



| 별책부록 |

# 흡입독성시험 결과 활용도 제고를 위한 Benchmark dose 도입 연구

OSHRI

산업재해예방

안전보건공단

산업안전보건연구원





연구보고서 부록. BMDS 사용자 매뉴얼

# 흡입독성시험 결과 활용도 제고를 위한 Benchmark dose 도입 연구

이 미 주

산업재해예방

안전보건공단

산업안전보건연구원







## 목 차

1. BMDS의 설치 .....	3
2. BMDS3의 기본 메뉴 .....	6
3. 분석 실행 및 분석 세부 조정 .....	17
4. BMDS의 모델링 .....	37
5. 각 모델 유형의 공통 출력물 .....	44
6. 연속형 데이터의 종말점 .....	49
7. 이분형 자료의 종말점 .....	80
8. 다중 종양 분석 .....	101
9. 고려해야 할 점 .....	111
10. BMDS 권장 및 결정 로직 .....	115



# BMDS 사용자 매뉴얼







## 〈부록〉 BMDS 사용 매뉴얼

\* 본 매뉴얼은 EPA에서 발행한 Benchmark dose software(BMDs) version 3.2 user guide(EPA/600/R-20/216, August 2020)를 한국어로 번역한 것입니다. 원본 매뉴얼에 있는 모든 내용을 다 포함하는 것은 아니며 등지형(nested) 자료 등 본 연구원에서 수행하는 흡입시험과 관련성이 없는 자료 분석방법 등은 제외하였습니다.

### 1. BMDS의 설치

#### 1) 필요 사양 및 프로그램

BMDS를 사용하려면 매크로가 활성화된 Windows용 Microsoft Excel 2010, 2013 또는 2016(32비트 또는 64비트) 이상의 데스크톱 버전이 필요하다. BMDS는 64비트 버전의 Microsoft Windows와 호환된다. BMDS는 macOS 시스템에서는 실행되지 않는다.

#### 2) 1단계: BMDS 다운로드

BMDS의 최신 버전은 항상 BMDS 다운로드 페이지에서 사용할 수 있다. BMDS가 포함된 .zip 파일을 다운로드한다.

#### 3) 2단계: 폴더에서 BMDS 압축풀기

BMDS는 자체 폴더에 압축을 풀어야 하며 zip 파일 내에서 BMDS를 실행하려고 하면 작동하지 않는다. BMDS는 사용자에게 읽기/쓰기 권한이 있는 모든 폴더에 압축을 풀 수 있으며 관리자 권한은 필요하지 않다. 다음 지침은 EPA 표준 데스크탑인 Windows 10을 기준으로 작성되었다.

- a. 1단계에서 다운로드한 압축 파일을 찾는다.
- b. 압축된 파일을 마우스 오른쪽 버튼으로 클릭하고 상황에 맞는 메뉴에서 모두 추출을 선택한다. "압축(압축) 폴더 추출" 대화 상자가 표시된다.
- c. 아래 그림 2와 같이 "압축된(Zipped) 폴더 압축 풀기" 대화 상자에서 BMDS 폴더의 압축을 풀어야 하는 폴더를 입력하거나 찾는다. (BMDS의 각 버전은 고유한 폴더에 있어야 한다.) 압축 해제가 완료된 후 BMDS 폴더를 열려면 "완료 시 압축 해제된 파일 표시" 확인을 선택한다.

파일 경로를 짧게 유지한다. 최상의 결과를 얻으려면 추출된 폴더(및 해당 하위 폴더)를 특수 문자나 공백 없이 가장 간단하고 짧은 디렉토리에 배치한다.  
예: C:\ Users\ C:\

#### 4) 3단계: 시작 시 EULA(End-User License Agreement, 최종 사용자 사용권 계약)에 동의하고 엑셀 매크로를 허용 및 디지털서명한다.

처음 시작할 때 BMDS는 EULA 패널이 표시된다. 내용을 확인하고 수락 또는 거절을 선택한다. 작은 "옵션 저장됨" 대화 상자가 나타난다. 확인을 선택한다.

- 필수: 처음 시작할 때 매크로 활성화

BMDS가 작동하려면 처음 시작할 때 Excel 매크로를 활성화해야 한다. Microsoft Office 웹 사이트에서 Excel 매크로 활성화 방법을 확인하고 수행하면 된다.

- 선택 사항: 디지털 서명 매크로

BMDS 웹사이트를 통해 배포되는 BMDS 버전에서는 BMDS Excel 매크로가 디지털 서명되지 않는다. 서명된 버전이 필요한 경우 BMDS 사본의 매크로에 디지털 자체 서명에 대한 Microsoft의 지침을 따라 수행한다.

조직에서 디지털 서명된 버전의 BMDS를 필요로 하는 사용자(BMDS Excel

매크로가 신뢰할 수 있는 개발자의 것임을 보여주기 위해)는 BMDS 전자 티켓 시스템을 통해 요청할 수 있다. 그러기 위해서는 VBA 기반 응용 프로그램에 대한 대부분의 조직의 보안 요구 사항을 충족해야 한다. 그러나 디지털 서명된 버전의 BMDS 3.2는 32비트 및 64비트 버전을 비롯한 다양한 Excel 버전에서 일관되게 작동하지 않을 수 있다.

### 5) 4단계: BMDS 데스크탑 아이콘 만들기(옵션)

바탕 화면 바로 가기 아이콘에서 BMDS를 실행할 수도 있다.

- 바탕 화면에서 이전 BMDS 바로 가기 아이콘을 삭제한다.
- Windows 탐색기에서 새로 설치된 BMDS 응용 프로그램 폴더로 이동한다.
- BMDS3.xlsm 파일(또는 시스템에 파일 확장자가 표시되지 않는 경우 "BMDS3")을 마우스 오른쪽 버튼으로 클릭하면 컨텍스트 메뉴가 나타난다.
- 보내기를 클릭하면 하위 메뉴가 나타나는데 "데스크톱(바로가기 만들기)"를 클릭한다. Windows는 바탕 화면에 파일에 대한 바로 가기를 만든다.

## 2. BMDS3의 기본 메뉴

### 1) 엑셀 기반의 사용자 인터페이스

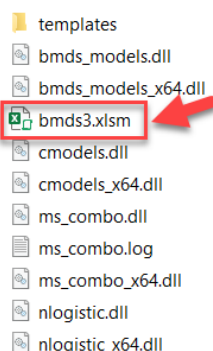
BMDS 2.7과 그 이전 버전은 주로 독립 실행형 Windows 응용 프로그램이었으나 BMDS 3는 분석 및 결과 표시를 위해 고도로 사용자 정의된 Microsoft Excel 통합 문서를 사용한다.

1. 읽기 전용 분석 통합 문서에는 데이터 세트 및 분석 옵션이 있다.
2. Analysis 통합 문서는 모델 분석을 위해 DLL(dynamic link libraries)이라는 사용자 지정 프로그램으로 데이터를 보낸다.
3. DLL의 결과는 별도의 결과 통합 문서에 표시됩니다.
4. 각 결과 워크북에는 분석 결과가 들어 있다. 단일 분석은 여러 데이터 세트로 구성될 수 있다.
5. 사용자는 Analysis 통합 문서에서 특정 결과 통합 문서를 선택하여 데이터 세트와 모델 옵션을 완전히 다시 로드하여 다시 실행하거나 추가 구성을 할 수 있다.

### 2) 워크북 분석(bmds3.xlsm)

BMDS 3를 시작하려면 BMDS 프로그램 폴더에서 bmds3.xlsm 파일을 더블클릭한다. bmds3.xlsm 파일은 BMDS 분석 워크북이다.

**Figure 1.** BMDS3 folder contents, with bmds3.xlsm workbook file.

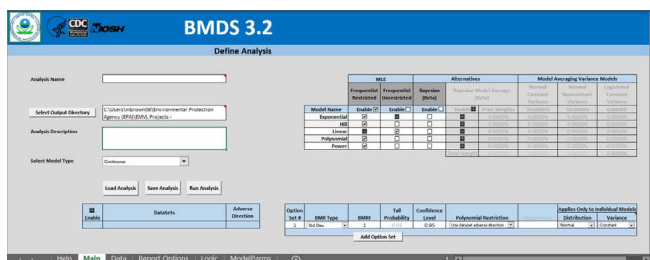


bmds3.xlsm/Analysis Workbook 파일에 대한 기본 정보:

- 여기에는 분석을 렌더링하고 결과 통합 문서를 만드는 등의 작업에 필요한 매크로, 사용자 양식 및 기타 데이터가 포함되어 있다.
- 파일이 쓰기 금지되어 있다. 어떤 사용자도 이 파일에 대한 변경 사항을 저장할 수 없다.
- BMDS 매크로도 보호되며 보거나 편집할 수 없다.

사용자는 BMDS 3 Analysis Workbook에서 분석을 위한 데이터 세트, 모델링 및 보고 옵션을 입력한다. 사용자는 직관적인 양식과 선택 목록에서 모델링 옵션을 지정할 수 있다. 모든 계산은 Analysis Workbook 내에서 수행된다.

**Figure 2.** The Analysis Workbook on first opening, with the Main tab displayed.



Analysis Workbook은 연속(continuous), 이분(dichotomous) 또는 중지 이분(nested dichotomous) 반응이 있는 여러 데이터 세트의 용량-반응

분석을 쉽게 수행하고 추적할 수 있도록 설계되었다. 복합 암(dichotomous) 종말점을 모델링하고 이들의 조합과 관련된 BMD를 도출하는 추가 기능이 있다.

위험 평가의 필요에 따라 사용자는 다음에 의해 연결된 데이터 세트에 대한 BMDS 3 분석에 집중할 수 있다.

- 연구(예: 연구 데이터베이스가 많은 화학 물질의 경우)
- 화학 물질(예: 잘 연구되지 않은 화학 물질의 경우)
- 건강 결과(예: 여러 연구 및/또는 여러 대응 조치에 의해 평가된 건강 결과가 있는 화학 물질의 경우)

### (1) Help 탭

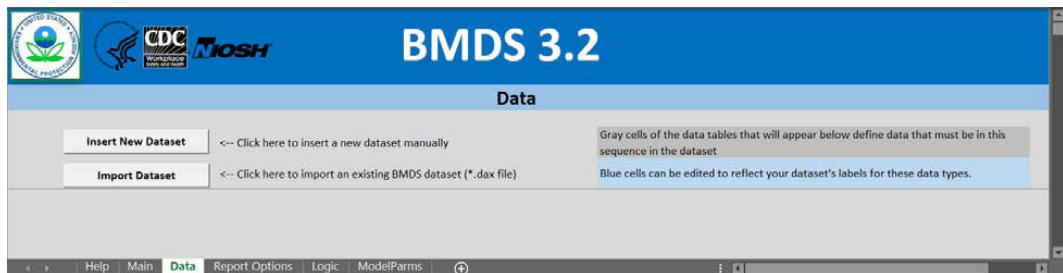
도움말 탭에는 BMDS의 온라인 지원 도구, 문서 다운로드, 용어집 및 기술 지침에 대한 링크가 포함되어 있다.

### (2) Main 탭

모델, 데이터 세트 및 옵션 세트가 분석을 위해 대기하는 BMDS의 기본 작업 공간이다. 이 탭의 컨트롤을 사용하는 방법은 “3. 분석 실행 및 분석 세부조정”에 설명되어 있다.

### (3) Data 탭

Figure 3. Data tab, on first opening.



