

食品健康影響評価技術研究  
二値反応の用量反応データを対象としたベンチマークドーズ計算ソフトウェア  
の開発研究（課題番号 1907）  
研究成果報告書 別添 2

# BMDMA

## マニュアル

### —実践編—

# 1. メインウィンドウやメニューバー、選択チェック項目等の説明

## (1) メインウィンドウの構成

R を起動し、

`bmdma::bmdma_gui()`

と入力すると、以下のメインウィンドウが表示される (BMDMA のインストール手順についてはインストールマニュアルを参照されたい)。



図 1. BMDMA 起動時の画面 (メインウィンドウ)

BMDMA のメインウィンドウの構成は図 2 のようになる (「設定」タブ選択時)。

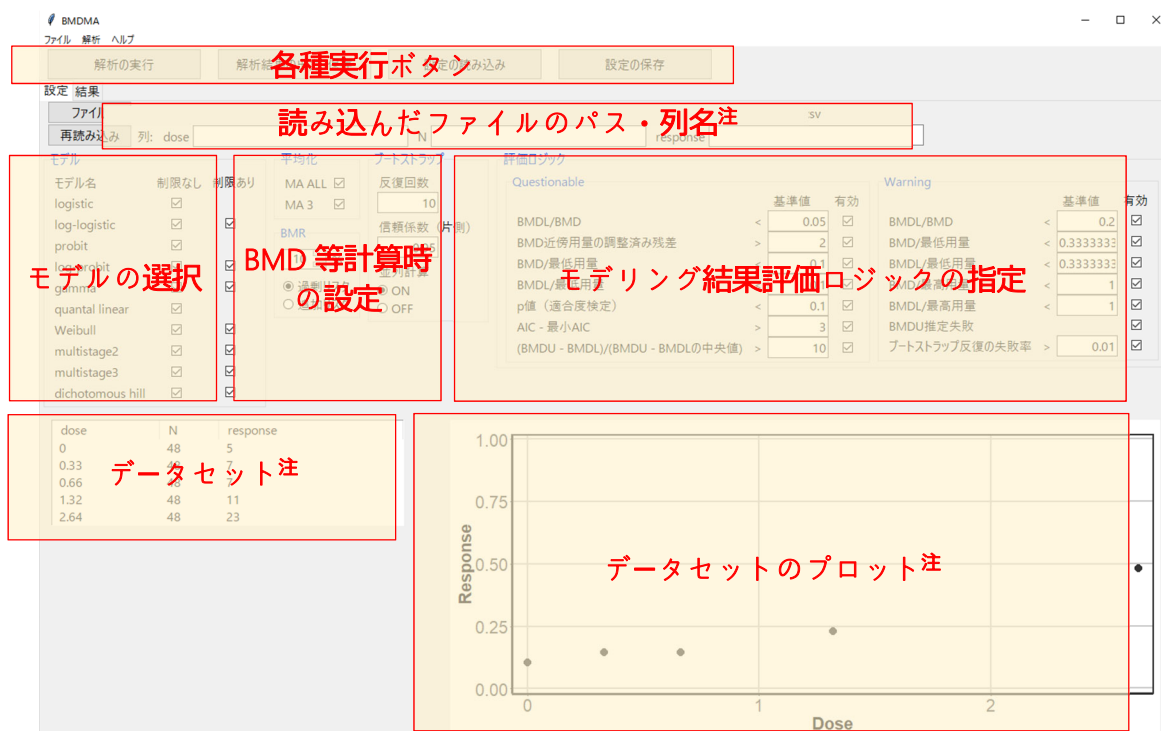


図 2. メインウィンドウの構成概略

(注：読み込んだファイルのパス・列名、データセット及びデータセットのプロットは用量反応データ読込後に表示される)

メインウィンドウの一番上のメニューバーには、「ファイル」、「解析」、「ヘルプ」の3つが存在する。メニューバーの下に、「解析の実行」や「解析結果の出力・保存」といった各種実行ボタンが並んでいる。その下に、表示内容を切り替えるための「設定」と「結果」の2つのタブがある。設定タブの上部には、読み込んだ用量反応データのファイルのパスとファイルに入力された「dose」や「response」といった列名が表示される。設定タブ中央には、解析とその評価に用いる条件等を設定可能にするチェックボックス等が並ぶ。左から、「モデル」では適合を試みるモデルの選択、「平均化」、「BMR」「ブートストラップ」ではBMD等計算時の設定、「評価ロジック」ではモデリング結果の評価に用いる各種評価ロジックの詳細な設定が可能である。設定タブ下部には、解析を行う用量反応データの読み込み後に当該データセットとそのプロットが表示される。以下で、これらの詳細について説明する。

