

打ち切りを含む多因子疼痛試験データの解析

嵯山陽二郎,^{*a} 小宮山 靖,^b 塚田秀夫^c

Analysis of Censored Data in Multi-factorial Analgesic Test

Yojiro SAKIYAMA,^{*a} Osamu KOMIYAMA,^b and Hideo TSUKADA^c

^aResearch Planning and Coordination, Nagoya Laboratories, Pfizer Global Research and Development, Pfizer Inc., 5-2 Taketoyo-cho, Chita-gun, Aichi 470-2393, Japan, ^bStatistics and Clinical Programming, Pfizer Global Research and Development, Tokyo Laboratories, Pfizer Japan Inc., 3-22-7, Yoyogi, Shibuya-ku, Tokyo 151-8589, Japan, and ^cPET Center, Central Research Laboratory, Hamamatsu Photonics K. K., 5000 Hirakuchi, Hamamatsu City 434-8601, Japan

(Received May 23, 2007; Accepted September 3, 2007)

Animal pain testing is essential for the development of new analgesic drugs, where appropriate data analyses as well as appropriate multi-factorial design of experiments are necessary to obtain meaningful results in an efficient fashion. The tail withdrawal experiment is one of the pain tests in which a rhesus monkey is restrained in a chair from which its tail hangs free by so it can be immersed in warm water. The monkeys consistently kept their tails in 38–40°C water for an extended period of time, and thus, the data were censored at 120 sec. The effect of temperature on the tail withdrawal latency was evaluated using three monkeys with a randomized block design. The effect of morphine on the thermal sensitivity was also evaluated. A Friedman-type two-way analysis of variance (Mack-Skillings test) demonstrated that the effects of both temperature and the animals were significant. The effect of repeated measurement in one animal was not significant using the Friedman test, indicating that the significance of the effect of animals could be attributed to the difference in the intrinsic thermal sensitivity between animals. This method, together with a graphical approach, may prove to be valuable for assessing the sensitivity and reproducibility of an experimental condition, as well as the pharmacological effects of analgesic drugs.

Key words—pain; analgesic; monkey; Friedman test; Mack-Skillings test

緒 言

優れた鎮痛薬の新規開発を行うためには、非臨床疼痛試験における適切な実験計画とデータ解析が必要不可欠である。非臨床疼痛試験の1つに、TWL (tail withdrawal latency: 尾退避反応潜時) 測定試験がある。TWL 試験は最初 Janssen らによりラットで考案され、¹⁾以降鎮痛の評価モデルとしてラット、サル²⁻⁵⁾などに広く利用されている。TWL とは、動物の尾を温浴に浸してから尾を温浴から引き上げるまでの時間で、侵害性疼痛に対する感受性の指標である。鎮痛薬により侵害性疼痛のしきい値が上昇し TWL は延長する。この延長は薬効のよい指

標となるが、これを得るためには実験系が高感度でかつ安定である必要がある。すなわち、温熱刺激に対する反応が十分に検出でき、かつ繰り返し測定や個体の変更に対する再現性が高くなければならない。そこで、まず薬剤を投与しない状態で、温度と個体の2つの因子の効果を事前に明らかにする必要がある。一方、今回測定した TWL 値の特徴としては、温度が高い場合は直ちに尾を温浴から引き上げるので数秒となるが、温度が体温付近の場合は温浴に際限なく尾を浸し続けるため、上限を持った(タイプ I の打ち切りのある)測定値となる。このように打ち切りを含む応答変数に対しては、正規性を想定した二元配置分散分析はかならずしも適切ではなく、Friedman 型のノンパラメトリック二元配置分散分析 (Mack-Skillings 検定) の方が望ましい。しかし、その適用事例についてはほとんど知られていない。本稿では、2つの因子を有する非臨床疼痛試

^aファイザー(株)中央研究所研究企画管理室, ^bファイザー(株)統計・解析部統計コンサルティング室, ^c浜松ホトニクス(株)中央研究所 PET センター

*e-mail: Yojiro.Sakiyama@pfizer.com.

