

### 3.4 SPBS (秋田大学)

---

#### 3.4.1 名称及びバージョン

---

SPBS V9.67 (2014年1月10日改訂版) (2019年2月末時点)

Microsoft Windows XP/Vista/7 で稼動

#### 3.4.2 SPBS の概要

---

##### (1) 概要

SPBS は、村田勝敬教授 (秋田大学) によって 1986 年に初代が開発された。無償で入手できる教育用統計ソフトウェアである (図 3.4-1)。

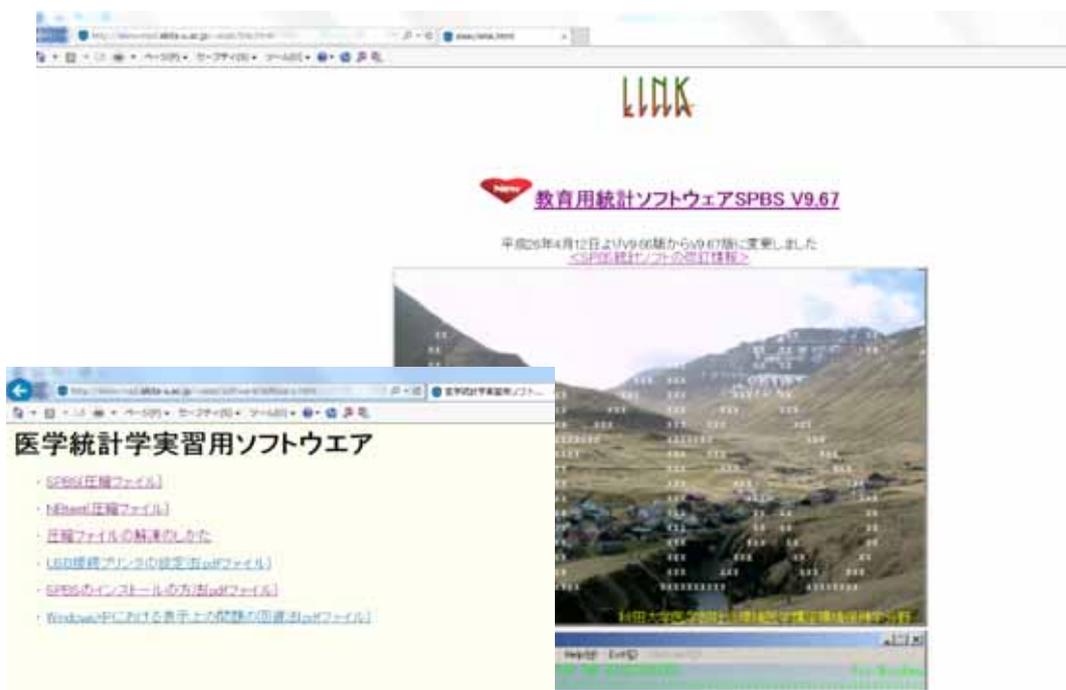


図 3.4-1 ダウンロード画面

##### (2) 主要な機能

###### (2)-1 利用可能なデータ及び使用されるモデル

主要な機能は、連続値データ (個体) を対象とした、BMD 及び BMD の信頼区間の下限

値 (BMDL) 算出である。収載されているモデルは、一次回帰モデル、べき関数モデル、 $\log(X+1)$  変換モデルであり、手動で選択する。

## (2)-2 BMR の設定

BMR は 0.02、0.05、0.10 の 3 種類から選択し、同一用量に対する反応量分布が正規分布と仮定している。分析方法は Budtz-Jørgensen ら (2001)<sup>57</sup>の方法に基づくハイブリッド法である。この手法は、非ばく露集団において設定された異常率  $P_0$  に対して、ばく露集団における  $\alpha\%$  (BMR) の異常増加をもたらすばく露濃度を BMD と定義し、この BMD の 95% 信頼区間の下限値を BMDL として求める手法である。分析方法の概要を以下に示す。

非ばく露群の異常率  $P_0$ 、有意な影響の反応レベル BMR、ばく露指標  $d$ 、影響指標  $\mu(d)$ 、用量反応関係  $\mu(d) = b_0 + b_1 g(d)$  とすると、 $b_1 > 0$  はばく露を受けると影響指数が増加することを示し、 $b_1 < 0$  はばく露を受けると影響指数が減少することを示す。

①  $b_1 > 0$  (ばく露を受けると影響指標が増加する) 場合：

$$P_0 = 1 - \Phi[(C - b_0)/\sigma] \text{ 及び } BMD = g^{-1}\{[\Phi^{-1}(1 - P_0) - \Phi^{-1}(1 - P_0 - BMR)] \sigma / b_1\} \text{ から計算}$$

②  $b_1 < 0$  (ばく露を受けると影響指標が減少する) 場合：

$$P_0 = \Phi[(C - b_0)/\sigma] \text{ 及び } BMD = g^{-1}\{[\Phi^{-1}(P_0) - \Phi^{-1}(P_0 + BMR)] \sigma / b_1\} \text{ から計算}$$

( $g(d) = d^k$  はべき関数、 $\Phi$  は正規累積分布関数、 $C$  はカットオフ値、 $b_0$  は期待反応関数の定数値、 $\sigma$  は集団の標準偏差値を示す。また  $K$  は  $\geq 1$ )

$P_0 = \Phi[(C - b_0)/\sigma]$  は Crump (1995)<sup>58</sup> が考案した式で、非ばく露群の  $P_0 = 5\%$  を入力すると、カットオフ値として非ばく露群の 95% 上限値が算出される。

非ばく露群において異常率を  $P_0$  とした場合に定義される影響指標のカットオフ値が、臨床上の基準値に近いかな否かでモデルあるいは対象集団の妥当性を判断する。

例えば、ばく露によって血圧が上昇し、カットオフ値が 140 mmHg と算出された場合、これを臨床上の基準値と比較しその妥当性を判断する。

---

<sup>57</sup> Budtz-Jørgensen E, Keiding N and Grandjean P: Benchmark dose calculation from epidemiological data. *Biometrics* Vol. 57, No. 3 (Sep., 2001), pp. 698-706.

<sup>58</sup> Crump KS. Calculation of benchmark doses from continuous data. *Risk Anal* 1995; 15: 78-89.

### 3.1.1. 3.4.3 利用方法に関する情報

#### (1) セットアップ方法

秋田大学のホームページからダウンロードできる<sup>59</sup>。

上記ホームページの教育用統計ソフトウェア SPBS V9.67 をクリックし、圧縮ファイルの解凍方法を参考に、SPBS (圧縮ファイル)、NBtest (圧縮ファイル) をダウンロードし解凍することにより利用できる。

操作方法に関する資料 (操作マニュアル、ガイダンス等)

- ・ダウンロードした spbs フォルダ内の spbshelp ファイルに操作方法が記載されている。
- ・ダウンロードした spbs フォルダ内の電卓マークの spbs ファイルのショートカットを作成し、デスクトップに移動させる。

#### (2) 操作の流れ

操作の流れを以下に示す。

##### 1) 入力データの作成

エクセルで作成し、保存する際に CSV ファイルの形式で選択する(タブ区切り等)(図 3.4-2)。



	A	B	C	D	E	F	G	H
1	ID番号	年齢	尿中Cd	毛髪Hg	喫煙	α1-MG	爪Hg	
2	1	20	0.662	1.94	0	2.28	0.54	
3	2	20	0.456	1.39	0	1.83	0.4	
4	3	19	0.464	1.31	1	1.61	0.47	
5	4	20	1.458	3.6	1	3.64	1.68	
6	5	20	0.638	0.76	0	1.41	0.3	
7	6	20	1.123	1.83	0	1.48	0.89	
8	7	20	0.259	2.03	0	0.78	0.81	
9	8	20	0.275	1.16	0	1.07	0.38	
10	9	20	0.34	0.97	0	0.78	0.41	
11	10	20	0.452	1.37	0	2.11	0.38	
12	11	20	0.192	1.72	0	0.75	0.82	
13	12	19	0.487	1.55	0	1.48	0.66	
14	13	20	0.614	1.12	0	2.19	0.5	
15	14	19	0.238	0.94	0	1.36	0.64	
16	15	20	0.304	1.14	0	0.73	0.38	
17	16	20	0.456	3.97	0	0.79	0.98	
18	17	20	0.259	0.55	0	0.42	0.21	
19	18	20	1.223	1.18	0	1.62	0.38	

図 3.4-2 入力データの例

<sup>59</sup> <http://www.med.akita-u.ac.jp/~eisei/link.html>

## 2) プログラムの起動

Spbsspbs (計算機のアイコン) をダブルクリックし起動する (図 3.4-3)。左上の Program(D)を選び、プルダウンして Several tests→Benchmark dose Calculation を選択する。

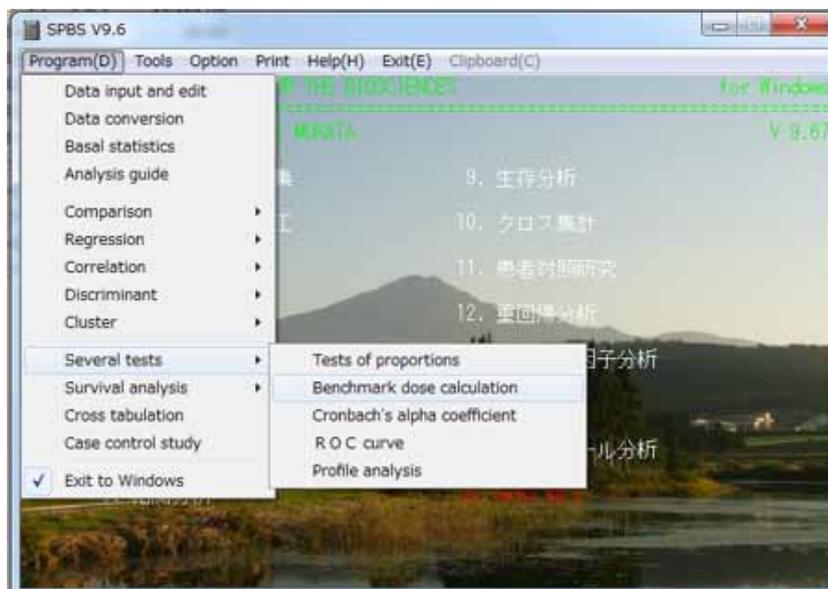


図 3.4-3 プログラムの起動画面

## 3) 分析するデータファイルの選択、 $P_0$ 設定、べき関数のフィッティング

### ① 分析するデータファイルの選択

- Benchmark dose Calculation 画面上 (図 3.4-4) で File 確認ボタンを押し、解析に用いる用量反応データファイルを選択し、確認する。(図 3.4-5、図 3.4-6)

- ファイル名にスペースが含まれている場合や、ファイル名称が半角 12 文字以上の場合は読み込めない場合がある旨が表示される



図 3.4-4 Benchmark dose Calculation 画面

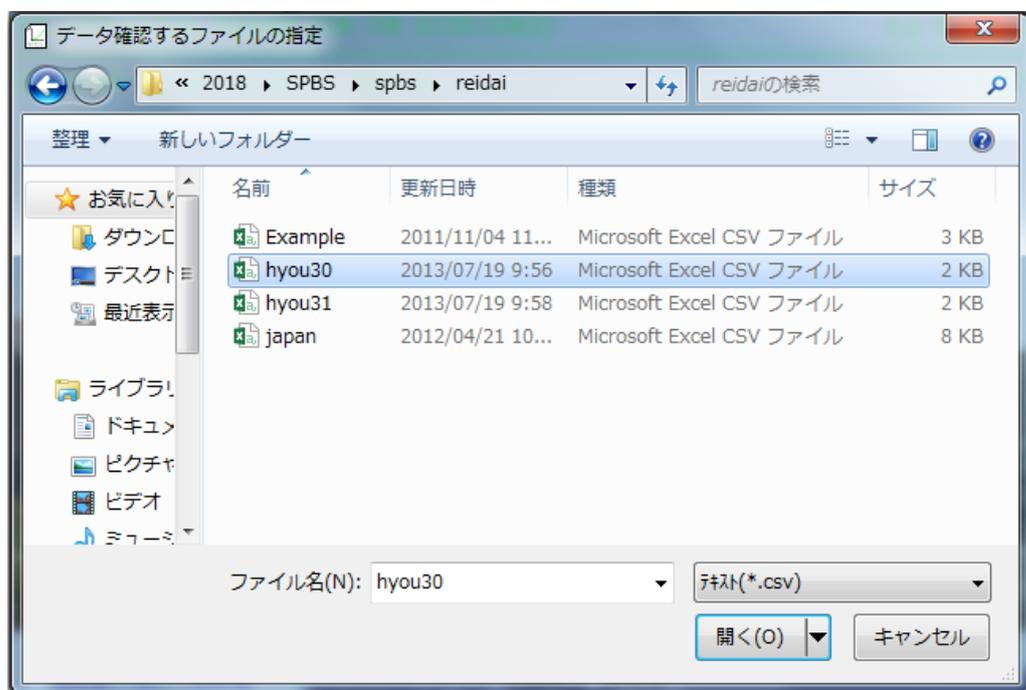


図 3.4-5 解析に用いる用量反応データファイルの選択

カラム/変数	1	2	3	4	5	6	7
1	20	0.662	1.94	0	2.28	0.54	
2	20	0.456	1.39	0	1.83	0.40	
3	19	0.464	1.31	1	1.61	0.47	
4	20	1.458	3.60	1	3.64	1.68	
5	20	0.638	0.76	0	1.41	0.30	
6	20	1.123	1.83	0	1.48	0.83	
7	20	0.253	2.03	0	0.78	0.81	
8	20	0.275	1.16	0	1.07	0.38	
9	20	0.340	0.97	0	0.78	0.41	
10	20	0.452	1.37	0	2.11	0.38	
11	20	0.192	1.72	0	0.75	0.82	
12	19	0.487	1.55	0	1.48	0.66	
13	20	0.614	1.12	0	2.19	0.50	
14	19	0.238	0.94	0	1.36	0.64	
15	20	0.304	1.14	0	0.73	0.38	

