

できる MPE も予測モデルとして評価尺度に加えられた。

- ・各施設における個々の被試験物質評価は、同一の試験を独立に 2 回行うことによるとされた。

### B-3 Phase III 研究

- ・基本的に Phase II 研究の SOP に従い試験が実施された。
- ・被験物質溶解のための溶媒の使用法と被験物質の適切な濃度に関する SOP が新たに作成された。
- ・被験物質はブラインド化された。(Phase II と同じ)

### C. バリデーション研究期間中のプロトコールの変更や逸脱

試験結果を評価する時、プロトコール違反や SOP 違反、また、データの不整合がないことに越したことはないが、どのようにプロトコールや SOP を工夫しても、施設間バリデーションにおける誤りや誤解の発生の可能性を完全になくすることはできないであろう。また、試験実施前にプロトコールに不備等が見つかった場合にプロトコールの変更がなされることもあるだろう。プロトコールの変更、誤り・不明な点・記録不足などに対するデータの扱い、結果としての防止できなかった逸脱状況を調べることは、研究結果の質を評価する際に重要な点の一つであり、OECD の報告<sup>4)</sup>でも GLP 基準に準じて試験結果の作成、実施、記録が必要とされている。

3つの研究の中で、この点について詳細な記載をしていたのは、Phase II 研究のみであった。これは Phase II 研究が検証試験として位置付けられた研究であったためかもしれない。しかし、施設間バリデーション研究は複数施設で試験が実施される研究であり、施設間の違いを評価することが研究の目的の一つであることを考慮すると、各施設から得られたデータの取扱いについての記載がなされていることが望ましい。以下に3つの研究におけるプロトコールの変更、データ収集時の手続き、プロトコールの逸脱について記載されている点をまとめた。

#### C-1 Phase I 研究

- ・除外されたのか、求められなかったのかは不明であるが、用量当たりの例数が物質によって異なっていた。
- ・評価尺度 (PIF) を計算できた物質は 20 物質中 15 物質であった。

#### C-2 Phase II 研究

- ・MS EXCEL スプレッドシートによりデータの提出がされた。(詳細なフォーマット形式が指定されていたかどうかは不明)
- ・規定通りのデータが完全に得られているかどうかチェックされた。
- ・すべてのデータシートに記入されている数値の形式がチェックされた。
- ・濃度単位誤入力の有無のチェックがされた。
- ・SOP からの逸脱の有無と結果への影響がチェックされた。
- ・評価した 9 施設中 1 施設について、30 物質中 14 物質が、照射の有無についての 2 つの試験が同時に行われなかったため、SOP 不適合となった。

- ・ vivo データの質が不十分であるとみなされたため、1物質のデータが vivo データとの対応性の評価の計算から除外された。(この物質を評価した施設が8施設あったため、256データ中8データが除外された。)
- ・ 被験物質の溶媒として、SOPにはEBSSを使用することが勧められていたが、必要な場合にはエタノール等の溶媒を用いることが記載されていたため、施設により使用された溶媒に差が生じた。(結果は大きな違いがなかったと考察された。)
- ・ UV照射器にSOL 3を用いた一施設あった。(両者のスペクトルは非常に似ているため違いは小さいと考察された。)
- ・ 試験の最適な濃度の選択についてSOPに規定されていたが、低い濃度を選択したラボがあった。
- ・ 同一実験を2回施行することになっていたが、2回の施行で異なる溶媒を用いた施設があった。

### C -3 Phase III 研究

- ・ 2回の試験のうちPIFまたはMPEが求まらないデータはなかった。
- ・ PIFとMPEの値は各施設で同一のソフトを用いて求められた。

### D. 研究間の違い

3つの研究は一連の研究であるために、基本的には同一の試験方法で実施されている。しかし、研究が進むことにより、新たな進展が見られる。以下に研究間の違いをまとめた。

- ・ 予測モデルとして、Phase I 研究ではPIFを用い、Phase II 試験以降はPIFと共にMPEも用いられている。
- ・ Phase III 研究では被験物質はブラインド化されているが、ブラインド化の各物質について適切な溶媒や濃度範囲が指定されている。

### E. データの監査

実施された試験のデータの質を確保することは、データの信頼性、再現性、客観性を保証することになる。そのためには、評価に用いるデータの源となる各施設に対する規範と第三者による監査が必要となるであろう。ここでは、文献による評価について議論を進めているが、文献による評価ではこの作業を我々が行うのは実際上不可能である。また、今回検討したいずれの文献にも監査が、第三者により行われたという記載はなかった。また、文献には、PIFやMPEを求めるために使用された生データ、試験情報までは示されておらず、このレベルにおける質の評価は不可能であった。

### F. 質の評価

ここでは、EC/COLIPAで実施されたBalb/c 3T3細胞を用いたNR法のバリデーション研

研究結果が記載されている文献に基づき、質の評価を行った。3つの研究を比べると、特に主要なバリデーション研究となった Phase II 研究は、データの質を確保するためにどのようなことが行われていたのかがよく記載されている文献であるといえよう。他の2つの研究は、質について判断するために必要な情報に関してあまり詳細な記載がない。このことは、質が良いとか悪いということではなく、質について評価ができないということの意味する。その意味では、Phase II 研究は、他の2つの研究に比べて信頼性の高い結果といえるであろう。代替法のバリデーションやその行政的受け入れのための基準については OECD<sup>4)</sup>で示されているが、バリデーション研究を文献評価のみで行う際に報告される文献中にどのような記載が必要であることを明確に記述した文書はない。後述するように、ここでは、これについてわずかながらの提言をしている。

また、項目 E で記述したように、文献による質の評価は、生データがどのように得られたかについてまでのことを知ることはできない。今後、動物実験代替法が化学物質の評価方法の一つとして行政的に受け入れられるようになるためには、バリデーション研究のデータの質を保証するために、何らかの形で第三者が生データや、試験の実施状況や研究の進行状況を確認するための記録等が必要となるかもしれない。

