

厚生労働科学研究費補助金（化学リスク研究事業）  
化学物質の動物個体レベルの免疫毒性データ集積とそれに基づくMulti-ImmunoTox  
assay（MITA）による予測性試験法の確立と国際標準化  
令和2年度分担研究報告書  
分担研究報告書

免疫毒性評価試験法Multi-ImmunoToxicity assayの予測性の検討

研究分担者 大森崇  
神戸大学医学研究科 生物統計学分野

### 研究要旨

**【背景と目的】** Multi-Immuno Tox Assay(以下、MITA)は、化学物質がヒトの免疫系に与える影響を評価することを目標とする *in vitro* の試験法である。現在、バリデーション研究が実施されている IL-2 Luc アッセイ及び IL-1 $\beta$  Luc アッセイは高い再現性を有する試験法であるが、両者を組み合わせた信頼性の検討は行われてこなかった。本検討では、2つの試験法を組み合わせた予測性を検討する。

**【方法】** 63物質を用いて、*in vivo* 試験の判定結果と IL-2 Luc アッセイ及び IL-1 $\beta$  Luc アッセイの判定結果を用いた5つの方法について、Receiver Operation Characteristics（ROC）曲線の曲線下面積である Area Under the Curve（AUC）の値とその95%信頼区間とともに感度、特異度を計算した。

**【結果】** 5つの方法のうち、最もAUCが高かった方法は2つのアッセイを組み合わせた方法において、AUCは0.59（95%信頼区間 [0.44, 0.75]）、感度77.6%、特異度14.3%であった。

**【結論】** この検討結果から、検討した物質において高い信頼性の結果は得られなかったが、結果は暫定的であり、評価を行う物資を含めて、信頼性の評価のためにさらなる検討が必要であることがわかった。

### A. 研究目的

Multi-Immuno Tox Assay(以下、MITA)は、化学物質が免疫系に与える影響を簡便に評価するための光レポーターを利用した *in vitro* 免疫毒性評価試験法である。現在、IL-2 Luc アッセイ及び IL-1 $\beta$  Luc アッセイを、経済協力開発機構（Organisation for Economic Co-operation and Development : OECD）の試験法ガイドライン（Test Guideline : TG）としての公定化を目指し、バリデーション研究が施行されている。

IL-2 Luc アッセイ及び IL-1 $\beta$  Luc アッセイでは、特定の化学物質の発現を評価は独立した3回の実験から得られる測定値を用いて行われる。IL-2 Luc アッセイでは、最終的には immunosuppression、

no effect、Immunoaugmentation のいずれかの判定がされる。一方、IL-1 $\beta$  Luc アッセイでは最終的には immunosuppression、no effect のいずれかの判定がされる。バリデーション研究を行う中で、このアッセイに関する実験方法だけでなく、判定方法についても検討が行われてきた。判定方法に関しては、これまでに、

- ・特定のある一濃度で発現が生じることがあるため、濃度別の統計的有意性のみで判定を行うと高感度になりすぎる
- ・濃度依存性のある程度考慮したい
- ・他の試験法の判定で行われてきたように3回の実験の判定は独立に行うことが望ましい
- ・特定の濃度以上は判定を行わない

という意見に基づき、それぞれのアッセイに対する判定方法が構築されてきた。

新たに開発された試験法が特定の毒性の判定を行うことが可能かどうかの判断を行う際に、再現性（信頼性）と予測性という異なる2種類の評価が必要になる。再現性は、同一の物質を同一の試験方法で適用したときの施設内再現性、施設間再現性の評価が求められる。予測性は、ターゲットとする毒性の判定と新たに開発された試験法の判定がどの程度一致するかの評価が求められる。

本研究では、IL-2 Luc アッセイ及び IL-1 $\beta$  Luc アッセイ単独および組み合わせた際の予測性についての評価を目的とした。

## B. 研究方法

### 利用したデータ

検討に利用したデータは、1施設で実施された IL-2 Luc アッセイ及び IL-1 $\beta$  Luc アッセイの結果及び、文献により集められた *in vivo* の判定結果を含めた 63 物質（陽性物質 49 物質、陰性物質 14 物質）のデータある。

### 検討した判定方法

予測性評価は、

1. IL-2 Luc アッセイ単独の判定
2. IL-1 $\beta$  Luc アッセイ単独の判定
3. IL-2 Luc アッセイと IL-1 $\beta$  Luc アッセイ単独がともに陽性と判定した場合に陽性、そうでない場合に陰性とした判定
4. IL-2 Luc アッセイと IL-1 $\beta$  Luc アッセイ単独のどちらか陽性と判定した場合に陽性、そうでない場合に陰性とした判定
5. ロジスティック回帰分析の説明変数に IL-2 Luc アッセイの判定結果と IL-1 $\beta$  Luc アッセイ単独の判定結果を含めて予測した場合の判定とした。