有効サンプル数	59	59	59	59	59	59	
欠損値数	0	0	0	0	0	0	
平均値	30	20	0.472	1.52	0	1.77	0.58
最小値	1	19	0.192	0.25	0	0.42	0.19
最大値	59	20	1.458	5.10	1	5.39	1.83

変数名	1	2	3	4	5	6	7
1	ID番号						
2	年齢						
3	尿中Cd						
4	毛髪Hg						
5	喫煙						
6	2-I-MG						
7	爪Hg						

図 3.4-6 解析に用いる用量反応データファイルの確認

② 非ばく露集団の異常率の設定

Benchmark dose Calculation 画面上 (図 3.4-4) で  $P_0 = 0.05$ 、 $P_0 = 0.10$  のいずれかを選択する。(通常  $P_0 = 0.05$  を選択)

③ ばく露群における用量のべき関数によるフィッティング

Benchmark dose Calculation 画面上 (図 3.4-4) でばく露量の  $\log(X+1)$ 変換にチェックを入れて実行する。**実行ボタン**を押し、再度、解析に用いる用量反応データファイルを選択する。

4) データ変数の設定及び解析

各変数の定義付けを行なう。「Dose (ばく露量) を表す変数」を指定し実行ボタンを押す。次に「Response (影響) を表す変数」を指定し実行ボタンを押す (図 3.4-7)。

共変数の設定については、交絡変数の指定を行う画面で設定を行う (図 3.4-8)。

設定に際しては、過去の論文を確認し、共変数として調整した方が良いことが確認された場合は使用し、確認されなかった場合は、一旦共変数として加えて、解析結果「量-影響関係」表の変数名と t 値に対する P 値が有意であれば使用する (この判断は、試行錯誤的に行う)。一方、統計学的に有意でない場合でも、“Response (影響)” 指標と関連することが研究報告等で知られている場合には含めることが望ましい。

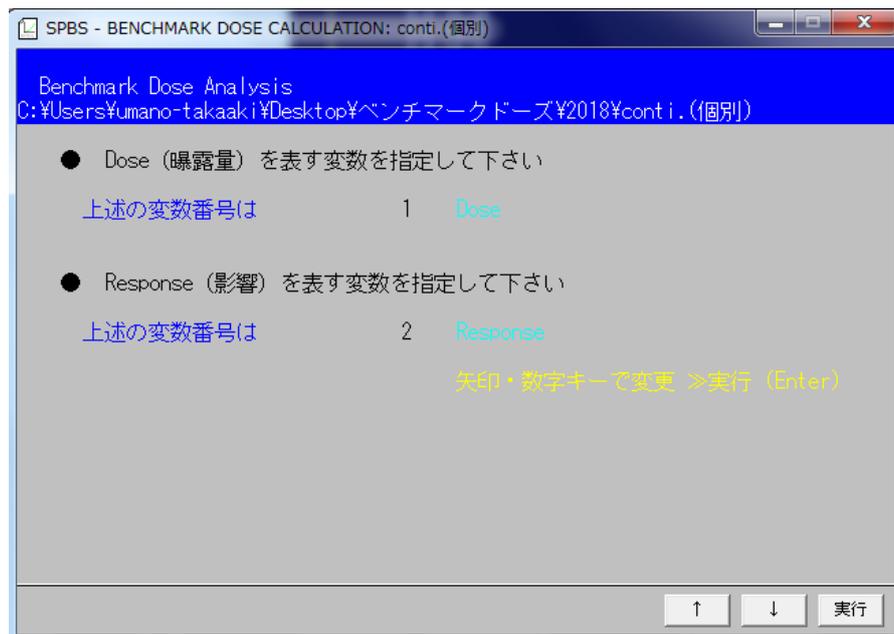
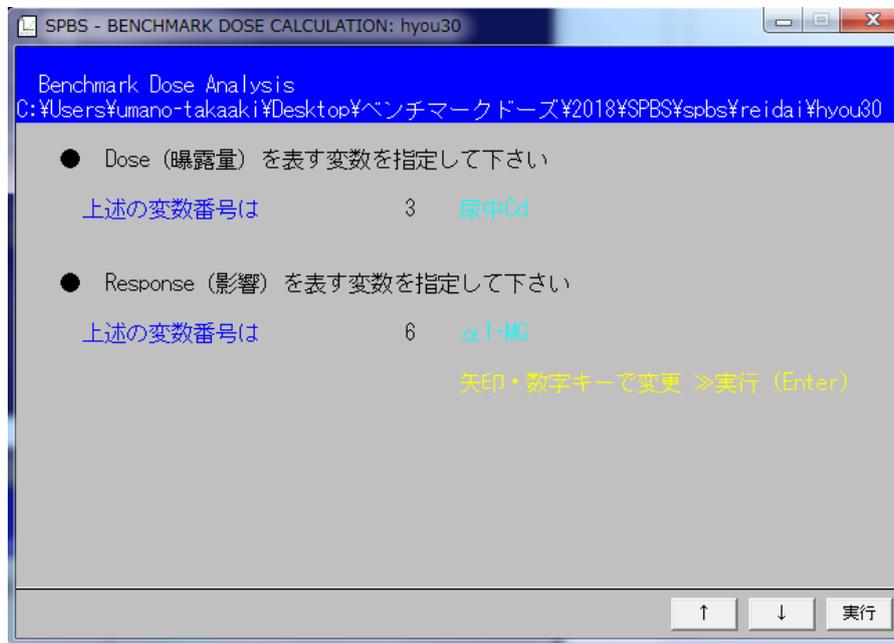


図 3.4-7 データ変数の設定 (ばく露量を表す変数、影響を表す変数) の設定場面



7) 解析結果の出力

解析に使用した用量反応データ (表 3.4-1) 及び解析結果の出力例を以下に示す。

表 3.4-1 使用した用量反応データ (個体データ)

ID 番号	年齢	尿中 Cd	毛髪 Hg	喫煙	$\alpha 1$ -MG	爪 Hg
1	20	0.662	1.94	0	2.28	0.54
2	20	0.456	1.39	0	1.83	0.4
3	19	0.464	1.31	1	1.61	0.47
4	20	1.458	3.6	1	3.64	1.68
5	20	0.638	0.76	0	1.41	0.3
6	20	1.123	1.83	0	1.48	0.89
7	20	0.259	2.03	0	0.78	0.81
8	20	0.275	1.16	0	1.07	0.38
9	20	0.34	0.97	0	0.78	0.41
10	20	0.452	1.37	0	2.11	0.38
11	20	0.192	1.72	0	0.75	0.82
12	19	0.487	1.55	0	1.48	0.66
13	20	0.614	1.12	0	2.19	0.5
14	19	0.238	0.94	0	1.36	0.64
15	20	0.304	1.14	0	0.73	0.38
16	20	0.456	3.97	0	0.79	0.98
17	20	0.259	0.55	0	0.42	0.21
18	20	1.273	1.18	0	1.52	0.38
19	20	0.41	1.44	0	1.24	0.46
20	20	0.389	1.06	0	0.97	0.39
21	20	0.214	2.05	0	1.49	0.71
22	19	0.406	1.16	0	2.07	0.44
23	20	0.575	3.54	0	2.05	1.02
24	20	0.307	1.74	0	1.29	0.94
25	20	0.561	1.29	0	2.9	0.41
26	20	0.252	1.89	0	1.03	0.58
27	20	0.268	1.96	0	1.22	0.73
28	20	0.678	1.52	0	1.7	0.47
29	20	0.278	1.43	0	0.7	0.49

ID 番号	年齢	尿中 Cd	毛髪 Hg	喫煙	$\alpha$ 1-MG	爪 Hg
30	20	0.395	1.73	0	2.12	0.9
31	20	0.336	0.98	0	1.46	0.3
32	20	0.554	5.1	1	5.39	1.83
33	19	0.302	0.74	0	1.96	0.22
34	20	0.722	2.39	0	1.99	0.72
35	20	0.344	0.59	0	2.31	0.38
36	20	0.297	0.72	1	1.81	0.31
37	20	0.51	1.34	0	1.27	0.45
38	20	0.36	0.78	0	1.92	0.33
39	20	1.259	1.08	0	3.15	0.35
40	20	0.352	3.44	1	1.6	0.83
41	19	0.868	1.57	1	2.17	0.72
42	20	0.371	1.5	1	2	0.65
43	20	1.017	1.7	0	5.31	1.08
44	20	0.299	1.32	0	3.51	0.58
45	20	0.245	1.39	0	1.28	0.52
46	19	0.533	2.21	0	2.07	0.75
47	20	0.3	0.49	0	1.44	0.19
48	20	0.391	0.69	0	0.78	0.22
49	20	0.43	1.63	0	1.67	0.47
50	20	0.609	0.6	0	1.65	0.34
51	20	0.318	1.53	0	0.66	0.75
52	20	0.209	0.25	0	1.32	0.26
53	20	0.192	1.13	0	1.46	0.49
54	20	0.382	0.76	0	0.72	0.51
55	20	0.647	1.14	1	3.56	0.4
56	20	0.37	2.49	0	2.14	1.06
57	20	0.368	0.39	0	1.36	0.58
58	20	0.213	0.82	1	1.79	0.29
59	20	0.374	1.3	0	1.43	0.74

(SPBS ソフトのダウンロードフォルダ内にあるデータから引用)

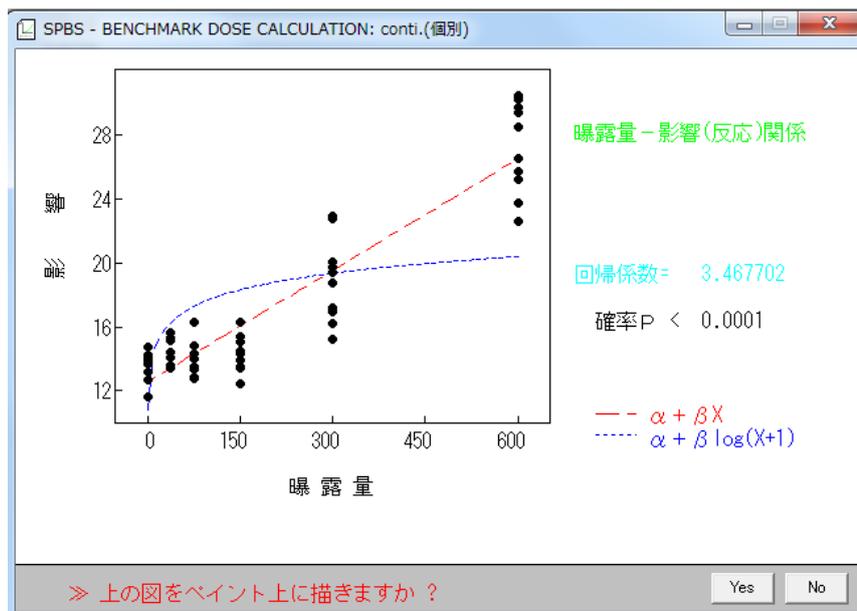
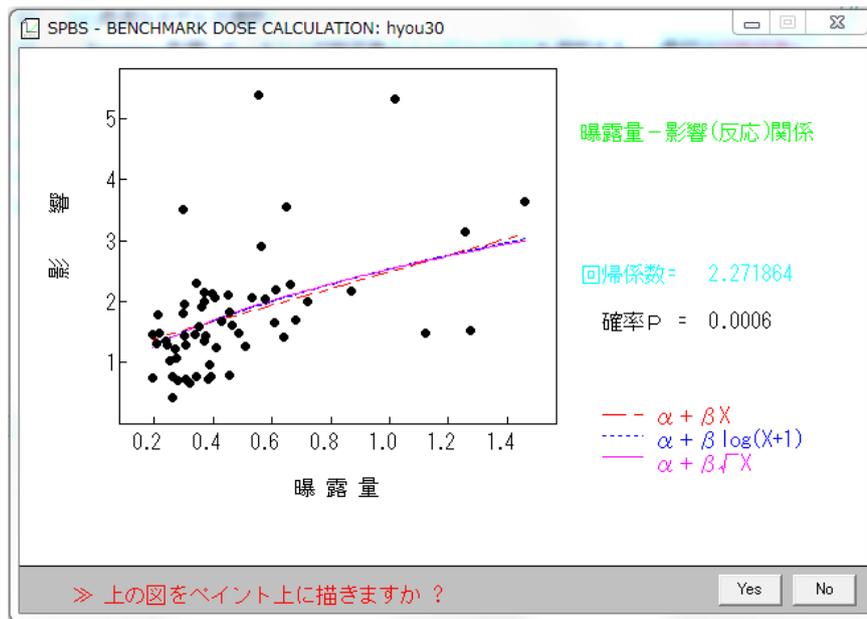


図 3.4-9 一次回帰モデル、ばく露量の  $\log(X+1)$  モデル及びべき関数モデルによる用量反応データ

① 一次回帰モデルによる解析結果 (基本的に画面の表示内容を転載)

■■ 基本データ ■■

解析の症例数 = 59

変数	平均値	標準偏差	最小値	最大値	変数名
1	0.472	0.277	0.192	1.458	尿中 Cd ← ばく露変数
2	1.515	0.907	0.250	5.100	毛髪 Hg ← 交絡変数
3	0.588	0.319	0.190	1.830	爪 Hg ← 交絡変数
4	1.766	0.983	0.420	5.390	α1-MG ← 影響変数

◇ 相関行列 [▲: P 値 ▼: 相関係数]

	1	2	3	4
1		0.248	0.316	0.498
2	.0580		0.857	0.368
3	.0148	<.0001		0.471
4	<.0001	.0042	.0002	