## G. 提言

試験の結果を評価する上で、データの質を論じることは重要であるが、文献評価の場合には、記載がされてなければ、どのようなことが行われているかは不明となってしまう。記述されていないことは、質がよいことを保証しているわけではなく、不明であるというだけであるので、第三者が文献を評価する場合には、その部分は不明であると判断するほかは無い。しかし、例えば、データの一致性のチェックが、その研究中で行われているということが研究結果を報告する文献中に記載されていればチェックがされたことは知ることができる。このような記載がないものとあるものとは、質の評価を変えるべきであろう。我々が今回評価を行う際には、文献レベルでのバリデーション研究の質の評価法を構築する必要が生じた。これらを解決する一つの方法として、データの質に関する記述がどこまでされているかをチェックするためのチェックリストの使用があげられる。その際に問題になるのはどのような項目を用いるかであるが、参考にすべきものが特になかったため、Phase II 研究を参考にチェックリストを作成した。これを付録に示す。Phase II 研究を参考にした理由は、この研究の報告が、データの質という点に関してよく記載されていたためである。このようなチェックリストを利用した時にチェックができない項目が多い場合は、試験の質について言及できない試験ということになるであろうし、そうでない場合は、試験の質が配慮された試験ということになるであろう。このチェックリストはまだ改善される余地が十分にあるが、リストが公になり、ここで記載されている項目が注意して論文に記載されるようになれば、試験の質はおのずと高くなるであろうし、求められる研究のあり方だと思われる。

上記のようなチェックリストを用いたとしても、文献のみによる質の検討は限界がある。例

えば、各施設のそれぞれの用量反応曲線を得るための生データまで文献に記載されることはないであろう。つまり、文献による質の検討では、この本当に記載されている予測モデルの値等が正しいかどうかを確認するというような評価は不可能である。

## H. 結論

文献のみによる質の評価は十分には行えない。従って必要に応じて独自のバリデーションを行い、試験法の妥当性を確認する必要がある。このバリデーションの規模と内容は事前に得られているデータの質の評価結果に依存している。

## 参考文献

- 1) Spielmann, H. et al., (1994) EEC/COLIPA Project on In Vitro Phototoxicity Testing: First Results Obtained with A BALB/C 3T3 Cell Phototoxicity Assay, *Toxicology in Vitro* 8, 793-796.
- 2) Spielmann, H. et al., (1998) The International EU/COLIPA In Vitro Phototoxicity Validation Study: Results of Phase II (Blind Trial). Part 1: The 3T3 NRU Phototoxicity Test, *Toxicology in Vitro* 12, 305-327.
- 3) Spielmann, H. et al., (1998), A Study on UV Filter Chemicals from Annex VII of European Union Directive 76/768/EEC, in the In Vitro 3T3 NRU Phototoxicity Test, *ATLA* 26, 679-708.
- 4) OECD, Final report of the OECD workshop on harmonization of validation and acceptance criteria for alternative toxicological test methods. ENV/MC/CHEM/TG(96)9

## 付録 バリデーションデータの質に関するチェックシート

●研究名：

文献名：

著者：

- 信頼できる機関の指導で行われたバリデーションか？
- 評価施設は何施設か？
- 評価物質は何物質か？
- プロトコル、SOP を入手することが可能か？
- 生データをどのように加工したデータが記載されているか？
  
- 質を良くするためのデザイン上の工夫がされているか？
  - 機器や SOP について記載がされているか？
  - 信頼性、妥当性の高い評価項目（Prediction Model : PM）が用いられているか？
  - in vivo のデータソースはどのようなものか記載がされているか？
  - 一つの物質を評価する際に、同じ用量で複数の測定が独立に行われているか？
  
- 提出されたデータのチェックが行われているか？
  - 提出されたデータの形式について記載されているか？
  - 提出されたデータのチェック項目について記載されているか？
  
- データ解析のためのデータの固定はどのように行われているか？
  - PM の計算方法は各施設で計算するか特定の施設で計算するか？
  
- 結果はどうであったか？
  - プロトコルや SOP の逸脱について記載されているか？
  - 解析から除外されたデータについて記載されているか？
  - 施設により実験条件が異なった場合の影響について記載されているか？
  - 施設内再現性は十分か？
  - 施設間再現性は十分か？
  
- その他注目すべきことはあるか？
  - データの質を低下させる可能性についての考察が記載されているか？

#### 4-2-2-2) 試験結果のピボ試験結果との対応性

担当：畑尾正人、大森 崇

##### A. データの説明

Balb/c 3T3 細胞を用いた光刺激性試験代替法は、被験物質存在下で UVB を減衰させるフィルターを通したソーラーシミュレーター光を照射して、Balb/c 3T3 細胞の生存をニュートラルレッドの細胞内取り込み量をエンドポイントとして評価する方法である。本試験法は光刺激性を定性的に評価する hazard identification のための方法として位置付けられる。判定に用いられるパラメーターは PIF (Photo Irritation Factor: 光刺激係数) および MPE (Mean Photo Effect: 平均光効果) の 2 種がバリデーションで使用されている。

本章ではこの 2 つのパラメーターを用いた試験法についてそれぞれ試験性能に対する評価を行った。この 2 種のパラメーターで光刺激性を施設間バリデーションで評価している公表論文は前項(4-2-2-1)で述べた 3 報<sup>1-3)</sup>のみであるため、それらの論文データを基に評価した。

なお、試験法の性能に関して適切な評価を行うためにそれらの論文の実験データを以下に示す基準で再検討し、データをまとめた (Table 4-2)。

- ・ 同一の化学物質が複数の論文で評価されている場合はまとめて一つの物質として解析を行った。
- ・ *in vivo* のデータは 3 論文を通じて、データの存在を明確に示す記述が一つでもあれば、それを採用した。
- ・ Phase I 研究での *in vivo* 試験結果は空欄、+/- (Inconclusive)、(+), +, -, A (光アレルギー性) というあいまいな書き方があったため、本解析では以下に示す基準を用いた。
  - 空欄：データの存在がないものとした
  - +/-：解析に使用できないのでデータなしとした
  - (+)：+とした
  - A (光アレルギー性)：試験データが存在するものとして扱った
- ・ *in vitro* の実験結果で施設間の判定結果に陽性および陰性が混在したものは多施設の示した判定に統一して解析した。その際、括弧内に異なる判定を示した施設数の全体に対する割合を示した。
- ・ 論文中では同一の化学物質とその塩が別の被験物質として扱われている場合と同一に扱われている場合が混在していた。論文データをまとめて記した Table 4-2 では公表論文そのものの形で表したが、*in vivo* との比較解析では同一の被験物質として扱った。

Table 4-2. 公表データのまとめ (PIF: Photo Irritation Factor, MPE: Mean Photo Effect,  
 +':in vivo陽性、'-':in vivo陰性、+/-:判定保留、  
 PI: photoirritant、NP: non-photoirritant、NA: not available、  
 空欄：試験実施なし、分数で示した値は異なる判定をした施設の割合)