## (2) メニューバー

ウィンドウ上部のメニューバーに「ファイル」、「解析」、「ヘルプ」の3つが存在する。それぞれのメニューには以下の項目がある。

### ファイル

- ・データのファイルを開く  
データファイル（CSV形式、Excel形式）の読み込みを行う。
- ・設定を読み込む  
「設定を保存する」で保存される設定ファイル（YAML形式）を読み込み、設定タブに反映する。
- ・設定を保存する  
設定タブで指定されている設定をファイル（YAML形式）に保存する。
- ・終了  
プログラムを終了する。

### 解析

- ・解析を実行する  
設定タブで指定されている設定で解析を実行する。
- ・解析結果を出力・保存する  
解析結果のレポート、推定値、グラフなどをフォルダに出力・保存する。
- ・解析結果のレポートを出力する  
解析結果のレポートを指定の形式で出力・保存する。

### ヘルプ

- ・このソフトウェアについて  
ソフトウェアのバージョンを表示する。

### (3) 実行ボタン

ボタンをクリックすると、解析の実行、解析結果の出力・保存、設定の読み込み、設定の保存がそれぞれ実行できるようになっている。これらの項目はすべて「ファイル」→「解析」のメニューバーでも操作可能である。設定タブのファイルボタンと再読み込みボタンについてはファイルの読み込みの仕様とともに 2 (1) と 2 (2) で説明する。

### (4) タブ

タブは、標準で最前面にある「設定」と、「結果」の 2 つがある。

「設定」タブで解析の設定を行う。「モデル」では 10 のモデル（パラメータの制限あり、なしの選択可能）から使用するモデルを選ぶ。「平均化」では、モデル平均化の設定として MAALL と MA3 の使用の有無を選択できる。「BMR」では BMR 値とリスクタイプを設定する。「ブートストラップ」では反復回数、信頼係数（片側）、並列計算について設定する。「評価ロジック」では、推奨モデルをきめるための、不適切なモデルを判定するための基準を設定する。

「結果」タブでは、解析の実行結果が表示されるようになっている（解析実行前は「解析結果がありません」と表示される）。解析実行後に「結果」タブを開くと、図 3 のような表が表示される。

解析の実行		解析結果の出力・保存		設定の読み込み		設定の保存		
設定		結果						
モデル名	自由度	リスクタイプ	BMR	収束	AIC	パラメータ1	パラメータ2	パラメータ3
logistic	2	extra	0.1	0	23.2497916022624	-2.171632798817	0.779295407324763	NA
log-logistic	3	extra	0.1	0	25.2834620974845	-2.51462022252893	2.19741368331274	0.11981454
log-logistic (restricted)	3	extra	0.1	0	25.2834620974845	-2.51462022252893	2.19741368331274	0.11981454
probit	2	extra	0.1	0	23.3475420757071	-1.28477439481377	0.455686908168623	NA
log-probit	3	extra	0.1	0	25.3661397762082	-1.51386564656165	1.30520477666192	0.12480099
log-probit (restricted)	3	extra	0.1	0	25.3661397762082	-1.51386564656165	1.30520477666192	0.12480099
gamma	3	extra	0.1	0	27.2212918685193	1.00533379928541	0.1583856071425	0.10697600
gamma (restricted)	3	extra	0.1	0	27.2212918685193	1.00533379928541	0.1583856071425	0.10697600
quantal linear	2	extra	0.1	0	24.8529016904829	0.174420151094051	0.0870298517213886	NA
Weibull	3	extra	0.1	0	25.2502827934984	1.89341687524621	0.0833707065465956	0.11729255
Weibull (restricted)	3	extra	0.1	0	25.2502827934984	1.89341687524621	0.0833707065465956	0.11729255
multistage2	3	extra	0.1	0	25.214000117261	0.0270111601573497	0.0651256972060126	0.11306044
multistage2 (restricted)	3	extra	0.1	0	25.214000117261	0.0270111601573497	0.0651256972060126	0.11306044
multistage3	4	extra	0.1	0	27.130895243976	0.0943588304029172	-0.023225107325321	0.02460649
multistage3 (restricted)	4	extra	0.1	0	27.1365058377312	0.0756913307564892	0	0.01832486
dichotomous hill	4	extra	0.1	0	27.2834620975962	-2.51460296622449	2.19739790745094	1
dichotomous hill (restricted)	4	extra	0.1	0	27.2834620975962	-2.51460296622449	2.19739790745094	1
MA ALL	NA	extra	0.1	NA	NA	NA	NA	NA
MA 3	NA	extra	0.1	NA	NA	NA	NA	NA