사용자는 data 탭에서 하나 이상의 데이터세트를 지정할 수 있다. 여기에 지정된 모든 데이터세트는 각 데이터세트에 대해 특정 옵션을 설정할 수 있는 기본 탭에 나열된다. 데이터 세트는 이전 BMDS 버전에서 가져오거나 수동으로 입력할 수 있다. 회색 셀 및 파란색 셀의 유용성을 설명하는 도움말 텍스트를 참고하면 좋다.

data 탭에 대한 자세한 내용은 “3.분석 실행 및 분석 세부 조정-2) 2단계: 데이터세트(dataset) 추가”를 참조한다.

#### (4) Report option 탭

Figure 4. Report Options tab, on first opening.

**BMDP 3.2**

**Output Options**

Save Default Output Options

**Export Options**

Output Directory: C:\Users\mbrown06\Environmental Protection Agency (EPA)\EMVL Projects - Documents\FY2020\

Choose Model/Analysis Type: Continuous

Export to Excel	Variable
<input checked="" type="checkbox"/>	Model Name
<input checked="" type="checkbox"/>	Data File Name
<input checked="" type="checkbox"/>	Description/Notes
<input checked="" type="checkbox"/>	Dose Response
<input checked="" type="checkbox"/>	Variance Model
<input checked="" type="checkbox"/>	Dataset
<input checked="" type="checkbox"/>	BMR Type
<input checked="" type="checkbox"/>	BMRF
<input checked="" type="checkbox"/>	Tail Probability
<input checked="" type="checkbox"/>	Confidence Level
<input checked="" type="checkbox"/>	Distribution Type
<input checked="" type="checkbox"/>	Variance Type
<input checked="" type="checkbox"/>	Total Number of Dose Groups/Observations
<input checked="" type="checkbox"/>	Dependent Variable
<input checked="" type="checkbox"/>	Independent Variable
<input checked="" type="checkbox"/>	Adverse Direction

Export to Excel	Variable
<input checked="" type="checkbox"/>	BMD, BMDL, BMDU
<input checked="" type="checkbox"/>	AIC
<input checked="" type="checkbox"/>	Test 4 P-value
<input checked="" type="checkbox"/>	Degrees Of Freedom (D.O.F.)
<input checked="" type="checkbox"/>	Number of Parameters in Model
<input checked="" type="checkbox"/>	Parameter Estimates (MLE)
<input checked="" type="checkbox"/>	Parameter Standard Errors
<input checked="" type="checkbox"/>	Goodness of Fit
<input checked="" type="checkbox"/>	Likelihoods of Interest
<input checked="" type="checkbox"/>	Tests of Interest
<input checked="" type="checkbox"/>	Cumulative Distribution Function (CDF)

**Word Report Options**

Help Main Data **Report Options** Logic ModelParms

Report option 탭에서 사용자는 BMDS 분석 결과에 포함될 데이터 값을 지정할 수 있다. 사용자는 Word 기반 보고서에 포함할 요소를 지정할 수도 있다.

## (5) Logic 탭

Figure 5. Model Recommendation Decision Logic

**Decision-Logic**

Recommend model in Viable Bin?	TRUE
Recommend model in Questionable Bin?	FALSE
BMDL range deemed "sufficiently close" to use lowest AIC instead of lowest BMDL in viable models:	3

**Reset To Default Logic**

**Model Recommendation/Bin Placement Logic**

Test Description	Test On/Off			Test Threshold (where appropriate)	Bin Placement if Test is Failed	Notes to Show
	Continuous	Dichotomous	Nested			
BMD calculated	On	On	On		Unusable Bin	BMD not estimated
BMDL calculated	On	On	On		Unusable Bin	BMDL not estimated
BMDU calculated	Off	Off	Off		No Bin Change (Warning)	BMDU not estimated
AIC calculated	On	On	On		Unusable Bin	AIC not estimated
Constant Variance	On			0.05	Questionable Bin	Constant variance test failed (Test 2 p-value < 0.05)
Non-Constant Variance	On			0.05	Questionable Bin	Non-constant variance test failed (Test 3 p-value < 0.05)
Goodness of fit p-test	On	On	On	0.1	Questionable Bin	Goodness of fit p-value < 0.1
Goodness of fit p-test (cancer)	On	On	On	0.05	Questionable Bin	Goodness of fit p-value < 0.05
Ratio of BMD/BMDL (serious)	On	On	On	20	Questionable Bin	BMD/BMDL ratio > 20
Ratio of BMD/BMDL (caution)	On	On	On	3	No Bin Change (Warning)	BMD/BMDL ratio > 3
Abs(Residual of interest) too large	On	On	On	2	Questionable Bin	Residual for Dose Group Near BMD > 2
BMD higher than highest dose	On	On	On	1	No Bin Change (Warning)	BMD higher than maximum dose
BMDL higher than highest dose	On	On	On	1	No Bin Change (Warning)	BMDL higher than maximum dose
BMD lower than lowest dose (warning)	On	On	On	3	No Bin Change (Warning)	BMD 3x lower than lowest non-zero dose
BMD lower than lowest dose (serious)	On	On	On	3	No Bin Change (Warning)	BMD 3x lower than lowest non-zero dose
BMD lower than lowest dose (warning)	On	On	On	10	Questionable Bin	BMD 10x lower than lowest non-zero dose
BMD lower than lowest dose (serious)	On	On	On	10	Questionable Bin	BMD 10x lower than lowest non-zero dose
Abs(Residual at control) too large	On	On	On	2	No Bin Change (Warning)	Residual at control > 2
Poor control response std. dev.	On	On	On	1.5	No Bin Change (Warning)	Modeled control response std. dev. > [1.5] actual response std. dev.
D.O.F. equals 0	On	On	On		Questionable Bin	d.f. = 0, saturated model (Goodness of fit test cannot be calculated)

Help Main Data Report Options **Logic** ModelParms

BMDS 3의 결과에는 2012 EPA Benchmark Dose Technical Guidance (U.S. EPA, 2012)에 부합하는 모델 선택에 관한 자동 권장 사항이 포함된다.

이러한 기준은 BMDS 3 Analysis Workbook의 Logic 탭에서 변경할 수 있다. 결정 로직(Decision logic)은 켜거나 끌 수 있으며 다양한 데이터 세트 유형에 대해 특정 기준을 활성화하거나 비활성화할 수 있다. 로직은 분석 중인 데이터 유형(연속, 이분법, 중첩)에 따라 다르다. 자세한 내용은 “10. BMDS 권장 사항 및 결정 로직”를 참조한다.



## (6) ModelParm 탭(Model Parameters)

Figure 6. Model Parameters tab.

**BMDS 3.2**

**Model Parameters**

**Frequentist Unrestricted**

**Log Logistic**

Type	Initial	N/A	Min	Max
g	0	-2	0	-18
a	0	-2	0	-18
b	0	1	0	1.00E-04

**Gamma**

Type	Initial	N/A	Min	Max
g	0	-2	0	-18
a	0	1	0	0.2
b	0	0.1	0	100

**Logistic**

Type	Initial	N/A	Min	Max
a	0	-2	0	-18
b	0	0.1	0	100

**Probit**

Type	Initial	N/A	Min	Max
a	0	-2	0	-18
b	0	0.1	0	0

**Quantal Linear**

Type	Initial	N/A	Min	Max
a	0	-2	0	-18
b	0	0.1	0	0

**Frequentist Restricted**

**Log Logistic**

Distribution	Type	Initial	N/A	Min	Max
Normal	g	0	-2	0	-18
Normal	a	0	-2	0	-18
LogNormal	b	0	1.2	0	1

**Gamma**

Distribution	Type	Initial	N/A	Min	Max
Normal	g	0	-2	0	-18
LogNormal	a	0	1	0	1
LogNormal	b	0	1	0	100

**Logistic**

**Probit**

**Quantal Linear**

**Bayesian**

**Log Logistic**

Distribution	Type	Mean	StdDev	Min	Max
Normal	g	1	0	2	-20
Normal	a	1	0	1	-40
LogNormal	b	2	0.693147	0.5	1.00E-04

**Gamma**

Distribution	Type	Mean	StdDev	Min	Max
Normal	g	1	0	2	-18
LogNormal	a	2	0.693147	0.424264	0.2
LogNormal	b	2	0	1	1.00E-04

**Logistic**

Distribution	Type	Mean	StdDev	Min	Max
Normal	a	1	0	2	-20
LogNormal	b	2	0.1	1	1.00E-12

**Probit**

Distribution	Type	Mean	StdDev	Min	Max
Normal	a	1	0	2	-8
LogNormal	b	2	0.1	1	0.00E+00

**Quantal Linear**

Distribution	Type	Mean	StdDev	Min	Max
Normal	a	1	0	2	-8
LogNormal	b	2	0.1	1	0.00E+00

ModelParms 탭에는 각 모델의 매개변수 및 "사양"과 함께 BMDS가 실행하는 모든 모델이 표시된다. ModelParms 탭에 나열된 값은 프로그램에 대한 실제 입력이다. 모델 실행 파일에서 읽고 사용한다. 값은 암호로 보호되며 사용자가 변경/편집할 수 없다. 자세한 내용은 "4. BMD 모델링-2) 모델 parameter"를 참조한다.

## 3) 워크북 세팅

BMDS의 "분석 저장" 기능을 사용하면 설정 워크북이 생성된다. "3. 분석 실행 및 분석 세부 조정-3) 3a단계: 모델링 옵션의 선택과 저장"을 참조한다.

Settings Workbook에는 사용자가 입력/선택/수락한 모든 모델링 옵션(Main 탭), 데이터 세트 정보(Data 탭), 보고 옵션(Report 탭) 및 모델 선택 로직 옵션(Logic 탭)이 포함되어 있다.

설정 통합 문서에는 분석에 대한 설정만 포함되어 있지만 결과는 포함되어 있지 않다. 분석이 실행되기 전에 분석 옵션이 저장되었기 때문이다. 설정 통합 문서의 명명 규칙은 다음과 같다.

[분석명].xlsx

여기서 분석명은 Main 탭의 "Analysis Name" 필드 내용이다. 설정 통합 문서 파일은 기본 탭의 "Select output directory(출력 디렉토리 선택)" 필드에서 사용자가 식별한 디렉토리 위치에 저장된다. 나중에 "Load Analysis(부하 분석)" 기능을 통해 BMDS로 가져올 수 있다. 자세한 내용은 "3. 분석 실행 및 분석 세부 조정-4) 3b단계: 불러오기(load), 저장(save), 분석실행(run)"을 참조한다.

#### 4) 워크북 결과

Analysis Workbook의 Main 탭에서 Run Analysis 버튼을 선택하면 Result Workbook이 생성된다.

모델링 상태를 보여주는 막대가 나타난다. Excel에서 결과 통합 문서를 만들 때 표시가 업데이트되지 않는다.

개별 데이터 세트 분석을 위해 BMDS는 분석된 각 데이터 세트에 대한 별도의 결과 통합 문서에 모든 모델 결과를 기록한다. 예를 들어 사용자가 5개의 데이터 세트를 분석하는 경우 BMDS는 5개의 개별 결과 통합 문서 파일을 생성한다.

주어진 데이터 세트에 대한 결과 통합 문서 파일 이름은 다음과 같이 구성된다.