験を事例として、Mack-Skillings 検定の適用方法について紹介する。

方 法

本実験は、浜松ホトニクス中央研究所動物実験倫理委員会の承認を得て実施した。また、サル飼育及び実験に関しては US National Institute of Health の推奨基準及び浜松ホトニクス中央研究所のガイドラインに準拠して実施した。

1. 材料 動物は浜松ホトニクス中央研究所が所有する雄性アカゲザル (*Macaca mulatta*) 3 匹 (体重 4.4–8.3 kg)、週 2 回 3 ヶ月以上のチェアトレーニングを施されたものを覚醒状態で使用した。覚醒状態におけるサルの使用に関しては塚田らの方法に従った。^{6,7)} サルの尾は先端から 10 cm 以上剃毛して、モンキーチェアに固定した。

薬剤は塩酸モルヒネ (morphine hydrochloride, 10 mg/ml, 塩野義) を用いた。

2. 尾退避潜時 (tail withdrawal latency) の測定 実験の概観を Fig. 1 に示す。一定温度に保った水道水の入った容器を用意し、チェアに覚醒状態のまま固定されたサルの尾の先端約 10 cm をその中へ浸し、直ちに手を離してから尾が自発的に退避するまでの時間を TWL とした。水道水の温度は 37 °C, 39 °C, 41 °C, 43 °C, 45 °C, 47 °C, 50 °C 及び 55 °C の 8 通りをそれぞれ 1 回ずつ実施しこれを 1 ブロックとし、個体当たり 4 ブロック行った。ブロック内では実験は無作為化した、すなわち乱塊法 (randomized block design) として実施した。測定の間

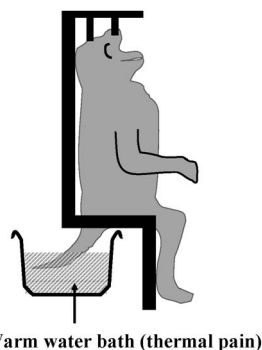


Fig. 1 Schematic Representation of a Tail Withdrawal Experiment in which a Monkey Sit on a Dedicated Monkey Chair

The tail was immersed in a warm water bath for thermal pain stimulation, followed by the measurement of tail withdrawal latency. The temperature was preliminarily adjusted with various ranges.

隔は 10 分空け、合計で $8 \times 4 = 32$ 回の刺激を 1 個体当たりに行った。実験は午前 4 時間、午後 4 時間、合計 8 時間実施しその日のうちに実験を終了した。別の日に、別の個体を同じ要領で実施した。これを 3 個体について実施した。打ち切りの cut off 値は 120 秒とした。

3. モルヒネによる鎮痛効果の検討 上記で用いた 3 個体につき、モルヒネによる鎮痛効果の検討を行った。個体は覚醒状態のままチェアに固定し、橈側皮静脈にカテーテルを留置した。塩酸モルヒネ 1 mg/kg を留置カテーテルから静脈内投与した。刺激温度は先の疼痛試験の結果より、尾を引く反応が十分速くかつ動物に過度の負荷を与えない温度として 47 °C が適切であると判断しこれを適用した。TWL の測定はモルヒネ投与の 30 分前、0 分 (薬剤投与直前)、30 分後、60 分後、120 分後、240 分後の 6 点で行った。

4. 繰り返しを含む 2 元配置 Friedman 型順位検定 (Mack-Skillings 検定) 本試験における応答変数である TWL には打ち切りがあるため、分布の正規性を想定することはかならずしも適切でない。このような場合、因子の効果を検定するにはノンパラメトリック法の方が望ましい。さらに本実験のように温度、個体と因子が 2 つありまた繰り返しを持つ場合は、以下に述べるような Mack と Skillings らにより考案された繰り返し有りの 2 元配置 Friedman 型順位検定 (Mack-Skillings 検定) を行えばよい。⁸⁾ まず、「 H_0 : 処置による変化がない」を帰無仮説とする。処置とはここではいろいろな温度の刺激を与えたり、個体を変えたりするなどの一連の実験処置に相当する。観測値を n 個のブロックに分け、ブロック毎に順位変換を行う。各ブロックには kc 個 (k は処置数, c は繰り返し数) の観測値がある。観測値の総数は $N = nkc$ 個となる。いま、 i 番目のブロック ($i = 1, \dots, n$) における j 番目の処置 ($j = 1, \dots, k$) において q 回の繰り返し ($q = 1, \dots, c$) を行ったとする。 r_{ijq} を X_{ijq} のブロック内順位とすると、 j 番目の処置における平均順位 S_j は以下の式で表される。

$$S_j = \sum_{i=1}^n \left[\sum_{q=1}^c r_{ijq} / c \right], \text{ for } j = 1, \dots, k. \quad (1)$$