図 3. データの解析結果の例

### (5) 解析とその評価に用いる条件設定

この項目では、解析に関する設定の詳細について説明をする。

#### 1. モデル

以下の 10 のモデルが利用できる。Logistic、Probit、Quantal linear 以外のモデルについてはオプションのパラメータの範囲の制限をつけたモデルも選択できる。二項分布によって尤度を定義し、準ニュートン法の一つである L-BFGS-B を用いて最尤法

で各モデルのパラメータは推定される。 $d$ は用量、 $x_i$  ( $i = 1, \dots, 4$ )はパラメータである。 $\epsilon$ は最尤法のために用いる L-BFGS-B のパラメータ制限で使用する小さな値で標準では $10^{-10}$ である。

### Logistic

$$p(d) = \frac{1}{e^{-x_1 - x_2 d} + 1}, \quad x_2 \in [0 + \epsilon, \infty)$$

### Log-logistic

$$p(d) = x_3 + \frac{1 - x_3}{e^{-x_1 - x_2 \log(d)} + 1}, \quad x_2 \in [0 + \epsilon, \infty), \quad x_3 \in [0, 1 - \epsilon]$$

オプションのパラメータの範囲の制限  $x_2 \in [1, \infty)$

### Probit

$$p(d) = \int_{-\infty}^{x_1 + x_2 d} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) dt, \quad x_2 \in [0 + \epsilon, \infty)$$

### Log-probit

$$p(d) = x_3 + (1 - x_3) \int_{-\infty}^{x_1 + x_2 \log(d)} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) dt, \quad x_2 \in [0 + \epsilon, \infty), \quad x_3 \in [0, 1 - \epsilon]$$

オプションのパラメータの範囲の制限  $x_2 \in [1, \infty)$

### Gamma

$$p(d) = x_3 + (1 - x_3) \frac{1}{\Gamma(x_1)} \int_0^{x_2 d} t^{x_1 - 1} e^{-t} dt, \quad x_1 \in [0.2 + \epsilon, \infty), \quad x_2 \in [0 + \epsilon, \infty), \quad x_3 \in [0, 1 - \epsilon]$$

オプションのパラメータの範囲の制限  $x_1 \in [1, \infty)$

### Quantal linear

$$p(d) = x_2 + (1 - x_2) (1 - e^{-x_1 d}), \quad x_1 \in [0 + \epsilon, \infty), \quad x_2 \in [0, 1 - \epsilon]$$

### Weibull

$$p(d) = x_3 + (1 - x_3) (1 - e^{-x_2 d^{x_1}}), \quad x_1, x_2 \in [0 + \epsilon, \infty), \quad x_3 \in [0, 1 - \epsilon]$$

オプションのパラメータの範囲の制限  $x_1 \in [1, \infty)$

### Multistage2

$$p(d) = x_3 + (1 - x_3)(1 - e^{-x_1 d - x_2 d^2}), \quad x_3 \in [0, 1 - \epsilon]$$

オプションのパラメータの範囲の制限  $x_1, x_2 \in [0 + \epsilon, \infty)$

### Multistage3

$$p(d) = x_4 + (1 - x_4)(1 - e^{-x_1 d - x_2 d^2 - x_3 d^3}), \quad x_4 \in [0, 1 - \epsilon]$$

オプションのパラメータの範囲の制限  $x_1, x_2, x_3 \in [0 + \epsilon, \infty)$

### Dichotomous hill

$$p(d) = x_3 \left( x_4 + \frac{(1-x_4)}{e^{-x_1 - x_2 \log d + 1}} \right), \quad x_1, x_2 \in [0 + \epsilon, \infty), \quad x_3 \in [0 + \epsilon, 1], \quad x_4 \in [0 + \epsilon, 1 - \epsilon]$$