### 効果の評価指標

上記の 1~5 について、効果の評価指標は、Receiver Operation Characteristics (ROC) 曲線の曲線下面積である Area Under the Curve (AUC) の値とその 95%信頼区間とともに感度、特異度とした。上記の方法 5 についての感度と特異度を求めるカットオフ値は、感度が 100%未満の最大値となる値とした。

### データの取り扱い

上記の評価において、IL-2 Luc アッセイの判定は、immunosuppression、Immunoaugmentation を「陽性」として、no effect を「陰性」とした。また、IL-1 $\beta$  Luc アッセイについては、immunosuppression、を「陽性」として、no effect を「陰性」とした。

### 解析に用いたソフトウェア

本研究の解析には統計ソフト SAS (version 9.4) を用いた。

## C. 研究結果

5つの方法のそれぞれにおける ROC 曲線の曲線下面積である AUC とその 95%信頼区間を表 1 に示す。いずれの方法においても AUC が 0.6 より大きな値となるものはなく、95%信頼区間の下限は 0.5 を下回っていた。

表 1 5つの方法の AUC

方法	AUC	95% Confidence	
		Lower	Upper
1	0.56	0.45	0.68
2	0.57	0.42	0.71
3	0.59	0.44	0.73
4	0.54	0.43	0.65
5	0.59	0.44	0.75

表 2 に 5 つの方法の感度および特異度を示す。80%以上の感度を得ることができなかった。

表 2 5 つの方法の感度 (%) と特異度 (%)

方法 5 は感度が 1 未満の最大となるようにカットオフ値を決めた場合

方法	感度 (%)	特異度 (%)
1	73.5	14.3
2	51.0	35.7
3	46.9	35.7
4	77.6	14.3
5	77.6	14.3

また、各方法の ROC 曲線を付録に示した。

#### D. 考察

本研究では、IL-2 Luc アッセイ及び IL-1 $\beta$  Luc アッセイの予測性の検討を行った。本検討では、評価の指標として AUC を用いた。予測性は *in vivo* 試験の結果と *in vitro* 試験の関連の強さを評価する指標である。関連が見いだせず、偶然誤差の変動である場合、AUC は 0.5 となることが知られている。つまり、AUC が 0.5 に近い評価結果で高い予測性を有するとはいえないことになる。本検討では、AUC の 95%信頼区間を算出したが、95%信頼区間が 0.5 を含むことは、本研究で検討したデータセットにおいて、AUC が 0.5 であることを否定できないことを意味する。評価した 5 つの方法のい

ずれにおいても 95%信頼区間が 0.5 を含んでいた。AUC は全体の評価を与えるが、感度と特異度は、*in vivo* 試験の結果が陽性の場合と陰性の場合のそれぞれについての予測能を示す指標である。どの方法であっても、感度は 80%を超えることはなく、特異度は 50%に満たなかった。検討した 5 つの方法のうち、方法 1 が IL-2 Luc アッセイ単独であり、方法 3 から 5 が IL-2 Luc アッセイに IL-1 $\beta$  Luc アッセイを組み合わせた方法となる。したがって、本検討の範囲では、IL-2 Luc アッセイと IL-1 $\beta$  Luc アッセイを組み合わせても、予測性の改善はあまり大きくないといえる。

なお、本研究の検討では、IL-2 Luc アッセイの判定は、Immunosuppression、Immunoaugmentation を「陽性」として、no effect を「陰性」とした。方法 5 ではロジスティック回帰分析を行っているため、Immunosuppression、Immunoaugmentation、no effect の 3 カテゴリーにした場合も検討したが、結果が改善することはなかった。

ここでの検討結果のみから 2 つアッセイの予測性が悪いと結論づけるべきではないと思われる。特異度が特に低かった理由は、2 つアッセイの結果が陽性と判定された物質が多かったことによる。陽性物質が 49 であったのに対し、陰性物質はわずか 14 物質である。このうちの多くの物質が 2 つのアッセイで陽性と判定されていた。予測性の性能評価の前に行われた再現性の検討では、2 つのアッセイとも、陰性物質は良好な施設間再現性の下、用量が増加しても反応を示さなかった。したがって、2 つのアッセイがどのような物質にも反応を示すアッセイであるとは考えにくい。より広い化学物質に対して、これらのアッセイを適用した場合の結果の特異度は異なるものになる可能性がある。*in vivo* 試験の陽性となる物質が何を検出しているのかということもさらなる検討が必要であると思わ

れるが、検討に用いた 63 物質より広い化学物質の中で *in vivo* 試験で確認するまでもないとして評価されてこなかった物質は、検討した物質には含めることができているだけなのかもしれない。

感度に注目した場合、感度も決して高いという結果ではなかった。これらは過去 2 年で検討した判定方法のカットオフ値と関係があるかもしれない。

本検討の範囲には含まれていないが、IL-2 Luc アッセイ及び IL-1 $\beta$  Luc アッセイのどちらも高い施設内、施設間再現性を示しているアッセイである。本検討から、予測能は現段階ではまだ試験法の課題を残しているといえるが、より広い物質を含めたさらなる検討が必要である。

2012, SAS Institute.