ついで、Mack-Skillings (MS) 統計量は以下の式で与えられる。

$$MS = \left[\frac{12}{k(N+n)} \right] \sum_{j=1}^k \left[S_j - \frac{N+n}{2} \right]^2$$

$$= \left[\frac{12}{k(N+n)} \right] \left\{ \sum_{j=1}^k S_j^2 \right\} - 3(N+n) \quad (2)$$

別途で、 k を処置数、 c を繰り返し数、 n をブロック数、 α を有意水準としたときの棄却限界値を付表^{8,9)}で求めこれを ms_α としたとき、 $MS \geq ms_\alpha$ のとき H_0 は棄却され、処置による変化が有意となる。なお、 $c=1$ とした繰り返しのない場合が Friedman 検定に相当する。言い換えれば Friedman 検定は広義の Mack-Skillings 検定である。

ここでは、個体と温度を 2 因子とした効果の検定を行った。「 H_{0a} : 個体による変化がない」を帰無仮説としたときの Mack-Skillings 検定 ($k=3, n=5, c=4$)、及び「 H_{0b} : 処置 (温度) による変化がない」を帰無仮説としたときの Mack-Skillings 検定 ($k=3, n=5, c=4$) を実施した。

また、同一個体でも測定 of 繰り返しによる効果があるかどうかを確認したい。そこで、用いた 3 匹の個体のそれぞれについても繰り返しと温度を 2 因子とした効果の検定を行った。「 $H_{0c}-H_{0e}$: それぞれの個体で繰り返しによる変化がない」を帰無仮説としたときの Friedman 検定 ($k=4, n=5, c=1$)、及び「 $H_{0f}-H_{0h}$: それぞれの個体で温度による変化がない」を帰無仮説としたときの Friedman 検定 ($k=5, n=4, c=1$) を実施した。

検定には、eq.(2) で計算した MS 値と、棄却限界値 (正確値)^{8,9)} を用い比較した。MS 値 > 棄却限界値であれば因子の効果は有意であるとした。また棄却限界値の近似値として $\chi^2(0.05, k-1)$ を計算しこれとも比較した。

結 果

1. 尾退避反応潜時 (TWL: tail withdrawal latency) の測定 実験の結果、37°C-45°C においては、ほとんどが 120 秒に達し、この範囲で最も高温の 45°C においても、12 回の測定中 11 回は 120 秒に達していた。したがって、37°C-41°C のデータは因子の効果の評価の上で冗長なデータであると判断し除外した。そして、Table 1 に示すように、43°C-55°C のデータのみを解析の対象とした。TWL と刺激温度との関係は Table 1 及び Fig. 1 に示すように、どの個体においても高温になるにつれ TWL

Table 1. Results TWL Measurement with Three Monkeys

Factor	43°C	45°C	47°C	50°C	55°C	
Monkey A	1st	120	120	120	5	3
	2nd	120	120	4	2	2
	3rd	120	120	60	3	2
	4th	120	120	63	6	1
Monkey B	1st	120	5	5	1	2
	2nd	120	120	6	4	2
	3rd	120	120	3	2	1
	4th	24	120	10	1	1
Monkey C	1st	120	120	120	19	2
	2nd	120	120	120	2	3
	3rd	120	120	14	3	3
	4th	120	120	6	100	3
Factor	43°C	45°C	47°C	50°C	55°C	
Mean	112.0	110.0	44.3	12.3	2.08	
S.D.	27.7	33.2	50.2	28.0	0.79	

が短くなり、50°C 以上ではほとんどが数秒の TWL であった。また、反応が著明に変化しようとする 47°C 付近においては大きな個体差が認められた。例えば 4 回の反復測定において、個体 B はどの測定も数秒で反応したが、個体 A 及び C では数秒-120 秒以上の大きな変動を示した。平均値 (Fig. 2 の線) でみても、個体 A、個体 C に比べ個体 B (点線) の高い温度感受性が認められた。標準偏差はどの個体においても温度により大きく異なり、また同じ温度でも個体毎に大きく異なっていた。

2. モルヒネによる鎮痛効果の検討 Figure 3 に示すように、上記いずれの個体においても、モルヒネ投与前は TWL は数秒であったのに対し、モルヒネ投与後は TWL が時間とともに増加し、1 時間以降は 2 匹の動物で 120 秒まで達した。

3. Mack-Skillings 検定 Mack-Skillings 検定の結果を Table 2 に示す。結果として、個体、温度のいずれの因子についても MS 値は棄却限界値よりも大きく、有意であることが判明した。また棄却限界値の正確値は近似値に比べ若干小さいがほぼ近い値となり、判定結果も同一であった。また、各個体について Friedman 検定を行った結果もまとめて Table 2 に示す。どの個体についても 4 回の繰り返しの効果は有意でないこと、また温度の効果は有意であることが判明した。