オプションのパラメータの範囲の制限  $x_2 \in [1, \infty)$

## 2. モデル平均化について

モデルの選択の設定のチェックボックスで選択されたオプションのパラメータの範囲の制限のないモデルだけがモデル平均化の対象となる。GUI では、対象のモデルすべてを使用した平均化である「MA ALL」と AIC (Akaike's Information Criterion) が小さい上位 3 つのモデルの平均化である「MA 3」を利用できる。また、BMDMA の R の関数を用いれば任意の数のモデル平均化を使うことができる。モデル平均化は

$$p_{MA}(d) = \sum_{i \in S} w_i p_i(d)$$

によって計算される。S はパラメータの範囲の制限がない全モデルのインデックスの集合、 $p_i$  はインデックス  $i$  のモデルである。 $w_i$  はモデル平均化における重みであり、MA ALL では

$$w_i = \frac{e^{-I_i/2}}{\sum_{i \in S} e^{-I_i/2}} \quad (1)$$

で計算される。ここで  $I_i$  はモデル  $i$  の AIC である。MA3 では、AIC が小さい上位 3 つのモデルのインデックスの集合を  $S'$  とし、

$$w_i = \begin{cases} \frac{e^{-I_i/2}}{\sum_{i \in S'} e^{-I_i/2}} & \text{if } i \in S' \\ 0 & \text{otherwise} \end{cases}$$

で計算される。MA3 では、1 回毎のブートストラップで AIC が低い順に 3 つのモデ

ルが選ばれてモデル平均化が行われることになる。

※MA3 では MAALL にあるようなモデル平均化における重みは表示されない。

### 3. BMR

BMR の値とリスクタイプ（過剰リスクまたは追加リスク）を設定できる。

### 4. ブートストラップ

BMDL や BMDU は BMD の信頼限界であるので、BMDMA ではブートストラップ法により求められる。ブートストラップ反復の回数、信頼係数（片側）、並列計算の使用の有無が設定できる。

### 5. 評価ロジック

データに適合したモデルは、図 4 のテスト項目にしたがって Unusable、Questionable、Viable (Warning あり)、Viable (Warning なし) のいずれかに分類される。BMDMA では MA3、MAALL の順番で最初に Unusable でないモデルを推奨モデルとする。

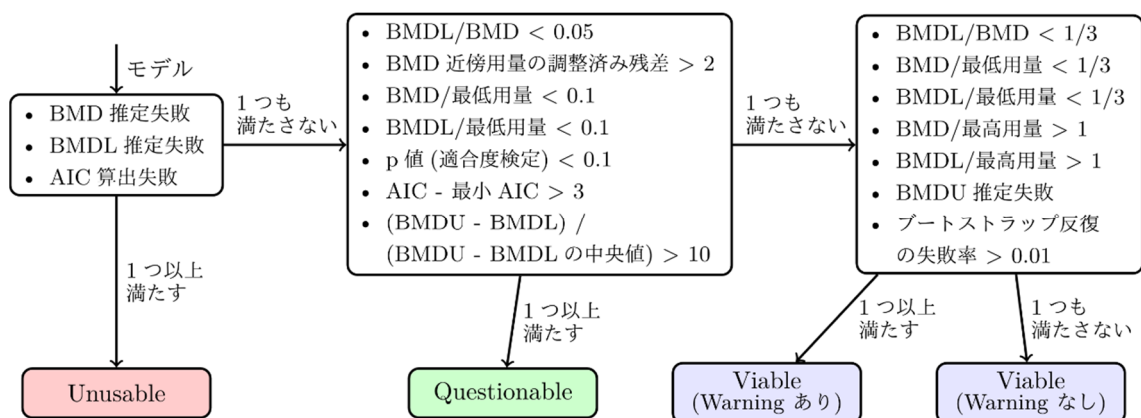
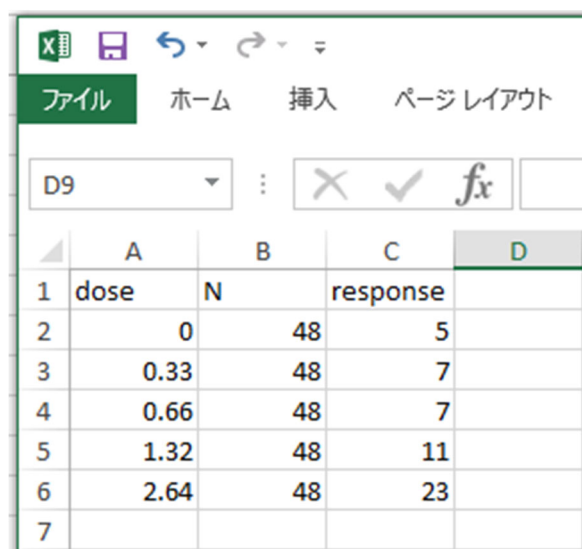


