平行線検定を利用した薬物の効力比較

原田 淳

要約：2種類の薬物の効力を比較する場合、平行線検定を用いて効力比を求めることがよく行われている。薬物間の効力比をその信頼区間とともに求めれば、その効果の違いを具体的な数値パラメータとして扱える上に、効力比の信頼区間が1を含まないならば、効果の有意な違いを統計学的に示すことができる。また、適当な同等性マージンを設定すれば2種の薬物の効力の同等性を示すこともできる。従来、平行線検定は直線をあてはめるものが主流であったが、近年ではシグモイド曲線を用いた非線形回帰分析での平行線検定が可能な統計解析ソフトウェアが利用可能で、シグモイド曲線にしたがうことが多い薬理学での利便性が高まった。一方、本手法の前提となる2つの用量反応直(曲)線の平行性に関しては、2つの直線の傾きの違いをみるF検定で有意差が無いことで検討してきたが、米国薬局方の生物学的アッセイ法の新しいガイドラインではそれに代わって、傾きの差や比の信頼区間を用いる方法が提唱されている。本稿ではこれらの新しい流れを紹介し、実地で活用する際の注意点に関して考察する。

1. 平行線検定・効力比算出について

ある薬物の薬理効果を調べる際に、別の薬物(標準薬など)とその効力を比較したい、ということは研究を進める上でよくあることである。単独の薬物の効果を調べる場合であれば、溶媒投与群を対照として、複数用量を用いるならばDunnett検定等の適当な検定手法によって群間比較の有意差を示すことができ、また用量反応曲線に特定のモデルが仮定できるならば、ED₅₀のような薬効パラメータを算出すれば、その効果の程度を定量的に表すことができる(図1)。しかしながら、2つの薬物についてそれぞれの用量反応曲線

図1 事例データ1

各群の平均値と標準誤差を示す (n=8). **はDunnett検定P値<0.01を示す。ED₅₀値とその95%信頼区間は、対照薬で2.73 mg/kg (1.68~4.42)、被験薬で1.04 mg/kg (0.69~1.57)であった。

が得られたときに、溶媒投与群との有意差検定の結果をもって2つの薬物の効力を直接比較することはできない。ED₅₀値の比較は有用ではあるが、その効力の差を得られたデータのばらつきを加味して評価するのは難しい。

平行線検定は、そのような場合に有用な解析手法である(1)。その原理自体は単純であり、2つの薬物の用量反応モデルを直線で近似すると、2つの平行な直線をあてはめ、そのx軸(用量軸)側の距離を測れば、それが2つの薬物の効力の違いを示すことになる(図2)。用量軸には対数用量が用いられることが多く、その場合効力の違いは比(効力比)で表される。従って、効力比が1であるということは2つの薬物が同じ効力を示すということであり、効力比とその95%信頼区間が1を含まなければ、2つの薬物の効力が有意に異なる結論できる。

これを利用して、逆に2つの薬物作用の同等性を調

キーワード：用量反応、平行線検定、効力比、シグモイド、非線形回帰

第一三共株式会社 データサイエンス部 (〒140-8710 品川区広町 1-2-58)

E-mail: harada.jun.ys@daiichisankyo.co.jp 原稿受領日：2017年4月6日、依頼原稿

Title: Estimation of relative potency using parallel line analysis Author: Jun Harada

図2 平行線検定のイメージ

べることもできる。まず効力比1を中心として作用が同等であるとみなせる範囲（同等性マージン）を設定し、効力比が信頼区間も含めて同等性マージン内に収まれば、2つの薬物活性の同等性を積極的に示すことができる。

2. 平行線検定の前提条件

直線モデルによる平行線検定が成り立つ前提条件は、2つの薬物の用量反応が直線で近似でき、なおかつ2つの直線が平行とみなせる、ということである。直線近似に関しては、例えば用量反応がシグモイド曲線を示す場合でも、そのED₅₀を中心とする反応域では直線で近似できることが多い。反応値を0から100の範囲で示すならば、例えば20から80の範囲を直線で近似する、といった具合に使われる。