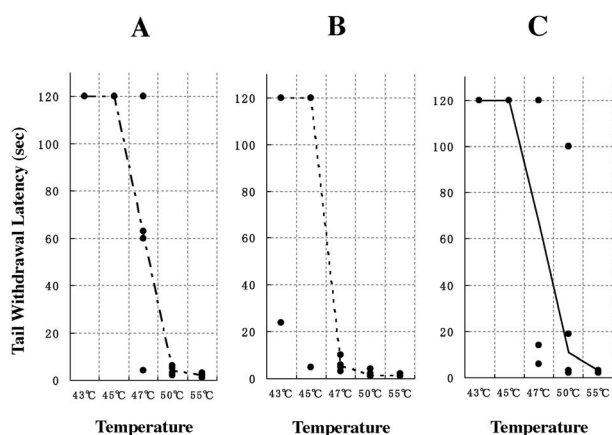


Fig. 2 Tail Withdrawal Latencies in Three Individual Monkeys, Monkey A, B and C
Dots denote individual values obtained, and lines denote medians within each temperature. Cutoff of TWL was set at 120 seconds.

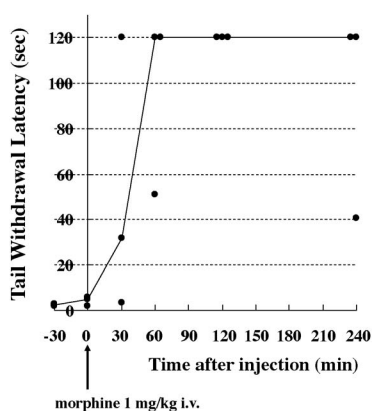


Fig. 3 The Effect of Intravenous Injection of Morphine (1 mg/kg) on the Tail Withdrawal Latencies in Three Monkeys
Dots denote individual values obtained, and a line denotes medians within each time.

考 察

TWL を痛みのエンドポイントとして鎮痛薬の薬効評価に利用するためには、まず薬を投与しない状態で TWL が刺激温度に応じて変化すること、また TWL が個体内での繰り返し及び個体間で再現性が保たれていることが必要である。Figure 2 のグラフで明らかな通り、いずれの個体においても温度を上げていくと 45°C まではほとんど反応がなく上限値 (120 秒) まで達したが、47°C 付近を境界に著明な尾の退避反応が出現し、50°C 以上ではほとんどが数秒で退避した。この変化はどの個体にも共通しており、また、各個体における 4 回の測定の変動も 47°C を除いては比較的小さかった。一方 47°C の結果をみると、個体 B は 4 回の測定がすべて数秒の

Table 2. Results of Two-way Layout Nonparametric Test

Null hypothesis	MS	Critical Value		Judgment
		Exact	Approximation (χ^2)	
H0a	10.27	5.98 ($\alpha=0.051$)	5.99 ($\alpha=0.05, df=2$)	Significant
H0b	43.26	9.30 ($\alpha=0.050$)	9.49 ($\alpha=0.05, df=4$)	Significant
H0c	2.70	7.32 ($\alpha=0.055$)	7.81 ($\alpha=0.05, df=3$)	Not significant
H0d	2.58	7.32 ($\alpha=0.055$)	7.81 ($\alpha=0.05, df=3$)	Not significant
H0e	0.06	7.32 ($\alpha=0.055$)	7.81 ($\alpha=0.05, df=3$)	Not significant
H0f	13.80	8.80 ($\alpha=0.049$)	9.49 ($\alpha=0.05, df=4$)	Significant
H0g	13.90	8.80 ($\alpha=0.049$)	9.49 ($\alpha=0.05, df=4$)	Significant
H0h	11.35	8.80 ($\alpha=0.049$)	9.49 ($\alpha=0.05, df=4$)	Significant

H0a: No differences among the effect of monkey, H0b: No differences among the effect of temperature, H0c: No differences among the effect of repeated measurement in monkey A, H0d: No differences among the effect of repeated measurement in monkey B, H0e: No differences among the effect of repeated measurement in monkey C, H0f: No differences among the effect of temperature in monkey A, H0g: No differences among the effect of temperature in monkey B, H0h: No differences among the effect of temperature in monkey C.

