

Table 13. Results of bacterial reverse mutation test of phenylboronic acid
(Activation method: +S9)

Compound	Dose (µg/plate)	Mean revertant colonies per plate				
		TA100	TA1538	WP2uvrA	TA98	TA1537
DMSO	0	114	11	28	39	17
Phenylboronic acid	39.1					16
	78.1					17
	156	121	13	24	38	17
	313	99	7	22	28	17
	625	110	9	25	34	24
	1250	129	8	26	37	31
2500	151	10	34	40	35	
5000	148 *	10	38	39	69	
Positive control compound	2-AA	2-AA	2-AA	2-AA	2-AA	2-AA
Dose (µg/plate)	1	2	10	0.5	2	
Mean revertant colonies per plate	1102	351	766	495	194	

DMSO: Negative control (Dimethyl sulfoxide, 100 µl/plate)
2-AA: 2-Aminanthracene
*: Growth inhibition was observed.

Phenylboronic acid 処理群の復帰変異コロニー数は、-S9 処理および+S9 処理の TA1537 株においてのみ、陰性対照群の 2 倍以上を示す増加が認められたが、明確な用量依存性は認められなかった。試験菌株に対する生育阻害作用は、-S9 処理の TA100 および TA1537 株の 5000 µg/プレートならびに +S9 処理の TA100 株の 5000 µg/プレートにおいて認められた。

変異原性強度の相対的比較値である比活性 (mg/プレート当たり) は下表の通り算出された。なお、用量設定試験の+S9 処理において、比活性の最高値が 332 と算出されたが、51.2 µg/プレート付近で認められた陽性反応は明確な用量依存性を伴わないことから、2000 µg/プレートで算出された比活性値 11 が妥当であると考えられる。

比活性値				
試験	S9	菌株	用量	比活性
用量設定試験	-	TA1537	128 µg/プレート	55
	+	TA1537	51.2 µg/プレート	332
本試験	-	TA1537	78.1 µg/プレート	51
	+	TA1537	5000 µg/プレート	10

d. 4-Pyridineboronic acid

用量設定試験結果を Figure 23、24、Table 14、15 に示す。

4-Pyridineboronic acid 処理群では、-S9 処理の TA1537 株および+S9 処理の WP2uvrA 株において、最高用量の 5000 µg/プレートで陰性対照群の 2 倍を超える復帰変異コロニー数の増加が認められた。-S9 処理の TA100 および WP2uvrA 株においても

復帰変異コロニー数の増加傾向が認められたが、どちらも陰性対照群値の 2 倍未満であった (1.9 倍)。-S9 および+S9 処理のその他の試験菌株においては、復帰変異コロニー数の明確な増加は認められなかった。また、試験菌株に対する生育阻害作用は、いずれの菌株においても認められなかった。

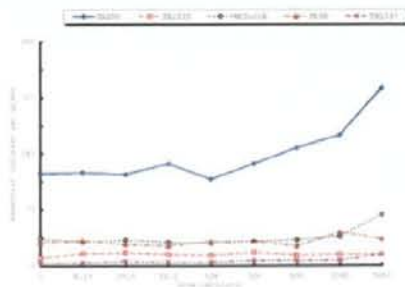


Figure 23. Dose-finding study of 4-pyridineboronic acid (Activation method: -S9)

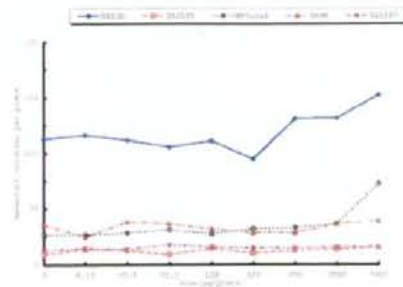


Figure 24. Dose-finding study of 4-pyridineboronic acid (Activation method: +S9)

Table 14. Results of dose-finding study of 4-pyridineboronic acid [Non-activation method: -59]

Compound	Dose (pg/plate)	Mean revertant colonies per plate				
		TA100	TA1535	WP2uvrA	TA98	TA1537
D.W.	0	115	10	33	28	4
4-Pyridineboronic acid	8.19	116	15	29	31	4
	20.5	114	16	32	26	6
	51.2	127	14	29	24	5
	128	108	13	28	30	5
	320	127	17	30	30	7
	800	147	13	32	24	7
	2000	163	14	36	41	7
5000	222	14	63	33	14	
Positive control compound		AF-2	NaN ₁	AF-2	AF-2	9-AA
Dose (pg/plate)		0.01	0.5	0.01	0.1	80
Mean revertant colonies per plate		610	637	166	662	266

D.W.: Negative control (Water for injection, 100 µL/plate)
 AF-2: 2-(2-furyl)-3-(5-nitro-2-furyl)acrylamide NaN₁: Sodium azide 9-AA: 9-Aminoacridine hydrochloride

Table 15. Results of dose-finding study of 4-pyridineboronic acid [Activation method: +59]

Compound	Dose (pg/plate)	Mean revertant colonies per plate				
		TA100	TA1535	WP2uvrA	TA98	TA1537
D.W.	0	124	10	29	39	14
4-Pyridineboronic acid	8.19	128	15	29	27	16
	20.5	123	14	31	42	15
	51.2	116	10	34	40	19
	128	122	16	30	35	17
	320	104	11	35	31	16
	800	144	14	36	31	16
	2000	145	15	40	40	17
5000	168	17	80	43	17	
Positive control compound		2-AA	2-AA	2-AA	2-AA	2-AA
Dose (pg/plate)		2	2	10	0.5	2
Mean revertant colonies per plate		801	392	890	434	199

D.W.: Negative control (Water for injection, 100 µL/plate)
 2-AA: 2-Aminoanthracene

本試験結果を Figure 25~29、Table 16、17 に示す。

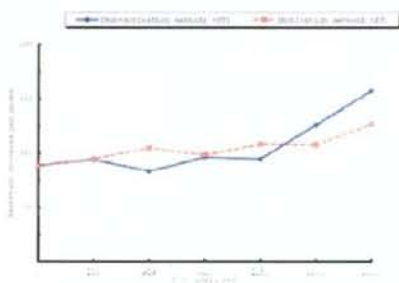


Figure 25. Bacterial reverse mutation test of 4-pyridineboronic acid in strain TA100

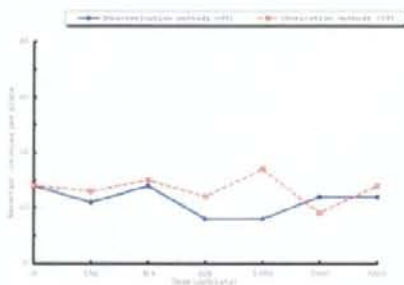


Figure 26. Bacterial reverse mutation test of 4-pyridineboronic acid in strain TA100

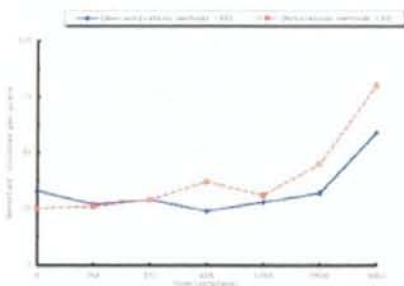


Figure 27. Bacterial reverse mutation test of 4-pyridineboronic acid in strain TA100

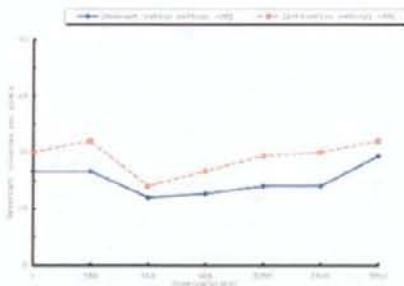


Figure 28. Bacterial reverse mutation test of 4-pyridineboronic acid in strain TA100

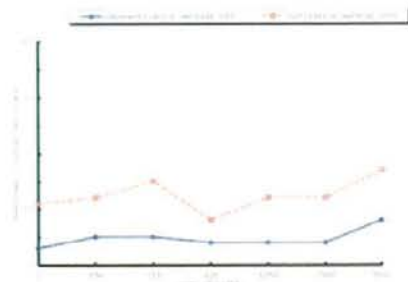


Figure 16. Bacterial reverse mutation test of 4-pyridineboronic acid to strain TA1537.