一方、平行性が成り立っているかどうかは、どのように検討したらよいだろうか。通常行われているのは、F検定によって傾きの有意差を見る方法である。2つの直線は、それぞれ別々に最小2乗法で直線回帰することができ、それにより回帰式から予測される値と実際の値の差の2乗の和（残差平方和）が計算できる。最小2乗法では、この残差平方和を最小にする回帰式を求めている。ここで、2つの直線に個々の傾きでなく共通の傾きを設定してその場合の残差平方和を計算する。個々の傾きと共通の傾きそれぞれの場合の残差平方和を比較（F検定）することにより、帰無仮説を「2つの直線の傾きは等しい」とする検定が行える。これが有意でなければ、傾きに有意差が無い、すなわち平行性が否定できなかったということになり、消極的な判断（平行性を直接示したわけでない）ではあるが、平行線検定の前提条件が成立したとみなされる。

図3 事例データ1を用いた平行線検定（直線のあてはめ）

対照薬のデータを●で、被験薬のデータを□で示す。また、それぞれ個別に最小2乗法を適用した傾きを実線で、共通に適用した傾きを点線で示す。回帰式はそれぞれ上が対照薬で下が被験薬のもの。

3. 直線回帰による平行線検定の例

実際に、架空データを用いて直線あてはめによる平行線検定を実施してみる。図3は、図1で示した架空データ（事例1）の用量反応を直線で近似したものである。図の左側が被験薬、右側が対照薬の用量－反応直線を示し、別々に回帰させたときの回帰式と、共通の傾きを設定した際の回帰式が右下に示されている。それぞれの回帰式から残差平方和を計算し、F値を算出すると3.663と計算され、検定P値は0.0597となる。すなわち検定結果は有意でなく、傾きの平行性が否定されないため平行線検定を適用する。その結果、得られた対照薬に対する被験薬の効力比は3.92であり、またその95%信頼区間は3.05～5.03であった。このことから、被験薬の効力は対照薬よりも3.92倍強い、別の言い方をすると、3.92分の1の用量で同じ効果を発揮できるということになる。また信頼区間が1を含まないことから、水準5%で有意に強い効果があることが示される。

4. シグモイド関数を用いた平行線検定について

近年、バイオ医薬品の評価機会が増え、またその活性の評価やロット間の同等性を示したりするため、効力の比較が必要になる機会も多い。それらのデータはELISAのように典型的なシグモイド曲線を示し、直線の近似では正確な比較が難しい場合も多い。一方で、多数の濃度点を設定できることから精度のよいシグモイド曲線のデータが取得できる。このようなデータの解析には、シグモイド曲線を想定した非線形回帰分析の適用が望ましい(2)。理論的には、平行線検定は直

線モデルを拡張し、シグモイド曲線の非線形回帰分析でも実施可能であり(1,3), 実際に近年は複数の統計解析ソフトウェアで、非線形回帰による平行線検定が可能になってきた。著者の知る範囲でも、EXSUS (CAC クロア), JMP (SAS Institute), SoftMax Pro (Molecular Devices), Prism (GraphPad), Gen5 (BioTek), PLA (Stegmann Systems) 等の市販ソフトウェアが挙げられる。また Excel (Microsoft) でも、ソルバーを使うことにより算出可能である。

シグモイド曲線は、例えばそれが 4-parameter logistic model で表される場合、その形状は 4 つのパラメータ値が決まれば自動的に決まる。すなわち、垂直方向での下側・上側の漸近線での反応値 (それぞれ Y_{min} , Y_{max} と表す) とシグモイド関数の傾きを表す Hill 係数、そして ED_{50} の濃度値である。回帰式がシグモイド関数であっても、傾きのパラメータ値を共通に設定することで平行線検定が可能になる。傾きの共通性も、直線回帰と同様に回帰式と実測値の残差平方和の違いから判断できる。

先の事例 1 のデータをシグモイド関数の平行線検定で解析するとどうなる結果になるだろうか。実際に平行線検定を適用する際には、シグモイド曲線特有の問題を考慮する必要がある。直線回帰の場合は直線を示すパラメータは傾きと切片の 2 つしかなく、2 つ

の直線の傾きを同じと仮定すれば自動的に切片のずれ、すなわち効力比が算出される。一方、シグモイド曲線のパラメータは前述のように 4 つ存在する。従って、さらに 2 つのパラメータ、下側・上側の漸近線が存在し、これら 2 つのパラメータを 2 つのシグモイド関数で個別に設定するか、共通で設定するかで、算出される効力比の値は異なってくる。