量-影響関係 (回帰式の F 値 10.322 P<0.0001)

回帰係数	標準回帰係数	標準誤差	t 値	変数名
1.366	0.38542	0.4033	3.387 (P=0.0013)	尿中 Cd
-0.1022	-0.10403	0.2061	-0.496 (P=0.6219)	毛髪 Hg
0.4309	0.43843	0.2105	2.047 (P=0.0454)	爪 Hg
1.121				定数項

Cutoff point = 2.535 ← この値の妥当性を検討する

○ Benchmark Dose Calculation

Po=0.05 BMR=0.02 BMD= 0.106 (BMDL= 0.069)

Po=0.05 **BMR=0.05** BMD= 0.229 (BMDL= 0.148) ← 通常この値を使用

Po=0.05 BMR=0.10 BMD= 0.383 (BMDL= 0.248)

全尤値 (定数項のみの対数尤度: 共変量によらない対数尤度) 171.43  
 偏尤値 (回帰モデルの対数尤度: 共変量からの最大対数尤度) 145.08  
 モデルの尤度比検定(-2 LOG L) 26.35 (自由度 = 3) P<0.0001

② ばく露量の log (X+1) モデルによる解析結果(基本的に画面の表示内容を転載)

■■ 基本データ ■■

解析の症例数 = 59

変数	平均値	標準偏差	最小値	最大値	変数名
1	0.162	0.072	0.076	0.391	尿中 Cd ← ばく露変数
2	1.515	0.907	0.250	5.100	毛髪 Hg ← 交絡変数
3	0.588	0.319	0.190	1.830	爪 Hg ← 交絡変数
4	1.766	0.983	0.420	5.390	α1-MG ← 影響変数

注) ばく露データは log(X+1)変換値

◇ 相関行列 [▲: P値 ▼: 相関係数]

	1	2	3	4
1		0.256	0.306	0.511
2	.0505		0.857	0.368
3	.0183	<.0001		0.471
4	<.0001	.0042	.0002	

量-影響関係 (回帰式の F 値 10.986 P<0.0001)

回帰係数	標準回帰係数	標準誤差	t 値	変数名
5.500	0.40409	1.525	3.607 (P=0.0007)	尿中 Cd
-0.1242	-0.12641	0.2036	-0.610 (P=0.5442)	毛髪 Hg
0.4477	0.45557	0.2068	2.165 (P=0.0347)	爪 Hg
0.8775				定数項

Cutoff point = 2.277

○ Benchmark Dose Calculation [log(X+1)変換値から算出]

Po=0.05 BMR=0.02 BMD= 0.062 (BMDL= 0.041)

Po=0.05 **BMR=0.05** BMD= 0.138 (BMDL= 0.089) ← 通常この値を使用

Po=0.05 BMR=0.10 BMD= 0.242 (BMDL= 0.154)

全尤値 (定数項のみの対数尤度: 共変量によらない対数尤度) 171.43

偏尤値 (回帰モデルの対数尤度: 共変量からの最大対数尤度) 143.73

モデルの尤度比検定(-2 LOG L) 27.70 (自由度 = 3) **P<0.0001**

③ ばく露量のべき関数モデルによる解析結果(基本的に画面の表示内容を転載)

■■ 基本データ ■■

解析の症例数 = 59

変数	平均値	標準偏差	最小値	最大値	変数名
1	0.664	0.176	0.438	1.207	尿中 Cd ← ばく露変数
2	1.515	0.907	0.250	5.100	毛髪 Hg ← 交絡変数
3	0.588	0.319	0.190	1.830	爪 Hg ← 交絡変数
4	1.766	0.983	0.420	5.390	α1-MG ← 影響変数

注) ばく露データは $\sqrt{X}$  変換値

◇ 相関行列 [▲: P値 ▼: 相関係数]

	1	2	3	4
1		0.259	0.303	0.513
2	.0475		0.857	0.368
3	.0195	<.0001		0.471
4	<.0001	.0042	.0002	

量-影響関係 (回帰式の F 値 11.129 P<0.0001)

回帰係数	標準回帰係数	標準誤差	t 値	変数名
2.272	0.40779	0.6219	3.653 (P=0.0006)	尿中 Cd
-0.1330	-0.13536	0.2031	-0.655 (P=0.5152)	毛髪 Hg
0.4553	0.46329	0.2058	2.212 (P=0.0311)	爪 Hg
0.2563				定数項

Cutoff point = 1.654

○ Benchmark Dose Calculation [√X 変換値から算出]

Po=0.05 BMR=0.02 BMD= 0.004 (BMDL= 0.002)

Po=0.05 **BMR=0.05** BMD= 0.018 (BMDL= 0.008) ← 通常この値を使用

Po=0.05 BMR=0.10 BMD= 0.052 (BMDL= 0.023)

全尤値 (定数項のみの対数尤度：共変量によらない対数尤度) 171.43

偏尤値 (回帰モデルの対数尤度：共変量からの最大対数尤度) 143.45

**モデルの尤度比検定(-2 LOG L) 27.99 (自由度 = 3) P<0.0001**

## 3.5 MADr-BMD

### 3.5.1 名称及びバージョン

MADr-BMD ver.1.0 (2019年2月末時点)

### 3.5.2 MADr-BMD の概要

#### (1) 概要

MADr-BMD は、Matthew W. Wheeler (CDC/NIOSH) と A. John Bailer (マイアミ大学と CDC/NIOSH) によって 2007 年に開発された、無償で入手できるソフトウェアである。本ソフトウェアは、Microsoft Windows オペレーションシステムにおけるシェルコマンドプロンプト60上で操作する他のプログラムを要せず単体で動作可能なプログラムである。操作は、マウスではなくキーボードでコマンドと呼ばれる命令文を入力して行う。

#### (2) 主要な機能

主要な機能は、二値の用量反応データを対象とした数理モデル平均化手法による BMD 及び BMD の信頼区間の下限値 (BMDL) の算出である。収載されている数理モデルを表 3.5-1 に示す。数理モデルの選択は手動で行う (単一モデルの選択も可能)。BMR は 1%、5%、10%等の任意の値を入力できる。モデル性能の評価法は AIC、BIC、KIC から選択する (その他 BICB、AICB、KICB も選択可能)。また、信頼区間の推定はブートストラップ法、パラメータ推定は最尤推定による。MADr-BMD に収載されている数理モデルを表 3.5-1 に示す。

表3.5-1 MADr-BMDに収載されている数理モデル

数理モデル	式	パラメータ
Logistic	$\pi_1(d) = \frac{1}{1 + \exp[-(\alpha + \beta \times d)]}$	
Log-logistic	$\pi_2(d) = \gamma + \frac{(1 - \gamma)}{1 + \exp[-(\alpha + \beta \times d)]}$	$0 \leq \gamma \leq 1$ $\beta \geq 0.5$

<sup>60</sup> コマンドと呼ばれる命令文を用いて Windows の操作や設定を行うためのツール

数理モデル	式	パラメータ
Gamma	$\pi_3(d) = \gamma + \frac{1-\gamma}{\Gamma(\alpha)} \int_0^{\beta d} t^{\alpha-1} e^{-t} dt$	$0 \leq \gamma \leq 1$ $\alpha \geq 1, \beta \geq 0.5$
Multistage	$\pi_4(d) = \gamma + (1-\gamma)(1 - \exp(-\theta_1 d - \theta_2 d^2 \dots))$	$0 \leq \gamma \leq 1$ $\theta_1 \geq 0, \theta_2 \geq 0, \dots$
Probit	$\pi_5(d) = \Phi(\alpha + \beta d)$	
Log-probit	$\pi_6(d) = \gamma + (1-\gamma)\Phi(\alpha + \beta \ln(d))$	$0 \leq \gamma \leq 1, \beta \geq 0$
Quantal Linear	$\pi_7(d) = \gamma + (1-\gamma)(1 - \exp(-\beta d))$	$0 \leq \gamma \leq 1, \beta \geq 0$
Quantal quadratic	$\pi_8(d) = \gamma + (1-\gamma)(1 - \exp(-\beta d^2))$	$0 \leq \gamma \leq 1, \beta \geq 0$
Weibull	$\pi_9(d) = \gamma + (1-\gamma)(1 - \exp(-\beta d^\alpha))$	$0 \leq \gamma \leq 1$ $\beta \geq 0.5$

(Wheeler and Bailer (2008)から引用)

### 3.5.3 利用方法に関する情報

#### (1) セットアップ方法

##### 1) MADr-BMD ソフト等のダウンロード

米国 EPA のサイトから入手できる<sup>61</sup>

また、操作方法に関する資料 (操作マニュアル、ガイダンス等) の入手方法は以下のとおりである。

- ・ 操作方法は EPA の上記のサイトに記載されている。
- ・ 統計学的なアプローチ、使用例、設定例については、Wheeler and Bailer (2008) の論文から入手できる。
- ・ ソースコードや解析に必要な入力ファイルの例 (input と表示されているテキストファ

<sup>61</sup> <https://www.epa.gov/tsca-screening-tools/model-averaging-dichotomous-response-benchmark-dose-madr-bmd-tool>

イル) も上記のサイトから入手できる。

## 2) madrbmd ソフトと入力ファイルの書き込み可能なフォルダへの移動

書き込み可能なフォルダに、ダウンロードした madrbmd ソフトと入力ファイルを移動する。

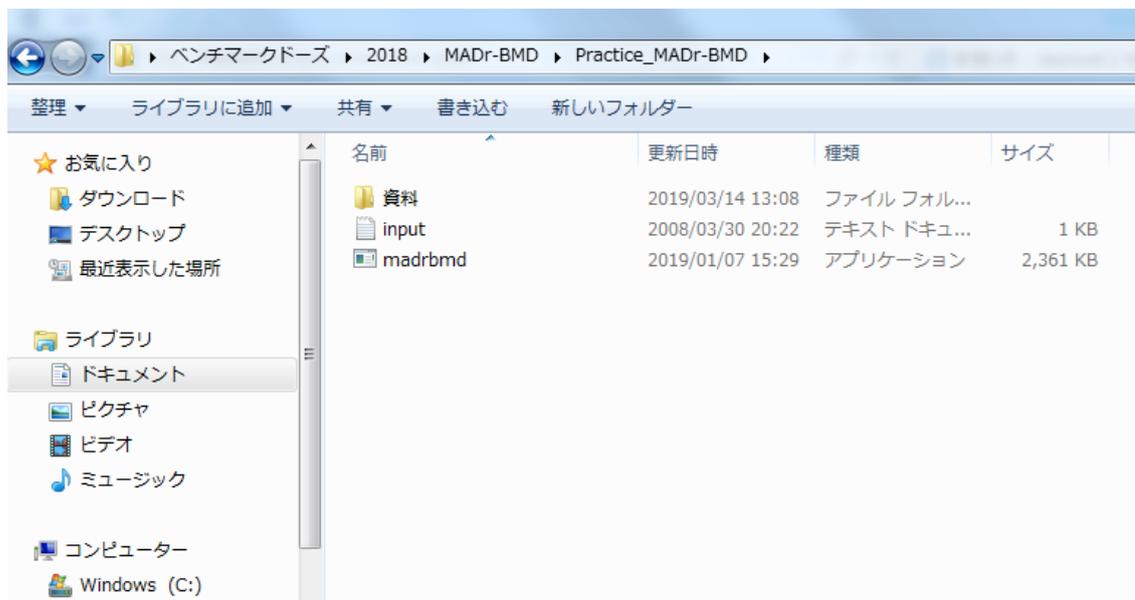


図 3.5-1 ダウンロードした madrbmd ソフトと入力ファイルの保管場所の例

## (2) 操作の流れ

MSDr-BMD では、入力ファイル上で、解析する用量反応データとともに、解析に当たっての各種条件を入力、指定する必要がある。

操作の流れを以下に示す。

### 1) データの入力

入力ファイルの入力例及び設定について以下に示す (図 3.5-2、表 3.5-2)。

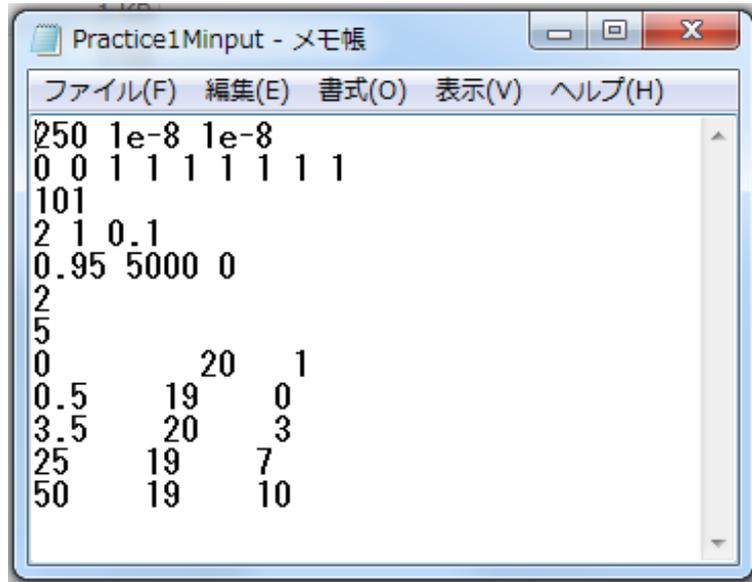


図 3.5-2 入力ファイルの入力例

表 3.5-2 入力ファイルで設定する解析条件

250 1e-8 1e-8	最大反復回数、関連する収束、一般的な収束
0 0 1 1 1 1 1 1	モデル平均化 (MA) に含める数理モデル (1 = 含める、0 = 含めない) 左から次の数理モデルを示している (Quantal-linear、Quantal-quadratic、Multistage、Logistic、Probit、Weibull、Log-probit、Log-logistic、Gamma) (単一モデルのみをフィッティングしたい場合は、当該モデルの順番の位置に 1 と入力し、その他の数理モデルは 0 を入力する。)
101	ランダムシード (0 を用いると現在の時刻を使用)
2 1 0.1 ① ② ③	①平均化の際に重み付けとして用いる指標 (1 = BIC、2 = AIC、3 = KIC、4 = BICB、5 = AICB、6 = KICB) ②BMR のタイプ (1 = 追加リスク (added risk)、2 = 過剰リスク (extra risk)) ③BMR の設定値
0.95 5000 0	①BMD の信頼区間の信頼水準 (BMD の信頼区間の下限値 (BMDL) 算出に反映される)