Table 16. Results of bacterial reverse mutation test of 4-pyridineboronic acid [Non-activation method: -S9]

Compound	Dose (µg/plate)	Mean revertant colonies per plate				
		TA100	TA1535	WP2uvrA	TA98	TA1537
D.W.	0	125	14	33	25	3
4-Pyridineboronic acid	156	132	11	27	25	5
	313	118	14	29	19	5
	625	135	9	24	19	4
	1250	133	8	28	21	4
	2500	177	12	32	21	4
5000	221	12	59	29	9	
Positive control compound		AF-2	NaN	AF-2	AF-2	9-AA
Dose (µg/plate)		0.01	0.5	0.01	0.1	80
Mean revertant colonies per plate		744	505	166	723	266

D.W.: Negative control (Water for injection, 100 µL/plate)
 AF-2: 2-(2-Furyl)-3-(5-nitro-2-furyl)acrylamide NaN: Sodium azide 9-AA: 9-Aminoacridine hydrochloride

Table 17. Results of bacterial reverse mutation test of 4-pyridineboronic acid [Activation method: +S9]

Compound	Dose (µg/plate)	Mean revertant colonies per plate				
		TA100	TA1535	WP2uvrA	TA98	TA1537
D.W.	0	124	14	25	30	11
4-Pyridineboronic acid	156	133	13	26	32	12
	313	147	15	29	21	15
	625	130	12	37	25	9
	1250	152	17	31	29	12
	2500	151	9	45	30	12
5000	170	14	80	33	17	
Positive control compound		2-AA	2-AA	2-AA	2-AA	2-AA
Dose (µg/plate)		1	2	10	0.5	2
Mean revertant colonies per plate		1149	324	833	397	170

D.W.: Negative control (Water for injection, 100 µL/plate)
 2-AA: 2-Aminanthracene

4-Pyridineboronic acid 処理群では、-S9 処理の TA1537 株および+S9 処理の WP2uvrA 株において、最高用量の 5000 µg/プレートで陰性対照群の 2 倍を超える復帰変異コロニー数の増加が認められた。-S9 処理の TA100 および WP2uvrA 株においても復帰変異コロニー数の増加傾向が認められたが、どちらも陰性対照群値の 2 倍未満であった (1.8 倍)。

-S9 および+S9 処理のその他の試験菌株においては、復帰変異コロニー数の明確な増加は認められなかった。また、試験菌株に対する生育阻害作用は、いずれの菌株においても認められなかった。

変異原性強度の相対的比較値である比活性 (mg/プレート当たり) は下表の通り算出された。

比活性値				
試験	S9	菌株	用量	比活性
用量設定試験	-	TA1537	5000 µg/プレート	2

本試験	+	WP2uvzA	5000 µg/プレート	10
	-	TA1537	5000 µg/プレート	1
	+	WP2uvzA	5000 µg/プレート	11

1.4. 考察及び結論

Hafnium(IV) chloride では、最高用量には生育阻害が認められると考えられる用量またはガイドラインに規定されている上限の 5000 µg/プレートを設定し、試験を行った。その結果、Hafnium (IV) chloride 処理群では、代謝活性化系非存在下 (-S9 処理) および代謝活性化系存在下 (+S9 処理) のいずれにおいても、陰性対照群の 2 倍以上を示す復帰変異コロニー数の増加は認められなかった。

これら両処理法での用量設定試験または同追加試験における試験結果は、本試験において再現された。

陰性対照群の平均復帰変異コロニー数は、無処理対照群と比較して明確な差が認められなかった。ことから、アセトニトリルの試験系への影響はないと考えられた。

以上の結果から、当該試験条件下において、Hafnium (IV) chloride の遺伝子突然変異誘発性は陰性と判定された。

Hafnocene dichloride では、最高用量にガイドラインに規定されている上限の 5000 µg/プレートを設定し、試験を行った結果、Hafnocene dichloride 処理群では、代謝活性化系非存在下 (-S9 処理) および代謝活性化系存在下 (+S9 処理) のいずれにおいても、陰性対照群の 2 倍以上を示す復帰変異コロニー数の増加は認められなかった。

以上の結果から、当該試験条件下において、Hafnocene dichloride の遺伝子突然変異誘発性は陰性と判定された。

Phenylboronic acid では、最高用量には、ガイドラインに規定されている上限の 5000 µg/プレートを設定し、試験を行った。その結果、Phenylboronic acid 処理群では、代謝活性化系非存在下 (-S9 処理) および代謝活性化系存在下 (+S9 処理) の TA1537 株においてのみ陰性対照群の 2 倍以上を示す復帰変異コロニー数の増加が認められた。

これら両処理法での用量設定試験における試験結果は、本試験において再現された。

以上の結果から、当該試験条件下において、Phenylboronic acid の遺伝子突然変異誘発性は陽性と判定された。

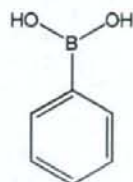
4-Pyridineboronic acid では、-S9 処理の TA1537 株および+S9 処理の WP2uvzA 株において、最高用

量の 5000 µg/プレートで陰性対照群の 2 倍を超える復帰変異コロニー数の増加が認められた。-S9 処理の TA100 および WP2uvzA 株においても復帰変異コロニー数の増加傾向が認められたが、どちらも陰性対照群値の 2 倍未満であった。(1.9 倍)

これら両処理法での用量設定試験における試験結果は、本試験において再現された。変異原性強度の相対的比較値である比活性の最大値は、+S9 処理、WP2uvzA 株の 5000 µg/プレートで 11 (mg/プレート当たり) と計算された。

以上の結果から、当該試験条件下において、4-Pyridineboronic acid は遺伝子突然変異誘発性を示すもの (陽性) と判定された。

上記 4 化合物の内、Hafnium に塩素が結合した 2 化合物とも陰性を示したが、ニトロ基を有した Hafnium 化合物では陽性(私信)となっていることから、更なる類似化合物の検索が必要と考えられた。また、ベンゼン核にホウ素が結合した 2 化合物とも陽性を示しており、更にフッ素が結合した化合物も陽性になっていること(私信)から、下記の構造を持つ化合物について AMES 試験陽性のアラートとして考えられた。



2.28 日間反復投与毒性試験データに基づく、肝毒性および腎毒性に対する Derek for Windows 予測精度改善のための Rapid Prototype 手法開発に関する研究 [和訳]

2.1. 概要

Derek for Windows は毒性予測のためのエキスパートシステムである。その知識ベースは、そのシステムにより行われる毒性予測にそれぞれ寄与するアラート、例証化合物および Reasoning rule から成り立っている。現行プロジェクトの目的は Derek for Windows の肝毒性および腎毒性の予測精度を改善するための Rapid Prototype 戦略を開発することである。

当プロジェクトにて Rapid Prototype 手法により開発されたアラートは、化合物構造と国立医薬品食品衛生研究所（以下、NIHS と記す）により提供された 28 日間反復投与毒性試験データから単純に導出され、それらの結果を支持するための更なる刊行済み文献調査は行っていない。もしも、肝臓病理組織欄において陽性を示す試験値があれば、その化合物は肝毒性に対して陽性であると考えられ、また腎臓病理組織欄において陽性を示す試験値があれば、腎毒性に対して陽性であると考えられる。

潜在的な Rapid Prototype アラートを特定するために三通りの手法が使用された。その三通りの手法とは、(i)化合物構造クラスタリングツールの使用、(ii)トレーニングデータセットの視覚的解析、(iii)関連する毒性エンドポイントに対しての既存の Derek for Windows アラート比較である。調査の後、肝毒性について 34 個の Rapid Prototype アラート、および腎毒性について 35 個の Rapid Prototype アラートが開発された。これら 69 個の Rapid Prototype アラートは Derek for Windows Version 11 の知識ベースにカスタマイズにて実装された。

この知識ベースは、トレーニングデータセット化合物に対して、肝毒性予測について 61%の感度および 80%の特異性を有し、また腎毒性予測についてはそれぞれ 62%、86%であった。