[DatasetName]\_analysis.xlsx

- DatasetName = Data 탭에서 데이터 세트의 사용자 지정 이름
- \_analysis = 파일 이름 끝에 BMDS프로그램에 의해 덧붙여짐

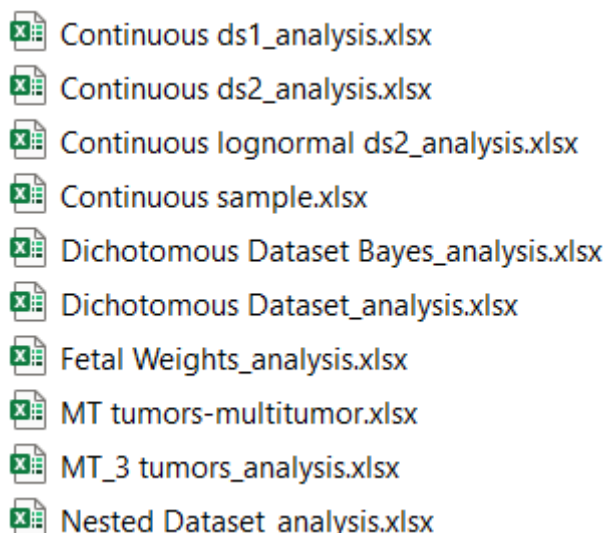
위의 명명 규칙은 연속형(continuous), 이분형(dichotomous) 및 중첩(nested) 모델링 결과 통합 문서에 사용된다.

명명 규칙은 일반적으로 여러 데이터 세트를 포함하는 다종양 분석의 경우 약간 다르다. 다중 종양 분석의 경우 결과 통합 문서 명명 규칙은 다음과 같다.

[Analysis Name]-multitumor.xlsx

- AnalysisName = 기본 탭의 분석 이름 필드 내용
- -multitumor = 파일 이름 끝에 BMDS 프로그램에 의해 덧붙여짐

**Figure 7. Results Workbook files (.xlsx).**

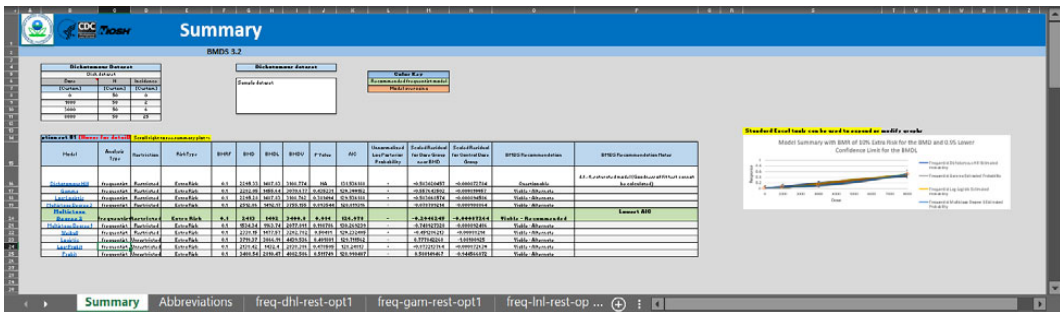


[참고] BMDS 설정 및 결과 관련 파일은 BMDS 3 프로그램 폴더가 아닌 다른 별도의 폴더를 만들어 저장하는 것을 권장한다.

분석에 사용된 모든 옵션은 결과 워크북에 저장되어 나중에 분석을 다시 실행할 수 있다. 자세한 내용은 “3. 분석 실행 및 분석 세부 조정-4) 3b단계: 불러오기(load), 저장(save), 분석실행(run)을 참조한다.”

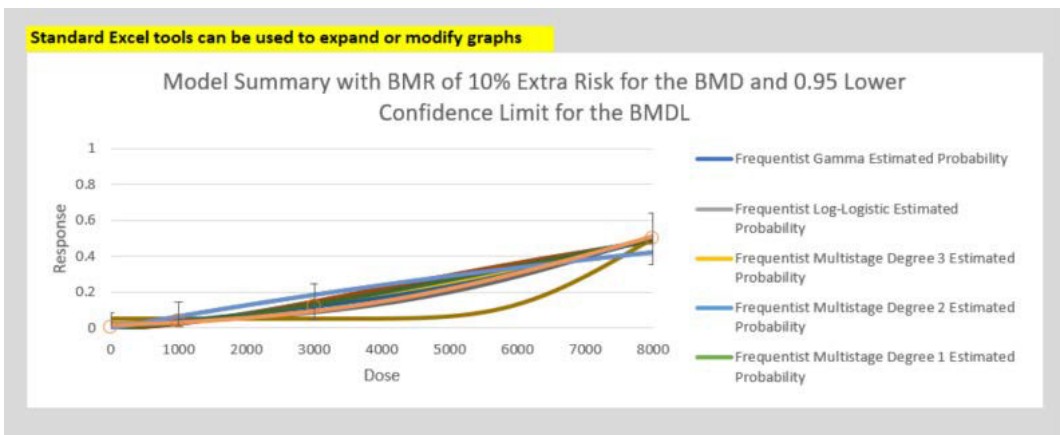
결과 통합 문서에는 데이터 집합의 복사본, 사용자가 분석 통합 문서에 입력한 데이터 집합 설명 및 지정된 각 옵션 집합에 대한 개별 결과 탭이 포함되어 있다.

Figure 8. A Results Workbook's Summary tab.



Summary 탭에서 오른쪽으로 스크롤하여 요약 결과 그래프를 확인한다. Excel의 모든 도구는 이미지와 그래프를 편집할 때 사용할 수 있다.

Figure 9. Summary model results graph, on the Summary tab.



BMD5 결과에 대한 자세한 내용은 “3. 분석 실행 및 분석 세부 조정-5) 4단계: 분석실행”을 참조한다.

## 5) Pre-BMD5 3 모델을 위한 업그레이드

BMD5 연속형 모델은 "Hybrid" 모델링 기능 및 로그정규 반응(lognormal response) 옵션을 포함하도록 업그레이드되었다. 자세한 내용은 표 5를 참조

한다. 또한 기존의 모든 모델을 다시 코딩하여 유지 관리를 용이하게 하고 안정성, 정확성, 신뢰성 및 속도 측면에서 성능을 개선했다.

[참고] BMDs 3은 Akaike 정보 기준(AIC) 계산을 BMDs 2.x와 다소 다르게 처리하여 다양한 가능성(즉, 정규 대 로그 정규)을 가진 모델을 쉽게 비교할 수 있습니다. 자세한 내용은 "9. 고려해야 할 점-1) 연속형 모델의 AIC"를 참조한다.

## 6) Backwards-Compatibility

BMDs 3.2는 BMDs 2.7 및 BMDs Wizard 1.11과의 역호환성이 있다. BMDs에 대한 모든 업그레이드의 목적은 BMD를 포함한 매개변수의 최적화 및 추정을 개선하는 것이다.

[참고] BMDs 3.x 및 2.x는 동일한 데이터에서 실행할 때 로그 가능성 및 AIC에 대해 서로 다른 값을 반환합니다. 자세한 내용은 "7. 이분형 자료의 종말점-3) Simple designs에서 이분형 종말점을 위한 모델의 수학적 상세-(2) AIC와 모델 비교"를 참조한다.

숙련자를 위해 BMDs 3은 다음과 같은 면에서 기존 BMDs 마법사와 유사하게 제작되었다.

- Excel 기반
- 사용자가 단일 탭에서 모델링 옵션을 보고 지정할 수 있음.
- EPA 권장 사항 또는 사용자 정의 로직에 따라 "최상의" 결과를 식별하기 위한 자동 선택 기능이 포함.
- 단일 결과 통합 문서에 모든 입력 및 출력 문서화

- 위험 평가에서 표시하기에 적합한 형식의 Microsoft Word 표로 결과를 표시하기 위한 유연한 인쇄 옵션을 제공.

## 7) BMDS 3에 포함되지 않는 모델

BMDS 3에는 BMDS 2.7 및 BMDS Wizard 1.11에서 사용할 수 있었던 모든 모델과 기능이 포함되어 있으나 다음의 모델은 없다.

- 이분법적 배경선량 모델
- Rai 및 Van Ryzin 중첩 이분법 모델
- ToxicoDiffusion 모델
- 기능은 동일하지만 기능과 옵션이 추가된 EPA의 범주형 회귀 소프트웨어 CatReg로 대체된 10개의 Berge 모델
- NCTR(National Center for Toxicological Research) 중첩 이분법 모델(향후 BMDS에 포함될 예정)

이러한 모델은 BMDS 2.7에 포함되어 있으며, BMDS의 아카이브 버전으로 BMDS 웹사이트에서 다운받을 수 있다.

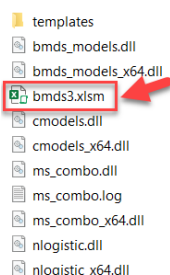
### 3. 분석 실행 및 분석 세부 조정

새로운 분석(용량-반응 모델 실행의 "세션")을 시작하려면 BMDS3 프로그램 폴더에서 bmds3.xlsm 파일을 연다. 파일이 저장된 디렉토리에서 파일을 열거나 해당 바로 가기 아이콘에서 파일을 선택한다.

[참고] 이 사용자 가이드 전체에서 bmds3.xlsm 파일을 Analysis Workbook 이라고 한다.

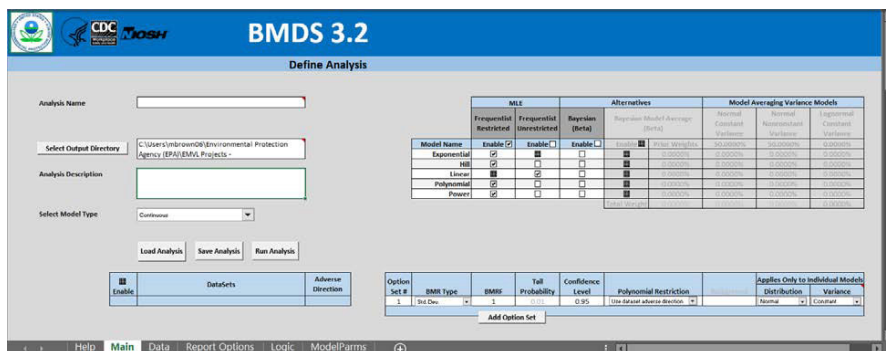
BMDS가 작동하려면 사용자가 처음 시작할 때 Excel 매크로를 활성화해야 한다.

**Figure 10.** BMDS3 folder contents, with bmds3.xlsm file highlighted



bmds3.xlsm이 열리면 Analysis Workbook의 Main 탭이 표시된다.

**Figure 11.** The Analysis Workbook on first opening.



Analysis Workbook에는 5개의 탭이 있다.

- Help - 사용자 가이드, 추가 정보, 버전 정보, Benchmark Dose Technical Guidance 및 사용자 지원 질문에 대한 eTicket을 포함한 BMDS 웹 사이트 지원 자료에 대한 링크가 포함되어 있다.
- Main - 실행할 분석을 정의하고, 분석 구성을 저장하고, 저장된 분석 구성을 로드하고, 분석을 실행합니다. 이러한 기능에 대한 자세한 내용은 "3. 분석 실행 및 분석 세부 조정-1) 1단계: 문서분석 및 -3) 3a단계: 모델링 옵션의 선택과 저장"을 참조한다.
- Data - 데이터를 수동으로 입력하거나 BMDS .dax 데이터 세트 파일 (BMDS 2.7 .dax 파일과 호환)을 로드하여 분석할 데이터 세트를 정의한다. 자세한 내용은 "3. 분석 실행 및 분석 세부 조정-2) 2단계: 데이터세트 (dataset) 추가"를 참조한다.
- Report Options - 결과 파일에 포함할 입력 및 분석 결과의 유형을 정의하고 Microsoft Word 기반 보고서 파일에 포함되어야 하는 항목도 정의합니다. 자세한 내용은 "3. 분석 실행 및 분석 세부 조정-7)5단계: 요약 Word 리포트 생성"을 참조한다.
- Logic - BMDS 모델 권장 사항 결정 logic 규칙을 검토하고 필요한 경우 사용자가 조정할 수 있다. 자세한 내용은 "10. BMDS 권장 및 결정 로직"을 참조한다.