	②ブートストラップサンプリング数 ③ブートストラップサンプルの出力 (0=しない、1=する)
2	Multistage の次数
5	データの行数
0 20 1 0.5 19 0 3.5 20 3 25 19 7 50 19 10	データの入力項目 1列目：用量、2列目：個体数 (n 数)、3列目：発生数

実際にデータの入力、解析時の条件を設定する場合は、ダウンロードした入力ファイル (input と表示されているテキストファイル) を基に入力し、入力ファイルを“任意の名称” .txt として保存する。この際、ファイル名にスペースは使用しない。

## 2) データの解析

### ① コマンドプロンプトの起動

使用するパソコンの OS に応じた方法でコマンドプロンプトウィンドウを開く (図 3.5-3) (例：Windows 7 の場合は、スタートメニューを開き、一番下のエントリーに “cmd” と入力、Enter キーを押す)。

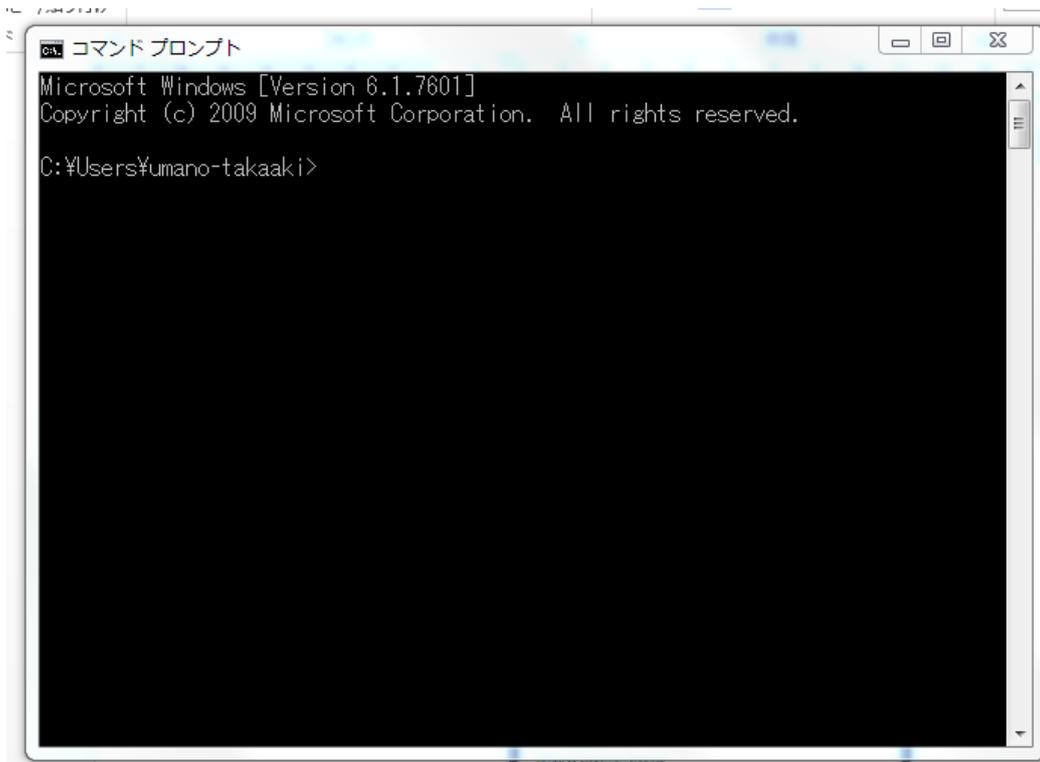


図 3.5-3 コマンドプロンプトウィンドウ

## ② ディレクトリの変更と確認

### ・ディレクトリの変更 (図 3.5-4)

**madrbmd** ソフトと解析に用いるデータ入力ファイルが保存されているディレクトリに変更するため、“**cd コマンド**”を用いて変更する。

コマンドプロンプトウィンドウに表示されているディレクトリ名の後に、**cd+半角スペース+「madrbmd ソフトとデータ入力ファイルが保存されているディレクトリ名」**を入力

ディレクトリの変更例：保存先が **Desktop¥ベンチマーク ドーズ¥2018¥MADr-BMD¥Practice\_MADr-BMD** の場合

**C:\Users\umano-takaaki>**の後に、**cd C:\Users\umano-takaaki¥Desktop¥ベンチマーク ドーズ¥2018¥MADr-BMD¥Practice\_MADr-BMD** と入力し **Enter**

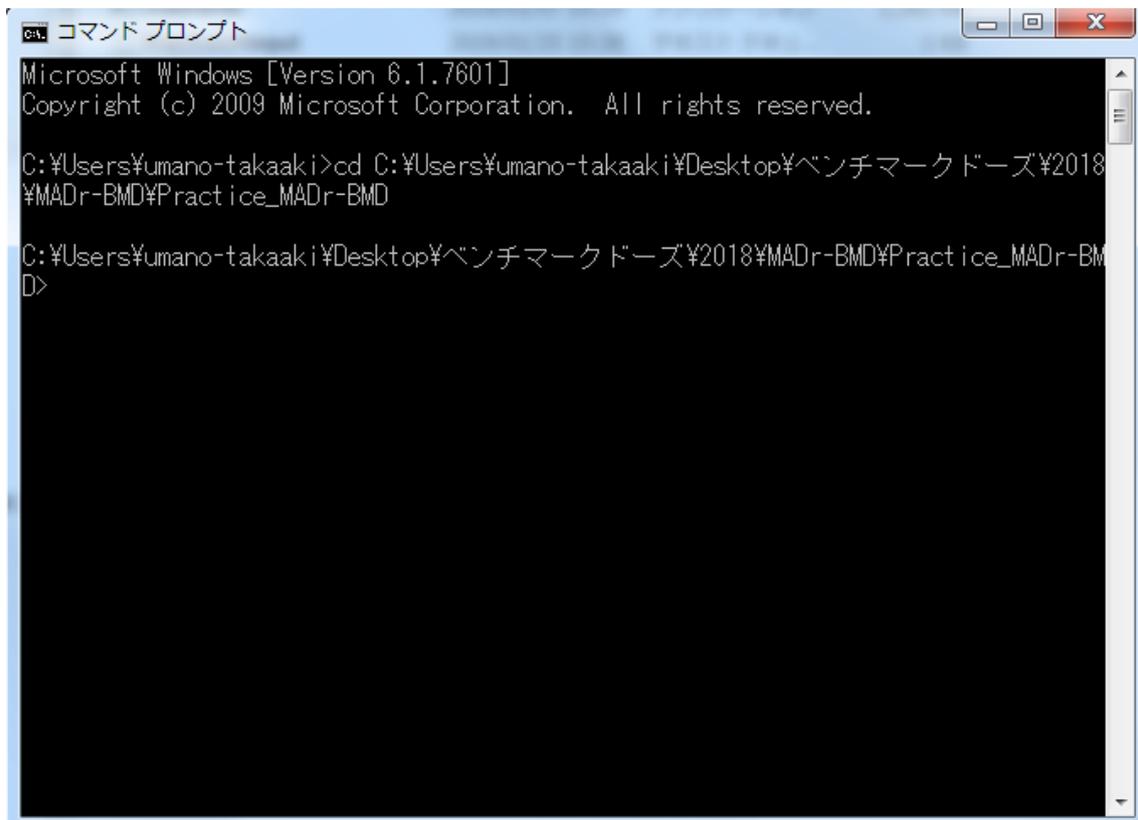


図 3.5-4 ディレクトリの変更操作画面

・ディレクトリの変更の確認

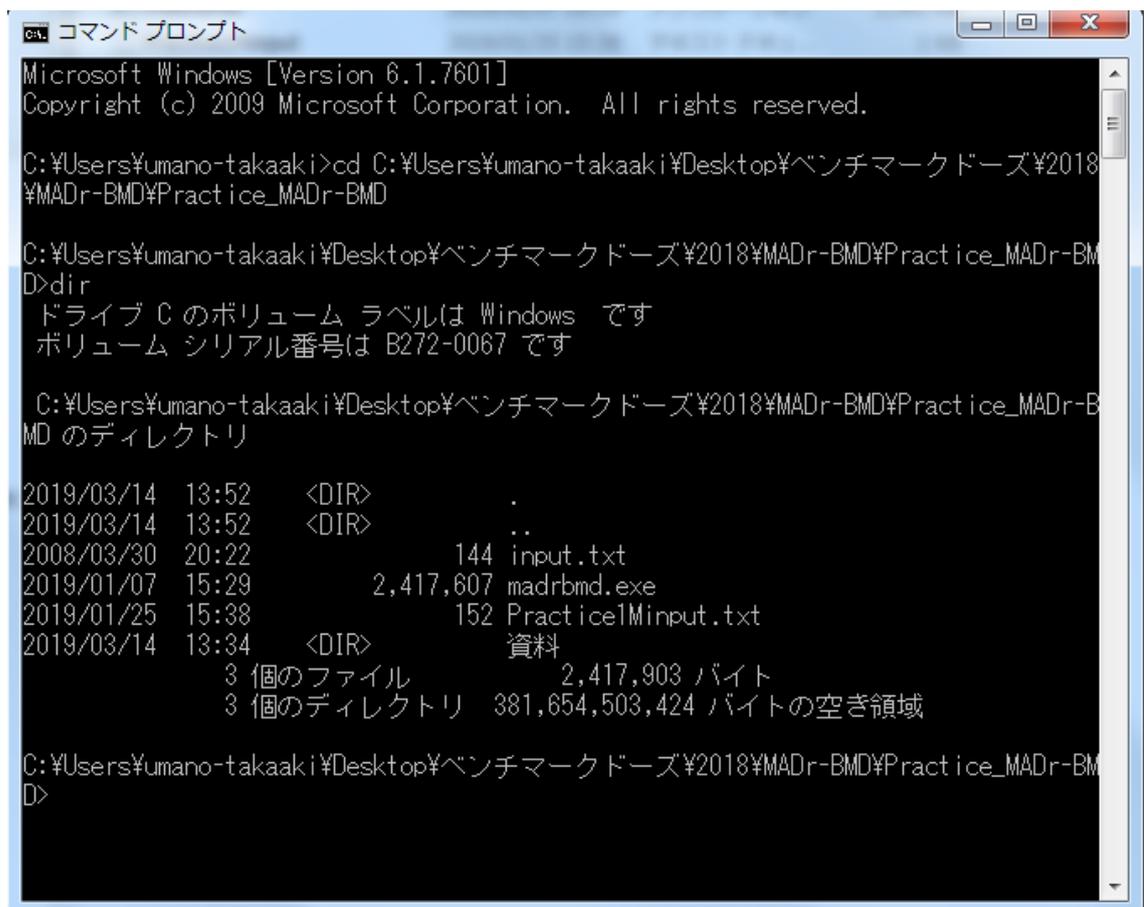
madrbmd ソフトとデータ入力ファイルは同じディレクトリにあることが必要なため、操作中のディレクトリにあることを次の操作により確認する。

“dir コマンド”を用いてディレクトリの変更を確認する。“dir”と入力するとディレクトリにあるファイルが表示され確認できる。

dir コマンドの例 :

C:\Users\umano-takaaki>cd C:\Users\umano-takaaki\Desktop\ベンチマークドーズ\2018\MADr-BMD\Practice\_MADr-BMD>の後に dir を入力

図 3.5-5 に示すように、解析対象のデータである Practice1Minput.txt ファイル、解析ソフトである madrbmd.exe ファイルがこのディレクトリにあることが確認できる。



```
ca. コマンド プロンプト
Microsoft Windows [Version 6.1.7601]
Copyright (c) 2009 Microsoft Corporation. All rights reserved.

C:\Users\umano-takaaki>cd C:\Users\umano-takaaki\Desktop\ベンチマークドーズ\2018
\MADr-BMD\Practice_MADr-BMD

C:\Users\umano-takaaki\Desktop\ベンチマークドーズ\2018\MADr-BMD\Practice_MADr-BM
D>dir
ドライブ C のボリューム ラベルは Windows です
ボリューム シリアル番号は B272-0067 です

C:\Users\umano-takaaki\Desktop\ベンチマークドーズ\2018\MADr-BMD\Practice_MADr-B
MD のディレクトリ

2019/03/14  13:52    <DIR>          .
2019/03/14  13:52    <DIR>          ..
2008/03/30  20:22                144  input.txt
2019/01/07  15:29           2,417,607  madrbmd.exe
2019/01/25  15:38           152  Practice1Minput.txt
2019/03/14  13:34    <DIR>          資料
                3 個のファイル          2,417,903 バイト
                3 個のディレクトリ 381,654,503,424 バイトの空き領域

C:\Users\umano-takaaki\Desktop\ベンチマークドーズ\2018\MADr-BMD\Practice_MADr-BM
D>
```

図 3.5-5 ディレクトリの変更の確認画面

③ MADr-BMD ソフトの起動・解析

解析に用いる入力データの読み込みは、次のように `madrbmd+半角スペース+入力ファイル名.txt` と入力する。例えば、「`madrbmd Practice1Minput.txt`」と入力すると、入力ファイルとして `Practice1Minput.txt` を使用するようにプログラムに指示がされる。数十秒程度で、解析結果がコマンドプロンプトウィンドウに表示される。

入力データの読み込み、解析例 (図 3.5-6) :