二つの外部データソースを使用して、このプロジェクトにて開発された Rapid Prototype アラートを評価したところ、肝毒性・腎毒性双方の感度・特異性の減少が認められた。この調査結果は、幾つかの Rapid Prototype アラートの適用範囲の改良、および/または、適用領域の定義が予測精度の改善に役立つかも知れないということを示唆している。加えて、病理組織学的所見とともに臨床的指標を考慮することが更なる見識を与えるかも知れない。

当プロジェクトは、ある適切なトレーニングデータセットが与えられると、Rapid Prototype 手法が Derek for Windows の知識ベース内の既存および新規の毒性エンドポイントに対するアラートの予測精度の向上に使用できるということを示した。新しいテストデータセットの有効性は、これらの Rapid Prototype アラートを更に保障もしくは改良するために使用でき、Derek for Windows の標準アラートに対して、それに続く開発を可能にする。適用領域の考慮といった追加的な技法が予測の信頼度を説明するために、また、各々の Rapid Prototype アラートの適合性および品質の調査を支援するために使用されるかも知れない。

2.2. 序論

Derek for Windows は毒性予測のための知識ベースエキスパートシステムである。その知識ベースは、そのシステムにより行われる毒性予測にそれぞれ寄与するアラート、例証化合物および Reasoning rule より成り立っている。国立医薬品食品衛生研究所(NIHS)との共同プロジェクトは、28 日間反復投与毒性試験データが肝毒性に対する予測精度を向上するための Derek for Windows のアラートを導出するために首尾よく使用できるということを示した。

さらに、染色体異常に対する Rapid Prototype 手法が、NIHS より提供された in vitro 染色体異常試験データに基づき、Derek for Windows における染色体異常の予測精度を改善するために、首尾よく応用された。現行プロジェクトの目的は Derek for Windows の肝毒性および腎毒性予測精度を改善するための 28 日間反復投与毒性試験データに基づく Rapid Prototype 手法を開発することである。

Derek for Windows における肝毒性の被覆率が、NIHS 反復投与毒性プロジェクトの一年目において、NIHS およびその他の組織との共同作業にしたがって、2008 年 11 月の最終リリース版 (Derek for Windows version 11) に対して大規模に開発された。

知識ベースは現在、40 個の肝毒性アラートを有している。しかしながら、腎毒性は系統的に詳しく調査されておらず、最新版の Derek for Windows 知識ベースにおいて一つしかない (α -2-mu-globulin nephropathy を除く)。私有のデータセットからの腎毒性アラート開発のための Rapid Prototype 手法は、腎毒性の予測精度を迅速かつ効果的に改善する可能性を秘めている。

2.3. 方法

NIHS はまず、808 個の化合物に対する毒性試験データを要約した ISIS/Base データベース (NIHS Ames database 1.db) を提供した。このデータは、NIHS 反復投与毒性試験プロジェクトの一年目に、肝毒性アラートを導出するために使用された。このデータベースはコピーされ、そして、一年目に開発された肝毒性アラートのために使用された。専門家要約報告書から得られたデータを使用して更新された (NIHS_database_UPDATED.db)。NIHS により提供されたこれら専門家要約報告書の幾つかの化合物は、評価されたことが無く、また、データベース内のそれら化合物のデータの欠落が実験的に陰性を意味しているわけではない、ということを示した。さらに、データベース内に数多くの重複した化合物構造が認められた。これら重複化合物は、Rapid Prototype 開発において Derek for Windows の予測精度調査に影響を及ぼすと思われ、従って、データベースから取り除かれた。残っている化合物に対す

る化学構造は、視覚的解析を容易にするために必要に応じて、調査および整理された結果として得られた。データベース (NIHS_UPDATED_Without_duplicates_FINAL_280408.db) は、全部で 776 個の化合物を有し、Rapid Prototype 開発のデータソースとしてまず使用された。このデータベースは 29 個の Rapid Prototype 肝毒性アラートを導出するために使用された。

そのデータベースはプロジェクト中に NIHS により新しい更新に従って更に変更された。その結果得られたデータベース (NIHS5_20080916.db) には 731 個の化合物が含まれている。既存の 29 個の Rapid Prototype 肝毒性アラートは、この新しいデータベースに対する再評価の後に改定された (Figure 30 を参照)。残っている肝毒性 Rapid Prototype アラートと全ての腎毒性 Rapid Prototype アラートが、このデータベースを使用して開発された。

肝毒性および腎毒性が陽性であるための定義は、

Table 18 ChemTK によるデータセット解析において使用された設定

手法	設定
環	Min/Max = 1/3、 2/3* and 3/3
立体構造形成能を持つフラグメント	Min/Max fragment size = 3/10
枝分かれしたフラグメント	Min/Max atoms in fragment = 4/8 Min/Max number of branched points = 0/2* Min/Max number of bonds between terminal atoms = 1/7*

SD ファイルが、化学構造・ID 番号および付加された。活性値欄を含むデータベース

(NIHS_UPDATED_Without_duplicates_FINAL_280408.db) から作成された。活性値欄は肝臓の病理組織学値に対応しており、毒性が陽性である化合物に対しては値"1"が与えられ、陰性である化合物に対しては"0"が与えられた。その後、その SD ファイルは ChemTK で処理された。

遮断値が、潜在的に Rapid Prototype アラートに変換することができる化合物クラスを選択するために選ばれ、この値は 55% もしくはより多くの陽性化合物が化学物質クラス内に含まれるように設定された。

もしも、特定されたクラスが 55% 以下の陽性化合物しか有していなかった場合、そのクラスはこの解析からは除外した。残っている化合物クラスに対して、トキシコファーの範囲を導出するための更なる解析 (例えば、ISIS/Base データベース内での部分構造検索など) が行われ、その後 Derek for Windows

臨床的指標や病理組織学的な効果が認められる投与量に関係なく、データベース内の肝臓と腎臓の病理組織学的値の有無に基づく。このことから、病理組織学的欄に陽性値が存在する場合には化合物は陽性に分類し、値が無い場合には陰性に分類した。

Rapid Prototype 手法に対する最初の候補クラスが、クラスタリングソフトウェア"ChemTK"を使用して特定された。そのソフトウェアは、定義済みの設定 (環、枝分かれした原子、立体構造形成能を持つフラグメント) に基づき与えられたデータセット中の化合物クラスの迅速な検索に対して有用なツールであり、また、各々の化合物クラスに対して平均的な活性に関する情報を用意することも可能である。解析において使用された設定は Table 18 に示されている。

に実装された。

また、データセットに対して 55% もしくはより大きな基準値に適合しなかった化合物クラスを更に特定するために視覚的解析が行われた。しかし、それはまだ潜在的に有効な Rapid Prototype アラートとなるかも知れない。候補クラスは、官能基や類似な化学的環境を考慮すること、および、Derek for Windows 知識ベース内の他の毒性エンドポイントに対する既存トキシコファーとのアナロジーによって特定された。特定された化合物クラスは先に記載した方法で処理され、Rapid Prototype アラートが Derek for Windows の知識ベース内に実装された。

上記とは異なる手法が、腎毒性 Rapid Prototype アラートを開発するために採用された。まず、既存の肝毒性アラート (Derek for Windows の先行バージョンである version10 と Rapid Prototype アラート) が腎毒性予測の判断材料として調査された。NIHS によって提供されたデータセットは、肝毒性 Rapid Prototype アラートを含む知識ベースに対し

て予測が実施され、腎毒性予測に対する肝毒性アラートの子測精度が調査された。最良の子測判断材料（化合物クラス内の 55%を超える化合物を正しく子測したもの）が選ばれ、腎毒性 Rapid Prototype アラートとして開発された。さらに、視覚的解析が、肝毒性アラートに対して使用されたのと同じ基準を使用して、腎毒性に対する Rapid Prototype アラートをさらに開発するために行われた。

Derek for Windows の知識ベース内の Rapid Prototype アラートの実装は、pattern、alert description、picture およびアラートの適用範囲を定義するための簡素なコメントから構成される。

2.4. 結果

全部で69個の Rapid Prototype アラートが開発され、Derek for Windows version11 知識ベースに

実装され、その知識ベースはこの報告書において "DfW11 NIHS 201108_1"と参照される34個の肝毒性 Rapid Prototype アラートと、35個の腎毒性アラートが実装された。