## 1) 1단계: 문서 분석

### (1) 분석명(Analysis Name) 및 분석설명(Analysis Description) 입력

**Figure 12.** Fields for Analysis metadata.

The figure shows a software interface for entering analysis metadata. It consists of a light gray rectangular container. Inside, there are three main sections. The top section is labeled 'Analysis Name' and contains a white text input field. The middle section is labeled 'Select Output Directory' and contains a button with the same text. Below the button, a text box displays the selected directory path: 'C:\Users\mbrown06\Environmental Protection Agency (EPA)\EMVL Projects -'. The bottom section is labeled 'Analysis Description' and contains a larger white text input field.

Analysis Name란에 분석명을 기입한다. BMDS는 이를 사용하여 결과 통합 문서 파일 이름을 생성한다.

Analysis Description란에는 데이터세트를 설명하는 더 자세한 메모를 기록한다. 필수 단계는 아니지만 보다 완전한 자유 텍스트 분석 설명을 기록해두면 나중에 사용하거나 검토시 유용하다. 이 메모는 설정 통합 문서, 결과 통합 문서 및 Word report 파일에 표시된다.

### (2) 사용자세팅 및 결과 경로지정(directory)

기본 설정은 BMDS가 처음 열릴 때 자동으로 로드된다. 대부분의 경우 초기 Main, Data, Report Options 및 Logic 탭 설정은 항상 동일하다. 유일한 예외는 BMDS가 마지막으로 닫힐 때의 값으로 설정되는 "Select Output Directory(출력 디렉토리 선택)" 필드이다.

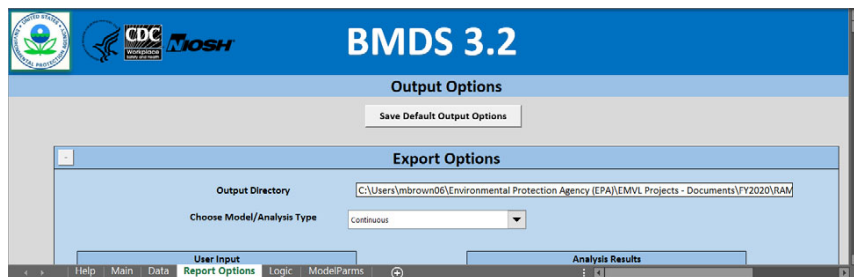
기본적으로 BMDS는 결과를 설치 디렉토리에 저장한다. Select Output Directory 버튼(Figure 12)을 누르면 다른 디렉토리를 선택하기 위한 파일 관리자 대화 상자가 나타난다. 사용자가 지정한 출력 디렉터리는 "Select Output Directory" 필드에 표시된다.

[참고] 필드 위로 마우스 커서를 가져가면 전체 출력 디렉토리 경로가 표시된다. 주소가 상자에 비해 너무 긴 경우에 유용하다.

출력 디렉터리는 설정 통합 문서(분석별로 분석 이름 사용), 결과 통합 문서(데이터세트당 하나, 데이터세트 이름 사용) 및 Word Report 파일(데이터세트 이름을 사용하여 데이터세트당 하나)을 저장한다.

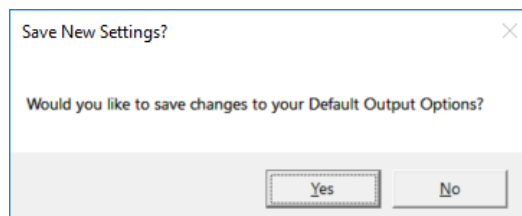
기본 탭에 지정된 출력 디렉토리는 Figure 13과 같이 Report Option 탭의 상단에도 표시된다.

**Figure 13.** Report Options tab, with the same Output Directory as specified on the Main tab.



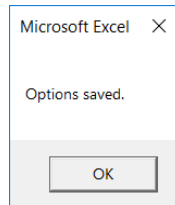
Main 탭에서 Select Output Directory 버튼을 사용하여 출력 디렉토리를 변경하면 BMDs는 "Save New setting?(새 설정을 저장하시겠습니까?)"라는 메시지를 표시한다(Figure 14). Yes를 선택하여 새 디렉토리를 기본값으로 설정한다(Figure 14).

**Figure 14.** Save New Settings dialog box.



그런 다음 BMDs는 옵션이 저장되었다는 메시지를 표시하는데 계속하려면 확인을 선택한다(Figure 15).

**Figure 15.** Confirmation dialog box for changing default output options.



## 2) 2단계: 데이터세트(dataset) 추가

Main 탭에 분석 문서 정보를 입력한 후 Data 탭에 용량 반응 데이터를 입력해야 한다.

사용자는 네 가지 유형의 여러 데이터세트를 추가할 수 있다.

- Summarized continuous(요약 연속형), 예: 평균 및 SD
- Individual continuous(개별 연속형), 예: 피험자별 용량 및 반응
- Dichotomous(이분형), 예: 병변 발생
- Nested dichotomous(등지 이분형 반응), 예: 생식 발달 연구

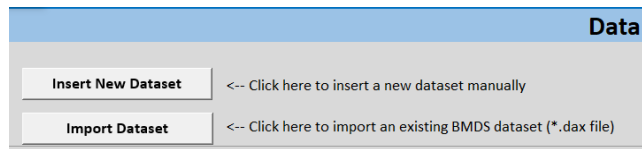
[참고] 사용자는 Data 탭에서 다양한 모델 유형(연속형, 이분형 및 등지)의 여러 데이터 세트를 입력할 수 있다. 그러나 Main 탭에서 선택한 모델 유형으로 모델링할 수 없는 데이터세트는 회색으로 표시되고 선택할 수 없다. 예를 들어 연속 및 이분법 데이터 세트는 별도의 분석 실행에서 분석해야 한다. 자세한 내용은 “3. 분석 실행 및 분석 세부 조정-3) 3a단계: 모델링 옵션의 선택과 저장”을 참조한다.

## (1) .dax 데이터세트 가져오기(importing)

이전 버전과 마찬가지로 BMDS 3 또한 데이터 세트를 .dax 형식으로 저장한다. 기존 .dax 데이터세트를 가져오려면

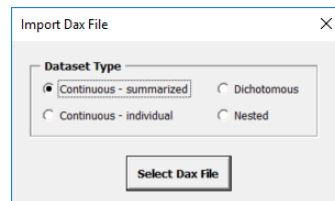
1. Data 탭에서 **Import Dataset** 버튼을 선택한다.

**Figure 16.** Data tab buttons for specifying datasets.



2. Import Dax File 대화 상자가 표시된다.

**Figure 17.** Import Dataset dialog



3. 데이터 세트 유형을 선택한 다음 **Select Dax File** 버튼을 선택한다. Dax 파일 선택 대화 상자가 표시된다.
4. 데이터 세트의 위치로 이동하여 .dax 파일을 선택하고 열기를 누른다.
5. **Map Data Columns** 대화 상자가 표시된다. 선택 목록에서 BMDS가 예상하는 변수 유형에 해당하는 적절한 **Dataset Header**를 선택한다.

**Figure 18.** Continuous-summarized dataset import dialog

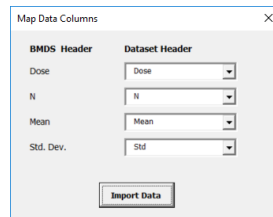


Figure 19. Continuous-individual dataset import dialog

Figure 20. Dichotomous dataset import dialog

Figure 21. Nested dataset import dialog

Figure 22. A .dax dataset imported into BMDS 3

Dose	N	Mean	Std. Dev.
0	20	6	1.2
25	20	5.2	1.1
50	19	2.4	0.81
100	20	1.1	0.74
200	20	0.75	0.66

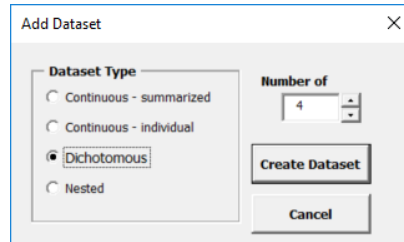
6. Import Data 버튼을 선택한다. BMDS는 데이터 세트를 Data 탭으로 가져온다.

[참고] 고유한 데이터 세트 이름을 입력하기 위한 가이드라인을 따른다. 또한 "3. 분석 실행 및 분석 세부 조정-2) 2단계: 데이터세트(dataset) 추가-(7) 회색 셀과 파란색 셀의 차이회색 셀과 파란색 셀의 차이점"을 참조한다.

## (2) 새로운 데이터세트 삽입(inserting)

1. Data 탭에서 **Insert New Dataset**를 선택한다. Add Dataset 대화 상자가 표시된다. 데이터 세트 유형 및 행 수를 지정한다.

**Figure 23.** Add Dataset dialog.



2. Create Dataset를 선택한다. BMDS는 지정된 수의 빈 행이 있는 테이블을 삽입한다.

**Figure 24.** Example of empty dataset table created for dichotomous data.

Edit	DataSet Name1		
	[Add user notes here]		
	Dose	N	Incidence
	[Custom]	[Custom]	[Custom]

3. 파란색 셀을 더블클릭하고 "DataSet Name1"을 의미 있는 설명으로 변경한다.
4. "[Add user notes here]" 셀을 편집하거나 비워 둔다.
5. 회색으로 표시된 셀은 편집할 수 없다. 적절한 모델 분석을 위해 BMDS가 허용하는 데이터 유형을 나타낸다. "[Custom]"으로 표시된 파란색 셀은 데이터세트가 각기 이름이 다르다면 사용자가 편집가능하다. 자세한 내용은 “3. 분석 실행 및 분석 세부 조정-2)2단계: 데이터세트(dataset) 추가-(7)회색 셀과 파란색 셀의 차이”를 참조한다.
6. 나머지 셀에 데이터를 입력한다. 또는 다른 테이블, 스프레드시트 또는 프로그램(예: 이전 BMDS 버전)에서 가져온 데이터를 복사하여 붙여 넣는다.

7. Main 탭을 클릭하여 표시한다. Dataset 테이블에는 입력한 Dataset 이름이 표시된다. 선택된 활성화 열은 BMDS가 선택한 데이터 세트에 대한 분석을 실행하도록 지시한다.

**Figure 25.** The Main tab after a dataset has been specified.

The screenshot shows the 'Main' tab interface. At the top, there are three buttons: 'Load Analysis', 'Save Analysis', and 'Run Analysis'. Below these is a table with the following structure:

<input checked="" type="checkbox"/>	DataSets
<input checked="" type="checkbox"/>	Dich dataset

### (3) 다중(multiple) 데이터세트 삽입

Data 탭에 더 많은 데이터세트를 추가하려면 **Import Dataset** 또는 **Insert New Dataset** 버튼을 선택하고 위에서 설명한 대로 데이터세트를 지정한다. BMDS는 자동으로 기존 데이터세트의 아래에 데이터 세트를 추가한다.

많은 데이터세트를 입력할 때 기본적으로 BMDS는 탭의 상단 부분을 "고정"하여 데이터 세트 가져오기 및 삽입 버튼이 항상 표시되도록 한다.

### (4) 데이터세트 편집(editing)

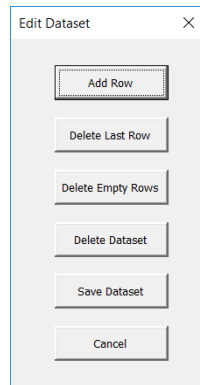
데이터세트를 입력하면 데이터세트 옆에 작은 **Edit** 버튼이 나타난다(Figure 26).