具体的に事例 1 のデータで解析を行ってみよう。下側・上側の漸近線 (Y_{min} , Y_{max}) と傾きの仮定に関して、ここでは 4 通りのパラメータ設定を行っている。 Y_{min} に関しては薬物なしの状況での応答であるから同一値のはずであり、今回は共通の値として設定する。一方、 Y_{max} に関しては、個別に設定したのがパターン A, B であり、共通に設定したのがパターン C, D である (図 4)。また平行線検定の定義上、傾きは共通で設定する (パターン B, D) わけだが、ここでは参考のため、傾きを個別に設定したパターン A, C の結果も表示している。また、ここでは ED_{50} の定義として、 $(Y_{min} + Y_{max})/2$ の反応値を得る薬物の用量として計算している。従って、2 つのシグモイド曲線の Y_{max} を個別に設定すると、その反応値も異なってくる。この問題は直線回帰の平行線検定では起こらないので、注意すべき点である。得られた効力比は、 Y_{min} , Y_{max} , 傾きのパラメータ値や ED_{50} を得る 50% 反応値

図 4 事例データ 1 を用いた平行線検定 (シグモイドカーブを用いた非線形回帰)

Y_{max} の設定は A, B では個別に、C, D では共通に設定している。 Y_{min} は共通に設定。また、傾きの設定は A, C では個別に、B, D では共通に設定している。

表1 事例データ1のシグモイド関数を用いた平行線検定の解析結果

	設定		対照薬					被験薬					平行性	効力比		
	Ymax	傾き	Ymin	Ymax	50%	傾き	ED ₅₀	Ymin	Ymax	50%	傾き	ED ₅₀		効力比	95%L	95%U
A	個別	個別	204	1521	863	1.11	2.86	204	1721	963	1.12	1.08	0.97	2.66	1.53	4.62
B	個別	共通	204	1519	862	1.11	2.85	204	1723	964	1.11	1.08	0.97	2.64	1.77	3.92
C	共通	個別	170	1681	926	0.88	3.68	170	1681	926	1.11	0.95	0.06	3.87	2.96	5.05
D	共通	共通	175	1663	929	1.02	3.55	175	1663	929	1.02	0.92	0.06	3.85	2.95	5.01

50%はED₅₀と効力比の基準となる反応値、平行性は傾きの平行性のF検定のP値。効力比の95%L、95%Uはそれぞれ下側、上側の95%信頼区間を表す。

表2 事例データ1のシグモイド関数を用いた平行線検定の解析結果

	設定		ED ₂₅			ED ₅₀			ED ₇₅		
	Ymax	傾き	対照薬	被験薬	効力比	対照薬	被験薬	効力比	対照薬	被験薬	効力比
A	個別	個別	1.06	0.40	2.63	2.86	1.08	2.66	7.72	2.88	2.68
B	個別	共通	1.06	0.40	2.64	2.85	1.08	2.64	7.65	2.90	2.64
C	共通	個別	1.06	0.36	2.98	3.68	0.95	3.87	12.8	2.55	5.02
D	共通	共通	1.20	0.31	3.85	3.55	0.92	3.85	10.5	2.73	3.85

50%に加えて25%、75%の反応値を基準とする効力比を示す。

とともに表1に示した。さらに比較のためにED₅₀に加えてED₂₅、ED₇₅の推定値とそれぞれにおける効力比を表2に示した。傾きを共通でなく個別に設定した場合は、反応値の取りようによって効力比が変わってくることに注意が必要である。

シグモイド曲線の平行性が成り立っている場合、2つの曲線の傾きに加えてYmin、Ymaxも共通とみなせる。従ってこれら3つのパラメータを共通の値に設定したパターンDの結果が、求める効力比となる。その値は3.85であり、95%信頼区間は2.95~5.01であった。この効力比の値は、傾きを個別に設定したパターンCでも3.87と、大きくは変わっていない。しかし、これは薬物作用の50%値を基準に効力比を算出した場合であって、薬物作用の25%や75%を反応値とした場合には、傾きを共通に設定したパターンDでは、効力比が変わらないのに対して、傾きを個別に設定したパターンCでは効力比が大きく変わってくる。これは図4を見て分かるように、傾きを個別に設定すると反応値の高低によって2つのシグモイド曲線間の左右(用量軸)の距離が異なってくるためである。