2.4.1. 肝毒性 rapid prototype アラート

新しい知識ベースの子測精度が ISIS/Base データベース "NIHS5_20080916.db"からのデータを使用して Derek for Windows version11 で調査された。Table 19 と 20 は、それぞれ、標準の Derek for Windows version11 および DfW11 NIHS 201108_1 知識ベースに対する肝毒性子測精度を示したものである。

Table 19 肝臓の病理組織学的な所見に準じた分類に基づく、NIHS データベースに収載された化合物に対する Derek for Windows version11 の子測精度

Derek for Windows \ 試験結果	陽性	陰性	合計
陽性	69	231	300
陰性	51	380	431

Table 20 肝臓の病理組織学的な所見に準じた分類に基づく、NIHS データベースに収載された化合物に対する DfW11 NIHS 201108_1 の子測精度

Derek for Windows \ 試験結果	陽性	陰性	合計
陽性	184	116	300
陰性	88	343	431

Table 20 より、肝毒性子測感度が 61%に向上したが、その一方で、特異性が 80%へと減少したことが分かる。特異性の減少は、Rapid Prototype アラートの実装に由来する追加的な 37 個の偽陽性子測によるものである。

2.4.2. 腎毒性 rapid prototype アラート

新しい知識ベースの子測精度を ISIS/Base データベース "NIHS5_20080916.db"からのデータを使用して Derek for Windows version11 と比較した。

Table 21 と 22 は、それぞれ、標準の Derek for Windows version11 と DfW11 NIHS 201108_1 知識ベースの腎毒性子測精度を示したものである。

Table 21 腎臓の病理組織学的な所見に準じた分類に基づく、NIHS データベースに収載された化合物に対する Derek for Windows version11 の子測精度

Derek for Windows \ 試験結果	陽性	陰性	合計
陽性	4	226	230
陰性	3	498	501

Table 22 腎臓の病理組織学的な所見に準じた分類に基づく、NIHS データベースに収載された化合物に対する DfW11 NIHS 201108_1 の予測精度

試験結果 \ Derek for Windows	陽性	陰性	合計
陽性	143	87	230
陰性	69	432	501

Table 22 から、腎毒性予測感度が 62%に向上したが、一方で特異性は 85%以上の値を維持していることが分かる。

知識ベースの予測精度を Rapid Prototype アラート実装過程を通じて調査した。Rapid Prototype 開発の進捗が各々の毒性エンドポイントについて、知識ベースのバージョンに対する感度と特異性をプロットすることにより視覚化できる (Figure 30 と Figure 31)。これらの散布図は、例えば、肝毒性予測の感度および特異性の双方が開発過程の初期において比較的大きな変化を示した後に平坦域が現れることを示している。その一方で、腎毒性 Rapid Prototype アラートの場合においては、そのような

平坦域は肝毒性の時ほど明確ではないことが分かる。さらに、肝毒性 Rapid Prototype アラート開発期間において、トレーニングデータセット

(NIHS_UPDATED_Without duplicates_FINAL_280408.db) を含むデータベースがさらに更新され (NIHS5_20080916.db)、それが知識ベースの予測精度に対して及ぼした影響が Figure 30 に示されている。腎毒性予測感度は、29 個の Rapid Prototype アラートを含む知識ベースにおいて減少しており、その後、更新された。データベースに対してアラートが再評価された後、再び感度が増加した。

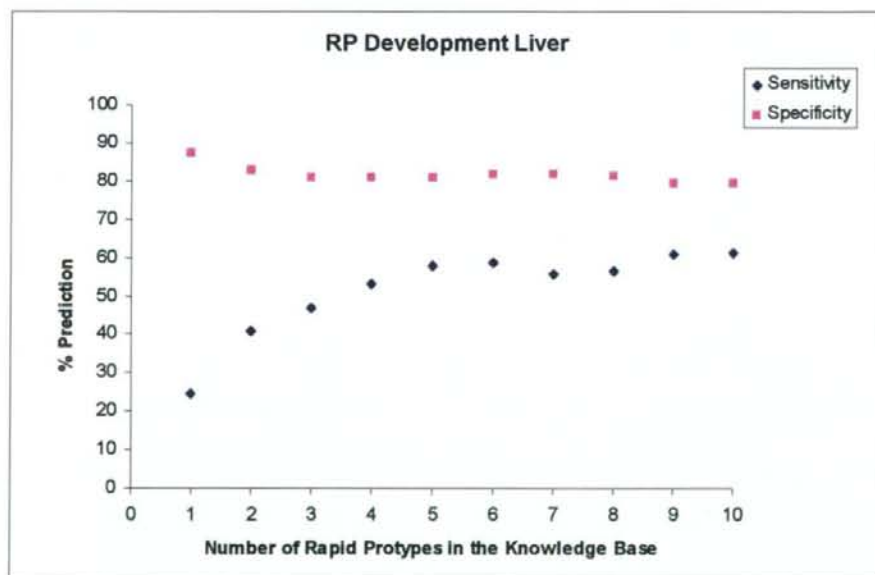


Figure 30 Rapid Prototype アラートを含む知識ベースのバージョンによって示された Rapid Prototype 開発の個々の段階における予測感度と特異性の散布図 *更新された NIHS5_20080916.db の影響

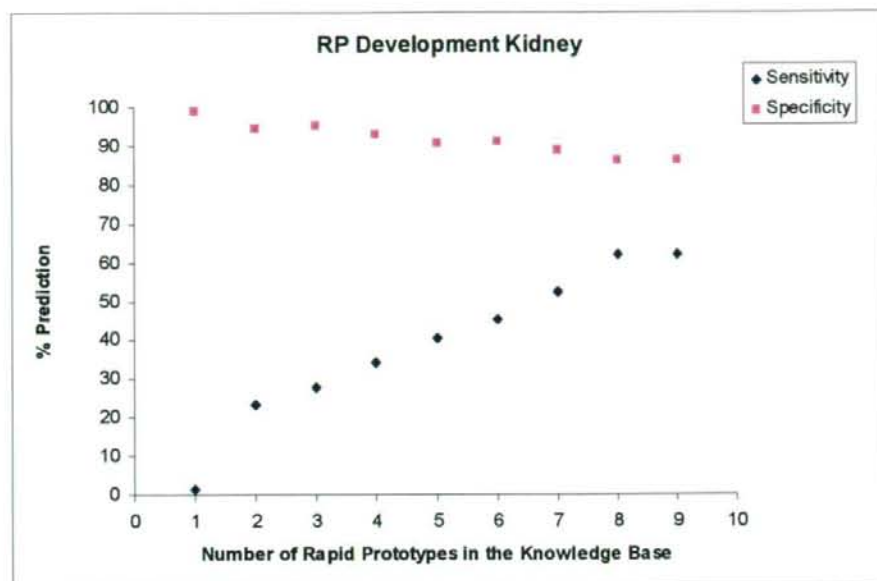


Figure 31 Rapid Prototype アラートを含む知識ベースのバージョンによって示された。Rapid Prototype 開発の個々の段階における予測感度と特異性の散布図 *知識ベース中の Rapid Prototype アラート数は、知識ベース内の既存の Rapid Prototype アラートの結合により、23 個に減っている

2.4.3. 二つの外部データセットに対する予備的な結果

DfW11 NIHS 201108_1 に実装された Rapid Prototype アラートの予測精度を調査するために、

NIHS により提供された二つの外部データセット ("A"と"B") に対して予測を実施し、肝毒性に対する予測精度 (Table 23 と Table 25) と腎毒性に対する予測精度 (Table 24 と Table 26) を算出した。