図 4. 評価ロジック

## 2. 用量反応データの準備と読み込み

### (1) 用量反応データの準備

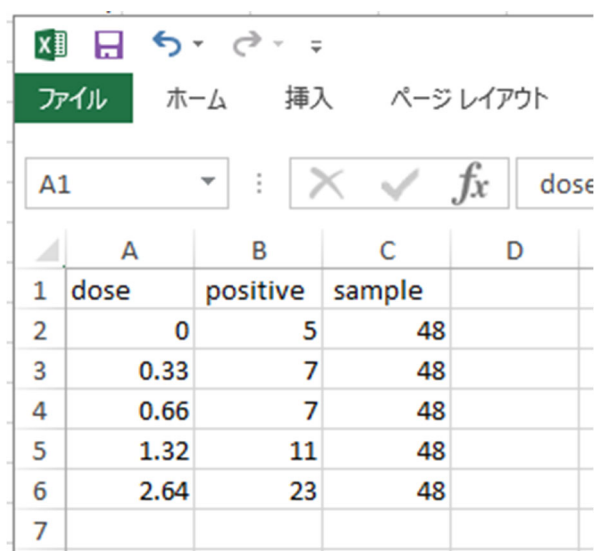
評価したい用量反応（毒性）データを CSV ファイル（拡張子 csv）もしくは Excel ファイル（拡張子 xlsx）で用意する。データのフォーマットに関して、第 1 行は列名であり、dose、N、response が必要である。それぞれ、用量、サンプルサイズ、陽性数である。



	A	B	C	D
1	dose	N	response	
2	0	48	5	
3	0.33	48	7	
4	0.66	48	7	
5	1.32	48	11	
6	2.64	48	23	
7				

図 5. Excel (CSV) ファイル例

標準の列名（dose、N、response）と異なる用量、サンプルサイズ、陽性数を指示する列名の場合、対応する列名を指定することが可能である（列名はアルファベットを推奨）。たとえば、以下のような標準とは異なる列名、列順でデータを整理したファイルであっても、次の（2）の手順で読み込むことが可能である。



	A	B	C	D
1	dose	positive	sample	
2	0	5	48	
3	0.33	7	48	
4	0.66	7	48	
5	1.32	11	48	
6	2.64	23	48	
7				

図 6. 列名及びその順番が標準とは異なる CSV ファイル例

## (2) データの読み込み

メインウィンドウのファイルボタン (図 7) を押して、評価したいデータファイルを読みこむ。左上のファイルタブから、「データファイルを開く」を選択しても同じである。



図 7. データを読み込むためのファイルボタン

標準の列名 (dose、N、response) のファイルを読み込むと以下ようになる。

dose	N	response
0	48	5
0.33	48	7
0.66	48	7
1.32	48	11
2.64	48	23

図 8. 推奨の順番で構成されたデータをアップロードした際の画面

「ファイル」ボタンの横に、読み込んだファイルのパスが表示される。ウィンドウの左下にはデータのテーブル表示 (赤色の四角枠) がある。dose (用量)、N (サンプル数)、response (陽性数) が正しく読み込まれているかを確認する。また、ウィンドウ右下にはデータのグラフが表示される。標準とは異なる列名 (例えば dose、positive、sample) のファイルを読み込む場合は以下のようにする。同じように、左下の取り込み画面をよく見ていただきたい。



図 9. 非推奨の順番で構成されたデータをアップロードした際の画面

用量を意味する **dose** は正しく認識されているが、**N** と **response** はそれぞれ逆に読み込まれているのがわかるだろう。その場合は、赤色の楕円で示した各項目の入力フォームに列名を指定する必要がある。**dose** は **dose** 列を、**N** は **positive** 列、**response** は **sample** 列が読み込まれているのがわかる。したがって、**N** は **sample** 列、**response** は **positive** 列として BMDMA に正しく認識させるためには以下の画面のように入力すればよい。



図 10. 非推奨のデータフォーマットに正しく列を認識させた際の画面

読み込む列を正しく書き直したら、左隣にある再読み込みボタンを押す。書き直した割り当てで再読み込みされたデータセットが左下に現れ、**dose**、**N**、**response** のそれぞれの列がきちんと認識されていることを確認する。