入力データが `madrbmd Practice1Minput.txt` の場合、

`C:\Users\umano-takaaki>cd C:\Users\umano-takaaki\Desktop\ベンチマークドーズ\2018\MADr-BMD\Practice_MADr-BMD>` の後に、「`madrbmd Practice1Minput.txt`」を入力

```

コマンド プロンプト
C:\Users\umano-takaaki\Desktop\ベンチマークドーズ\2018\MADr-BMD\Practice_MADr-BM
D>madrbmd Practice1Minput.txt
MABMD VERSION 1.0
Thu Mar 14 15:34:03 2019

This program's results are given "as is", without warranty, either expressed
or implied by the National Institute for Occupational Safety and Health.

INPUT DATA
Dose          Count      Observed
-----
0.000000     20         1
0.500000     19         0
3.500000     20         3
25.000000    19         7
50.000000    19         10
-----

Model Fit Statistics
-----
Model          Weight  -2log(L)      AIC      BIC
-----
Multistage     0.162   78.87        84.87    92.59
Logistic       0.111   81.63        85.63    90.78
Probit         0.137   81.21        85.21    90.36
Weibull        0.162   78.87        84.87    92.59
Log-Probit     0.068   80.61        86.61    94.33
Log-Logistic   0.199   78.45        84.45    92.18
Gamma          0.162   78.87        84.87    92.59
-----

'Average-Model' Benchmark Dose Estimate
-----
Nominally Specified Confidence Level:0.950
Weighting Criterion: AIC
BMD Calculation: Added Risk
BMR: 0.200000
BMD: 16.991535812414
BMDL(BCa):7.163513952322
BMDL(Percentile):11.500404572231
Acceleration: 0.030539
Bootstrap Resamples: 5000
Random Seed: 101
-----

```

図 3.5-6 解析結果画面

### 3) 解析結果の出力

ソフト起動時に、**madrbrmd** 解析に用いる入力ファイル.txt ユーザー指定の出力ファイル名.txt とすると、解析結果がユーザー指定の出力ファイル名の付いたテキストとして、同じディレクトリに自動で保存される。(この際、出力ファイル名は必ず独立したファイル名とするように注意する)(図 3.5-7)

例：

```
C:\Users\umano-takaaki>cd C:\Users\umano-takaaki\Desktop\ベンチマーク ドーズ
¥2018¥MADr-BMD¥Practice_MADr-BMD> madrbrmd Practice1Minput.txt> Practice1Moutput.txt
```

なお、出力ファイルを設定しなかった場合、結果が表示されているコマンドプロンプトウィンドウを右クリックして全てを選択した後コピーし、ワードファイル等にペーストすることで、解析結果を出力することもできる。

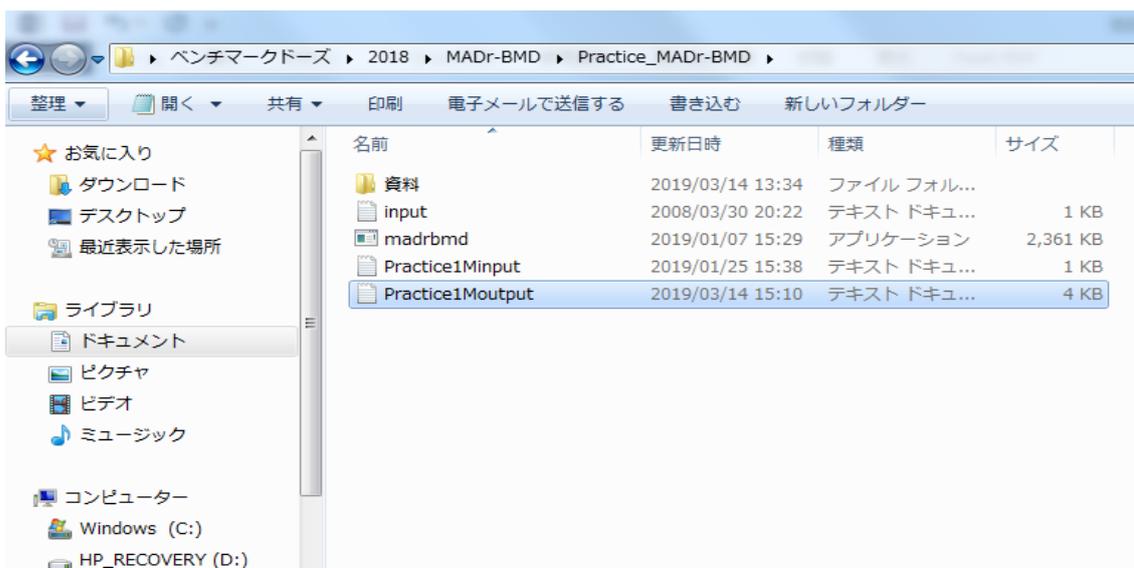


図 3.5-7 出力ファイルの自動保存

解析結果の出力例を以下に示す。

MABMD VERSION 1.0

Thu Mar 14 17:00:07 2019

This program's results are given "as is", without warranty, either expressed or implied by the National Institute for Occupational Safety and Health.

INPUT DATA

Dose	Count	Observed
0.000000	20	1
0.500000	19	0
3.500000	20	3
25.000000	19	7
50.000000	19	10

Model Fit Statistics

Model	Weight	-2log(L)	AIC	BIC
Multistage	0.162	78.87	84.87	92.59
Logistic	0.111	81.63	85.63	90.78
Probit	0.137	81.21	85.21	90.36
Weibull	0.162	78.87	84.87	92.59
Log-Probit	0.068	80.61	86.61	94.33
Log-Logistic	0.199	78.45	84.45	92.18
Gamma	0.162	78.87	84.87	92.59

'Average-Model' Benchmark Dose Estimate

-----  
Nominally Specified Confidence Level: 0.950

Weighting Criterion: AIC

BMD Calculation: Added Risk

BMR: 0.100000

BMD: 8.212328013576

BMDL (BCa):3.041899204839

BMDL (Percentile):5.574802713355

Acceleration: 0.027111

Bootstrap Resamples: 5000

Random Seed: 101  
-----

-----  
'Average-Model' Goodness of Fit Test  
-----

Test Statistic : 2.135780

Bootstrap P-Value: 0.444600  
-----

MODEL: Multistage, 2-degree polynomial:  
-----

Parameters	Estimate	StdErr
GAMMA:	3.852992e-002	3.188900e-002
BETA(1):	1.573000e-002	4.292851e-003
BETA(2):	0.000000e+000	N/A

Optimization Succeeded

-----  
MODEL: Logistic:  
-----

Parameters	Estimate	StdErr
ALPHA:	-2.460499	0.446256

BETA: 0.056080 0.013585

Optimization Succeeded

---

MODEL: Probit:

---

Parameters	Estimate	StdErr
------------	----------	--------

---

ALPHA:	-1.455329	0.234799
--------	-----------	----------

BETA:	0.033341	0.007740
-------	----------	----------

Optimization Succeeded

---

MODEL: Weibull:

---

Parameters	Estimate	StdErr
------------	----------	--------

---

GAMMA:	0.038530	0.031884
--------	----------	----------

ALPHA:	1.000000	N/A
--------	----------	-----

BETA:	0.015730	0.004293
-------	----------	----------

Optimization Succeeded

---

MODEL: Log-Probit:

---

Parameters	Estimate	StdErr
------------	----------	--------

---

GAMMA:	0.064740004794	0.033896
--------	----------------	----------

ALPHA:	-3.794321	0.245232
--------	-----------	----------

BETA:	1.000000	N/A
-------	----------	-----

Optimization Succeeded

---

MODEL: Log-Logistic:

---

Parameters	Estimate	StdErr
GAMMA:	0.032534	0.029626
ALPHA:	-3.806055	0.348835
BETA:	1.000000	N/A

Optimization Succeeded

MODEL: Gamma:

Parameters	Estimate	StdErr
GAMMA:	0.038530	0.031891
ALPHA:	1.000000	N/A
BETA:	0.015730	0.004293

Optimization Succeeded

引用

Matthew W. Wheeler, A. John Bailer (2008) Model Averaging Software for Dichotomous Dose Response Risk Estimation. Journal of Statistical Software. Vol. 26, Issue 5.

<https://www.jstatsoft.org/article/view/v026i05>

## 3.6 BMDEExpress

---

### 3.6.1 名称及びバージョン

---

最新バージョン：BMDEExpress 2.2 (2019年2月末時点)

2018年10月30日にリリースされた最新バージョン2.2にはLinux版、Mac版、Windows 64bit版が存在する。但し、Linux版及びMac版では、BMDEExpressプロジェクトファイルの保存形式である.bm2ファイルの閲覧や既存のBMD解析における機能分類のみ可能で、BMD解析や一方向ANOVAフィルタリングは実行できない。

### 3.6.2 BMDEExpress の概要

---

#### (1) 概要

BMDEExpressは、毒性学、医薬品設計、リスクアセスメント及びトランスレーショナルリサーチ<sup>62</sup>の分野で増加しているトランスクリプトミクスによる用量反応データ<sup>63</sup>の解析を目的に開発された無償のデスクトップアプリケーションである。このアプリケーションでは、トキシコゲノミクスデータからのBMDの推計と、ユーザーが定義したカテゴリー(シグナル経路、遺伝子オントロジー(GO)等)に基づく機能分類解析を組み合わせたワークフローが提供される。その開発は米国の国立環境衛生科学研究所(NIEHS)、カナダ保健省、米国EPA及びSciome社によって行われ、ソフトウェアの設計や実装はSciome社が担当した。

BMDEExpressでは、個々の遺伝子プローブや関連する統計量の用量反応データを近似したパラメータ曲線と、視覚化された一連の分析結果が得られる。トランスクリプトミクスは、ガイドラインに従った毒性試験と比べ、リソースが集約されていないが、BMDEExpressアプローチによってトランスクリプトミクスデータを用いた用量反応モデリングが可能になってきた。リスクアセスメントで考慮する必要がある生物学的因子の適用範囲が広がったと認知されるようになった。

BMDEExpressを用いた解析の特徴として、下記の解析フロー(特に機能分類解析モジュール)が挙げられる。その他、視覚化による解析効率の向上も本ソフトの特徴といえる。

BMDEExpressで採用されている全ての用量反応曲線のフィッティングモデルは、米国EPAのBMDSソフトウェア<sup>64</sup>に記載されている連続値に対応したパラメトリックモデル

---

<sup>62</sup> 大学等における基礎研究の優れた成果を次世代の革新的な診断・治療法の開発につなげるための橋渡し研究。

<sup>63</sup> 薬物等を投与又は適用した場合の各用量におけるトランスクリプトーム(特定の条件下において細胞中に存在する全てのmRNA(又は一次転写産物)の総体)をDNAマイクロアレイ等の技術で網羅的に解析したデータ。

<sup>64</sup> <https://www.epa.gov/bmnds>

(continuous parametric models) である。個々のプローブ<sup>65</sup>や遺伝子レベルのモデル化プロセスは、米国 EPA が毒性評価 (US EPA, 2012) において推奨しているものと概ね一致する。

## (2) 主要な機能

### (2)-1 プラットフォーム及びアノテーション

BMDEExpress では、現在までに各メーカーのプラットフォーム (Affymetrix : 20 種、Agilent : 8 種、Illumina : 4 種及び BioSpyder : 3 種) と複数の RNA-Seq アノテーション<sup>66</sup> (RefSeq : 5 種及び Ensembl : 5 種) 合計 40 種がサポートされている (表 3.6-1)。

また、GO アノテーションは最新のものが反映され、パスウェイ分類<sup>67</sup>は REACTOME の定義を採用している。アノテーションはカスタムで定義することも可能である。

表 3.6-1 BMDEExpress でサポートされているマイクロアレイ  
及び次世代シーケンスプラットフォーム

ID	名称	プロバイダー	Species
GPL1319	Zebrafish Genome Array	Affymetrix	<i>Danio rerio</i>
GPL1322	Drosophila Genome 2.0 Array	Affymetrix	<i>Drosophila melanogaster</i>
GPL72	Drosophila Genome Array	Affymetrix	<i>Drosophila melanogaster</i>
GPL13158	HT HG-U133+ PM Array Plate	Affymetrix	<i>Homo sapiens</i>
GPL570	Human Genome U133 Plus 2.0 Array	Affymetrix	<i>Homo sapiens</i>
GPL571	Human Genome U133A 2.0 Array	Affymetrix	<i>Homo sapiens</i>
GPL96	Human Genome U133A Array	Affymetrix	<i>Homo sapiens</i>
GPL341	Rat Expression 230A Array	Affymetrix	<i>Homo sapiens</i>
GPL11180	HT MG-430 PM Array Plate	Affymetrix	<i>Mus musculus</i>
GPL16985	HT RG-230 PM Array Plate	Affymetrix	<i>Mus musculus</i>
GPL201	Human HG-Focus Target Array	Affymetrix	<i>Mus musculus</i>
GPL1261	Mouse Genome 430 2.0 Array	Affymetrix	<i>Mus musculus</i>
GPL8321	Mouse Genome 430A 2.0 Array	Affymetrix	<i>Mus musculus</i>
GPL81	Murine Genome U74A 2.0 Array	Affymetrix	<i>Mus musculus</i>
GPL32	Murine Genome U74A Array	Affymetrix	<i>Mus musculus</i>
GPL339	Mouse Expression 430A Array	Affymetrix	<i>Rattus</i>

<sup>65</sup> DNA マイクロアレイ (多数の DNA 断片をプラスチック等の基板上に配置した分析器具) の基板上に固定された遺伝子断片。プローブの塩基配列が相補関係にある mRNA と水素結合を形成して 2 本鎖を形成 (ハイブリダイゼーション) する性質を利用し、mRNA の発現の有無や増減を把握できる。