**Figure 26.** Example dataset with Edit button.

Edit	Dich dataset		
	[Add user notes here]		
	Dose	N	Incidence
	[Custom]	[Custom]	[Custom]
	0	50	0
	1000	50	2
	3000	50	6
	8000	50	25

**Edit** 버튼을 클릭하여 데이터 세트 편집 명령 팔레트를 표시한다. 이 명령은 선택한 데이터 세트에서만 작동한다.

**Figure 27. Edit Dataset commands**



- Add Row - 테이블 끝에 마지막 행의 복제본을 삽입한다.
- Delete Last Row - 테이블의 마지막 행을 삭제한다.
- Delete Empty Row - 모든 빈 행을 삭제한다. (용량 값이 없는 경우 BMDS는 행을 건너뛰다.)
- Delete Dataset - 전체 데이터 세트를 삭제한다.
- Save Dataset - 데이터 세트를 .dax 형식으로 저장한다. 버튼을 클릭하고 데이터세트를 저장할 위치로 이동한 다음 저장을 클릭한다.
- Cancel — 팔레트 창을 닫는다.

#### (5) 각 데이터세트의 고유이름 입력

각 데이터세트의 고유한 이름을 입력한다.

BMDS는 이 이름을 사용하여 기본 탭(사용자가 모델링 분석에 포함할 데이터 세트를 선택할 수 있음)에서 데이터세트를 참조하고 데이터세트 모델링에서 생성된 모든 결과 통합 문서 및 Word Report 파일의 이름을 지정한다.

[참고] 미만 기호(<), 수직 파이프(|) 등과 같은 Windows 특수문자를 사용하여 데이터 세트의 이름을 지정할 수 없다. 피해야 할 특수 문자는 아래와 같다.

< (less than)

> (greater than)



: (colon)  
 “ (double quote)  
 / (forward slash)  
 \ (backslash)  
 | (vertical bar or pipe)  
 ? (question mark)  
 \* (asterisk)

#### (6) 데이터세트는 빈 열(rows)을 가질 수 있다

데이터세트에는 빈 행이 있을 수 있다. BMDS 분석 실행은 용량 값이 누락된 모든 행을 건너뛰는다.

#### (7) 회색 셀과 파란색 셀의 차이

**Figure 28.** Explanatory text on the Data tab for the gray and blue cells.

Gray cells of the data tables that will appear below define data that must be in this sequence in the dataset

Blue cells can be edited to reflect your dataset's labels for these data types.

데이터세트의 회색 셀은 편집할 수 없다. 회색셀은 더 적절한 모델 분석을 위한 BMDS 시퀀스의 요구조건에 데이터 유형이 적합하다는 것을 의미한다.

파란색 셀은 데이터세트에 다른 변수 이름을 적용하고자 할 때 편집할 수 있다.

### 3) 3a단계: 모델링 옵션의 선택과 저장

분석에 사용할 수 있는 모든 모델 및 모델링 옵션은 Main 탭에서 선택할 수 있다.

분석을 실행하기 전이나 후에 언제든지 옵션을 저장하고 다시 로드할 수 있다.

분석에는 다음 네 가지 모델 유형 중 하나를 사용하는 것이 포함될 수 있습니다.

- 연속형(continuous)
- 이분형(dichotomous)
- 이분형 - 다종양(multi-tumor)
- 이분형 - 중첩(nested)

이전에 언급했듯이 Analysis 통합 문서에는 Data 탭의 모든 데이터 유형이 혼합되어 포함될 수 있다. 그러나 단일 분석에서는 한 가지 유형의 데이터만 실행할 수 있다. 예를 들어, Analysis Workbook은 Data 탭에 정의된 Continuous, Dichotomous 및 Nested 데이터 세트를 포함할 수 있지만, 분석을 실행할 때 사용자는 각 데이터 유형에 대해 별도의 분석을 실행해야 한다.

물론 사용자는 단일 분석 통합 문서에서 두 개(또는 그 이상)의 분석을 자유롭게 실행할 수 있으며, 각각은 잠재적으로 다른 데이터 유형의 분석을 구현할 수 있다.

사용자는 여러 사용자 정의 모델링 "Options Set" 및 여러 데이터 세트에 대해 선택한 모델을 실행할 수 있다.

BMDs Main 탭에는 Data 탭에 입력된 데이터세트가 나열되어 사용자가 선택한 모델 및 옵션 세트를 사용하여 분석할 적절한 모델 유형의 데이터세트를 선택할 수 있다.

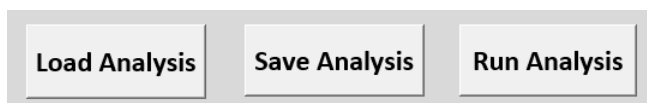
각 모델 유형은 다양한 모델 및/또는 모델링 옵션 세트를 제공한다.

- 연속 끝점 모델 옵션에 대한 자세한 내용은 "6. 연속형 데이터의 종말점-3) 옵션(options)"을 참조한다.
- 이분형 자료의 종말점 모델 옵션에 대한 자세한 내용은 "7. 이분형 자료의 종말점-2) 옵션"을 참조한다.
- 중첩형 이분형 종말점 모델 옵션에 대한 내용은 본 매뉴얼에서는 다루지 않았으며 원본의 매뉴얼을 참고하도록 한다.

#### 4) 3b단계: 불러오기(load), 저장(save), 분석실행(run)

[기억하기] 분석 통합 문서에는 데이터가 저장되어 있지 않다. bmds3.xslm에 대한 쓰기 방지의 부작용은 사용자가 통합 문서 자체에 데이터를 저장할 수 없다는 것이다. 그러나 모든 데이터세트, 모델링, 리포팅 옵션, 로직 설정 등은 나중에 사용하기 위해 저장하고 다시 로드할 수 있다.

**Figure 29.** Analysis function buttons.



**Save Analysis:** 분석을 저장(실행하지 않고)한다. 결과가 생성되지 않은 상태로 통합 문서를 저장하려면 이 버튼을 선택한다.

분석을 실행하지 않고 분석 옵션을 저장하려면

1. 분석을 위한 옵션, 데이터세트 등을 지정한다. Analysis Name 필드는 파일 이름으로 사용된다.
2. Save Analysis 버튼을 선택한다. BMDs는 Output Directory에 Settings Workbook 파일을 생성한다. 파일을 열고 Setting Workbook 파일의 Summary 탭에 다음의 메시지를 표시한다(Figure 34).

**Figure 30.** BMDs displays the following message on the Settings Workbook's Summary tab.

This analysis was saved without running. Please open the BMDs analysis workbook (BMDs3.xlsm) then click the 'Load Analysis' button to load this analysis.

**Run Analysis:** 분석을 실행한다. 분석을 지정한 후 이 버튼을 선택한다. BMDs는 Results Workbook 파일과 (선택한 경우) 사용자 지정 Output Directory에 Word Report 파일을 자동으로 생성한다.

모든 Analysis Workbook 사양은 Setting Workbook 또는 Results Workbook 파일에 저장되므로 해당 Analysis Workbook을 다시 생성할 수 있다. **Load Analysis** 버튼을 선택하면 BMDS가 Analysis Workbook을 다시 생성한다. 모델링과 마찬가지로 원래 Analysis Workbook의 모델선택로직, 리포팅옵션 모델링, 모델 선택 논리 및 보고 옵션은 물론 모든 데이터세트가 Analysis Workbook으로 다시 로드된다. 그런 다음 사용자는 다른 옵션을 사용하여 분석을 다시 실행할 수 있다.

#### 분석을 로드하고 다시 실행하려면

나중에 분석을 다시 실행하거나 다른 파라미터 또는 추가 데이터로 분석을 실행하려면

1. bmds3.xslm 파일을 연다.
2. Main 탭에서 Load Analysis를 선택한다.
3. Open file manager 대화 상자가 표시된다. 저장된 BMDS Results Workbook 파일 또는 이전에 저장된 분석 파일로 이동하여 선택하고 열기를 선택한다.
4. BMDS는 선택한 통합 문서의 저장된 데이터세트와 파라미터를 bmds3.xslm으로 불러온다. 거기에서 새 파라미터를 지정하고, 데이터 또는 데이터 세트를 추가하고, 추가적으로 Model-Option Set을 만드는 등의 작업을 수행한다.

[참고] Analysis Workbook이 다시 생성되면 열려 있는 Analysis Workbook의 모든 옵션과 데이터세트를 덮어쓴다. BMDS는 사용자가 열려 있는 Analysis Workbook 파일을 저장하거나 덮어쓸 수 있도록 경고를 표시한다.

## 5) 4단계: 분석실행

Main 탭의 Dataset 표에서 분석을 실행할 데이터세트를 하나 이상 선택한다.

헤더에서 Enable checkbox를 클릭하여 로드된 모든 데이터 세트의 선택/선택 취소를 결정한다(Figure 31). 이를 통해 선택한 모델 유형에 해당하는 데이터 세트만 선택할 수 있다.

**Figure 31.** With Continuous as the selected Model Type, selecting the Enable checkbox selects the continuous dataset(s) that can be run in the current analysis.

Enable	DataSets	Adverse Direction
<input type="checkbox"/>	Dich dataset	
<input checked="" type="checkbox"/>	Cont dataset	automatic
<input type="checkbox"/>	Nested dataset	

Analysis Workbook의 Main 탭에서 Run Analysis 버튼을 선택하여 모델링 실행을 시작한다. 모델링 상태를 보여주는 막대가 나타난다. Excel에서 Results Workbook을 편집할 때 디스플레이 업데이트가 일시 중단된다.

사용자는 원하는 대로 파일의 이름을 바꿀 수 있다. BMDs 결과 통합 문서 파일 명명 규칙은 “2.BMDs3의 기본 메뉴-3) 워크북 세팅”에 설명되어 있다.

각 데이터세트는 자체의 Result Workbook에 저장된다. 분석의 일부로 지정된 모든 모델 옵션도 새 Results Workbook에 저장된다.

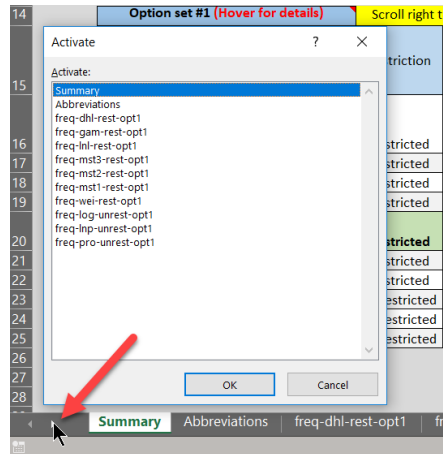
## 6) 5단계: 결과 리뷰

사용자가 Run Analysis 버튼을 선택하면 BMDs는 지정된 분석에 대한 Results Workbook을 생성하고 파일을 연다.

Results Workbook 창 하단에는 완료된 분석에 대한 Summary 및 model/

option별 결과가 포함된 탭이 있다. Excel의 탭 선택 삼각형을 마우스 오른쪽 버튼으로 클릭하여 파일의 모든 탭 목록을 본다(Figure 36).

**Figure 32.** Viewing the list of Results Workbook tabs in Excel by right-clicking the tab-selection triangle.



각 모델의 결과 탭에는 다른 결과 중에서 기준용량(BMD, benchmark dose), 모델 파라미터(model parameters), 적합도(goodness of fit) 및 이탈 분석(analysis of deviance)과 관련된 보다 자세한 보고서와 함께 자체 BMD 그래프가 포함되어 있다.