Table 23 肝臓の病理組織学的な所見に準じた分類に基づく、"A"データベースに収載された化合物に対する DfW11 NIHS 201108_1 の予測精度

試験結果 \ Derek for Windows		陽性	陰性	合計
陽性		73	165	238
陰性		140	340	480

感度: $73/238 = 31\%$; 特異性: $340/480 = 71\%$; 一致率: $413/718 = 58\%$

Table 24 腎臓の病理組織学的な所見に準じた分類に基づく、"A"データベースに収載された化合物に対する DfW11 NIHS 201108_1 の予測精度

試験結果 \ Derek for Windows		陽性	陰性	合計
陽性		47	119	166
陰性		143	409	552

感度: $47/166 = 28\%$; 特異性: $409/552 = 74\%$; 一致率: $456/718 = 64\%$

Table 25 肝臓の病理組織学的な所見に準じた分類に基づく、"B"データベースに収載された化合物に対する DfW11 NIHS 201108_1 の予測精度

試験結果 \ Derek for Windows	陽性	陰性	合計
	陽性	2	13
陰性	4	11	15

感度: 2/15 = 13%; 特異性: 11/15 = 73%; 一致率: 13/30 = 43%

Table 26 腎臓の病理組織学的な所見に準じた分類に基づく、“B”データベースに収録された化合物に対する DfW11 NIHS 201108_1 の予測精度

試験結果 \ Derek for Windows	陽性	陰性	合計
	陽性	3	8
陰性	7	12	19

感度: 3/11 = 27%; 特異性: 12/19 = 63%; 一致率: 15/30 = 50%

"A"データセットに対する肝毒性 (Table 23) と腎毒性 (Table 24) の予測感度は30%前後に落ち込んでおり、その一方で特異性は70%以上を維持している。Table 23 からは、73 個の真陽性化合物と 140 個の偽陽性化合物が肝毒性アラートにヒットし、そのうち51 個の真陽性化合物と 84 個の偽陽性化合物がそれぞれこのプロジェクトの一部として開発された Rapid Prototype アラートと関連していることが分かる。腎毒性に関しては、Table 24 から 47 個の真陽性化合物と 143 個の偽陽性化合物が認められ、そのうち 44 個の真陽性化合物および 140 個の偽陽性化合物がそれぞれ Rapid Prototype アラートに関連している。肝毒性に対する標準 Derek for Windows アラートと Rapid Prototype アラート双方の予測精度に対するより詳細な調査により、Rapid Prototype アラートは既存の Derek for

Windows アラートよりもよい予測精度を有することが分かった。

54%の肝毒性 Rapid Prototype アラートが、“A”データセットに対して 50%以上の陽性予測率を有している一方で、たった 39%の肝毒性 Derek for Windows アラートしか 50%を超える予測精度を有していない。この結果は、肝毒性および腎毒性双方に対する“DfW11 NIHS 201108_1”の予測感度の向上は、既存の Derek for Windows version11 のアラートと比較した場合、Rapid Prototype アラートの存在に依るところが大きく、しかしながら、それはまた偽陽性予測の増加も招くということを示している。

それぞれの Rapid Prototype の予測精度を解析し、その結果を Table 27 に要約した。

Table 27 “A”データセットに対する“DfW11 NIHS 201108_1”の陽性予測率の点から見た Rapid Prototype アラートの予測精度の要約

エンドポイント	Rapid prototype アラートの予測精度	
肝毒性 (34 個の Rapid Prototype アラート)	ヒットせず	8
	0%正しい	6
	1-49%正しい	6
	>50%正しい	14
腎毒性 (35 個の Rapid Prototype アラート)	ヒットせず	12
	0%正しい	5
	1-49%正しい	11
	>50%正しい	7

予測精度毎に分解したアラートの要約 (Table 27) から、8 個の肝毒性 Rapid Prototype アラートと 12

個の腎毒性 Rapid Prototype アラートがヒットしなかった。一方で、双方の毒性エンドポイント内のお

よそ 15 個の Rapid Prototype アラートはより低い (<50%) 予測精度を示しているということが分かる。これら予測精度の低いアラートをより詳細に解析すると、幾つかの Rapid Prototype アラートはトレーニングデータセット内の比較的少量の化合物に基づいて作成されていることが分かる。例えば、Hep16、Hep07、Nephro17、Nephro27 がそれ

に当たる。特定の Rapid Prototype アラートに対する、トレーニングデータセットからの偽陽性化合物と真陽性化合物の個数の間の関係を調査し、その結果が Figure 32 (肝毒性 Rapid Prototype アラート) と Figure 33 (腎毒性 Rapid Prototype アラート) に示されている。

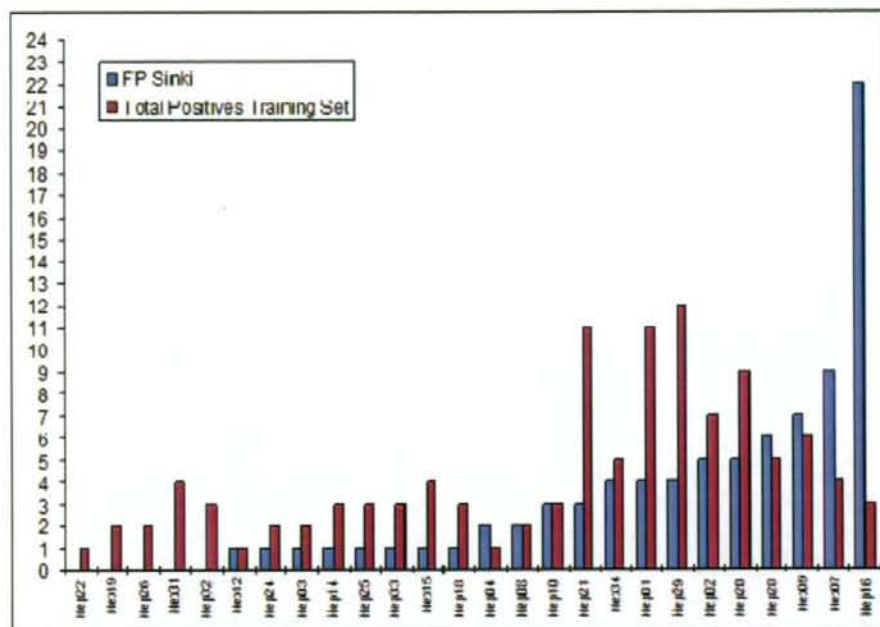


Figure 32 個々の肝毒性 Rapid Prototype アラートに対する、"A"データセット中の偽陽性化合物 (FP) とオリジナルトレーニングデータセット内の全真陽性化合物との関連

両方のグラフを比較すると、一般に腎毒性 Rapid Prototype アラートはより少ない真陽性化合物 (Figure 33) のトレーニングデータセットに基づいている一方で、殆どの肝毒性 Rapid Prototype アラートはトレーニングデータセット中のよりばらついた個数の化合物に基づいていることが分かる (Figure 32)。Rapid Prototype アラートの予測精

度は、アラートを開発するために使用されたトレーニングデータセット中の化合物の個数が適用範囲に影響を及ぼすので、これが部分的には関連しているかも知れず、そして、これは将来の Rapid Prototype 手法に対して考慮されるかも知れない。

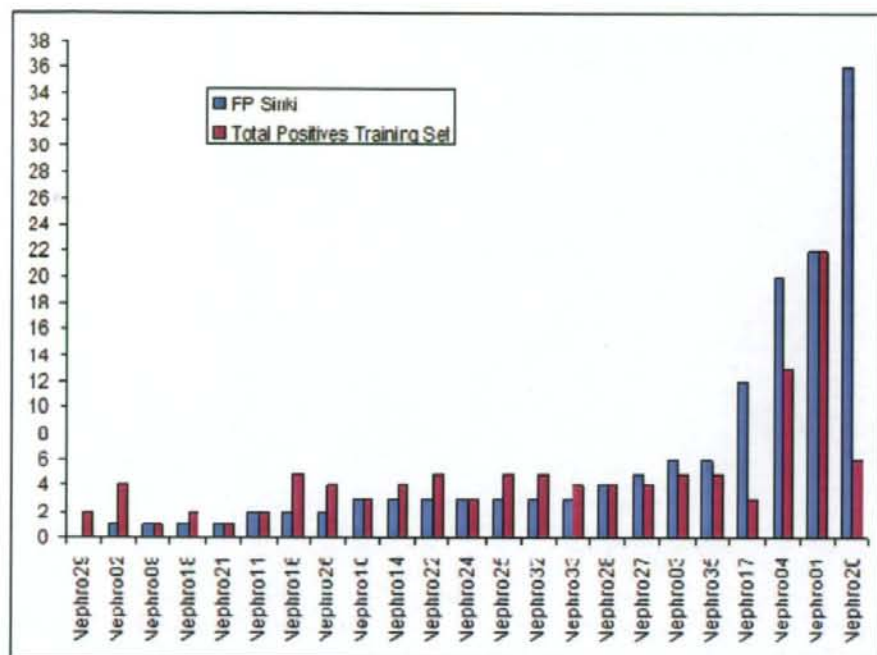


Figure 33 個々の腎毒性 Rapid Prototype アラートに対する、“A”データセット中の偽陽性化合物 (FP) とオリジナルデータセット内の全真陽性化合物との関連