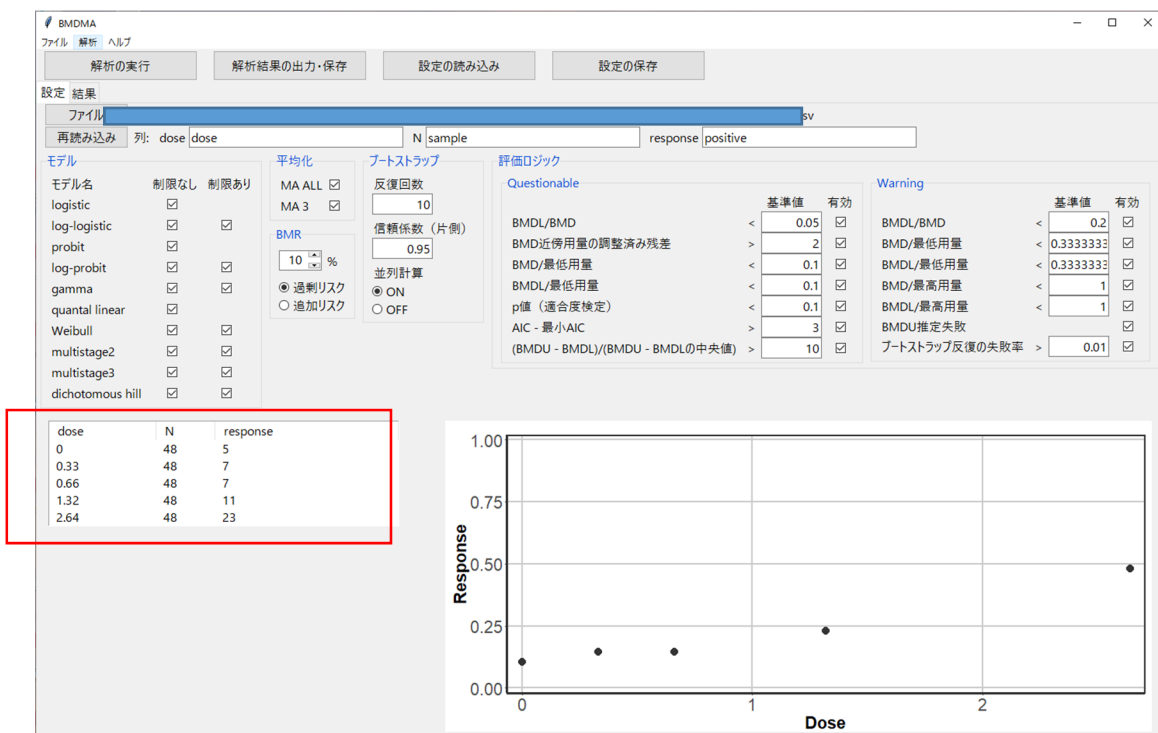


図 11. 非推奨のデータフォーマットを正しく再読み込みした際の画面

きちんとデータが読み込まれているかは、左下のテーブルとともに、右下のグラフでも確認することができる。横軸が **Dose**、縦軸が **Response** の割合を示しており、正しい値が描かれているかを確認する。

### 3. 解析の実行と保存方法

#### (1) 計算の実施と所要時間

ファイルを読み込み、データの dose、N、response がきちんと認識されていることを確認したら、「解析の実行」ボタンを押すとモデルの適合や BMDLなどを求めるための計算が始まる。計算が終わると、図 12 のように各モデルにフィットさせた推定結果が描かれる。