化合物名	In vivo data		1st validation		2nd validation		3rd validation	
	Animal	Human	PIF	MPE	PIF	MPE	PIF	MPE
Promethazine	+/-	+	PI		PI	PI	PI	PI
Chlorpromazine	+	+	PI		PI	PI	PI	PI
6-Methylcoumarin	+	+	PI		PI	PI		
TCSA	+	+	PI					
Doxycycline	+	+	PI					
8-MOP	+	+	PI					
Tetracycline	+	+	PI					
Piroxicam	-	+	NP					
Amiodarone	+	+	PI		PI-3/9	PI-1/9	PI-1/4	PI
Bithionol	+	+	PI		PI	PI-1/9	PI	PI
Neutral red	+	+	PI		PI	PI		
Rose Bengal (2)	-	+	PI		PI	PI		
Cinnamic Ald	-	-	NP					
Chlorhexidine (2)	NA	-	NP		NP-3/9	NP		
Uvinul MS40	NA	+/-	NP					
PABA	-	+/-	NP		NP-2/9	NP-4/9		
Penicilin G	NA	NA	NP		NP-2/9	NP-2/9		
L-histidine	NA	NA	NP				NP	NP
Thiourea	NA	-	NP					
Lauryl sulfate	NA	NA	NP					
5-MOP	+	+			PI	PI-1/8		
Acridine hydrochloride	+	+			PI	PI	PI	PI
Acridine -free base	+	+			PI	PI		
Anthracene	+	+			PI-1/8	PI-2/8	PI	PI
Bergamot oil	+	+			PI-4/9	PI		
Demecloxycline	+	+			PI	PI	PI-1/4	PI
Fenofibrate	NA	+			PI-1/8	PI-1/8		
Furosemide	NA	+			NP-2/8	NP-3/8		
Hexachlorophene	-	+/-			NP-1/9	NP		
Ketoprophen	-	+/-			PI	PI	PI	PI
SLS	NA	NA			NP	NP	NP	NP
Musk ambrette	-	+/-			PI-1/8	PI-2/8	PI	PI
Nalidixic acid Na	+	+			PI-2/8	PI		
Nalidixic acid	+	+			PI-2/8	PI-1/8		
Norfloxacin	+	+			PI	PI		
Ofloxacin	+	+			PI-1/8	PI		
Protoporphyrin F	+	NA			PI	PI		
Protoporphyrin S	+	NA			PI	PI	PI	PI
Tiaprofenic acid	+	+			PI	PI		
Octyl salicylate	-	-					NP	NP
Octyl methoxycinnamate	-	-					NP	NP-1/4
Benzilidene Camphor sulphonic acid	-	-					NP	NP
4-Methyl benzilidene camphor	-	-					NP	NP
3-Benzilidene camphor	-	-					NP	NP
Terephthalidine dicamphor sulphonic acid	-	-					PI-1/4	PI-1/4
Polyacrylamido methy benzilidene-camphor	-	-					NP	NP
Benzophenon-4 (2)	NA	+/-			NP-1/8	NP-1/8	NP-1/4	NP-1/4

TCSA: 3,3'4'5-tetrachlorosalicylonilid

8-MOP: Mothoxsalen

PABA: p-Aminobenzoic acid

5-MOP: Bergapten (5-Methoxypsoralene)

SLS: sodium lauryl sulfate

Benzophenon-4 (2): 5-benzoyl-4-hydroxy-2-methoxybenzene sulfonic acid

## B. 評価された化学物質の種類と性状の適切性

### B-1) 被験物質のカテゴリー

論文に掲載された評価の対象となった被験物質数は 44 化学物質であり、以下に示す 5 つのカテゴリーに分類された。このうち、*in vivo* での光毒性の有無が不明だったものを除くと、紫外線吸収剤 10 種、香料 4 種、抗生物質・抗菌薬 8 種、抗炎症薬・鎮痛薬 3 種、その他の薬剤 11 種、その他の化学物質 4 種となり、計 40 種となる。被験物質の選択としては紫外線吸収剤の比率が高いが、これは光毒性試験の評価という観点から適切といえることができる。また、それ以外は比較的広い領域から選択されていて、被験物質の選択は適切であったと考えられる。

I) 紫外線吸収剤 (10 種) : Uvinul MS40, p-Amino benzoic acid (PABA), Benzophenone-4, Octyl salicylate, Octyl methoxycinnamate, Benzilidene camphor sulphonic acid, 4-Methyl benzilidene camphor, 3-Benziliden camphor, Terephthalidine dicamphor sulphonic acid, Polyacrylamidomethyl benzilidene camphor

II) 香料 (4 種) : 6-Mehtylcoumarin, Cinnamic aldehyde, Bergamot oil, Musk ambrette

III) 抗生物質、抗菌薬 (9 種) : Tetrachlorosalicylanilide, Doxycycline, Tetracycline, Chlorohexidine dihydrochloride, Penicillin G, Demeclocycline, Nalidixic acid, Norfloxacin, Ofloxacin

IV) 抗炎症薬、消炎鎮痛薬 (3 種) : Piroxicam, Ketoprophen, Tiaprofenic acid

V) その他の薬剤 (11 種) : Promethazine (抗ヒスタミン剤)、Chlorpromazine (抗精神薬)、8-MOP (白斑治療剤)、Amiodarone (抗不整脈剤)、Bithionol (抗菌剤)、Thiourea (防腐剤)、5-MOP (白斑治療剤)、Fenofibrate (抗高脂血症剤)、Furosemide (ループ利尿剤)、Hexachlorophene (防腐剤)、Protoporphyrine IX (肝疾患治療剤)

VI) その他 (7 種) : Neutral red (pH 指示薬)、Rose bengal (色素)、L-Histidine (アミノ酸)、Lauryl sulfate (界面活性剤)、Acridine (蛍光色素)、Anthracene (芳香族炭化水素)、Sodium lauryl Sulfate (界面活性剤)

### B-2) 被験物質の性状

被験物質の性状という点では溶解性の観点が挙げられる。溶解性について Phase I 研究の論文では水に不溶性の被験物質については溶解補助剤として DMSO を使用することが推奨されているが、個々の被験物質の溶解性やその試験結果に対する影響は記述されていない。

Phase II 研究の論文では被験物質の水への溶解性を 7 段階に分類し、その水溶性と試験結果の不一致率に統計的に有意な相関性がないと議論している。しかし実際にバリデーションの過程では各施設の使用している溶解補助剤は DMSO、エタノール、リン酸緩衝液 (PBS) 等と統一がとれていないこと、また試験した被験物質濃度も施設間で大きく異なることから、この試験結果から溶解補助剤の影響を正當に評価することは難しい。なお、この論文のディスカッション



