Period or Shelf life**  
 394 **Estimation for Drug Substances or Products (excluding frozen products)**  
 395  
 396



409 **Appendix B: Examples of Statistical Approaches to Stability Data Analysis**

410  
411 Linear regression, poolability tests, and statistical modeling, described below, are  
412 examples of statistical methods and procedures that can be used in the analysis of  
413 stability data for a quantitative attribute that are amenable to statistical analysis  
414 and for which there is a proposed acceptance criterion.  
415

416 **B.1. Data Analysis for a Single Batch**

417  
418 In certain cases, the relationship between an attribute and time is assumed to be  
419 linear.<sup>1</sup> Figure 1a shows the regression line for assay of a product with upper and  
420 lower acceptance criteria of 105 percent and 95 percent of label claim, respectively,  
421 with 12 months of long-term data. In this example, two-sided 95 percent  
422 confidence limits for the mean are applied because it is not known ahead of time  
423 whether the assay would increase or decrease with time, e.g., in the case of an  
424 aqueous formulation packaged in a semi-permeable container. The lower confidence  
425 limit intersects the lower acceptance criterion at 30 months, while the upper  
426 confidence limit does not intersect with the upper acceptance criterion until later.  
427 Therefore, a proposed shelf life of 24 months can be supported by the statistical  
428 analysis. provided the recommendations in Sections 2.4 and 2.5 are followed.

430  
431 When data for an attribute with only an upper or a lower acceptance criterion are  
432 analyzed, the corresponding one-sided 95 percent confidence limit for the mean is  
433 recommended. Figure 1b shows the regression line for a degradation product with  
434 12 months of long-term data, where the acceptance criterion is not more than 1.4  
435 percent. The upper one-sided 95 percent confidence limit for the mean intersects  
436 the acceptance criterion at 31 months. Therefore, a proposed retest period or shelf  
437 life of 24 months can be supported by statistical analysis of the degradation data  
438 provided the recommendations in Sections 2.4 and 2.5 are followed.

440  
441 If the above approach is used, the mean value of the quantitative attribute (e.g.,  
442 assay, degradation products) can be expected to remain within acceptance criteria  
443 through the end of the retest period or shelf life at a confidence level of 95 percent.  
444 If, however, the acceptance criterion for the quantitative attribute calls for  
445 individual values (e.g., dissolution for individual dosage units), confidence limits for  
446 the individual values should be used.  
447

448 The approach described above can be used to estimate the retest period or shelf life  
449 for a single batch, individual batches, or multiple batches when combined after  
450 appropriate statistical tests described in Sections B.2 through B.5.  
451

452 **B.2 Data Analysis for One-Factor, Full-Design Studies**

453  
454 For a drug substance or for a drug product available in a single strength and a  
455 single container size and/or fill, the retest period or shelf life is generally estimated  
456 based on the stability data from a minimum of three batches. When analyzing data  
457 from such one-factor, full-design studies, two statistical approaches can be  
458 considered. The objective of the first approach is to determine whether the data  
459 from all batches support the proposed shelf life,shelf life. The objective of the  
460 second approach, testing for poolability, is to determine whether the data from  
461 different batches can be combined for an overall estimate of a single shelf life.shelf  
462 life.

463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516

## **B.2.1 Evaluating whether all batches support proposed retest period or shelf life**

The objective of this approach is to evaluate whether some of the batches have retest periods or shelf lives shorter than the one proposed. Retest periods or shelf lives for individual batches should first be estimated using the procedure described in Section B.1 with individual intercepts, individual slopes, and the pooled mean square error calculated from all batches. If each batch has an estimated retest period or shelf life longer than that proposed, the proposed retest period or shelf life will generally be considered appropriate, as long as the guidance for extrapolation in Section 2.4-2.5 is followed. There is generally no need to perform poolability tests or identify the most reduced model. If, however, one or more of the estimated retest periods or shelf lives are shorter than that proposed, poolability tests can be performed to determine whether the batches can be combined to estimate a longer retest period or shelf life.

Alternatively, this approach can be taken during the pooling process described in the following Section B.2.2.2. If the regression lines for the batches are found to have a common slope and the estimated shelf lives based on the common slope and individual intercepts are all longer than the proposed shelf life, there is generally no need to continue to test the intercepts for poolability.

## **B.2.2 Testing for poolability of batches**

### **B.2.2.1 Analysis of covariance**

Before pooling the data from several batches to estimate a retest period or shelf life, a preliminary statistical test should be performed to determine whether the regression lines from different batches have a common slope and a common time-zero intercept. Analysis of covariance (ANCOVA) can be employed, where time is considered the covariate, to test the differences in slopes and intercepts of the regression lines among batches. Each of these tests should be conducted using a significance level of 0.25 to compensate for the expected low power of the design due to the relatively limited sample size in a typical formal stability study.

If the test rejects the hypothesis of equality of slopes (i.e., there is a significant difference in slopes among batches), it is considered inappropriate to combine the data from all batches. The retest periods or shelf lives for individual batches in the stability study can then be estimated by applying the approach as described in Section B.1 using individual intercepts and individual slopes and the pooled mean square error calculated from all batches. The shortest estimate among the batches should be chosen as the retest period or shelf life for all batches.

If the test rejects the hypothesis of equality of intercepts but fails to reject that the slopes are equal (i.e., there is a significant difference in intercepts but no significant difference in slopes among the batches), the data can be combined for the purpose of estimating the common slope. The retest periods or shelf lives for individual batches in the stability study should then be estimated by applying the approach as described in Section B.1, using the common slope and individual intercepts. The shortest estimate among the batches should be chosen as the retest period or shelf life for all batches.

If the tests for equality of slopes and equality of intercepts do not result in rejection at a level of significance of 0.25 (i.e., there is no significant difference in slopes and