<sup>66</sup> RNA の塩基配列に対して付与される、公式の名称や遺伝子の機能の説明、データベース上の ID 等の「注釈」。

<sup>67</sup> 生体内での遺伝子やタンパク質、化合物の分子間相互作用を「経路」として表現した「パスウェイ」を、その機能等に応じて分類したもの。

ID	名称	プロバイダー	Species
			<i>norvegicus</i>
GPL340	Mouse Expression 430B Array	Affymetrix	<i>Rattus norvegicus</i>
GPL342	Rat Expression 230B Array	Affymetrix	<i>Rattus norvegicus</i>
GPL1355	Rat Genome 230 2.0 Array	Affymetrix	<i>Rattus norvegicus</i>
GPL85	Rat Genome U34 Array	Affymetrix	<i>Rattus norvegicus</i>
GPL7300	018972 D. melanogaster Oligo Microarray 4x44K G2519F	Agilent	<i>Drosophila melanogaster</i>
GPL6480	014850 Whole Human Genome Microarray 4x44K G4112F	Agilent	<i>Homo sapiens</i>
GPL14746	028282 Whole Rat Genome Microarray 4x44K v3	Agilent	<i>Homo sapiens</i>
GPL7202	014868 Whole Mouse Genome Microarray 4x44K G4122F	Agilent	<i>Mus musculus</i>
GPL11202	026655 Whole Mouse Genome Microarray 4x44K v2	Agilent	<i>Mus musculus</i>
GPL21163	074809 SurePrint G3 Mouse GE v2 8x60K Microarray	Agilent	<i>Mus musculus</i>
GPL7294	014879 Whole Rat Genome Microarray 4x44K G4131F	Agilent	<i>Rattus norvegicus</i>
GPL14664	026437 D. rerio (Zebrafish) Oligo Microarray V3	Agilent	<i>Rattus norvegicus</i>
EPA_BSP_Whole Transcriptome_170424	EPA_BSP_WholeTranscriptome_170424	BioSpyder	<i>Homo sapiens</i>
S1500_Human	S1500_Human	BioSpyder	<i>Homo sapiens</i>
S1500_Rat	S1500_Rat	BioSpyder	<i>Rattus norvegicus</i>
danRer10_Ensembl	danRer10_Ensembl	Ensembl	<i>Danio rerio</i>
dm6_Ensembl	dm6_Ensembl	Ensembl	<i>Drosophila melanogaster</i>
hg19_Ensembl	hg19_Ensembl	Ensembl	<i>Homo sapiens</i>
mm10_Ensembl	mm10_Ensembl	Ensembl	<i>Mus musculus</i>
rn6_Ensembl	rn6_Ensembl	Ensembl	<i>Rattus norvegicus</i>
GPL10558	HumanHT-12 V4.0 expression beadchip	Illumina	<i>Homo sapiens</i>
GPL6885	MouseRef-8 v2.0 expression beadchip	Illumina	<i>Mus musculus</i>
GPL6887	MouseWG-6 v2.0 expression beadchip	Illumina	<i>Mus musculus</i>
GPL6101	ratRef-12 v1.0 expression beadchip	Illumina	<i>Rattus norvegicus</i>
danRer10_RefSeq	danRer10_RefSeq	RefSeq	<i>Danio rerio</i>

ID	名称	プロバイダー	Species
dm6_RefSeq	dm6_RefSeq	RefSeq	<i>Drosophila melanogaster</i>
hg19_RefSeq	hg19_RefSeq	RefSeq	<i>Homo sapiens</i>
mm10_RefSeq	mm10_RefSeq	RefSeq	<i>Mus musculus</i>
rn6_RefSeq	rn6_RefSeq	RefSeq	<i>Rattus norvegicus</i>

## (2)-2 フィルタリングと可視化

ANOVA、ウィリアムズの両側傾向検定 (two-sided William's trend test)、ORIOGEN (Order Restricted Inference for Ordered Gene Expression) や変動倍率フィルタリング<sup>68</sup>によって統計学的に有意ではない発現変化を示すプローブを除去する。各手法の概要とデフォルト設定値を表 3.6-2 にまとめた。

なお、BMDEExpress には、リアルタイムで表示データのフィルタリングを行うためのユーザーインターフェースが用意されている。パラメータに様々な制限を加えた場合に、結果を可視化情報とデータの両方からリアルタイムで確認できる。

表 3.6-2 統計学的フィルタリング手法の概要とデフォルト設定値

手法	概要	デフォルト設定値
ANOVA 検定	「異なる用量での反応が全て同じである」という帰無仮説の検定。対立仮説は「応答がすべて同じではなく、応答の変化の方向に制限がない」。	P-Value Cutoff : 0.05 (0.1、0.01 もデフォルトの選択肢を含む) Multiple Testing Correction : False
ウィリアムズの両側傾向検定	「用量反応がない」という帰無仮説と「単調に増加又は減少する」という対立仮説の比較。等張回帰 (データに対する単調応答にフィットするノンパラメトリック回帰) を使用し、用量特異的応答の推定値及び入手可能な試験統計量を得る。	Filter Out Control Genes : True
ORIOGEN	ノンパラメトリックを利用して重要な遺伝子を同時に識別し、さまざまな不等式パターンに従ってそれらをグループ化する。BMDEExpress 2 では、遺伝子の全体的な有意性の p 値は、別の用量反応プロファイルの和集合に対して用量反応なしの帰無仮説を検定することによって計算する。	P-Value Cutoff : 0.05 (0.1、0.01 もデフォルトの選択肢を含む) Number of Initial Bootstrap Samples : 500 Number of Maximum Bootstrap Samples : 1000 Shrinkage Adjustment Percentile : 5.0

<sup>68</sup> 遺伝子発現の膨大な変動データには、わずかな変動や統計学的に有意でない用量依存性変動も含まれる。そのため一般的に、トランスクリプトミクスデータを扱う際は予め設定した変動倍率 (例えば 2 倍) に満たない変動データを除去する。

### (2)-3 BMDの算出

数理モデルのフィッティングの確認は米国 EPA の BMDS に実装されている手法<sup>69</sup>で行われる。

トランスクリプトミクスデータで得られる連続値モデルの場合、モデル式の項数は用量反応データの用量群の数によって自動的に決定される。なお、指数モデルを使用する場合には、負の値を生成しないようデータを予め正規化することが望ましい。

BMD の算出は、Linux 版、Mac 版、Windows 版のいずれの BMD Express でも可能であるが、BMDS モデル実行ファイルで採用されている浮動小数点演算<sup>70</sup>の影響により、OS 間で算出結果に若干、差が出る可能性がある。経験的に、プラットフォームから得られる結果の差異は限定的であるが、少数のプロープから得られる結果は大きな差となる可能性がある。なお、モデル実行ファイルの検証は Windows でのみ行われている。

BMDExpress で実装されているモデル式を下記に示す。

多項式モデル (Polynomial model) :

$$\mu(\text{dose}) = \beta_0 + \beta_1 \text{dose} + \beta_2 \text{dose}^2 + \dots + \beta_n \text{dose}^n$$

(ここで n は多項式の次数)

線形モデル (Linear model) :

多項式モデルで n=1 とする場合

累乗モデル (Power model) :

$$\mu(\text{dose}) = \gamma + \beta \text{dose}^\delta$$

(ここで  $0 < \gamma < 1$ ,  $\beta \geq 0$ , and  $18 \geq \delta > 0$ )

Hill モデル (Hill model) :

$$\mu(\text{dose}) = \gamma + (v \text{dose}^n) / (k^n + \text{dose}^n)$$

指数 2 (Exponential 2) :

$$\mu(\text{dose}) = a \exp(\text{sign } b * \text{dose})$$

<sup>69</sup> 3.6.3 (2)-3 項及び米国 EPA 作成の Benchmark Dose Technical Guidance の項を参照。

<sup>70</sup> コンピュータ演算における小数の扱い方の 1 つで、小数点の位置を固定せず、仮数、基数、指数の要素で表現することで有限桁数の 2 進数として扱う。例えば、「1.23」は「 $123 \times 10^{-2}$ 」と表すことができる。このとき、仮数は「123」、基数は「10」、指数は「-2」である。浮動小数点式を用いると、桁数の制約により誤差 (丸め誤差) が生じ、数値の精度が落ちることがある。さらに、演算時の処理オプションが OS によって異なるため、浮動小数点演算の結果は OS 間で異なることがある。

指数 3 (Exponential 3) :

$$\mu(\text{dose}) = a \exp(\text{sign}(b \cdot \text{dose})^d)$$

指数 4 (Exponential 4) :

$$\mu(\text{dose}) = a \cdot (c - (c - 1) \exp(-1 \cdot b \cdot \text{dose}))$$

指数 5 (Exponential 5) :

$$\mu(\text{dose}) = a \cdot (c - (c - 1) \exp(-1 \cdot (b \cdot \text{dose})^d))$$

## (2)-4 機能分類

プローブセットは、遺伝子に紐付けられた NCBI Entrez Gene ID に基づいて統合される。複数のプローブセットが単一遺伝子に関連する場合、BMD は平均されて Entrez ID に対応する単一の値が得られる。BMDEpress では、Entrez ID と各種機能分類の照合に基づいた遺伝子のカテゴリー分類が行われ、デフォルトでは下記の機能分類が採用されている。

- ・ GO (生物学的プロセス / 分子機能 / 細胞の構成要素 / 全て)
- ・ シグナル経路 (Reactome)
- ・ ユーザー定義カテゴリー

## (2)-5 組み込まれているデータベース

BMDEpress には、表 3.6-3 に示すデータベース及びソフトウェア情報が組み込まれている。

表 3.6-3 BMDEpress に組み込まれているデータベース及びソフトウェア

名称	内容	URL
Reactome Pathway Database	シグナル経路	<a href="https://www.reactome.org/">https://www.reactome.org/</a>
Gene Ontology	遺伝子オントロジー	<a href="http://www.geneontology.org/">http://www.geneontology.org/</a>
DONLP2 optimizer	非線形適合	<a href="https://www2.mathematik.tu-darmstadt.de/fbereiche/numerik/staff/spellucci/DONLP2/">https://www2.mathematik.tu-darmstadt.de/fbereiche/numerik/staff/spellucci/DONLP2/</a>

## (3) 今後の展開

今後、以下の機能の更新・リリースが計画されている。

- ・ モデル平均化

- ・ノンパラメトリック用量反応モデリング
- ・自動レポート生成
- ・IVIVE (in vitro から in vivo への外挿) 機能
- ・データ閲覧ビューの更新
- ・より安定した接続マッピングを可能にする wAUC (加重曲線下面積) メトリクス
- ・GUI のパフォーマンス改善

その他の統計学的手法、バイオインフォマティクス的手法についても現在開発が進められており、今後ワークフローに導入される可能性がある。

### 3.6.3 利用方法に関する情報

#### (1) セットアップ方法

BMDEExpress は Sciome 社のホームページからダウンロードできる<sup>71</sup>。

#### (2) 操作の流れ

BMDEExpress による解析の流れを示す。図は下記 WEB サイトから引用した。操作の詳細も WEB サイトを参照のこと。解析画面例を図 3.6-1 に示す。

<https://github.com/auerbachs/BMDEExpress-2/wiki>



図 3.6-1 解析画面例

<sup>71</sup> <https://www.sciome.com/bmdexpress/> (2019 年 1 月現在)

## (2)-1 データ入力

遺伝子発現データファイル (タブ区切り.txt ファイル) をインポートし、アノテーションを付与する。なお、BMDEExpress では遺伝子発現データ以外の連続値データの用量反応モデリングも実行可能である。データファイルの読み込み画面を図 3.6-2 に示す。

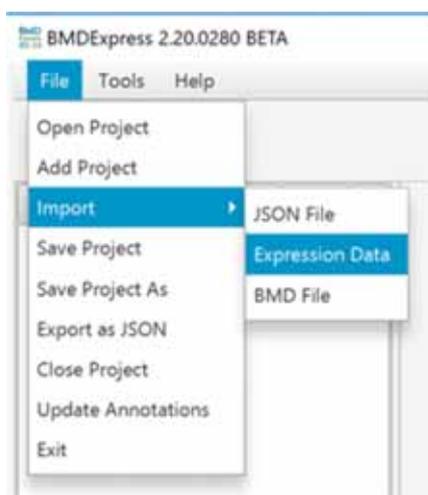


図 3.6-2 データファイルの読み込み画面

データファイルの読み込み後、アレイプラットフォームを指定する。データ取得に使用したプラットフォームがドロップダウンリストに含まれない場合は "generic" を選択する。さらに、データの前処理に使用したログ変換の種類を選択する。データファイルのプラットフォーム及びログ変換の設定画面を図 3.6-3 に示す。



図 3.6-3 データファイルのプラットフォーム及びログ変換の設定画面

## (2)-2 フィルタリング

変動倍率フィルタリングと ANOVA や傾向検定を組み合わせ、分析対象とするプローブの数を減らす (任意)。データフィルタリングの設定画面を図 3.6-4 に示す。

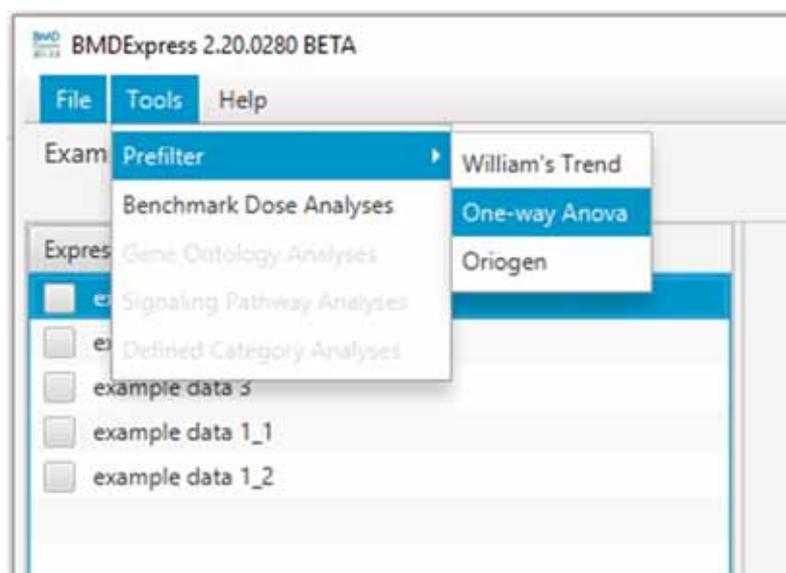


図 3.6-4 データフィルタリングの設定画面

一元配置分散分析 (One-Way ANOVA) 及びウィリアムズの両側傾向検定では以下のオプションを設定する。一元配置分散分析及びウィリアムズの両側傾向検定のオプション設定画面を図 3.6-5 に示す。