#### (1) 모델 약어표(Model Abbreviations Table)

Results Workbook의 각 탭은 명명 규칙을 따라 각 결과를 고유하게 식별한다.

Abbreviation 탭에는 분석에 사용된 모델 유형에 대한 세 개 문자로 된 약어가 모두 요약되어 있다. 예를 들어, 이분형 모델이 있는 분석의 Abbreviation 탭에는 이분형 모델에 대한 약어만 포함된다.

다음 표에는 BMDS 3에서 사용할 수 있는 모든 모델, 해당 약어(결과 탭 이름 및 프로그램 내에서 기타 레이블 지정 목적으로 사용됨) 및 버전 번호를 나타내었다(Table 1-4).

모델 버전 번호는 BMDS 2.7에 사용된 버전 번호에서 재설정되었다. BMDS 3 이전에는 모델이 고유한 버전 번호가 있는 별도의 실행 파일로 코딩되었다.

BMDS 3의 경우 모델은 이제 단일 DLL로 코딩되므로 달리 명시되지 않는 한 동일한 버전 번호를 사용한다.

일부 사용자는 결과의 출처를 밝히기 위해 출판물의 버전 번호를 인용하고자 할 수 있다.

**Table 1.** *Frequentist Continuous Models: Abbreviations & Versions.*

Name	Abbreviation	Version (Date)
Exponential	exp	1.1 (07/19/2019)
Hill	hil	1.1 (07/19/2019)
Linear	lin	1.1 (07/19/2019)
Polynomial	ply	1.1 (07/19/2019)
Power	pow	1.1 (07/19/2019)

**Table 2.** *Frequentist Dichotomous Models: Abbreviations & Versions.*

Name	Abbreviation	Version (Date)
Gamma	gam	1.1 (07/19/2019)
Logistic	log	1.1 (07/19/2019)
Log-Logistic	lnl	1.1 (07/19/2019)
Log-Probit	lnp	1.1 (07/19/2019)
Multistage	mst	1.1 (07/19/2019)
Probit	pro	1.1 (07/19/2019)
Weibull	wei	1.1 (07/19/2019)
Quantal Linear	qln	1.1 (07/19/2019)
Dichotomous Hill	dhl	1.1 (07/19/2019)

**Table 3.** *Frequentist Nested Dichotomous Models: Abbreviations & Versions.*

Name	Abbreviation	Version (Date)
Nested Logistic	nlm	2.20 (04/27/2015)

**Table 4.** *Other Models: Abbreviations & Versions.*

Name	Abbreviation	Version (Date)
Bayesian model averaging (dichotomous)	bma	1.1 (07/19/2019)
Multi-tumor (MS_Combo)	msc	1.8 (04/27/2015)

## (2) 탭에서의 그 외 공통 약어

결과 통합 문서 탭 이름에 사용되는 다른 일반적인 약어는 다음과 같다.

- Frequentist("freq") 또는 Bayesian("bayes").
- 모델 평균이 선택된 경우 Dichotomous Model Averaging ("DichoMA"), if Model Averaging was selected.
- Restricted ("rest"), Unrestricted ("unrest"). 사용자는 Analysis Workbook의 Main 탭에서 restricted 및 unrestricted 실행을 선택할 수 있다.
- "dsetn" = dataset1, dataset2, 등등.
- "optn" = 여러 Option Sets이 지정된 경우 결과를 생성하는 데 사용된 Analysis Workbook의 Main 탭에서 Option Set을 식별한다. 만약 하나의 옵션 세트만 식별된 경우 항상 "opt1"이 된다.

[예] "freq-gam-rest-opt1"이라는 결과 탭에는 Analysis Workbook에 정의된 첫 번째 Option Set을 사용하고, 제한된 실행, frequentist 감마 모델에 대한 결과가 포함되어 있다.

- Dichotomous-Multi-tumor(MS\_Combo) 약어

**Figure 33. Multi-tumor Results Workbook tabs.**

Abbreviations	<b>frequentist-msc-option1</b>	freq-mst3-rest-dset1-opt1	freq-mst2-rest-dset1-opt1	freq-mst1-rest-dset1-opt1
---------------	--------------------------------	---------------------------	---------------------------	---------------------------

다음 약어는 다종양 결과에서만 사용된다.

"msc" = MS\_Combo.

"mstn" = 개별 multistage ("mst") 결과 탭은 MS\_Combo 결과의 일부로 간주된다.

- Dichotomous-Nested 약어

**Figure 34. Nested Results Workbook tabs.**

<b>freq-nln-rest_isc+ilc+_opt1</b>	freq-nln-rest_isc+ilc-_opt1	freq-nln-rest_isc-ilc+_opt1	freq-nln-rest_isc-ilc-_opt1
------------------------------------	-----------------------------	-----------------------------	-----------------------------



다음 약어는 등지형 결과에서만 사용된다.

“nlm” = Nested Logistic model. 이 경우는 단일 등지형 모델에만 있다.

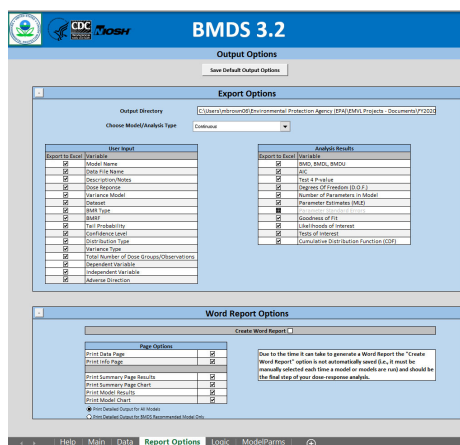
“lsc” = Litter Specific Covariate. “+”는 LSC가 포함되었음을 의미한다. “-”는 LSC가 포함되지 않았음을 의미한다.

“ilc” = Intralitter Correlation. “+”는 ILC 평가가 포함되었음을 의미한다. “-”는 ILC가 zero 라고 추정한다.

## 7) 5단계: 요약 Word 리포트 생성

BMDs 3의 Analysis Workbook's Report 탭에서 사용자는 각 모델 유형을 리포트하기 위해 "Export Options" 및 "Word Report Options"에서 리포팅할 모델링 입력 및 결과를 선택할 수 있다.

Figure 35. Analysis Workbook's Report Options tab.



- 선택한 Export Options은 분석시 생성된 Result Workbook 및 Word Report 파일 모두에 영향을 준다.
- Export Options("User Input" 및 "Analysis Results")은 서로 다른 모델/분석 유형 각각에 대해 별도로 설정된다. 따라서 "Continuous", "Dichotomous-Nested" 등은 자신의 자체 Export options 세트가 있다. Model/Analysis 유형 드롭다운 메뉴에서 모델 유형을 선택한다.

- Word Report Options은 Microsoft Word에서 모델링 결과의 표 형식 문서를 생성할 때 적용된다.

[참고] Word 보고서를 편집하는 데 몇 분 정도 걸릴 수 있으므로 분석 결과를 확인한 후 Word Report Options을 실행하는 것이 좋다. 분석 결과가 만족스러우면 "Create Word Report" 옵션을 선택한 상태에서 다시 실행한다.

[참고] Word Report 옵션을 사용할 때 사용자는 Main 탭에서 단일 Model-Option 세트 조합만 지정하는 것이 좋다. 지정된 데이터 세트와 옵션이 많을수록 BMDS가 보고서를 생성하는 데 더 오래 걸린다. Word Report 생성 개선 사항은 향후 BMDS 버전에서 해결할 예정이다.

## 4. BMDS의 모델링

BMDS가 모델링할 수 있는 세 가지 종말점(endpoint) 유형은 연속, 이분형 및 등지 이분형이다. 이들 이외에 별도로 고려되는 한 가지 추가 상황이 있는데 이는 "MS\_Combo" 옵션이다. 그러나 이 옵션은 다중 종말점(multiple endpoint)의 이분형 모델링의 특별한 경우일 뿐이다. MS\_Combo 모델링 옵션에 대한 자세한 내용은 “8. 다중 종양 분석-1) 이분형 자료: multi-tumor models과 옵션”에서 설명하였다. 여기에 제시된 것은 모든 모델 유형에 공통적인 일반적인 모델링 고려 사항이다.

### 1) Frequentist and Bayesian

이제 BMDS를 통해 사용자는 Bayesian 또는 non-Bayesian 분석 중 하나 (또는 둘 다)를 수행할 수 있다.

BMDS는 non-Bayesian 접근 방식을 "Frequentist" 또는 "maximum-likelihood estimation(MLE), 최대 우도 추정"이라고 한다. 이 접근 방식은 가능성 계산을 기반으로 한다. 이러한 방법으로 적합된 모델(models fit)은 프로필 가능성 접근 방식에 의해 결정된 최대 가능성 추정치 및 관련 경계를 보고한다. p-값 등의 표시(예: 적합도(goodness-of-fit) 평가)는 전통적인 frequentist와 일치한다.

대조적으로 Bayesian 분석은 모수 추정치(parameter estimates)를 업데이트 한다. 모수 값(parameter value)의 선험적 불확실성을 설명하는 분포(소위 사전 분포(prior distributions))는 사후 분포를 산출하기 위해 고려 중인 데이터를 사용하여 업데이트된다. 이들로부터 사후 가능성(*posteriori* distribution)을 최대화하는 BMD 추정치(소위 최대 사후 확률(maximum a *posteriori* probability(MAP) 추정치)가 BMD를 신뢰할 수 있는 구간으로 보고된다.

[참고] 현재 EPA는 베이지안 모델링 또는 베이지안 모델 평균화에 대한 기술 지침을 제공하지 않는다.

## 2) 모델 매개변수(parameters)

Figure 36. Model Parameters tab.

The screenshot displays the 'Model Parameters' tab in BMDS 3.2. It is divided into three main sections: Frequentist Unrestricted, Frequentist Restricted, and Bayesian. Each section contains sub-tables for different models: Log Logistic, Gamma, Logistic, Probit, and Quantal Linear. Each model sub-table lists parameters (e.g.,  $\mu$ ,  $\sigma$ ,  $\alpha$ ,  $\beta$ ) with their initial values, distributions, and settings. The 'Logistic' model in the Frequentist Restricted section is highlighted in the center.

ModelParms 탭에는 각 모델의 매개변수(parameter) 및 사양(specifications)과 함께 BMDS가 실행하는 모든 모델이 표시된다. ModelParms 탭에 나열된 값은 프로그램에 대한 실제 입력값이다. 모델 실행 파일에서 읽고 사용한다. 값은 암호로 보호되며 사용자가 변경/편집할 수 없다.

사양(specification)은 모델링 접근 방식에 따라 빈도주의(Frequentist)와 베이지안(Bayesian)으로 구분된다.

빈도주의적 접근 방식의 경우 사양은 초기 값(우도(likelihood) 최적화를 시작하는 데 사용되는 값)과 매개변수 값에 대한 제약(최소값 및 최대값)으로 구성된다. 제한된 형태(restricted form)와 제한되지 않은 형태(unrestricted form)의 모델이 있는 경우, 두 형태에 대한 별도의 빈도주의적 사양이 있으며, bound(최대값 또는 최소값)와는 다르다.

베이지안 접근 방식의 경우 사양은 각 모수에 대한 사전 분포를 정의한다.

분포 유형(예: 정규분포)과 해당 사전을 정의하는 매개변수(예: 평균 및 표준 편차)가 제공된다. prior도 수적 안정성을 위해 제한된다. bound는 "최소" 및 "최대" 열 아래에 나열된다.