中で Acridine、Nalidixic acid、Protoporphyrin IX の3種についてはそれぞれの単体とその塩も同時に試験されており、それぞれいずれも光毒性物質として判定されたことから、PIF、MPEとも溶解性が試験結果に影響を及ぼさないと考察されている。しかし、Protoporphyrin IX については単体とその塩の溶解性の違いについてはデータが示されていない。

Phase III 研究においても被験物質の溶解性の影響が評価されていない。

以上の実験結果から被験物質の溶解剤が本試験法に与える影響について我々が議論できるような被験物質の選択とはなっていないし、情報提供がなされているとはいうことができない。ただし、Phase II で用いられた被験物質の範囲では水溶性の程度が光毒性の判定に大きな影響を与えておらず、ここでの結論が溶解性の影響を含めた本試験法の性能を否定するものではない。

### C. 評価された化学物質数・製品数に関する適切性

公表論文のデータをパラメーター別に *in vivo* のデータと比較した結果を Table 4-3~5 にまとめた。論文中で *in vivo* データの陽性、陰性が明確でないもの、あるいは動物試験、ヒト試験それぞれの存在が明らかでない被験物質は解析から除いた。同じ化学物質でもヒト試験結果と動物試験結果が異なる場合は、対照表(Table 42-3、4-4)では *in vivo* 判定結果は異なるものとして扱ったが、Table 4-5 ではヒト試験の判定結果を優先して結果をまとめた。

全体で評価している被験物質の数は 26~38 とそう多くはないが、評価を行う上で必要最小限のレベルには達していると考ええる。ただ紫外線吸収があつて *in vivo* で光毒性がない化学物質の数がやや不足していると考ええる。紫外線吸収があつてヒト試験および動物試験双方とも陰性となっている被験物質数は 7 に過ぎないため、*in vivo* で光毒性がある被験物質数 18 と比べて半分以下となる。この点は false positive を過小評価する可能性が考えられる。

また、これまでに公表された論文の中では Balb/c 3T3 を用いた本 *in vitro* 光毒性試験法で製品を評価している例はなく、製品系への適切性は議論できない。

Table 4-3. 論文中の *in vivo* vs. *in vitro* 対応被験物質数 (Human vs. Balb/c 3T3)

Human	Balb/c 3T3	PIF	MPE
+	+	21	17
+	-	2	1
-	+	1	1
-	-	9	7
評価被験物質数合計		33	26

+:陽性、-:陰性、数字は左2列で示されたカテゴリーに入る被験物質数

Table 4-4. 論文中の *in vivo* vs. *in vitro* 対応被験物質数 (Animal vs. Balb/c 3T3)

Animal	Balb/c 3T3	PIF	MPE
+	+	19	15
+	-	0	0
-	+	4	4
-	-	10	8
評価被験物質数合計		33	27

+:陽性、-:陰性、数字は左2列で示されたカテゴリーに入る被験物質数

Table 4-5. 論文中の *in vivo* vs. *in vitro* 対応被験物質数 (Human (or Animal) vs. Balb/c 3T3)

Human or (Animal)	Balb/c 3T3	PIF	MPE
+	+	22	18
+	-	2	1
-	+	3	3
-	-	11	9
評価被験物質数合計		38	31

+:陽性、-:陰性、数字は左2列で示されたカテゴリーに入る被験物質数

#### D. 試験法の性能データに関する適切性

##### D-1) 感度、特異性、陽性予測率、陰性予測率、一致率

前節で得られた結果を基に *in vivo* および *in vitro* の判定結果をパラメーターと *in vivo* データ別に Table 4-6~4-11 にまとめ、感度、特異性、陽性予測率、陰性予測率、一致率を計算した。この結果からは動物試験との対応性において疑陽性がやや多い傾向がある。しかし、これらの化合物にはヒト試験で陽性反応がでている Rose bengal、Ketoprophen、Musk ambrette が含まれているため、ヒト試験結果が真の光毒性を反映していると考えれば疑陽性率は必ずしも高くない。ヒト試験結果との比較で疑陰性となった3被験物質のうち2つは動物試験では陰性となっている。*In vitro* 試験の *in vivo* 試験に対する境界領域にある可能性が考えられる。

Table4-6. *In vivo*, *in vitro* 判定対応表 (Human vs. PIF)

Vitro \ Vivo	Phototoxic	Non-phototoxic	Total
Phototoxic	21	1	22
Non-phototoxic	2	9	11
Total	23	10	33

Sensitivity: 91.3%, Specificity: 90%, Positive predictivity: 95.5%,  
Negative predictivity: 81.8%, Accuracy: 90.9%

Table4-7. *In vivo*, *in vitro* 判定対応表 (Human vs MPE)

Vitro \ Vivo	Phototoxic	Non-phototoxic	Total
Phototoxic	17	1	18
Non-phototoxic	1	7	8
Total	18	8	26

Sensitivity: 94.4%, Specificity: 87.5%, Positive predictivity: 94.4%,  
Negative predictivity: 87.5%, Accuracy: 92.3%

Table 4-8. *In vivo*, *in vitro* 判定対応表 (Animal vs. PIF)

Vitro \ Vivo	Phototoxic	Non-phototoxic	Total
Phototoxic	19	4	23
Non-phototoxic	0	10	10
Total	19	14	33

Sensitivity: 100%, Specificity: 71.4%, Positive predictivity: 82.6%,  
Negative predictivity: 100%, Accuracy: 87.9%

Table 4-9. *In vivo*, *in vitro* 判定対応表 (Animal vs. MPE)

Vitro \ Vivo	Phototoxic	Non-phototoxic	Total
Phototoxic	15	4	19
Non-phototoxic	0	8	8
Total	15	12	27

Sensitivity: 100%, Specificity: 66.7%, Positive predictivity: 78.9%,  
Negative predictivity: 100%, Accuracy: 85.2%

Table 4-10. *In vivo*, *in vitro* 判定対応表 (Human (or animal) vs. PIF)

Vitro \ Vivo	Phototoxic	Non-phototoxic	Total
Phototoxic	22	3	25
Non-phototoxic	2	11	13
Total	24	14	38

Sensitivity: 91.7%, Specificity: 78.6%, Positive predictivity: 88.0%,  
Negative predictivity: 84.6%, Accuracy: 86.8%