調査に使用されたテストデータセット中の化合物がこのプロジェクトにて開発された Rapid Prototype アラートの適用領域の外側に位置するのではないかと、これらの結果から追加的に解釈できるかもしれない。Rapid Prototype アラート Hep09 を例にとると、分子量 (MW) と LogP 値の範囲に基づく単純な適用領域を、このアラートを含むトレーニングデータセット中の試験的に陽性な化合物に対して定義することができる (Figure 34)。アラートにヒットしたテストデータセット中の化合物に対する上記の物性値の比較は、適用領域内に化合物が位置するような予測に対してより大きな信頼を置くことができるということを示している。特に、アラートにヒット

するようなより高い分子量を有する化合物や適用領域外にあるような化合物は、極まれにしか陽性を示さないと思われる。

他の Rapid Prototype に対して、アラートそのものの適用範囲はあまりに広く、多くの偽陽性予測を導くことは大いにあり得る。例えば、もしも Nephro20 に対する適用領域が Hep09 と同様の方法で決定されるならば、アラートにヒットしてしまうテストデータセット中の多くの試験的に陰性な化合物が、そのアラートに対する適用領域内にあり続けることが分かる (Figure 35)。このことは、この場合においてはアラートの定義を更に見直す必要があることを示唆している。

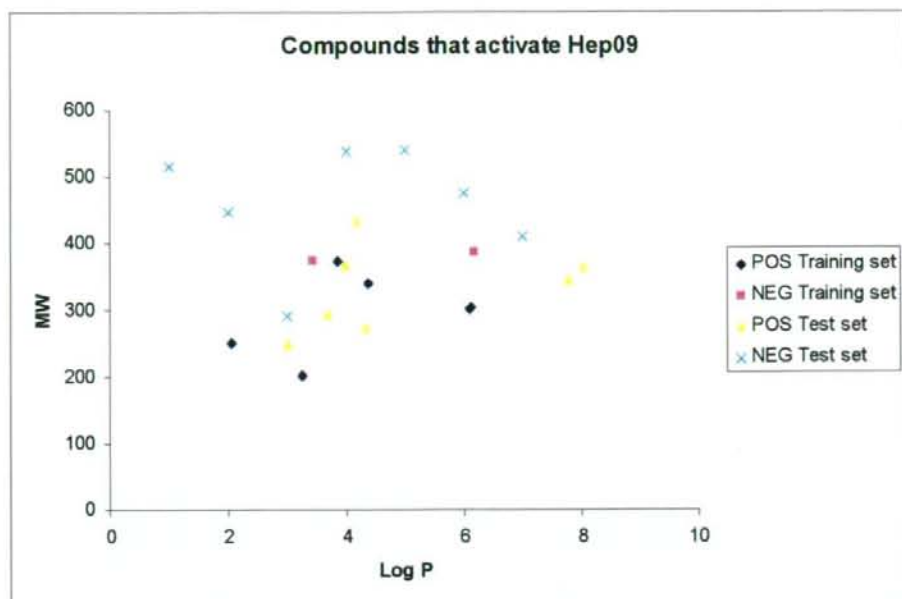


Figure 34 Rapid Prototype アラート Hep09 に対する、トレーニングデータセットとテストデータセット双方の MW と LogP の散布図

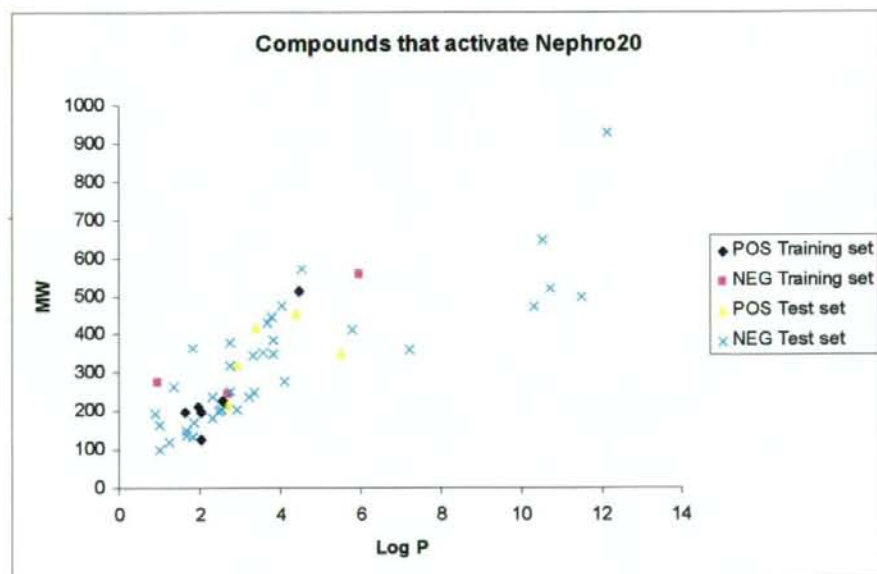


Figure 35 Rapid Prototype アラート Nephro20 に対して、トレーニングデータセットとテストデータセット双方の MW と LogP の散布図

最後に、全体を通して使用された肝毒性もしくは腎毒性の定義は、ただ単に肝臓や腎臓における病理組織学的所見に基づくということを中心に留めておくべきである。酵素レベル、臓器重量、コレステロール/トリグリセリドや血中尿素窒素といった付加的な変数を考慮すると、ここで示された結果の解釈が大きく影響を受けるかもしれない。

2.5. 結論

34 個の肝毒性および 35 個の腎毒性 Rapid Prototype アラートが Derek for Windows に実装された。全ての Rapid Prototype アラート実装後のトレーニングデータセットに対する予測精度は、肝毒性に対して感度 61%、特異性 80%であり、腎毒性に対して感度 62%、特異性 86%である。

このプロジェクトの開始前には Derek for Windows 知識ベースは 40 個の肝毒性アラートを含んでおり、Rapid Prototype アラート開発により 74 個まで増加した。このプロジェクトの開始前までは系統的に開発されていなかった腎毒性に対しては、35 個の Rapid Prototype アラートの実装により、Derek for Windows の予測精度が目覚ましく向上した。NIHS により提供された私有データセットからの Rapid Prototype 開発の程度は NIHS 保有のデータセット中のケミカルスペースに依存しており、そして、更なる開発は、双方の予測感度がおおよそ 60%の水準に達した後に、限定的であることが分かった。

二つの外部データソースを使用しての、当プロジェクトで開発された Rapid Prototype アラートの評価により、肝毒性および腎毒性双方の予測感度および特異性の減少が認められた。幾つかの標準 Derek for Windows アラートに対しても同様の減少が見られたので、これらの所見は Rapid Prototype 手法の使用に直接起因するものではあり得ない。

Rapid Prototype アラートの適用範囲を精密に定義することができる程度は、トレーニングデータセット中のそれを支援する化合物の個数と特徴によって制約される。テストデータセットでの結果は、この適用範囲の改良がいくつかの Rapid

Prototype アラートに対して求められることを示唆している。その他の場合において、予測精度の改善は、個々のアラートの適用領域の考慮の結果として予想されるかもしれない。最後に、病理組織学的な結果に準じた分類を越えた、予測精度の更なる調査は実施に値するかも知れない。

当プロジェクトは、ある適切なトレーニングデータセットが与えられた場合、Rapid Prototype 手法を Derek for Windows 知識ベース中の既存および新たな毒性エンドポイントに対するアラートの予測精度を改善するために使用することができることを示した。新しいテストデータセットを、これらの Rapid Prototype アラートをさらに保障もしくは改良するために使用できる可能性があり、それらに続く標準 Derek for Windows アラートの開発を可能にする。適用領域の考慮といった追加的な技法が、予測精度の信頼度を解釈したり、Rapid Prototype アラートの適合性や品質の調査を手助けしたりするために使用されるかもしれない。

3.有機化合物における反復投与毒性の予測(肝毒性について)に関する研究

3.1.はじめに

有機化合物の反復投与がヒトに及ぼす毒性を予測することを目的として、実験動物による毒性試験の結果を用いた反復投与毒性予測モデルを構築し、その予測モデルの検証を行った。