一方、Ymaxを個別に設定したパターンA、Bはどうであろうか。パターンDと比較して、ED₅₀による効力比の値は、2.66、2.64といずれも低め、すなわち差が小さい結果となっている。ここで注目すべきなのはYmaxの推定値である。2つのパラメータをどちらも個別に設定したパターンAの場合で、Ymaxの値は1521、1721であり、被験薬の方がやや高い値となっ

ている。そのためED₅₀の反応値も対照薬と被験薬でやや異なっており、それが効力比の値に影響している。一方、傾きを個別に設定したAでは薬物作用の25%や75%を反応値とした場合でも効力比はそれほど変わらないが、これは2つの曲線の傾きの差が小さいためである。

上記のように、シグモイド関数の平行線検定の場合はパラメータの設定方法によって求められる効力比の値が異なっているので注意が必要である。特にYmin、Ymaxの設定は、薬理効果のpotencyとefficacyの問題と関わってくる。直線回帰の平行線検定では、2つの直線の平行性を前提にすれば、後はpotencyだけを考慮すればよかった。そのため、そこから得られる効力比も、potency ratioで表された。一方、シグモイド曲線の場合にはefficacyの違いもパラメータ値から検出されてくるので、それを考慮に入れなければならない。

生物学的同等性試験や薬物ロット間のコンパラビリティ試験のような場合には、2つの薬物はpotencyに加えてefficacyも同等であるべきであり、Ymin、Ymaxは共通(パターンD)で設定すべきであろう。しかし、比較する2つの薬物が異なる場合には、本来比較したいpotencyに加えてefficacyの違いも考慮する必要があるかもしれない。今回紹介した事例1データも、Ymaxの推定値が約10%ほど異なっているように見える。このデータでYmaxを共通で設定した場合、傾きの平行性の検定P値は0.06と5%水準でぎりぎり有意になっていないが、これ以上のefficacyの開きがあった場合には、有意差がついて傾きの平行性の前提が成

り立たない。傾きの平行性の検定が有意であった場合には、個々の傾きを用いた効力比が採用される（繰り返しになるが、その場合には本来の意味での効力比算出の前提条件が成り立っていない）。その場合でも ED_{50} を基準とした効力比は、傾きを共通にした場合とそれほど大きな違いはない。しかしながら、それ以外の反応値（ ED_{25} , ED_{75} など）では効力比が異なってくることに注意して、総合的に結果を解釈する必要がある。Ymin, Ymax を個別に設定した場合の解釈はさらに注意を要する。Efficacy が異なる薬物間の比較でシグモイド曲線のあてはまりを良くするには、こちらを選択せざるを得ない場合も考えられるが、その場合は効力比の基準となる反応値のずれに注意する必要があるし、そもそも効力比を算出する目的・意味についてもよく検討する必要があると思われる。一方、阻害率のように 0 と 100% の間で表せるデータであれば、Ymin, Ymax を 0, 100% で常に固定できるため扱いやすい。ただし厳密にはこの場合、基準となる control 値などの観測値も誤差を含むため、それをモデルに組み込む考えもある (1, 2)。

5. 直線（曲線）の平行性をどう確認するか

先述のように、直線や曲線の傾きの平行性については、傾きの差の F 検定で有意でないことをもって判断することが従来行われてきた。この判断基準は、いわゆる「有意差無し同等」と同じであり、示そうとしている仮説（2本の線が平行）が帰無仮説になっているという点で、検定で有意でない場合でも、積極的に平行性を示すことはできない。この方法ではデータのばらつきが小さい時には検出力が高くなり、僅かな傾きの違いで有意となり平行でないと判断される一方、ばらつきが大きいときには傾きの差が大きくても検出されにくい、という問題がある（図 5）。この点に関して、