Table 4-11. *In vivo*, *in vitro* 判定対応表 (Human (or animal) vs MPE)

Vitro \ Vivo	Phototoxic	Non-phototoxic	Total
Phototoxic	18	3	21
Non-phototoxic	1	9	10
Total	19	12	31

Sensitivity: 94.7%, Specificity: 75.0%, Positive predictivity: 85.7%,  
Negative predictivity: 90.0%, Accuracy: 87.1%

次に感度、特異性、陽性予測力、陰性予測力、一致率を Table 4-12 にまとめた。全体として、これらの数値は特異性を除き 90%前後の値が多く、悪くない値を示している。特にヒトで陽性とされている物質を陽性する感度は PIF で 91.3%、MPE で 94.0%と高い値を示した。動物試験に対する specificity がやや低いが、前節に記載した事実からも実態として考えると疑陽性は多くないと考えられる。また PIF および MPE の2種のパラメーターによる違いは実験として Phase I 研究では MPE についてのデータを取らなかったということによるものであって、他の2つの論文で双方のパラメーターが異なる判定結果を導いた例はない。すなわち実験結果を解析する意味においては等価と考えて差し支えない。

Table 4-12. 感度、特異性、一致率、陽性予測力、陰性予測力一覧表

	vs. Human		vs. Animal		vs. Human (or animal)	
	PIF	MPE	PIF	MPE	PIF	MPE
Sensitivity	91.3 (21/23)	94.4 (17/18)	100 (19/19)	100 (15/15)	91.7 (22/24)	94.7 (18/19)
Specificity	90.0 (9/10)	87.5 (7/8)	71.4 (10/14)	66.7 (8/12)	78.6 (11/14)	75.0 (9/12)
Positive predictivity	95.5 (21/22)	94.4 (17/18)	82.6 (19/23)	78.9 (15/19)	88.0 (22/25)	85.7 (18/21)
Negative predictivity	81.8 (9/11)	87.5 (7/8)	100 (10/10)	100 (8/8)	84.6 (11/13)	90.0 (9/10)
Accuracy	90.9 (30/33)	92.3 (24/26)	87.9 (29/33)	85.2 (23/27)	86.8 (33/38)	87.1 (27/31)

括弧の中の数字はそれぞれ該当する被験物質数

## E. 試験法の有用性に関する結論とその科学的妥当性

### 1) 臨床との関連性及びヒトへの予測性

ヒトへの対応を考えた場合、本試験法は若干偽陰性が多い傾向にあるが、予測性全体としては十分許容内にあると考えられるため、光毒性の hazard identification を行うために有効であると考えられる。しかし、全体として評価した被験物質数は十分とはいえない。特に *in vivo* で陰性と判定される物質について検討が不足している。本解析では *in vivo* データの存在が明らかでないことから、Phase I 研究で紫外線吸収がなく光毒性がない被験物質に分類されている3種の化学物質 (Penicilin G, L-Histidine, Lauryl sulfate) を除いているため、これも解析を行った被験物質数を減らしている一因となっているが、それを含めても被験物質数が少ないと考えられる。今後さらなる検討をするのであれば、紫外線吸収性で非光毒性である化合物の中にどのくらいの疑陽性が含まれるかを追加検討することが望ましい。

また論文中の記載についても不明瞭な部分があるため、今回の解析が妥当であったかどうか判断できない部分もあった。例えば Thiourea は紫外線を吸収しないにもかかわらず光感作性があると分類されているが、これが妥当な分類であるのか、cinnamic aldehyde は中程度の感作性を持つが、これを紫外線吸収性の非光毒性物質として分類して良いのか、あるいは PABA は光感作性物質だが、非光毒性として扱って良いのか等のあいまいな点も多い。これらはヒト試験のデータソースを既存公表文献に依存しているために起きてくると考えられる。現実的な問題として光毒性を生じるというヒトデータは臨床試験から得られることが多く、この意味では光毒性がないというヒト試験データはやはり少なからざるを得ないだろう。これは今後のバリデーションスタディのヒトデータとの相関性を見る場合のデータソースをどうするかという課題にもつながる。

## F. まとめ

- Balb/c 3T3 細胞を用いた Neutral red 取り込みによる光毒性試験の感度、特異性、陽性予測力、陰性予測力、一致率は 90%前後の値であり全体として高く、化学物質の光毒性ポテンシャルを予測する上で有効である。特にヒトで陽性とされている物質を陽性とする感度

は PIF で 91.3%、MPE で 94.0%と高い値を示したことは本試験法をスクリーニング法として用いることの妥当性を示している。

- ・ 使用されている 2つのパラメーター（PIF、MPE）に予測性の点からの差異は認められなかった。

被験物質の水に対する溶解性の判定結果に及ぼす特異的な影響については評価できないが、プロトコールの範囲で使われる溶解助剤を用いた判定結果は大きくは変動しなかった。

#### 参考文献

- 1) Spielmann, H. et al., (1994) EEC/COLIPA Project on In Vitro Phototoxicity Testing: First Results Obtained with A BALB/C 3T3 Cell Phototoxicity Assay, *Toxicology in Vitro* 8, 793-796.
- 2) Spielmann, H. et al., (1998) The International EU/COLIPA In Vitro Phototoxicity Validation Study: Results of Phase II (Blind Trial). Part 1: The 3T3 NRU Phototoxicity Test, *Toxicology in Vitro* 12, 305-327.
- 3) Spielmann, H. et al., (1998), A Study on UV Filter Chemicals from Annex VII of European Union Directive 76/768/EEC, in the In Vitro 3T3 NRU Phototoxicity Test, *ATLA* 26, 679-708.