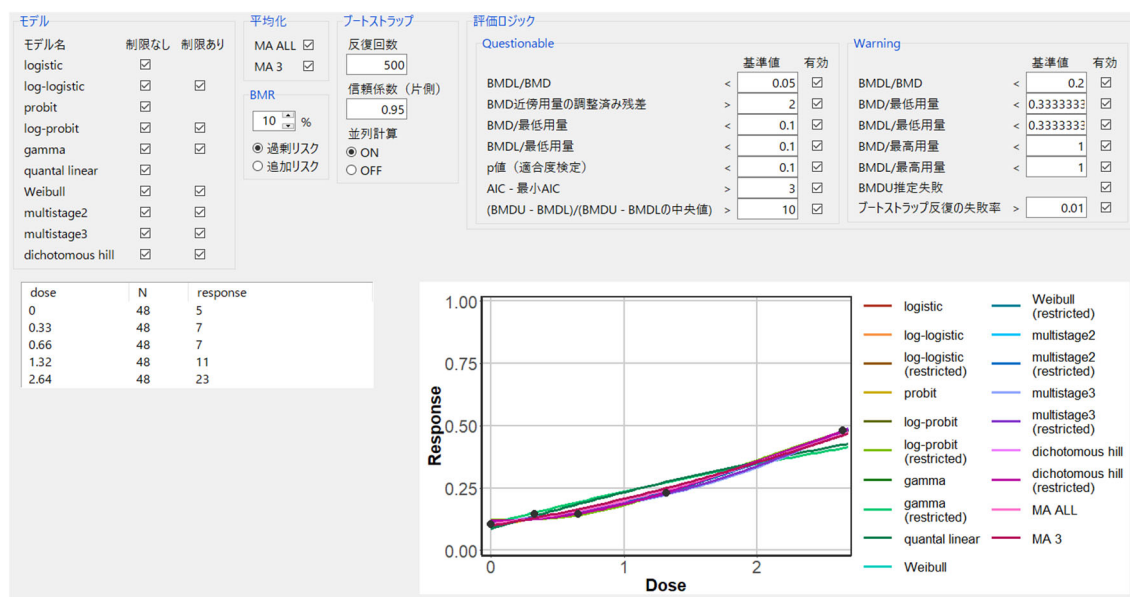


図 12. 解析が正しく実行された際の画面

図 12 ではデフォルトの設定で解析を行っているが、モデルやブートストラップの反復回数は必要に応じて変更していただきたい。

なお、以下は PC のスペックと解析の所要時間である。参考にされたい。

タイプ： デスクトップパソコン

メモリ： 64GB

プロセッサ： Intel(R) Xeon(R) Silver 4280 CPU @ 2.10GHz

コア数： 16

全モデル選択、反復回数 500 回 (N=48) の解析を実施 → 約 4 分

全モデル選択、反復回数 1000 回 (N=48) の解析を実施 → 約 8 分

※ブートストラップ法での反復回数について、反復すべき回数は事前に決定していない (慣例では 1000 回実施されることが多い)。しかし、反復回数が少ないと、観察データによっては BMDL などの出力値が安定しないことがあり、推定精度が劣る (例. 同じデータなのに実行毎に小数点第 2 位の値が異なる) という問題が生じ得る。それに備えるため、作成者は、十分な計算機能を備えた計算機を使用すべきである。また、以下のことを推奨する：

A) 同じデータで必ず 3~4 回の実行を行い、高い精度で BMDL が一致することを確認する

B) もしも実行毎に結果が一致しない場合はブートストラップの回数を 10 倍して、同様に実行結

果の比較を繰り返し、一致を確認することを推奨する。例えば、反復回数 1000 回で一致しない場合は 10000 回、のように反復を増やして再比較することを推奨する。

## (2) 結果の保存 - フォルダの作成

解析結果は、「解析結果の出力・保存」ボタンから、任意の空のフォルダに出力・保存することができる。この時には、あらかじめ空のフォルダを作成しておくことで円滑に進む。Windows 10 ではデスクトップやエクスプローラで指定したフォルダで右クリックから、新規作成→フォルダを選択することで、新しいフォルダを作成できる。