추가 정보를 위해 공식 수학적/통계적 세부 사항과 관련하여 빈도주의 및 베이지안 모델 방정식에 대한 설명은 다음 섹션에 기술되었다.

- 6. 연속데이터의 종말점-4) Simple design에서 연속적인 종말점을 위한 모델의 수학적 상세
- 7. 이분형 자료의 종말점-3) Simple design에서 이분형 종말점을 위한 모델의 수학적 상세

또한 특히 priors 및 모델 제약 조건이 표시되는 다음 표를 참조한다.

- Table 7: The individual Bayesian continuous models(preview) and their parameter priors
- Table 12: Bayesian dichotomous models and their respective parameter priors

모델 매개변수 추정치가 BMDS 결과에 보고되는 방식을 보려면 “5. 각 모델 유형의 공통 출력물-3) 모델 매개변수(parameter) 표”를 참조한다.

### 3) BMDS에서 사용되는 최적화 알고리즘

#### (Optimization Algorithms used in BMDS)

빈도 분석 및 일부 베이지안 계산의 경우 NLOpt 최적화 라이브러리가 BMDS 3.0에 사용된다. 라이브러리에서 사용할 수 있는 몇 가지 최적화 알고리즘은 추정의 신뢰성을 보장하기 위해 사용된다.

- 최대 우도(maximum likelihood) 또는 최대 사후 추정을 포함하는 전역 최적화의 경우 L-BFGS1 방법이 먼저 시도된다. 수렴에 실패하면 그래디언트 프리 알고리즘 "subplex" 및 "BOBYQA"2 알고리즘이 시도된다.

- 프로파일링의 경우 비선형 부등식 제약 조건만 필요한 경우 COBYLA3 및 MMA4 접근 방식을 사용하여 비교한다. 방법이 다른 최적값을 반환하는 경우 둘 중 더 큰 값을 생성하는 값이 사용된다.
- 등식 제약 최적화의 경우 증강 Lagrangian 알고리즘이 사용되며 L-BFGS, BOBYQA 또는 "subplex" 알고리즘이 지역적인 최적화 단계에서 사용된다. 두 가지 접근 방식이 서로 다른 결과를 생성하면 더 큰 최적값을 생성하는 값이 사용된다.

NLopt 2.4.1은 BMDS 3 코드를 개발할 때 사용되었다. 이 버전은 NLopt GitHub 사이트에서 다운로드할 수 있다. 알고리즘에 대한 자세한 내용은 NLopt 설명서 사이트를 참조한다.

#### 4) Model averaging을 포함한 베이지안 분석

(Bayesian analysis, including Model averaging)

BMDS 모델 averaging은 데이터세트를 사용하여 모델의 매개변수가 갱신되는 베이지안 분석을 기반으로 진행된다.

매개변수에 대한 사전 설정은 연속형(미리 보기에서) 및 이분형 용량 반응 모델에 대해 각각 섹션 “6. 연속데이터의 종말점-4) Simple design에서 연속적인 종말점을 위한 모델의 수학적 상세-(7) Bayesian 연속모델(미리보기) 설명” 및 “7. 이분형 자료의 종말점-3) Simple designs에서 이분형 종말점을 위한 모델의 수학적 상세-(7) Bayesian 이분형 모델 설명”에 정의되어 있다. BMDS 3.2에서는 이분형 모델 averaging만 사용할 수 있다.

각 모델  $M$ 에 대해 데이터 생성 메커니즘(이분형 종말점의 경우 이항 샘플링, 연속 데이터의 경우 정규 또는 로그 정규 분포)을 기반으로 하는 데이터  $(D|M)$ 의 가능성이 있다.

하나 이상의 모델에 관심이 있을 때 모델 선택 대신에 모델 averaging을 진지하게 고려해보아야 한다(예: "best"로 간주되는 한 모델에 기반한 추론).

이러한 상황에서  $K$  모델( $M_k, k=1, \dots, K$ )을 고려하고 있다고 가정한다. 각 모델에 대해 BMD는 Laplacian 근사치를 사용하여 BMD의 사후 밀도 (posterior density)를 근사화한다. 이 밀도를 모델  $K$ 의  $g_k(\text{BMD}|M_k, D)$ 라고 한다. 모델  $k$ 에 대한 매개변수 벡터가  $\theta_k$ 로 표시하면  $\hat{\theta}_k$ 는 사후 우도를 최대화 하는 벡터의 값을 지정한다(최대 사후 추정 또는 MAP 추정). 모델 평균 BMD의 후방 밀도는 다음과 같다.

$$g_{ma}(\text{BMD}|D) = \sum_{k=1}^K \pi_k(M_k|D)g_k(\text{BMD}|M_k, D),$$

여기서  $\pi_k$ 는 주어진 데이터에서 모델  $M_k$ 의 사후 확률이다.

이 접근 방식은 고려되는 각 모델에 대한 사후 확률의 추정을 필요로 한다. 이는 averaging 과정의 가중치이다. 다른 곳에서 사용된 접근 방식과 달리 정보(information)-기준(criteria)-기반(based) 가중치(예: 베이지안 정보 기준 또는 Akaike 정보 기준 기반)의 사용을 피한다. 오히려 BMD는 데이터의 한계 밀도에 대한 Laplace 근사치를 사용하여 가중치를 생성한다. 즉, 모델  $M_k$ 의 경우  $1 \leq k \leq K$ 이고 길이가  $s$ 인 매개변수 벡터  $\theta_k$ 에서 주변 밀도 (marginal density)는 다음과 같이 근사치를 구한다.

$$I_k = (2\pi)^{s/2} |\hat{\Sigma}_k|^{-1/2} \ell(D|M_k, \hat{\theta}_k) g(\hat{\theta}_k|M_k)$$

$\hat{\theta}_k$  is the MAP estimate,

$\hat{\Sigma}_k$  is the negative inverse Hessian matrix evaluated at  $\hat{\theta}_k$ ,

$\ell(D|M_k, \hat{\theta}_k)$  is the likelihood of the data, for model  $k$  evaluated at the MAP, and

$g(\hat{\theta}_k|M_k)$  is the value of the prior density for  $M_k$  evaluated at the MAP parameter estimates.

$M_k$ 에 대한 사후 모델 확률을 계산하기 위해 MAP를 계산한 다음 앞의 방정식을 사용하여  $I_k$ 를 계산한다. 모델의 사후 확률은

$$\pi_k(M_k|D) = \frac{w_k I_k}{\sum_{i=1}^K w_i I_i},$$

여기서  $w_k$ 가 모델  $M_k$ 의 사전 가능성이다. BMDS에서 사용자는 이들 가중치를 지정할 수 있다. 기본값은 각 모델에서 동일한 가중치를 둔다.

이 근사치는 Fletcher와 Turek(2012)의 MAPL(Model Averaged Profile Likelihood) 접근 방식과 유사하다. 그러나 MAPL은 가능성에만 의존하지만 BMDS에서의 접근 방식은 BMD의 한계 프로파일 밀도를 계산할 때 사전 정보를 통합한다. 즉, 우도와 prior가 모두 사용된다. 모델별 밀도는 프로파일 밀도 경계를 관심 매개변수에 대한 한계 사후 밀도의 분위수로 처리하여 정의되며, 현재 접근 방식과 MAPL 접근 방식과의 관계는 점근적으로(asymptotically) 정당화된다.

이 접근법은 각 단계의 가능성에 대한 사후 밀도를 대체하여 MAPL 프레임워크와 관련될 수 있다. 이 방법은 사후 MAP 추정값과 로그 사후(log-posterior)의 Hessian 행렬을 사용하여 한계 가능성을 근사화한다.

[참고] 모델 평균 접근 방식의 경우 이분형 베이지안 모델( 7)이분형 자료의 종말점-(1)이분형 반응 모델에서 설명)을 모델 평균화에서 사용할 수 있다. 이분법 모델 평균화의 경우 다단계 모델은 최대 2로 제한된다. 이에 대한 추론은 Nitcheva 등의 작업에 따른다. 모델 평균화 모음의 다른 모델(예: dichotomous Hill)이 증가된 곡률(curvature)을 제공할 수 있다는 사실을 고려할 때 고차 다항식이 필요하지 않음을 보여준다.

BMDS 모델 평균 BMD 포인트 추정치는 개별 모델의 BMD MAP 추정치의 가중 평균이며 사후 가중치  $\pi_k(M_k|D)$ 는 이것은  $\theta$ 의 대략적인 사후 밀도의 중앙값과 같다. BMDL 또는 BMDU 추정의 경우  $g_{ma}$ 를 정의하는 방정식이 통합된다. 100( $\alpha$ )% BMDU 추정 또는 100(1 -  $\alpha$ )% BMDL 추정은 다음과 같은  $BMD\alpha$  값이다.



$$\begin{aligned}\alpha &= \int_{-\infty}^{BMD_{\alpha}} g_{m,\alpha}(BMD|D) dBMD, \\ &= \sum_{k=1}^K \pi_k(M_k|D) \int_{-\infty}^{BMD_{\alpha}} g_k(BMD|M_k, D) dBMD.\end{aligned}$$

The quantity  $\int_{-\infty}^{BMD_{\alpha}} g_k(BMD|M_k, D) dBMD$  is approximated by,

$$\begin{aligned}&\int_{-\infty}^{BMD_{\alpha}} g_k(BMD|M_k, D) dBMD \\ &\approx \frac{1}{2} \Pr(-2 \log[\hat{g}_k(\widehat{BMD}|M_k, D)] - 2 \log[\hat{g}_k(BMD_{\alpha}|M_k, D)] < \chi^2_{1,\alpha}),\end{aligned}$$

여기서  $\hat{g}_k(x|M_k, D)$ 는  $x$ 에서 사후 측정된 최대값이다.  $\widehat{BMD}$ 는 BMD의 MAP 추정값이고,  $\chi^2_{1,\alpha}$ 는 자유도가 1인 카이제곱 확률변수의  $\alpha$ 변위치 (quantile)이다. 위의 근사값은  $BMD_{\alpha} < \widehat{BMD}$ 을 가정한다.

$\widehat{BMD} < BMD_{\alpha}$  일 때 이 방정식의 오른쪽은

$$\approx 1 - \frac{1}{2} \Pr(-2 \log[\hat{g}_k(\widehat{BMD}|M_k, D)] - 2 \log[\hat{g}_k(BMD_{\gamma}|M_k, D)] < \chi^2_{1,\gamma})$$

으로 대체된다.

이 근사값은 최대우도법을 사용하여 BMDL과 BMDU를 추정할 때 사용하는 profile-likelihood와 비슷하지만 이 경우에는  $\hat{g}_k(x|M_k, D)$ 가 사후 밀도 (posterior density)이고, 우도와 prior를 모두 포함한다.

## 5. 각 모델 유형의 공통 출력물

BMDs 3 Analysis Workbook에 의해 생성된 데이터 세트별 과 Results Workbook에는 별도의 탭에 결과 Model-Option Set 조합이 포함되어 있다. 각 탭은 모델링 입력 및 결과의 표 및 그래픽 요약으로 구성된다.

이 결과의 목적은 사용자에게 적합성 기준과 모델 결과를 제공하여 BMD 도출을 위해 적절한 모델 및 옵션 세트를 결정하는 데 도움을 주고자 하는 것이다.

이 섹션에서는 모든 모델 유형에 공통적인 BMDs 모델 출력에 대해 설명한다. 각 모델 유형별 출력에 대한 자세한 내용은 다음을 참조한다.