#### F. 健康危険情報

なし。

#### G. 研究発表

Kimura Y, Yasuno R, Watanabe M, Kobayashi M, Iwaki T, Fujimura C, Ohmiya Y, Yamakage K, Nakajima Y, Kobayashi M, Mashimo N, Takagi Y, Omori T, Corsini E, Germolec D, Inoue T, Rogen EL, Kojima H, Aiba S. An international validation study of the IL-2 Luc assay for evaluating the potential immunotoxic effects of chemicals on T cells and a proposal for reference data for immunotoxic chemicals. *Toxicol In Vitro*. 2020. doi: 10.1016/j.tiv.2020.104832.

#### H. 知的財産権の出願・登録状況

なし。

#### 参考文献

1) Paul D. A. Logistic Regression Using SAS: Theory and Application, Second Edition SAS,

付録

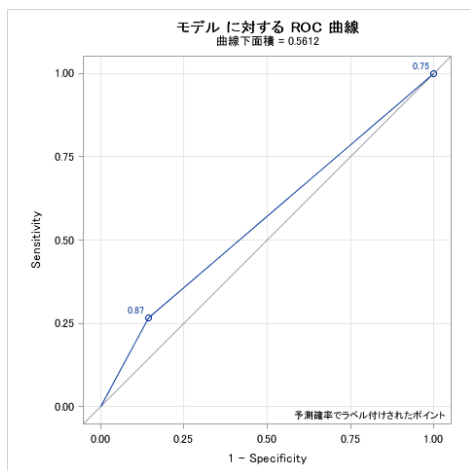


図 A1 方法 1 の ROC 曲線

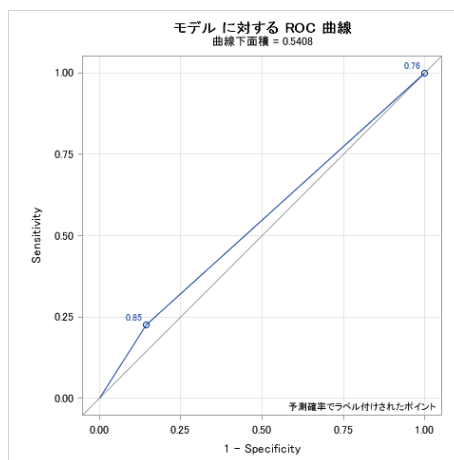


図 A4 方法 4 の ROC 曲線

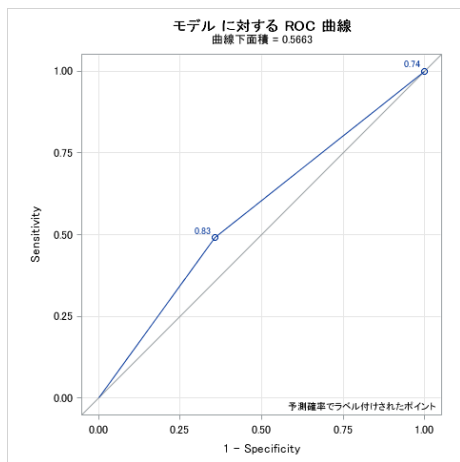


図 A2 方法 2 の ROC 曲線

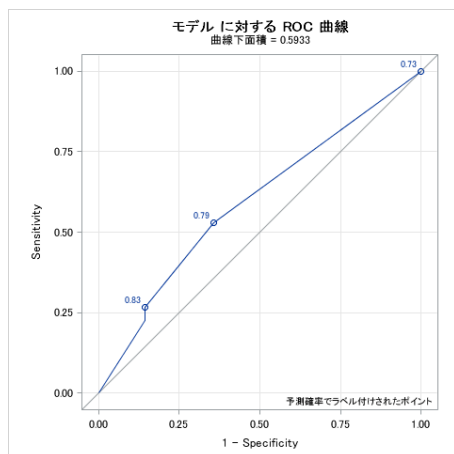


図 A5 方法 5 の ROC 曲線

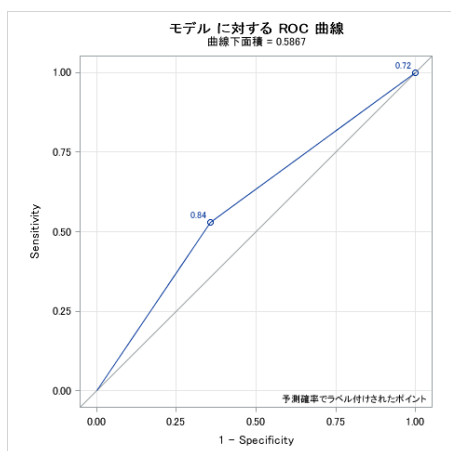


図 A3 方法 3 の ROC 曲線