### 4-2-2-3) 試験法の信頼性

担当：小島肇夫、金子豊蔵

#### A. はじめに

試験の信頼性は、試験方法の評価に当たり最も重要な点である。開発された試験方法が、生体毒性検出における作用機構の解明に有用で、操作性、経済性に優れたスクリーニング試験方法であり、さらに代替法として動物やヒトに対する毒性検出に優れた試験法であったとしても、試験の再現性が明確でなければ信頼性の高い方法であるとは言えない。その信頼性を把握する上で、バリデーション試験の再現性を物質毎に施設内、施設間で比較することが一般的である。このデータの解析には、変動係数や散布図、箱ヒゲ図などの統計手法を用いて評価されるが、どのような評価基準で判断するかは統計学者により意見が分かれるところである。物質の毒性強度や特性によっても値は異なり、ケースバイケースの対応が必要と考えられる。

本章では、EC/COLIPA で光毒性検出のために Phase I からⅢまで 3 段階で実施された Balb/c3T3 細胞を用いた Neutral red 取込み試験のバリデーション研究結果が記載されている Spielman et al の文献<sup>1-4)</sup>に基づき、バリデーション研究における試験法の信頼性について、種々の視点からまとめた。ただし、本来、施設毎の生データを解析することにより、再現性を調べることが常道であり、これがもっとも重要な点であると考えられる。今回は残念ながら、以下に示す 4 報の文献結果からの解析であり、その評価に限界があることを認識する必要がある。

#### Phase I

- ・ EEC/COLIPA Project on In Vitro Phototoxicity Validation Study: First Results Obtained with a BALB/C 3T3 Cell Phototoxicity Assay, H. Spielmann et al., *Toxicology in Vitro* 8(1994) 793-796.
- ・ EEC/COLIPA In Vitro Phototoxicity Program: Results of the first Stage of Validation , H. Spielmann et al., Elsner P, Maibach HI(eds):Irritant Dermatitis, New Clinical and Experimental Aspects. *Cuur Probl Dermatol. Basel, Krager*(1995) 23, pp.256-264.

#### Phase II

- ・ The International EC/COLIPA In Vitro Phototoxicity Validation Study: Results of Phase II (Brind Trial). Part 1: The 3T3 NRU Phototoxicity Test, H. Spielmann et al., *Toxicology in Vitro* 12(1998) 305-327.

#### UV filter

- ・ A Study on UV Filter Chemicals from Annex VII of European Union Directive 76/768/EEC, in the In Vitro 3T3 NRU Phototoxicity Test, H. Spielmann et al., *ATLA* 26(1998) 679-708.

#### B. 施設内再現性

施設内の複数回の濃度依存曲線が記載されたデータは、Phase II の報告<sup>2)</sup>の Fig.4 に示された施設 9 の結果のみである。参加 9 施設の内、なぜこの施設のグラフを用いたか不明である。一般的に考えれば、著者の施設またはもっとも再現性の高かったデータが公表されるであろうと考えられる。この結果を見る限り、Chemical No.15 (Hexachlorophen) を除き、施設内の再現性は高いと思われる。Phase I においては、Spielmann et al (1994)<sup>1)</sup>や Spielmann (1995)

4)の報告に一機関のみの IC50 が記載されている他は、すべて PIF や MPE に変換された数値となり、IC50 を見るができない。変換された数値でも Phase II までは、施設内の繰り返しデータが記載されていない。

Phase III の報告<sup>3)</sup>では、4 施設の 2 回の結果が書かれている。それぞれの数値を見る限り、PIF が 3 倍を越える物質は 4 施設、20 被験物質、計 80 データの内、施設 1 の No.11 (Protoporphyrin IX disodium) 666.67 と 109.59、No.18 (Demeclocycline hydrochloride) 57.33 と 393、施設 3 の No.13 (Anthracene) 19.01 と 69.69 と悪い結果は少ない。

これらの結果と過去の試験経験から、本試験における施設内再現性は高いと考えられる。しかし、論文の結果からのみでは、施設内再現性を正確に判断できない。

### C. 施設間再現性

施設間の再現性については、Phase I において Irritant Dermatitis (1995)<sup>4)</sup> に箱ヒゲ図にて記載されている (Fig. 4-1)。施設間のバラツキの 95%信頼限界の幅が 5 倍を越えた物は promethadine のみであり、全体的なバラツキは小さいと考える。しかし、得られた結果の最大値と最小値で比較すると、20 被験物質中 6 物質 (No. 7:Tetracycline、1:Promethazine、12:Rose bengal、4:TSCA、10:Bithionol、17:Penicillin G) の最大値、最少値の差が大きいと考える。施設によっては、光毒性判定に戸惑う物質 (No.9:Amiodarone、10:Bithionol、13:Cinnamic aldehyde) が 3 つはあり、その中でばらつきの大きい No.10 の判定が施設により分かれていると考えられる。

Phase II の Toxicology in Vitro (1998)<sup>2)</sup> による物質毎のデータのバラツキを示す CV (Classification variability: 変動係数にあたる coefficients of variation ではない。)値は PIF で最高 18.8%、MPE で 20%である。値としては、大きくない。しかし、これは数字が求まったものの変換値より求めたものであり、IC50 の値ではない。PIF や MPE を求める段階で、紫外線照射の有無における IC50 のバラツキが非顕在化されている可能性もあるが、この論文の Table 3 において数字が求まっていない結果を見る限りでは、逆に紫外線未照射の IC50 が求まらないものとして計算に利用していることもあり、バラツキはより大きいと考える。これは、PIF が評価基準の 5 を越えるか否かを判断材料としているためか、紫外線未照射の IC50 を真剣に求めようと考えていないことにより、最高濃度を十分に上げないことに起因するのではないかと考える。PIF に絞ってみても、約半数の物質において、他の施設の結果から飛び離れた値が現れている。論文の Fig.5 からわかるように、Anthracene や Musk ambrette のようにバラツキが大きく、評価に影響している物質もある。

論文中でも、この原因は溶媒や設定濃度の問題、プロトコールの不備によるものと記載されているものの、すべての施設が誤評価する物質は試験の特性で済まされるが、1~3 施設による誤評価を見過ごしていいものか疑問である。

これらの点を考慮して実施された Phase III である、UV filter 物質のみを用いた試験<sup>3)</sup>では、施設内より施設間において変動係数(coefficients of variation)が悪いとされている。もっとも、この論文では data variability を問題にしていないと記載されている。この記載の意味は不明である (我々は大変重大な過失であると思うが...)。論文の Table III では、PIF が大きく (光毒性物質である)、その差が施設間で 10 倍以上である No.11 (Protoporphyrin IX disodium)、13 (Anthracene)、18 (Demeclocycline hydrochloride)、20 (Musk ambrette)のデータに問題

を感じる。また、Fig. 2 を見る限りでは、施設 1 (No. 1 : Octyl salicylate、2 : Octyl methoxycinnamate、7 : Polyacrylamidomethylbenzylidene camphor、18: Demeclocycline hydrochloride)及び施設 2 (No. 1、11、14: Acridine hydrochloride、18) の PIF、MPE のバラツキが 10 倍以上である。

以上の結果から、変換された数字による評価では、施設間の再現性は正確に把握できない。しかし、この数字を用いて結論をつけるならば、施設間の再現性は低いという印象がある。その原因も抽象的で、具体的には記載されておらず、たぶんデータ検討会やマネージメントチーム (MT) では議論されたであろうが、論文のみの評価ではこれ以上は無理であると感じられる。