3.2.予測対象とする毒性

3.2.1. 肝毒性

- ・組織病理学的検査における肝臓への影響(項目 L とする。以下同じ)
 - ・血液生化学検査における GOT 値または GPT 値の高値(項目 G)
 - ・臓器重量における肝臓(絶対または相対)重量の高値(項目 W)
- 各項目のデータは、動物試験における最小毒性量(LOAEL)である。

3.2.1.1. 予測モデル構築のためのデータ項目

これらの項目を単独で (L)、または、組み合わせて (LGW)、肝毒性を示すデータ項目として用いた。

3.2.1.2 予測するデータ項目

構築したモデルで予測する値は、項目 L とする。Features の選択については、LGW を用いたモデル、および L を用いたモデルについて評価した。

3.3. 予測モデルとその検証

3.3.1. 予測モデル構築に使用するソフト

LeadScope (LeadScope 社製) を用いて (Q)SAR モデルの構築と (Q)SAR モデルによる化合物の毒性予測を行った。

3.3.2. 予測モデル構築のためのデータセット

予測モデル構築には、既知の化合物の反復投与毒性試験データを用いた。反復投与毒性試験の種類とデータ数は、Table28 の通りである。

Table28. 既知化合物の反復投与毒性試験の種類とデータ数

28 日間試験	1232 データ
90 日間 (13 週間) 試験	28 データ
併合試験	2 データ
その他 (2 年間: 2、21 日間: 1)	3 データ
計	1265 データ

3.3.3. 予測モデル検証のためのデータセット

予測モデルの検証には、既存化合物の反復投与毒性試験データを用いた。また、予測モデル構築のためのデータセットとの重複はない。既存化合

物の反復投与毒性試験の種類とデータ数は、Table29 の通りである。

Table29. 既存化合物の反復投与毒性試験の種類とデータ数

28 日間試験	136 データ
90 日間 (13 週間) 試験	2 データ
併合試験	113 データ
簡易試験	6 データ
計	258 データ

3.3.4. 予測精度に関する用語

予測精度に関する用語とその定義は以下の通りである。

- Sensitivity 感度：動物試験の陽性結果を陽性と判定した割合 (%)
- Specificity 特異性：動物試験の陰性結果を陰性と判定した割合 (%)
- Concordance 全体の一致率：動物試験の陽性結果を陽性、陰性結果を陰性と判定した割合 (%)

その他、

- 予測可能な化合物の割合：適用させるデータ全数に対して実際に予測したデータ数の割合 (%)

3.3.5. 予測モデルの構築

3.3.5.1 データのバイナリ化

動物試験における肝毒性の強さは、全身的な症状を考慮した上で、その毒性の種類や症状の程度によって表現されるものである。したがって、化合物間の肝毒性の強さは、各試験における項目ごとの LOAEL 値の大小で単純には比較できないことから、各項目のデータを全身的な症状について考慮せず扱う場合には、連続的数値として捉

えるのではなく、そのデータをバイナリ化 (0: 毒性なし、1: 毒性あり、の二極化) して使用することは妥当であると考えられた。

各動物試験において、最高用量まで影響を示さなかった項目を 0 (毒性なし)、それ以外の項目を 1 (毒性あり) とした。

3.3.5.2 Features の選択

記述子となる Feature 数が多いとモデルへの当てはめが良くなる (つまり、モデル作成に使用したデータが作成したモデルによく当てはまる) が、モデル作成に含まれていないデータに対する予測は、モデルの複雑さが増すにつれ、あるところから精度が低くなることが予想される。このことから、次に示す方法で、自動作成した Feature の数を減らすモデルの作成を試みた。

(1)自動選択を用いる方法

(2)Feature の平均値が 0.5 付近 (0.400~0.600)

の Feature を除外する方法:陽性と陰性の化合物が等量ずつ含まれる Feature は重要ではないと考え、平均値が 0.5 付近の Feature を除外し、モデルを作成した。

(3)類似 Feature を除外する方法:類似 Feature を LeadScope で自動的に除外し、モデルを作成した。

(4)Feature の Z スコアの絶対値が 2 より小さい Feature を除外する方法: Feature の Z スコアの絶対値が 2 より小さい Feature を有意ではな

いと考え除外し、モデルを作成した。

3.3.6. 適用範囲の設定

予測モデルにより算出される Predicted Probabilities (予測確率) は 0~1 の値である。Predicted Probabilities が 0.5 以上の場合を「陽性」、0.5 より小さい場合を「陰性」とした。この範囲を変化させることで、モデルへの当てはめが良くなるかについて検討した。

3.3.7. 種々のモデル検討

モデル LGW、モデル L の他に、それらのモデルを組み合わせたモデルについても検討した。

3.3.8. 予測モデルの検証

予測モデルの構築と同様にデータをバイナリ化して、モデルの検証に使用した。検証には、病理組織学的データ (L) を用いた。

3.4. 結果と考察

3.4.1. 予測モデルの当てはめ結果および Features の選択による影響

構築データセットにおける陽性データ数、陰性データ数、および、その割合を Table30 に示す。構築データセットにおける項目 L と項目 LGW の関係を Figure 36 に示す。

Table30. 構築データセットにおける陽性数、陰性数、その割合

	陽性	陰性	計	陽性(%)	陰性(%)
L	453	812	1265	35.8	64.2
LGW	712	553	1265	56.3	43.7

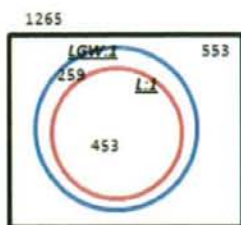


Figure 36. 構築データセットにおける項目 L と項目 LGW の関係

選択された Features ごとの当てはめ結果を Table 31 (モデル LGW)、Table 32 (モデル L) に示す。モデルの名称は、例えば、モデル L では Feature 選択を自動で行った場合を"モデル、L_自動"、0.5 付近の Feature を除外した場合を"モデル L_0.5"、

類似 Feature を除外した場合を"モデル L_類似"とした。

また、モデル L、モデル LGW 共に Z スコアの絶対値が 2 より小さい Feature はみられなかった。

Table 31. モデル LGW の当てはめ結果 (Features の選択方法ごと)

Feature 選択方法 (Feature 数)	予測項目	予測値 (化合物数)			
自動 (387)	LGW	1 (863)	0 (402)	75.7	Concordance
実際の LGW	1	634	78	89.0	Sensitivity
	0	229	324	58.6	Specificity
0.5 付近の Feature 除外 (262)	LGW	1 (842)	0 (423)	76.0	Concordance
実際の LGW	1	636	76	89.3	Sensitivity
	0	227	326	59.0	Specificity
類似 Feature 除外 (327)	LGW	1 (842)	0 (423)	77.9	Concordance
実際の LGW	1	637	75	89.5	Sensitivity
	0	205	348	62.9	Specificity
予測可能な化合物の割合 (1265/1265*100)				100.0 ^{a)}	

^{a)} Probability threshold を「陽性 \geq 0.5、0.5 $>$ 陰性」とした場合

Table 32. モデル L の当てはめ結果 (Features の選択方法ごと)

Feature 選択方法 (Feature 数)	予測項目	予測値 (化合物数)			
自動 (352)	L	1 (233)	0 (1032)	72.8	Concordance
実際の L	1	171	282	37.7	Sensitivity
	0	62	750	92.4	Specificity
0.5 付近の Feature 除外 (345)	L	1 (220)	0 (1045)	71.5	Concordance
実際の L	1	156	297	34.4	Sensitivity
	0	64	748	92.1	Specificity
類似 Feature 除外 (293)	L	1 (196)	0 (1069)	71.3	Concordance

実際の L	1	143	310	31.6	Sensitivity
	0	53	759	93.5	Specificity
予測可能な化合物の割合 (1265/1265*100)				100.0 ^{a)}	

a) Probability threshold を「陽性 \geq 0.5、0.5 $>$ 陰性」とした場合

3.4.2. 適用範囲の検討結果

図 2~4 に、Predicted Probabilities の各区間における化合物の陽性数、陰性数を示す。Figure 37 はモデル LGW、Figure 38 はモデル L、Figure 39