反応であったが、個体 A と C ではデータが数秒と 120 秒の両極に分かれた。これらの結果から、個体 B は個体 A, C に比べ刺激に対する感受性の高い個体であることが示唆された。こうした示唆は、それぞれ 1 回きりの測定では見逃す可能性があり、その意味で個体内の繰り返し測定を行うことは重要である。

このように、グラフ化して個々の観測値の挙動を考察することはまず重要であるが、次に全体としての因子の効果はどのように評価したらよいか。通常、2つの因子を含む場合は 2 元配置の (繰り返し有りの) 分散分析を行えばこれらの誤差の評価は可能である。しかし今回のデータの特徴として、上限値 (120 秒) と下限値 (数秒) 付近は変動が比較的小さく、反応が切り替わる 47°C 付近では変動が特に大きい。それぞれの処置 (温度) での平均値からの残差から全体の尖度と歪度を計算すると、尖度 = 4.0, 歪度 = -0.1 となった。したがって、データの分布は全体として大きな歪みはないものの、120 秒及び数秒に値が集中した尖りの大きい分布であると考えられる。このような分布では正規性の想定は困

難であるため、分散分析などのパラメトリックな解析手法はかならずしも適切ではなく、分布を仮定しないノンパラメトリックな手法を適用する方が望ましいと考えられる。そして今回のように2因子で繰り返しも含まれている場合は、2元配置の Mack-Skillings 検定が有効である。検定の結果は Table 2 に示した通り、MS 値が棄却限界値を超え有意となり、全体として個体の効果、温度の効果いずれも有意という結論になった。また各個体で繰り返しと温度を2因子とする Friedman 検定（繰り返しなしの Mack-Skillings 検定）を行ったところ、どの個体についても温度の効果は有意となったが繰り返しの効果は有意とならなかった。よって、繰り返しに伴う系統誤差などの変動は十分に制御されており、用いた3匹の個体間の変動は、温度に対する個々の個体固有の感受性の相違に基づくものであると推測される。このように、打ち切りを含むデータにおいて複数の因子の効果を検定するのに、Mack-Skillings 検定は有効な手法である。Mack-Skillings 検定は広く知られた手法ではないが、最近になって、マイクロアレイ実験における発現遺伝子の特定などに使われ始めている。¹⁰⁾

精度の高い定量計測を実現するためには、個々のサルで繰り返し測定を行うことが重要であることは言うまでもないが、それに加え今回のように各ブロック内で各温度1回ずつの乱塊法の実験デザインにすることにより、高温の刺激がある時間に集中し感作・脱感作が起こるといった問題を回避することができる。実験手技上も刺激と刺激の間を10分空けるなどの工夫を行うとより望ましい。例えば47°C付近の変動に関して、Table 1の Monkey Aでは、1回目は120秒、2回目は4秒であったが3回目以降は60秒以上であった。したがって、この変動は感作などの影響を含んだ系統的なものではないと考えられる。逆に次第にTWLが短くなるなど系統的な変化が認められた場合には、誤差の原因を考慮した実験デザインの見直しが必要となるであろう。今回の例では、末梢における侵害受容の自由神経終末からの入力レベルの差異によって生じ得る誤差は回避できたと考えられるが、それ以外にも、内因性の誤差としてはより高位中枢での痛覚伝導メカニズムの差異、嫌悪条件付け（刺激による不快を学習し尾が退避し易くなる）などが考えられる。これらは制

御が非常に困難な変動因子であるから、本実験を行う前に、実験動物選定時に簡単な予備試験を行い、反応の過敏な個体を除去するなどの工夫が必要となるであろう。また外因性の誤差の考慮も重要である。実験時の室温、照明などは言うまでもないが、注意すべきは日内変動、日間変動の影響である。今回は1匹当たり32回の計測で丸1日を要しているため、個体を変更すると日間変動の影響は不可避となる。仮に計測回数を半分にできるとしたら、午前と午後に分けた実験を行うことになり、日間変動は回避できるが今度は日内変動の問題が起こる。このような場合は、クロスオーバーデザインにするなどの工夫が必要であろう。そして何より重要なことは、想定される変動因子をできる限り制御することである。今回の3日間の実験に際しては、実験者や分担、手順等は固定し、手技や実験室環境等は3日も極力同じ条件になるように努めた。例えば刺激に用いる温浴を準備して尾に刺激を与えるまでの操作に手間取ると、その間に温浴の温度が低下するので、これらを常に同じ時間間隔で実行する、など細かい実験上の配慮の積み重ねが重要となることは言うまでもない。