#### D. 対照物質の再現性

論文中には、対照物質については記載されていない。OECD ガイドラインドラフトには Chlorpromazine を陽性対照とする記載がある。この物質はヒトにおける光毒性が明確であるとともに、本試験法でも紫外線照射の有無で明確な細胞毒性を示すこと、水溶性であり、扱いやすいことなどを考慮し、選ばれたと考えられる。この結果の再現性については、他の物質とともに次に示す。

#### E. 予測モデルによる評価結果の再現性

SLS、Promethazine、Chlorpromazine、Aminodarone、Bithionol の 5 物質が 3 回のバリデーションにおいて共通物質として検討されていることから、これらの物質のデータから評価結果の再現性を検討した。光毒性物質によっては、施設によるバラツキ大きいのが、被験物質の中ではいずれも PIF 値が落ち着いている。SLS の結果は極めてバラツキが小さい。ただ、これも PIF であり、試験毎の SLS の IC50 のバラツキはわからない。

Table 4-13 PIF の推移

被験物質名	Phase I の PIF ( ):データ数	Phase II の PIF ( ):データ数	Phase III の PIF ( ):データ数	光毒性
SLS	平均 1.5 (14)	1.2, 1.6, 1.0, 1.2, 1.3, 1.0, 1.2, 1.5, 1.1 (9)	1.78, 1.22, 0.93, 0.98 (4)	NP*
Aminodarone	平均 6 (9)	5.2, >2.7, 3.3, 7.7, 1.8, 14.9, 9.2, 1.6, >2.4 (9)	11.37, 12.61, 3.08, 7.07 (4)	PI**
Bitionol	平均 7 (13)	10.1, 59.8, 9.2, 7.2, >12.2, 13.2, 7.9, 16.7, 6.1 (9)	24.53, 17.97, 11.98, 11.14 (4)	PI
Chlorpromazine	平均 46.6 (13)	21.8, 42.7, 26.7, 60.9, 20.2, 35.3, 308.2, 18.9, 28.4 (9)	65.29, 38.40, 24.93, 19.02 (4)	PI
Promethazine	平均 78.5 (13)	17.3, 86.2, 53.1, 44.8, 13.6, 24.0, 47.5, 20.5, 496.6	72.41, 196.27, 29.18, 27.39 (4)	PI



		(9)		
--	--	-----	--	--

\* NP：非光毒性物質、\*\* PI：光毒性物質

#### F. 動物実験やヒト実験に対する in vitro 試験の再現性と信頼性

動物実験のバラツキが不明であることから、再現性を比較できない。ただ、施設内の動物実験の結果は、過去の経験からバラツキは大きくないを考える。これも施設内での比較データを検証したことがないため不明である。

また、結果を見る限り、動物実験によるヒト予測性は高い。明確なデータの食違いは、動物陰性でヒト陽性の Rose bengal のみである。

信頼性を確認するため、本試験とビボ試験の不一致数を調べてみた。Phase II にて、3 施設以上(3/11: 27%)食い違った被験物質数は PIF (Table3<sup>2)</sup>) で 4 物質(Chlorohexidine dihydrochloride, Furosemide, Bergamot oil, Amiodarone, 4/29: 13.8%)、MPE (Table5<sup>2)</sup>) で 2 物質(Hexachlorophene, Chlorohexidine dihydrochloride , 2/29: 6.9%)であり、光毒性の有無の判定での相違が起こる確率は少なかった。なお、これを両方の指標で判定し、conservative な立場から強い方を採用して判定すれば、MPE で 2 物質(2/29: 6.8%)の判定の間違いということになる。また、Phase III では、4 施設しか参加していないことから、2 施設以上(2/4: 50%)食い違った物質は、PIF、MPE とも 1 物質 (Tetraptalidene dicamphor sulphonic acid) のみであり (Table III、IV<sup>3)</sup>)、光毒性の有無の判定での相違が起こる確率は少なかった。

さらに、vivo との真度 (accuracy) を用い場合、Phase\_II では、PIF で 88% (219/256 (Table4<sup>2)</sup>)、MPE で 92% (227/256 (Table6<sup>2)</sup>) であった。Phase\_III でも、PIF で 93% (74/80)、MPE で 94% (75/80) (Table V<sup>3)</sup>) と高い数字を認めた。この数字をみると、MPE、PIF 値の施設による差は大きい場合もあるが、間違った評価をする確率は 1 割ほどであり、試験法の光毒性の有無の予知に関する信頼性は高いと判断できる。

#### G. 結論

施設内の再現性は高いが、施設間の再現性は高いとは言えない。しかし、バラツキが大きいからといって、ビボ試験との一致性が低いわけではなく、被験物質毎の不一致率や真度を見る限り、信頼性は高い試験方法であると思われる。

#### H. その他

繰り返しになるが、信頼性の評価は論文の表や図からはできない。生データの存在が不可欠である。また、そのデータをどのように扱うかを検討したデータ検討会や MT 会議の議事録からバリデーション研究のレベルを把握できる。生データの記録のない論文のみによる評価は、今後実施すべきでないを考える。

#### I. 要望

施設間のバラツキが大きく、弱い光毒性物質の評価には慎重な取り扱いを要する。このバラツキにより施設によっては誤った評価をする場合がある。Spielmann et al の論文<sup>2)</sup>の Table 4 から、PIF を用いた場合、false positive 数 8/45 はともかく、false negative 数 21/203 は多すぎると思われる。しかし、このうち 7 例は in vivo での光毒性が陽性であるとされた報告に

問題がある可能性が強い Furosemide についての結果であること、3例は MPE 法では9施設中1施設を除き、いずれも陽性と判断された非水溶性である Nalidixic acid-free acid であったこと、2例は同様に MPE 法ではいずれの施設でも陽性と判断された Ofloxacin の結果であったことから、両者を合わせて評価することが望ましい。また、強い光毒性物質である Anthracene について2施設がいずれの方法でも false negative となったがこれらの施設では EBSS を溶媒として用いていることによるものと思われたことから、溶媒の選択などの試験プロトコルや評価基準の見直し（グレーゾーンがあっても良い）が必要ではないかと考える。

#### 参考文献

- 1) Spielmann, H. et al., (1994) EEC/COLIPA Project on In Vitro Phototoxicity Testing: First Results Obtained with A BALB/C 3T3 Cell Phototoxicity Assay, *Toxicology in Vitro* 8, 793-796.
- 2) Spielmann, H. et al., (1998) The International EU/COLIPA In Vitro Phototoxicity Validation Study: Results of Phase II (Blind Trial). Part 1: The 3T3 NRU Phototoxicity Test, *Toxicology in Vitro* 12, 305-327.
- 3) Spielmann, H. et al., (1998), A Study on UV Filter Chemicals from Annex VII of European Union Directive 76/768/EEC, in the In Vitro 3T3 NRU Phototoxicity Test, *ATLA* 26, 679-708.
- 4) H. Spielmann et al. (1995) EEC/COLIPA In Vitro Phototoxicity Program: Results of the first Stage of Validation, Elsner P, Maibach HI (eds): *Irritant Dermatitis, New Clinical and Experimental Aspects*. *Curr Probl Dermatol*. Basel, Krager 23, pp.256-264.