は組合せモデルを考える場合の Predicted Probabilities の各区間における化合物の陽性数、陰性数である。

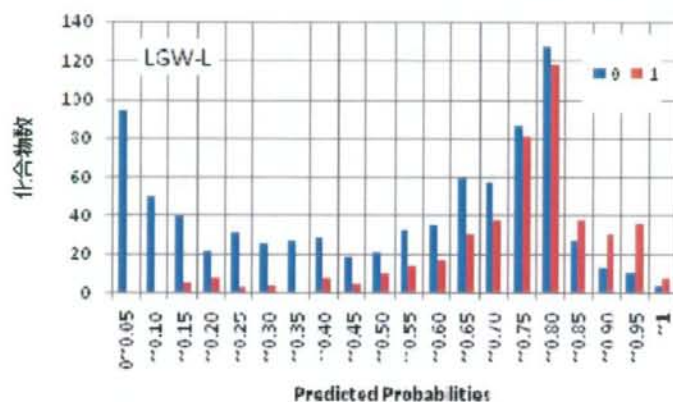


Figure 37. Predicted Probabilities の各区間における陽性、陰性の分布 (モデル LGW 自動-予測項目 L)

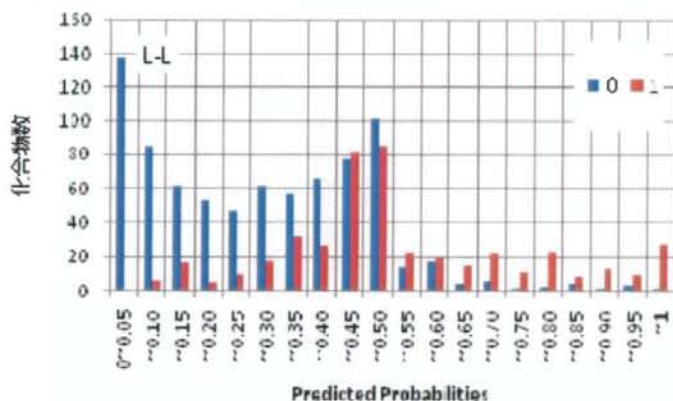


Figure 38. Predicted Probabilities の各区間における陽性、陰性の分布 (モデル L 自動-予測項目 L)

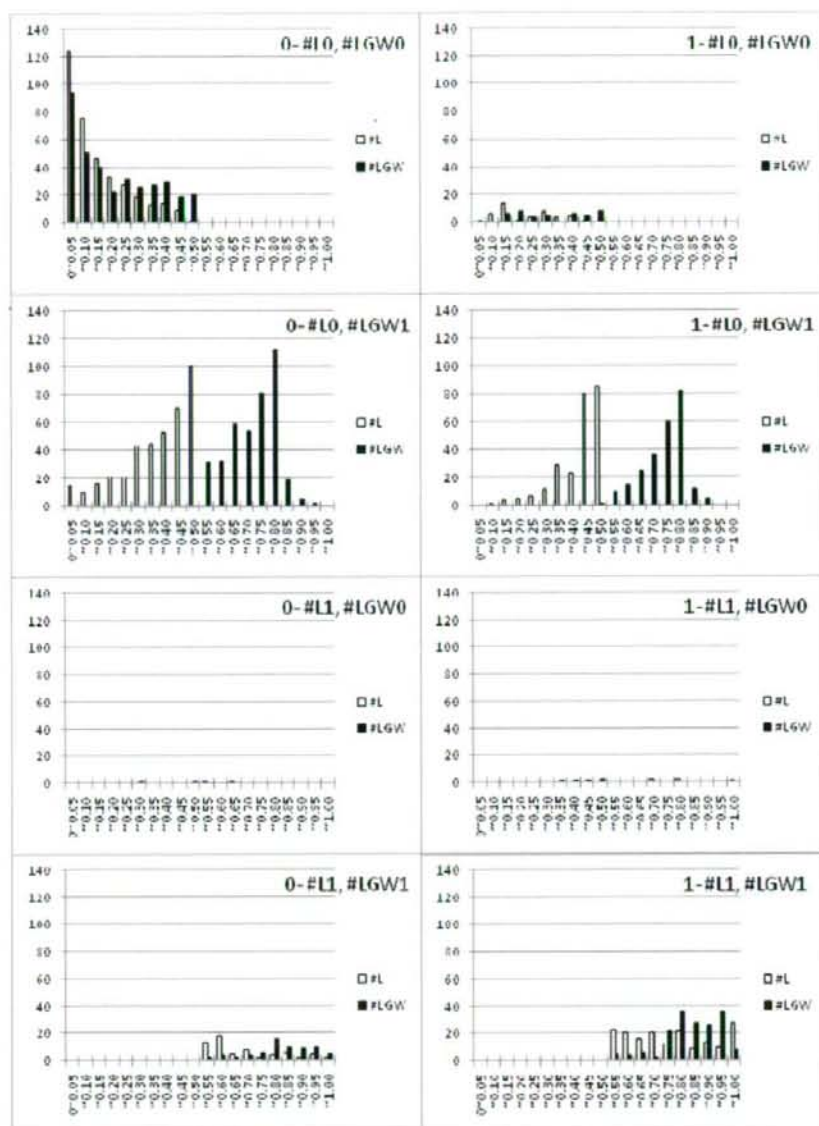


Figure 39. Predicted Probabilities の各区間における陽性、陰性の詳細な分布

Figure 37 より、モデル LGW_自動では Predicted Probabilities の陽性/陰性基準を 0.5 から移動させてもバランスよく精度をあげるにはあまり効果がないように思われた。Figure 38 より、モデル L_自動では Predicted Probabilities が 0.4 以上を陽性と判定した場合、効果的に陽性予測が増加して感度が上がり、一方、陰性予測は減少して特異性が低下することが考えられた。

Figure 39 より、組合せモデルにおいてモデル L_自動では Predicted Probabilities が 0.4 以上を陽性と判定した場合、効果的に陽性予測が増加して感度が上がり、一方、陰性予測は減少して特異性が低下することが考えられた (モデル Lとして検討)。モデル LGW_自動については単独でモデルを使用する場合に Predicted Probabilities の基準を移動させても効果が弱いと考えられたことから、

組合せモデルにおいてもモデル LGW_自動については基準の移動は検討しないこととした。モデル L_自動とモデル LGW_自動の当てはめ結果を Table 33 および 34 に示す。

同様に他のモデル L についても Predicted Probabilities の検討を行った。各モデルにおいて最も効果的と判断した。修正 Predicted Probabilities を Table 35 に示す。モデル L_0.5

およびモデル L_類似は、モデル L_自動から派生しており、モデル L_自動 (Figure 39 参照) と同様の分布を示した。結果として、モデル L_自動、モデル L_0.5、モデル L_類似ともに 0.4 以上を陽性、0.4 未満を陰性とした。修正 Predicted Probabilities を用いたモデルの名称は、モデル L とした。

Table 33. モデル L_自動とモデル L_自動の当てはめ結果 (Probability threshold 変更の影響)

	予測項目	予測値 (化合物数)			
陽性 \geq 0.5、陰性 $<$ 0.5 (初期設定)	L	1 (233)	0 (1032)	72.8	Concordance
実際の L	1	171	282	37.7	Sensitivity
	0	62	750	92.4	Specificity
陽性 \geq 0.4、陰性 $<$ 0.4	L	1 (578)	0 (687)	71.8	Concordance
実際の L	1	337	116	74.4	Sensitivity
	0	241	571	70.3	Specificity
予測可能な化合物の割合 (1265/1265*100)				100.0	

Table 34. モデル LGW_自動の当てはめ結果

	予測項目	予測値 (化合物数)			
陽性 \geq 0.5、陰性 $<$ 0.5 (初期設定)	L	1 (863)	0 (402)	60.8	Concordance
実際の L	1	410	43	90.5	Sensitivity
	0	453	359	44.2	Specificity
予測可能な化合物の割合 (1265/1265*100)				100.0	

Table 35. モデル L の修正 Predicted Probabilities

	Predicted Probabilities
モデル L_自動	1: 0.4 以上、0: 0.4 未満
モデル L_0.5	1: 0.4 以上、0: 0.4 未満
モデル L_類似	1: 0.4 以上、0: 0.4 未満