**P-Value Cutoff** :  $p$  値に基づくフィルター。デフォルトでは 0.05 に設定。

**Multiple Testing Correction** : 選択した  $p$  値に偽陽性率 (False Discovery Rate ; FDR) 補正を適用する。

**Filter Out Control Genes** : プラットフォームに特有の内部標準遺伝子を除去する。

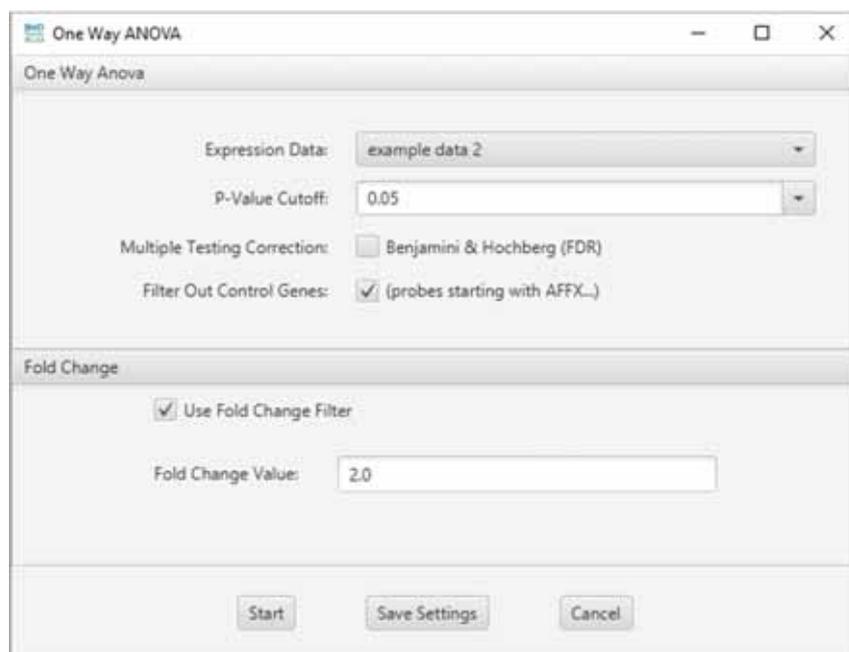


図 3.6-5 一元配置分散分析及ウィリアムズの両側傾向検定のオプション設定画面

ORIOGEN では以下のオプションを設定する。ORIOGEN のオプション設定画面を図 3.6-6 に示す。

**P-Value Cutoff :**  $p$  値に基づくフィルター。デフォルトでは 0.05 に設定。フィルタリング結果画面を図 3.6-7 に示す。

**Number of Initial Bootstrap Samples :**

計算効率の最大化のため、ORIOGEN では設定の範囲内でサンプル数を徐々に増やしながらブートストラップで  $p$  値を算出する。このオプションでは、全プローブの  $p$  値の計算に用いるブートストラップの開始時のサンプル数を設定する。

**Number of Maximum Bootstrap Samples :**

全プローブの  $p$  値の計算に用いるブートストラップの最大サンプル数。

**Shrinkage Adjustment Percentile :**

偽陽性の識別制御に用いる。デフォルト設定の 5.0 は、データセット内の全プローブの 5 パーセンタイル標準偏差を変動閾値に反映する。値を小さくすると、フィルターを通過するプローブの数が減少する可能性がある。

**Multiple Testing Correction :** 選択した  $p$  値に偽陽性率 (FDR) 補正を適用する。

**Filter Out Control Genes :** プラットフォームに特有の内部標準遺伝子を除去する。

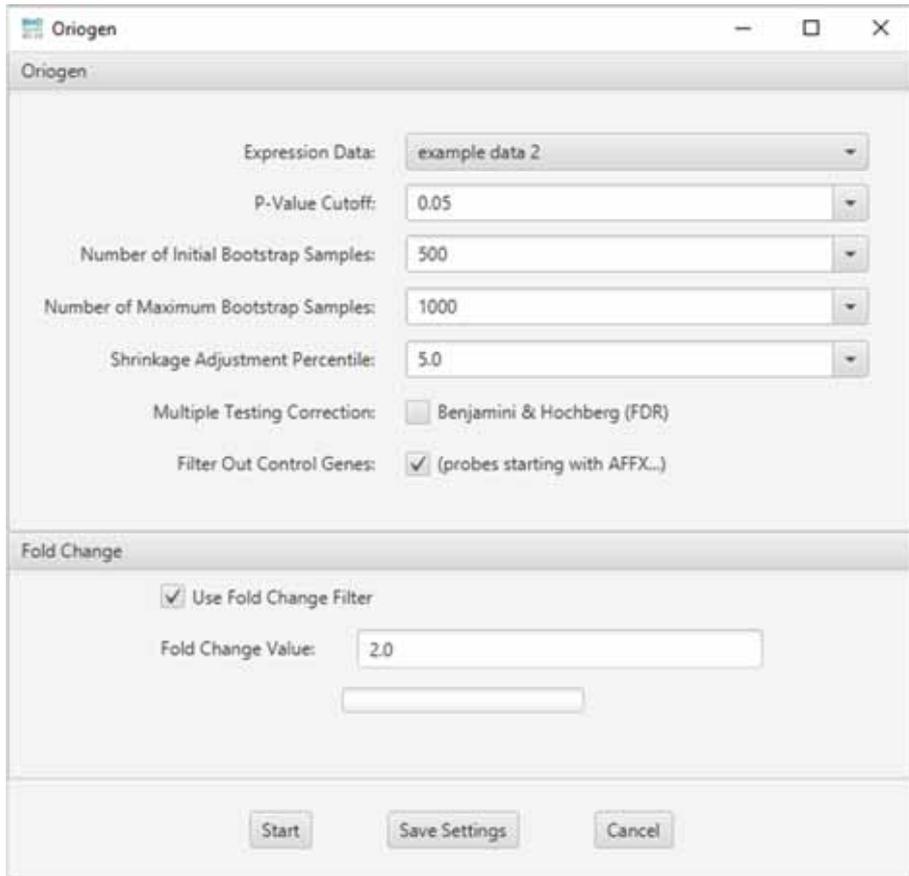


図 3.6-6 ORIOPEN のオプション設定画面



図 3.6-7 フィルタリング結果画面

### (2)-3 曲線のフィッティング(**Tools** > **Benchmark Dose Analysis**)

数理モデルサブセットに基づき、曲線のフィッティングを計算する。各プローブに対して最もよくフィッティングする数理モデルの BMD とその信頼区間の上限値 (BMDU) 及び下限値 (BMDL) を算出する。BMD オプション設定画面を図 3.6-8 に示す。BMD データオプションは以下の通り設定する。

#### Continuous Models :

曲線フィッティングの数理モデルを単一又は組合せとして選択する。一部の数理モデルはデフォルトで選択される。また、用量反応データを構成する用量数によって利用可能な数理モデルは異なる。BMDExpress で実装されているモデルは 3.6.2 (2)-3 に記載。

#### Parameters :

モデルフィッティングは米国 EPA の BMDS に実装されている方法 (アルゴリズム) で実行される。

Maximum Iterations : 数理モデルの収束基準。

Confidence Level : BMD の信頼区間算出に用いる信頼水準。

Restrict Power : 累乗モデルにのみ適用。1 以上の値を指定する。

Constant Variance : 選択すると数理モデルに定数分散が適用される。

BMR Factor : BMD の定義に用いる標準偏差の値。Benchmark response (BMR) や critical effect size と同義。デフォルト値は EPA の推奨値でもある 1 (1 標準偏差に相当)。

#### Model Selection :

##### Best Poly Model Test :

- Nested Chi Square : ネストしたモデルに対して尤度比検定を行い線形モデルと多項式モデルのいずれかを選択した後、選択されたモデル、Hill モデル、累乗モデルの各モデル分布族間で AIC の比較を行い、AIC が最も低い数理モデルを選択する。

- Lowest AIC : AIC に基づく選択プロセスが実行される。

P-Value Cutoff : 最良の線形/多項式モデルの選択における、ネストしたモデルに対して行う尤度比検定法 (カイ二乗検定) で用いる統計学的閾値 (有意水準)。

##### Flag Hill Model with 'k' Parameter < :

データにフィットする数理モデルの 1 つとして Hill モデルが選択されており、その 'k' パラメータが最低陽性用量又はその 1/2 や 1/3 よりも小さい場合に、Hill モデルにフラグを立てる。Hill モデルでは非現実的な BMD 及び BMDL 値を算出することがあるため、このオプションが含まれる。

#### Multiple Threads :

Number of threads : 同時に実行するモデルフィット計算の数。

Model Execution Timeout (secs) : 個々のプローブに対する数理モデルの実行時間。デフォルトの設定は 600 秒。

BMD 結果画面を図 3.6-9 に示す。

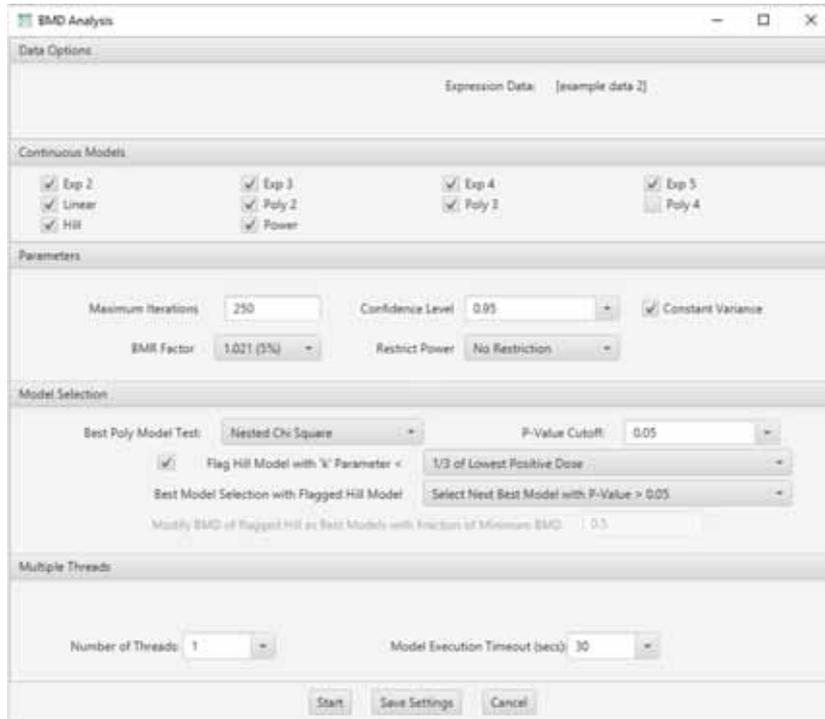


図 3.6-8 BMD オプション設定画面



図 3.6-9 BMD 結果画面

Curve Viewer で各プローブ ID について用量反応プロットとモデルフィット曲線を確認できる。プローブや数理モデルの切替えは、それぞれ「ID」「Model」のプルダウン選択で行う。Curve Viewer 画面を図 3.6-10 に示す。

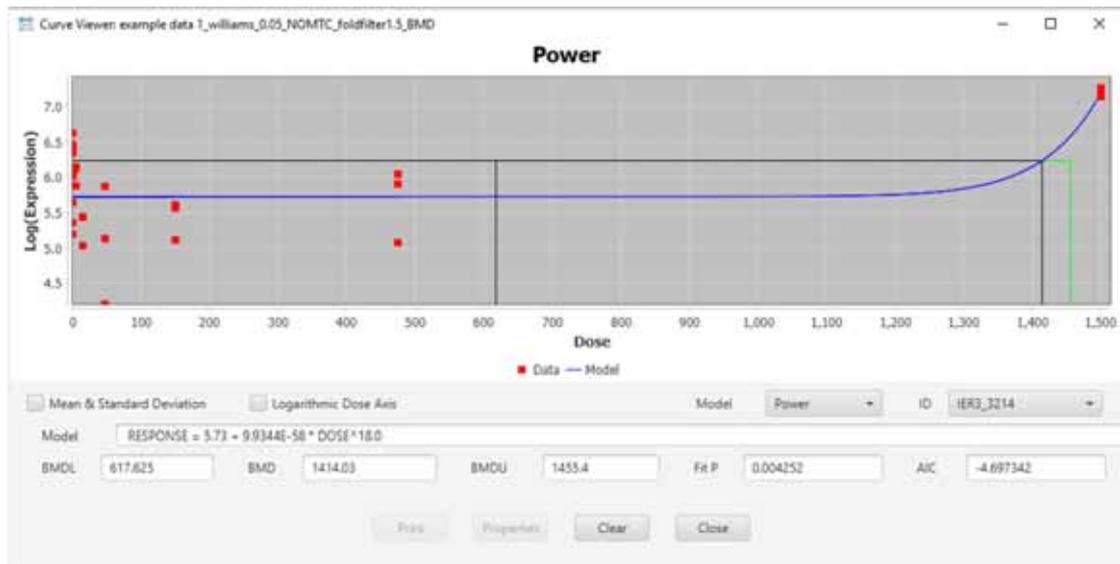


図 3.6-10 Curve Viewer 画面

#### (2)-4 遺伝子セット BMD/BMDL/BMDU の算出

各プローブ/プローブセットを固有の遺伝子にマッピングし、遺伝子を GO やシグナリングパスウェイ等に基づいてカテゴリーに分類する。

データ選択エリアで目的の BMD 結果セットを選択して、次のいずれかを選択する。

BMD 結果セットの選択画面を図 3.6-11 に示す。

**Tools** > **Gene Ontology Analysis**

**Tools** > **Signaling Pathway Analysis**

**Tools** > **Defined Category Analysis**

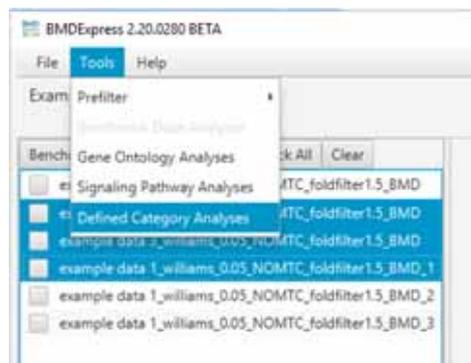


図 3.6-11 BMD 結果セットの選択画面

解析時のオプション (解析対象から除くデータの条件等) を設定し、各カテゴリーの全遺伝子のデータから、カテゴリー遺伝子セットの BMD/BMDL/BMDU の中心傾向とそれに関連する変動性を示す要約値を算出する。解析オプションの設定画面を図 3.6-12 に示す。