中型・大型動物を用いた薬理試験においては、完全ランダム化実験では個体の変更に手間が掛かるゆえに分割実験の型式が取られる場合が多い。今回のモルヒネを用いた検討でも、まずある個体で薬効評価を含む実験を一通り済ませてから、次の個体で実験を行った。これにより、薬物処理の前値と後値の同一個体での比較が可能となり、Fig. 3に示したような薬物効果の経時的追跡が可能となる。ここでも時間と個体を因子とするFriedman検定を行った結果、検定統計量 $MS=11.67$ 、棄却限界値 $\chi^2(0.05,5)=11.07$ で、時間の効果が有意となりモルヒネによる鎮痛作用の発現が確認された。今回はモルヒネ1用量での検討であったが、複数の鎮痛薬で反応を比較したり用量反応関係を検討したりする場合は、薬物若しくは用量という因子がさらに1つ増えることになる。その場合もMack-Skillings検定により因子の効果を評価することが可能となる。

複数個体で繰り返しを含む動物実験では、通常繰り返しありの2元配置分散分析が適用されるが、今回のように正規性を仮定できない動物実験データにおけるMack-Skillings検定の適用は、応用事例とし

ては新しいといえる。特に大型・中型動物を扱う薬理研究者は同一個体での繰り返し測定を行うことが多く、これと類似の問題に遭遇する可能性があり、その際の指針を与え得るものである。また、今回 Mack-Skillings 検定に加え各個体について Friedman 検定を適用し、繰り返しと温度の効果の判定を行った。そして、どの個体についても繰り返しの効果が有意でないことから、繰り返しに伴う系統誤差などの変動は十分に制御されており、温度に対する個々の個体の感受性の相違が全体としての個体の効果に反映していることを示唆することができた。こうした 2 本立てのアプローチにより因子の効果を考察する試みも十分応用可能と考えられる。

今回のような打ち切りを含んだデータを解析するには、正規性を想定したパラメトリックな解析はかならずしも適切でなく、ノンパラメトリック解析が望ましい。一方、ノンパラメトリック解析は簡便ではあるものの、観測値を順序に変換するため生物学的に重要と思われる量的情報までもが削ぎ落とされるという欠点もある。また仮説検定という方法は、「反応が因子によって変わらない」という仮説が、実際に観測したデータから否定できるかどうかを判定するものである。これに対し、反応が因子によりどのように変化するかを定量的客観的に示したいのであれば、薬理的洞察に基づき非線形関数（非線形モデル）を当てはめパラメータを推定する方法が有効である。今回の例でも、打ち切りを考慮した非線形モデルによる解析が可能であれば、得られた結果をより定量的客観的に示すことができると考えられ、今後の検討課題である。

謝辞 本実験の推進に御協力頂いた浜松ホトニクス中央研究所 PET センターの皆様に感謝申し上げます。

REFERENCES

- 1) Janssen P. A., Niemegeers C. J., Dony J. G., *Arzneimittelforschung*, **13**, 502–507 (1963).
- 2) Dykstra L. A., Gmerek D. E., Winger G., Woods J. H., *J. Pharmacol. Exp. Ther.*, **242**, 413–420 (1987).
- 3) Dykstra L. A., Woods J. H., *J. Pharmacol. Methods*, **15**, 263–269 (1986).
- 4) Sewell R. D., Spencer P. S., *Neuropharmacology*, **15**, 683–688 (1976).
- 5) Van Bever W. F., Niemegeers C. J., Schellekens K. H., Janssen P. A., *Arzneimittelforschung*, **26**, 1548–1551 (1976).
- 6) Tsukada H., Miyasato K., Kakiuchi T., Nishiyama S., Harada N., Domino E. F., *Synapse*, **45**, 207–212 (2002).
- 7) Tsukada H., Miyasato K., Nishiyama S., Fukumoto D., Sato K., Kakiuchi T., Domino E. F., *Neuropsychopharmacology*, **30**, 2144–2153 (2005).
- 8) Mack G. A., Skillings J. H., *J. Am. Statist. Assoc.*, **75**, 947–951 (1980).
- 9) Hollander M., Wolfe D. A., *Nonparametric Statistical Methods*, Wiley Series in Probability and Statistics, 2nd ed.: John Wiley & Sons, Inc., 328–332 (1999).
- 10) Barrera L., Benner C., Tao Y. C., Winzeler E., Zhou Y., *BMC Bioinformatics*, **5**, 42 (2004).