有意差検定ではなく信頼区間を用いる方法 (4, 5) が提唱され、最近発表された米国薬局方の生物アッセイの General chapter の改訂版（USP chapter 1032, 1034）において採用されている (6-9)。すなわち、2つの線の傾きの差や比に関して、それが平行と判断してよい範囲（goal post）を設定し、計算された差や比がその 90% 信頼区間まで含めて goal post に収まっていれば、平行とみなす、というものである（図 6）。これは後発医薬品の生物学的同等性試験に用いられているものと (10) 同じ概念であり、2つの線の平行性を積極的に示すことができる。Ymin, Ymax についても、傾きと同様に、goal post を設定してパラメータ値の同等性を示していくことができる。PLA のように、平行線検定に特化し、USP chapter 1032 準拠を謳っているソフトもある。

実際に、架空データを用いて直線の傾きやばらつきの異なるデータで平行線検定を実施してみる。図 7 は、架空データ（事例 2）の用量反応を直線で近似したも

図 6 goal post（傾きの比の信頼区間）に基づいた平行性の判定
傾きの比が 1 であれば完全に平行である。ここでは goal post を仮に 0.77~1.3 と設定した。算出された傾きの比がその信頼区間（図の横のバー）を含めて goal post に含まれれば平行と判定する。この場合は例 1, 2 が平行とされる。2 のように傾きに有意差が付いてもよい。

図 5 仮説検定（F 検定）による平行性判定の問題点

A ではばらつきが少ないため、僅かな傾きの違いも非平行と判定される一方、B ではばらつきが大きいため、明らかな非平行でも帰無仮説を棄却できず、非平行と判定されない。

図7 事例データ2を用いた平行線検定（直線のあてはめ）：平行性の判定

対照薬のデータを●で、被験薬のデータを□で示す。A～Dはそれぞれデータのばらつき具合や傾きの違いが異なり、それぞれ個別に最小2乗法を適用した傾きを実線で示している。

表3 事例データ2を用いた直線の平行性の解析結果

	データの特徴		平行性		効力比			傾きの比と信頼区間		
	平行性	ばらつき	P値	判定	効力比	95%L	95%U	比	90%L	90%U
A	高い	大きい	0.665	平行	3.53	1.97	6.82	1.084	0.791	1.499
B	高い	小さい	0.027	非平行	3.53	3.03	4.11	1.084	1.021	1.151
C	低い	大きい	0.158	平行	3.17	1.84	5.80	1.278	0.959	1.740
D	低い	小さい	< 0.001	非平行	3.17	2.65	3.81	1.278	1.210	1.351

平行性は傾きの平行性のF検定のP値。傾きの比の90%L、90%UはそれぞれUSP chapter 1032に基づいて算出した下側、上側の90%信頼区間を表す。

のである。図の左側が被験薬、右側が対照薬の用量－反応直線を示し、4つのデータ（A～D）で対照薬の傾きは等しい。一方で、A、Bは平行性が比較的高いデータであり、またB、Dがばらつきの比較的小さいデータとなっている。表3には、それぞれのデータを用いた傾きの平行性の検定（F検定）結果と求められた効力比、および傾きの比の点推定値と90%信頼区間を示している。F検定を用いた場合、平行性が高くばらつきも小さいパターンBの場合に有意となって非平行と判定されてしまっている。一方で平行性が低い上にばらつきが大きいパターンCでは有意にならず非平行とは判定できない。これらのデータで傾きの比とその90%信頼区間を算出すると、パターンBでは小さい幅に収まっていることが分かる。なお、これらパターンA～Dは、図6の3、2、5、4にそれぞれ対応している。また、この方法であれば、パターンBのように平行性が高いのにばらつきが小さすぎて有意差がついて、F検定で平行とみなされないケースでも、goal postに収まれば平行とみなせるわけである。

以上のような信頼区間を用いた判定法は、論理的には大変優れているが、実際に適用する上での問題点は

goal postをどう設定するかが難しいことである。後発医薬品の生物学的同等性試験では薬物動態パラメータの比に関して同等性マージン〔80～125%〕が決められているが、バイオアッセイや薬理データの場合には、同等とみなせる区間は対象となる生物活性やアッセイ系によって個別に設定する必要がある。USP chapter 1032でも基準は一律には決められないとした上で、1つの考え方として〔30%以下の傾きの差や5%以下の漸近線の差〕はあまり意味のある差とはみなされないが、〔50%以上の傾きの差や15%以上の漸近線の差〕を許容するには相当な注意が必要であろうと述べている。また、USPではgoal postの設定に際して、蓄積された背景データ（historical data）の活用を勧めていることでも分かるように、ルーチンで実施される品質試験の精度管理を主に想定しており、薬理データの解析においてはその違いも考慮する必要があると考えられる。シミュレーションを用いたF検定と信頼区間法の比較も報告されている(11)が、データの条件設定（傾きの差やばらつき）によって両者の優劣は異なり、どちらが優れているとは一概に言えないようである。