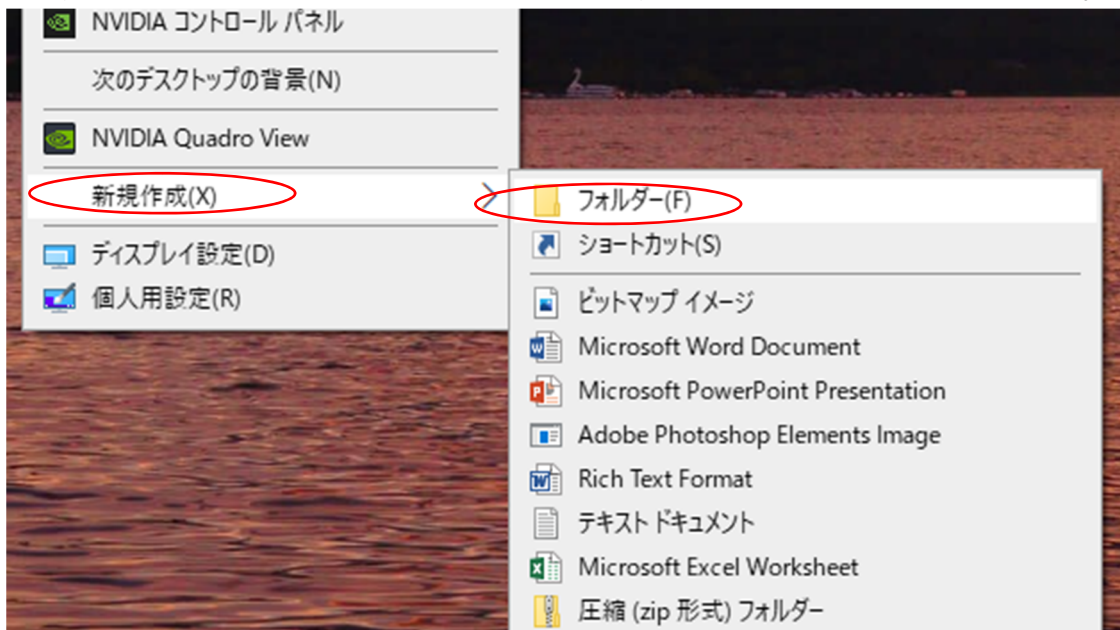


図 13. 空フォルダを右クリックから作成する際の画面

解析終了後、「解析結果の出力・保存」ボタンを押すことで、フォルダの選択ウィンドウが表示されるので、あらかじめ作成しておいた空のフォルダを選択する（あらかじめ空のフォルダを作成していなくても、フォルダ選択ウィンドウで右クリックから空のフォルダを作成することも可能である）。出力先のフォルダには、解析に用いたデータの CSV、結果のテーブルの CSV が保存され、各モデルのフィッティングのグラフは **graph** という名前のフォルダに PNG ファイルとして保存される。また、結果のテーブルや各モデルのグラフを含むレポートが **report.docx** という名前の Word ファイルとして保存される。

A screenshot of a Windows File Explorer window showing the contents of a folder named 'BMDMA結果'. The folder contains several files and subfolders. The table below represents the data shown in the screenshot.

名前	更新日時	種類	サイズ
graph	2021/03/04 13:47	ファイルフォルダ	
data	2021/03/04 13:47	Microsoft Excel C...	1 KB
report	2021/03/04 13:47	Microsoft Word D...	406 KB
result	2021/03/04 13:47	Microsoft Excel C...	7 KB
settings.yaml	2021/03/04 13:47	YAML ファイル	2 KB

図 14. 出力された結果を保存したフォルダを開いた際の画面

### (3) 結果の見方

「結果」タブには、推定結果や評価ロジックに使用された評価値を含む表が表示される。

- モデル名
- 自由度  
パラメータの自由度
- リスクタイプ  
設定した BMR のリスクタイプ。added または extra
- BMR  
設定した BMR の値
- 収束  
推定で用いた L-BFGS-B が収束していれば 0 (0 以外は L-BFGS-B が収束していないことを意味する)
- AIC  
適合したモデルの赤池情報量規準
- パラメータ 1、パラメータ 2、パラメータ 3、パラメータ 4  
モデルのパラメータ  $x_i$  ( $i = 1, \dots, 4$ ) の推定値
- BMD、BMDL、BMDU  
得られた BMD、BMDL、BMDU の値
- ブートストラップ反復の成功回数  
BMD を求めることができたブートストラップ反復の回数
- ブートストラップ反復  
設定したブートストラップ反復の回数
- 信頼係数 (片側)  
BMDL と BMDU を求めるために設定した信頼係数 (片側)
- BMDL/BMD  
BMDL を BMD で割った値
- BMDU/BMDL  
BMDU を BMDL で割った値
- BMDU 推定成功  
BMDU が推定できた場合 0、推定できなければ 1
- BMD/最低用量  
BMD を最低用量で割った値
- BMDL/最低用量  
BMDL を最低用量で割った値
- P 値 (適合度検定)  
適合度検定の P 値

- BMD 近傍用量の調整済み残差  
BMD に最も近い用量での調整済み残差
- ブートストラップ反復の失敗率  
BMD が求められなかったブートストラップ反復の割合
- モデル平均化 (MA ALL) における重み  
MA ALL の (1) 式の  $w_i$
- AIC – 最小 AIC  
各モデルの推定から得られた AIC と、それらの AIC の最小値の差
- (BMDU-BMDL)/(BMDU-BMDL の中央値)  
各モデルの(BMDU-BMDL)を(BMDU-BMDL)の中央値で割った値
- モデリング評価結果  
評価ロジックにより viable、warning、questionable のいずれかに分類される。
- BMDMA 推奨モデル  
推奨モデルには「Recommended」と表示される。
- 注記  
評価ロジックによる注記が表示される。