図 3.6-12 解析オプションの設定画面 (GO 解析の例)

解析結果は表形式での出力に加え、視覚化された形でも入手できる。図 3.6-13、図 3.6-14、図 3.6-15 に、その例を示す。

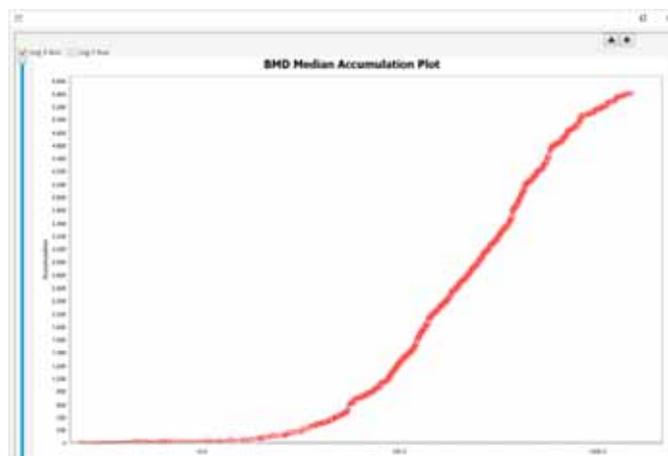


図 3.6-13 BMD 中央値の累積プロット

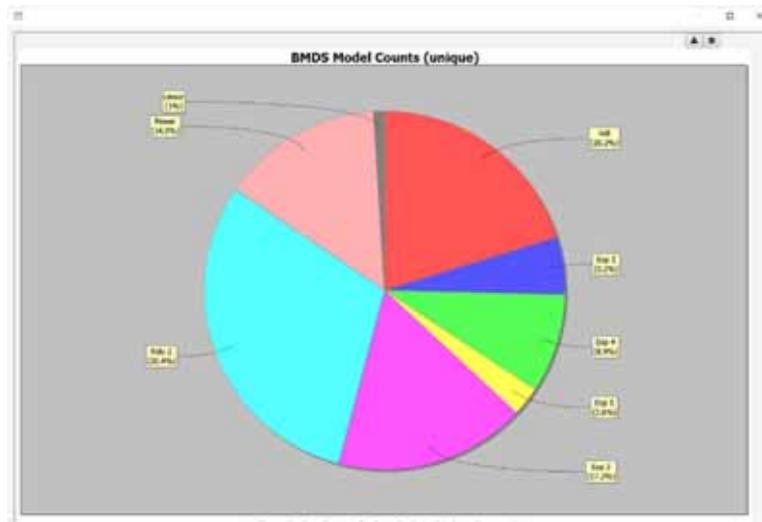


図 3.6-14 最適モデルの円グラフ

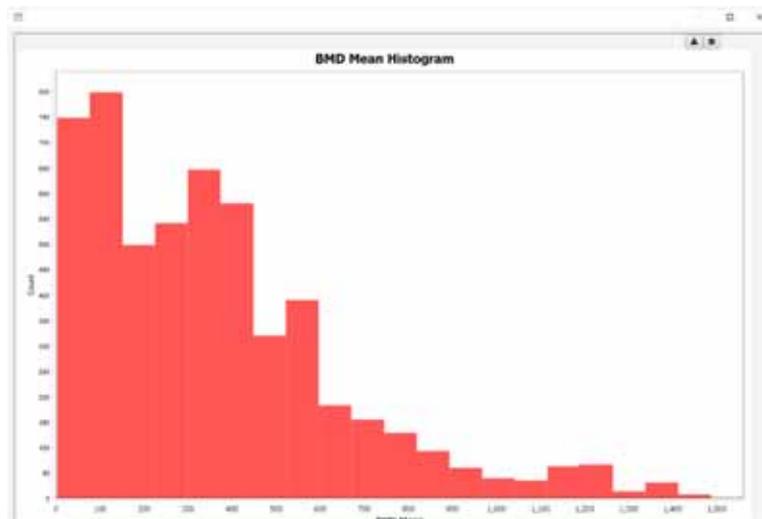


図 3.6-15 BMD 平均値のヒストグラム

3.7 ソフトウェアの主要な機能の比較

		BMDS		PROAST*	EFSA_Benchmark Dose Modelling	MADr-BMD	SPBS	BBMDS**
		Version 3.1	Version 2.7					
解析可能なデータセットの種類		連続値データ (要約、個体)、二値データ、二値データ (ネスト)	連続値データ (要約、個体)、二値データ、二値データ (ネスト)、一般データ (Generic Data)	連続値データ (要約、個体)、二値データ、順序カテゴリーデータ、クラスター化データ	連続値データ (要約、個体)、二値データ、2進データ、順序カテゴリーデータ	二値データ	連続値データ (個体)	連続値データ、二値データ
モデルの種類	連続値	Exponential, Hill, Linear, Polynomial, Power	Exponential, Hill, Linear, Polynomial, Power	Exponential model 3, Exponential model 5, Hill model 3, Hill model 5	Exponential model 3, Exponential model 5, Hill model 3, Hill model 5, Inv.Expon. m3, Inv.Expon. m5, Logn. DR m3, Logn. DR m5		一次回帰モデル べき関数モデル log (X-1)変換モデル	Linear, Power, Michaelis-Menten model, Hill, Exponential model 2~5
	二値	頻度論	Gamma, Logistic, Log-Logistic, Probit Log Probit, Multistage Weibull, Quantal Linear, Dichotomous Hill  【ネストしたモデル】 Nested Logistic, NCTR***  【Multi-Tumor モデル】 MS_Combo	Gamma, Logistic, Log-Logistic, Probit, Log Probit, Mutistage, Multistage-Cancer, Quantal Linear, Weibull Dichotomous Hill  【ネストしたモデル】 Nested Logistic, NCTR Rai and van Ryzin, ToxicoDiffusion, Ten Berge  【Multi-Tumor モデル】 MS_Combo	Logistic, Probit, Log-logistic, Log-Probit, Weibull, Gammma, Multistage, Latent variable models (LVMs)	Logistic, Probit, Log-logistic, Log-Probit, Weibull, Gammma, Multistage, Latent variable models (LVMs)	Logistic, Probit, Log-logistic, Log-Probit, Weibull, Gammma, Multistage, Quantal-linear, Quantal-quadratic	
		ベイズ法	Gamma, Logistic, Log-Logistic, Probit, Log Probit, Multistage, Weibull, Quantal Linear, Dichotomous Hill					
パラメータ制限の可否		ユーザーによる設定が可能	ユーザーによる設定が可能	ユーザーによる設定が可能	デフォルトで設定されている	デフォルトで設定されている	無し	ユーザーによる設定が可能
パラメータ推定法		最尤推定/ベイズ法	最尤推定	最尤推定	最尤推定	最尤推定	最尤推定	マルコフ連鎖モンテカルロサンプリング (MCMC) を適用したベイズ法

	BMDS		PROAST*	EFSA_Benchmark Dose Modelling	MADr-BMD	SPBS	BBMDS**
	Version 3.1	Version 2.7					
連続値データが表示確率分布の初期設定	正規分布/対数正規分布の選択が可能	Exponential 以外のモデルは正規分布のみ (Exponential モデルのみ正規分布/対数正規分布の選択が可能)	対数正規分布 (対数変換の省略か、又は平方根変換を適用することが可能)	対数正規分布		正規分布 (対数正規分布の場合は影響指数を対数(底 10)変換して解析)	対数正規分布
連続値データの BMR 設定方法	中心傾向に基づく方法/ハイブリッド法を選択	中心傾向に基づく方法	中心傾向に基づく方法	中心傾向に基づく方法		ハイブリッド法	中心傾向に基づく方法/ハイブリッド法を選択
使用するモデルの選択	ユーザーによる設定が可能	ユーザーによる設定が可能	ユーザーによる設定が可能	ユーザーによる設定が可能	ユーザーによる設定が可能	ユーザーによる設定が可能	ユーザーによる設定が可能
BMR 値	任意の値を入力 (EPA の基準: 0.1)	任意の値を入力 (EPA の基準: 0.1)	任意の値を入力 (デフォルトは連続値データ 5%、二値データで 10%)	任意の値を入力 (デフォルトは連続値データ 5%、二値データで 10%)	任意の値を入力 (デフォルト: 0.1 (10%))	0.02、0.05、0.10 が自動で算出 (通常 0.05 を選択)	任意の値を入力 (連続値データ、二値データ共にデフォルトは 10%)
モデリング結果の評価	尤度比検定の p 値 > 0.1 AIC が小さいモデルを推奨 BMD/BMDL 比 < 20 (推奨モデルの自動選択機能有り)	尤度比検定の p 値 > 0.1 AIC BMD/BMDL 比	AIC に基づく (AIC 間の 2 単位差を臨界値とする) 尤度比検定も選択可能	AIC に基づく (最小値を採用) モデル平均化: (AIC ≤ AICMin + 2 のモデルのみ平均化に用いる)	AIC、BIC、KIC から選択 (デフォルト: AIC) (その他 BICB、AICB、KICB もあるが有効ではない)	・非ばく露群における影響指標の cutoff 値と臨床上の基準値との近似性の確認 ・目視による確認 ・尤度比 (-2logL) 検定、P 値が有意なもの	事後予測 p 値 (Posterior predictive p-value: PPP) : 0.05 から 0.95 の範囲内 BIC モデルの重み近似法
モデル平均化機能	有り (二値データのみ) 各モデルの BMD 推定値を各モデルの事後確率密度で重み付けして加重平均する	無し	有り (二値データのみ) ブートストラップリサンプリング数: 任意に変更 (デフォルト: 200)	有り (二値データ/連続値データ) ブートストラップリサンプリング数: 任意に変更 (デフォルト: 200)	有り (二値データのみ) ブートストラップリサンプリング数: 任意に変更 (デフォルト: 5000)	無し	有り (二値データ/連続値データ)
信頼区間の設定方法	基本的な設定方法	プロファイル尤度法	プロファイル尤度法	プロファイル尤度法	プロファイル尤度法	-	MCMC を適用したベイズ法により算出した事後分布から両端 (100-a) / 2 を切り取って a% 信用区間**** とする
	モデル平均化	設定された信頼水準 100(1-α)パーセンタイル値を BMDL として算出		ブートストラップ法	ブートストラップ法	ブートストラップ法	事後分布 = (事前分布 × 尤度) / データの分布

\*Menu バージョンについて記載

\*\*ソフトウェアの調査対象ではないが、参考として記載

\*\*\*現時点では収載されていない

AIC: Akaike information criterion, BIC: Bayesian information criterion, KIC: Kashyap's information criterion

\*\*\*\*信頼区間は頻度主義の考えに基づき、信用区間はベイズ主義の考えに基づく

## 4. BMD 法に関連する論文等の収集

---

BMD 法に関する論文のうち以下の項目に該当する情報を収集した。検索は、データベース内でキーワードを用いることにより実施した。検索対象とするデータベースを以下に示す。

### ・検索したデータベース

PubMed、TOXLINE、BIOSIS Previews、Web of Science Core Collection 等

### ・収集対象とした文献

- ① BMD に関するガイドンスに引用されている論文(168 報)
- ② BMD 法で用いるモデル平均化手法を取り扱った論文 BBMD 法に関する論文、分位変動閾値法に関する論文 (47 報)
- ③ 2016 年以降に公表された BMD 法を取り扱った論文(上記①及び②に該当するものを除く)(100 報)

## 5. 資料の和訳

1.~3.の調査で収集した資料について、重要な資料を事務局との協議の上、和訳した。和訳対象とした文書を表5に示す。

表5 和訳対象の文書

資料名	発行年	作成機関 作成者
<b>BMD 法に関するガイダンス</b>		
Update: use of the benchmark dose approach in risk assessment	2017	EFSA
The use of advanced risk assessment methods in answering various types of risk management questions	2006	RIVM
<b>ソフトウェアの操作方法等に関する資料</b>		
PROAST MANUAL Menu version versions 65.6	2018	RIVM
Joint project on Benchmark Dose modelling with RIVM	2018	Wout Slob
Benchmark dose software (BMDS) user manual version 2.6	2016	EPA
Benchmark dose software (BMDS) ver 3.0 user guide	2018	
Manual for BMD Modeling	2017	EFSA
User Manual and Technical Guidance for the Bayesian Benchmark Dose (BBMD) Analysis System Version 1.1.0.0	2016	Shao K and Shapiro AJ