6. おわりに

以上のように、平行線検定は薬物間の効力比較を行ううえで有用な解析手法であり、また群間の有意差検定と異なり、効力の違いを何倍という具体的な定量値としてその信頼区間とともに算出できる。そしてその適用は直線のあてはめに留まらず、シグモイド型のデータの非線形回帰にも拡張可能である。ただし平行性検定の前提となる平行性の示し方はシグモイド型のデータの場合、推定パラメータが多いため複雑である。本稿で示した事例1のデータは敢えてYmaxを10%ほど変化させ、ぎりぎり適用可能だが解釈の難しい結果となっている。適当な解析手法を用いることで、そのデータの意味の解釈が容易になるが、解析モデルに組み込むことで、データの持っている情報を一部失うこともあるし、そのモデルへの適用の妥当性が問題となるケースもある。統計解析ソフトウェアの普及で計算自体は容易であるが、データの性質を十分吟味し、不適切な解析によって、データの意味を見誤ることのないように注意が必要である。平行線検定の解析理論と手順については成書(12)に詳細な解説があり、Excel SolverやJMPを用いた計算演習の例題も載っているので、理解を深めたい方は是非参照されたい。

著者の利益相反：原田 淳（第一三共株式会社）。

文 献

- 1) 福島慎二. 多剤用量反応実験の新たな統計解析—薬理実験における効力比の直接算出—. 第2期医薬安全性研究会 第8回. 2011. https://biostat.jp/archive_teireikai_2.php
- 2) 崑山陽二郎, 他. 日薬理誌. 2008;132:199-206.
- 3) Vølund A. Biometrics. 1978;34:357-365.
- 4) Callahan JD, et al. Testing the null hypothesis for a specified difference — The right way to test for parallelism. Bioprocess J. 2003;2:71-77.
- 5) Jonkman JN, et al. J Biopharm Stat. 2009;19:818-837.
- 6) The United States Pharmacopeial Convention. USP <1032> Design and development of biological assays. 2013.
- 7) The United States Pharmacopeial Convention. USP <1034> Analysis of biological assays. 2013.
- 8) Singer R. Statistical analysis using parallel lines or curves, CaSSS Bioassays 2009. https://c.ymcdn.com/sites/casss.site-ym.com/resource/resmgr/Bioassays_Speaker_Slides/2009_BIOA_SingerRobert.pdf
- 9) Bortolotto E, et al. Assessing similarity with parallel-line and parallel-curve models: implementing the USP development/validation approach to a relative potency assay. Bioprocess Int. 2015;13(6).
- 10) 後発医薬品の生物学的同等性試験ガイドラインの一部改正について. 2012. <https://www.pmda.go.jp/files/000160026.pdf>
- 11) Fleetwood K, et al. PDA J. Pharm Sci Technol. 2015;69:248-263.
- 12) 芳賀敏郎. 医薬品開発のための統計解析 第3部 非線形モデル. サイエンティスト社; 2010.

[事例1, 2のデータはJ-STAGE (<http://www.jstage.jst.go.jp/browse/fpj/-char/ja>) の電子付録に掲載しています.]

著者プロフィール

原田 淳 (はらだ じゅん)

第一三共株式会社 データサイエンス部 統計解析グループ, 主査, 博士 (薬学)

1988年京都大学薬学部卒業, '90年京都大学大学院薬学研究科修士課程修了 (薬理学). 同年, 三共株式会社 (現第一三共) 入社, 神経系作用薬の探索研究に従事. '99年ハーバード大学・マサチューセッツ総合病院, 博士研究員 (Michael A. Moskowitz 教授), '01年に帰国後, 糖尿病治療薬の探索研究に従事, '12年より現職.