Fig 4-1: 3T3-NR 法における光照射の有無による IC50 値の比 (PIF 値, 4-2-3 参照)の  
バラツキ (H. Spielmann et al. 1995)

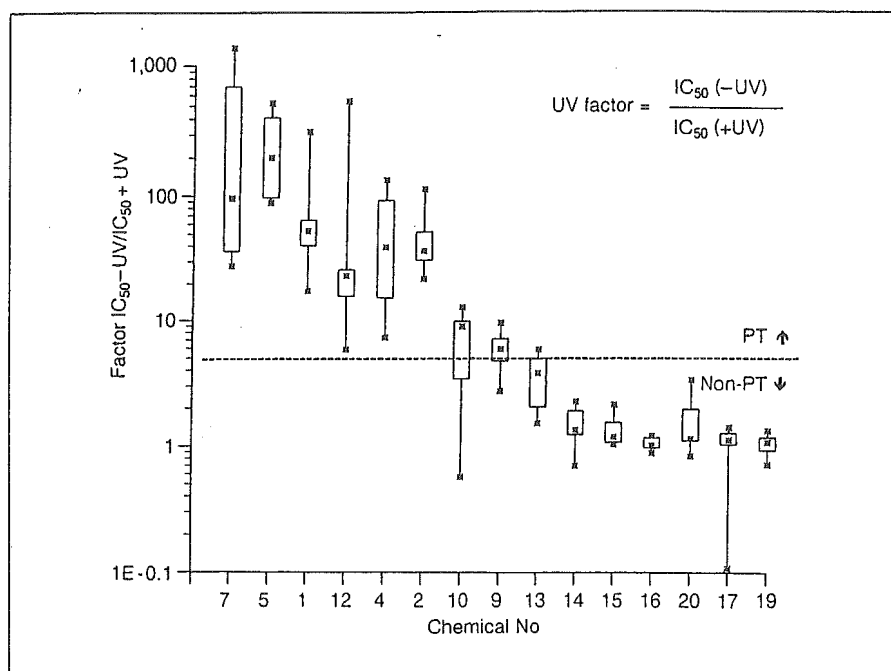


Fig. 2: 3T3 NRU phototoxicity assay: determination of UV photoactivation factors for 15 chemicals by discriminant analysis. UVA factors ( $IC_{50-UV}/IC_{50+UV}$ ) were determined with 15 chemicals, which were cytotoxic both, with and without UVA exposure. The numbers of the chemicals correspond to tables 1 and 2. The box plots show medians, 95% confidence limits (boxes) and minima/maxima of factors determined in all of the laboratories. The dashed line indicates the cut-off value determined by discriminant analysis to discriminate between nonphototoxic and phototoxic chemicals.

#### 4-2-2-4) 総合評価

担当：大森 崇

##### A. はじめに

ここでは、4-2-2-1)から 4-2-2-3)で個別に論じられた議論に基づき、光毒性試験の代替法として、Balb/c 3T3 細胞を用いた NR 法が妥当であるかどうかについて総合的な評価を行うことにする。

なお、必要な場合には補足を加えている。B 節ではデータの質について、C 節ではピボ試験結果との対応性について、D 節では試験間の信頼性について要約し、E 節で総合評価をまとめる。

##### B. データの質について

4-2-2-1)では、データの質について議論をした。3つの文献の中で、Phase II 研究はデータの質を評価する上で多くのことが記載されているが、他の2つの研究では、質を評価する上での記載は非常に少ない。データの質という点からは、文献のみで確認することは限界があるので、報告されている結果を確認することを目的に含めた研究を実施することが必要となるかもしれない。

##### C. ピボ試験結果との対応性について

ピボ試験結果との対応性として、我々は3つの研究のデータをまとめて、ピボ（ヒト、動物）と予測モデル（PIF、MPE）の関係を検討した。組み合わせにより異なるが、評価した被験物質は、26～38であった。評価指標として、感度、特異度、陽性予測力、陰性予測力、一致率を求めた。この結果、ヒト試験を優先した場合のヒトまたは動物の結果と PIF の関係は、感度 91.6%、特異度 78.6%、陽性予測力 88.0%、陰性予測力 84.6%、一致率 86.8%であり、MPE では、感度 94.7%、特異度 75.0%、陽性予測力 85.7%、陰性予測力 90.0%、一致率 87.0%と全体として高かった。この結果は、Balb/c 3T3 細胞を用いた NR 法は、化学物質の光毒性ポテンシャルを予測する上で有効な試験法であると思われる。

ただし、これらのバリデーション研究で用いられたデータは多くはなく、被験物質の溶解性の影響等について本試験方法の課題もある。

##### D. 試験結果の信頼性について

信頼性については、報告により記載が様々であった。施設内再現性については、UV Filter に関する Phase III 研究で評価することが望ましいと思われる。この研究は4つの施設で 20 物質を評価しており、報告には各施設で行われた2回の試験について個々の用量反応曲線から測定値から算出された PIF、MPE の値がそれぞれ記載されている。2つの PIF 値の比でみた場合、3倍を越える物質は、のべ 80 物質中わずか4物質であり（80 物質中1施設1物質が1回しか測定されていない）、再現性は悪くなかった。

施設間再現性については、Phase II 研究と UV Filter 研究で評価がされているが、両者で評価方法は異なっている。Phase II 研究では、得られたデータを復元抽出することにより用量反応曲線を繰り返し再現し、その変動を評価する CV（classification variability）という指標が導入され評価されている。11 施設で 30 物質を評価した Phase II 研究では、CV の最高値は、PIF で 18.8%、MPE で 20.0%となっており、それほど大きな値とはなっていない。一方、UV Filter