- 6. 연속데이터의 종말점-5)빈도주의 연속형 모델에 대한 특이적인 출력
- 6. 연속데이터의 종말점-6)베이지안 연속형 모델(미리보기)에 대한 특이적인 출력
- 7. 이분형 자료의 종말점-4)빈도주의 이분형 모델에 대한 특이적인 출력
- 7. 이분형 자료의 종말점-5)베이지안 이분형 모델에 대한 특이적인 출력

### 1) 모델 실행 문서(사용자 입력 표)

Figure 37. User Input table

User Input	
<b>Info</b>	
Model	frequentist Dichotomous Hill v1.1
Dataset Name	Inflammation, Chronic Active
User notes	[Add user notes here]
Dose-Response Model	$P[\text{dose}] = g + (v - v * g) / [1 + \exp(-a - b * \text{Log}(\text{dose}))]$
<b>Model Options</b>	
Risk Type	Extra Risk
BMR	0.1
Confidence Level	0.95
Background	Estimated
<b>Model Data</b>	
Dependent Variable	Dose
Independent Variable	Effect
Total # of Observations	5

각 Model-Options Set에 대해 생성된 Results Workbook 탭에는 User Input 표가 포함되어 있다. User Input 표에는 해당 Model-Option Set에 대해 선택된 옵션이 나열된다.

예를 들어, 두 명의 사용자가 결과를 비교하고 다른 답변을 얻은 경우 각각의 User Input 표를 참조하여 설정이 동일한지 또는 동일한(또는 가장 최신) 버전의 모델을 사용했는지 확인할 수 있다.

각 Model-Options Set의 Results Wrkbook 탭에 있는 User Input 표에는 모델 이름과 버전 번호, 데이터세트 이름, 데이터세트 사용자 메모를 비롯하여 Analysis Workbook의 Main 및 Data 탭에서 사용자가 입력하고 지정한 모델링 옵션이 포함되어 있다.

## 2) BMD 평가와 key fit 통계(benchmark dose table)

**Figure 38. Benchmark Dose table**

Benchmark Dose	
BMD	2295.279512
BMDL	1487.825333
BMDU	3108.76178
AIC	129.5361884
P-value	0.38149439
D.O.F.	1
Chi <sup>2</sup>	0.765880782

결과 탭의 Benchmark Dose 표에는 Model-Option 세트에 대한 BMD, BMDL 및 BMDU 추정치, AIC 및 전체 적합도 검정(goodness-of-fit test) p-값, 자유도(degrees of freedom, D.O.F) 및 카이제곱(chi-square)이 포함된다.

### (1) AIC

BMDS Results Workbook 탭의 Akaike Information Criterion(AIC) 값은 다음과 같이 계산되었다.

$$AIC = -2 \times LL + 2p$$

여기서 LL은 모수에 대한 최대 우도 추정치에서의 로그 우도(log-likelihood)이고  $p$ 는 추정된 모델 모수(제한 경계(restriction boundary)가 아님)의 수이다.

AIC는 동일한 데이터 세트에 대해 서로 다른 모델 적합(동일한 적합 방법, 예: 최소 제곱 또는 최대 우도 사용)을 비교하는 데 사용할 수 있다. AIC 값이 작을수록 적합도가 더 높음을 나타낸다. AIC 비교는 정확하지 않지만(AIC 차이를 해석하기 위해 경험 법칙에 의존) 모델 선택에 유용한 지침을 제공할 수 있다.

AIC에 대한 모델 유형별 세부 정보는 아래의 섹션에서 설명하였다.

- 연속형 종말점은 “6. 연속형 데이터의 종말점-4)Simple design에서 연속적인 종말점을 위한 모델의 수학적 상세-(4) AIC와 모델 비교”를 참조한다.
- 이분형 종말점의 경우는 “7. 이분형 자료의 종말점-3)Simple design에서 이분형 종말점을 위한 모델의 수학적 상세-(2)AIC와 model비교”를 참조한다.

## (2) P-value

$p$ -값은 자유도(D.O.F) 및 Chi2 값을 기반으로 계산된다(Chi2는 D.O.F와 같은 자유도를 갖는 카이제곱 분포로 분포된다고 가정함).  $p$ -값은 관찰된 데이터에 대한 모델 예측의 "근접성(closeness)"을 측정한다. 전체  $p$ -값이 미리 결정된 임계  $p$ -값보다 크면 사용자는 모델이 관찰된 용량-반응 패턴을 적절하게 설명한다고 추론할 수 있다. EPA에서 사용하는 임계  $p$ -값은 일반적으로 0.1이지만 암 데이터에 적용하는 다단계 모델의 경우 때때로 0.05로 완화하여 적용한다.

### 3) 모델 파라미터 표

**Figure 39.** Model Parameters Table, with hover tip explaining the Bounded estimate.

Model Parameters	
# of Parameters	4
Variable	Estimate
g	1.65207E-08
b1	3.70531E-05
b2	Bounded
b3	7.721E-13

The value of this parameter, 0, is within the tolerance of the bound (see user guide for tolerance limits)

Model Parameter 표에는 모형 적합(model fit)을 "최적화(optimize)"하는 매개변수 값에 대한 추정값이 포함된다.

매개변수 추정치는 매개변수 경계(parameter boundary)의 주어진 허용 오차(tolerance)( $1.0e^{-6}$ ) 내에 속하는지 알기 위해 확인된다. 허용 오차내에 속하는 경우에는 "Bounded"로 표시되며 이 값은 모든 매개변수에 적용된다.

### 4) 누적 분포 함수(cumulative distributive function, CDF) 표

이 블록은 BMDS 3이후 버전의 새로운 기능이다. CDF는 "누적 분포 함수"를 의미하며 BMD 추정치를 나타낸다. 추정되는 BMD에 대한 CDF와 관련된 백분위수를 나열한다.

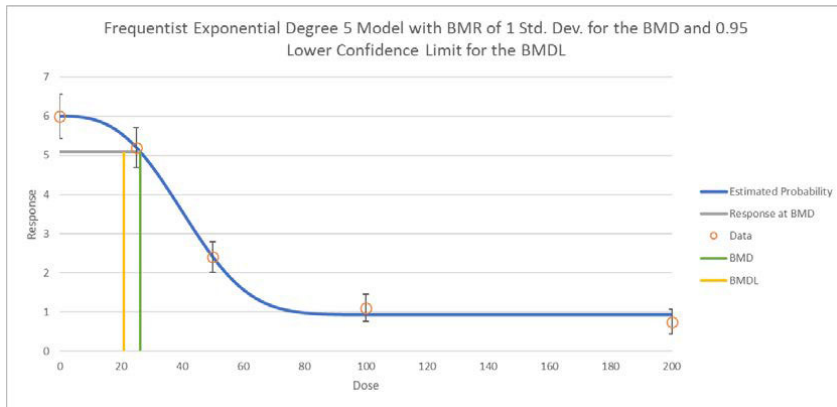
0.5의 CDF 값과 관련된 BMD 값이 BMD의 MLE이다(Benchmark dose 표에서 BMD에 대해 보고된 값과 일치함).

CDF 블록은 BMDL 및 BMDU 값의 관점에서 Benchmark Dose 표와 부합할 수 있다. 옵션에서 사용자가 지정한 신뢰 수준은 단측 신뢰 수준이다. 따라서 해당 신뢰 수준이 CDF 블록의 누적 백분위수 중 하나와 관련된 경우 BMD 값이 일치한다. 예를 들어 사용자가 지정한 신뢰 수준이 0.95(95% 단측 신뢰 한계 요청)인 경우 Benchmark Dose 표의 BMDU는 CDF 블록의 0.95에 대해 나열된 BMD 값과 일치한다. 그리고 BMDL은 CDF 블록의 0.05에 대해 나열된 BMD 값과 일치한다.

## 5) 그래픽 출력

그래픽 출력(plots)은 Summary 탭과 Results Workbook의 특정 Model-Option 탭에 표시된다. Summary 탭에는 모든 모델에 대한 plot이 표시된다. Option Sets는 주어진 분석을 실행한다. 개별 모델 결과는 해당 Model-Option Set에 해당하는 탭에 표시된다.

**Figure 40. Results plot for an individual model.**



- 녹색 및 노란색 수직선은 각각 BMD 및 BMDL이며, 수평의 회색선은 사용자-선택 기준 반응(Benchmark Response, BMR)을 나타낸다.
- 모델에 의해 추정된 용량-반응 곡선은 파란색 선으로 표시된다.
- 그래픽 표시 기능은 Excel 편집 기능을 사용하여 수정할 수 있다.
- 데이터 포인트는 개별 그룹 신뢰 구간(confidence intervals)과 함께 주황색 원으로 표시된다.
- 오차 막대(error bar) 계산은 종말점에 따라 약간 다르다.
  - a. 연속형 종말점의 경우 “6. 연속형 데이터의 종말점-5)빈도주의 연속형 모델에 대한 특이적인 출력-(4)Plot과 error bar 계산”을 참조한다.
  - b. 이분형 종말점의 경우 “7. 이분형 자료의 종말점-3)Simple design에서 이분형 종말점을 위한 모델의 수학적 상세-(3)Plot과 error bar 계산”을 참조한다.

## 6. 연속형 데이터의 종말점(Continuous Endpoints)

연속형 데이터의 종말점은 실수(정수와 반대)의 값을 취하여 연속적으로 변할 수 있는 항목(무게, 농도 등)을 측정한다.

BMD를 추정하기 위해 지정해야 하는 이러한 측정의 세 가지 주요 기능은 다음과 같다.

- A. 독성 반응을 나타내는 변화 방향(adverse direction)
- B. 반응의 변화와 관련하여 BMD를 정의하는 방식
- C. 반응이 배포되는 방식

분포의 관점에서는 분포의 유형과 분포 중심을 중심으로 한 변동성의 특성을 고려해야 한다. 섹션 “6. 연속형 데이터의 종말점-옵션(options)”에서 논의된 사용자가 사용할 수 있는 옵션은 이러한 모든 선택과 관련이 있다.

이 섹션에서는 다음 주제에 대한 세부 정보를 제공한다.

- BMDS 3에서 연속형 모델 구현
- 연속형 모델 데이터 입력
- 연속형 모델 옵션
- 연속형 모델 특이적인 출력
- 특정 모델 매개변수의 값을 제한하는 옵션
- 연속형 반응 모델링에 대한 베이지안 접근 방식, 사전 정의 방법 지정

## 1) 연속형 반응모델

이전 버전의 BMDS에서 연속형 응답 데이터를 분석하는 데 사용할 수 있었던 모든 기존 빈도 모델 및 옵션을 BMDS 3에서 사용할 수 있다.

**Figure 41.** Default selection of BMDS 3 continuous models, as they appear in the Analysis Workbook. Note that model averaging is disabled for continuous models in BMDS 3.2.

Model Name	MLE		Alternatives		Model Averaging Variance Models			
	Frequentist Restricted	Frequentist Unrestricted	Bayesian (Beta)	Bayesian Model Average (Beta)	Prior Weights	Normal Constant Variance	Normal Nonconstant Variance	Lognormal Constant Variance
Exponential	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	0.0000%	0.0000%	0.0000%	0.0000%
Hill	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	0.0000%	0.0000%	0.0000%	0.0000%
Linear	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	0.0000%	0.0000%	0.0000%	0.0000%
Polynomial	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	0.0000%	0.0000%	0.0000%	0.0000%
Power	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	0.0000%	0.0000%	0.0000%	0.0000%
Total Weight					0.0000%	0.0000%	0.0000%	0.0000%

또한 사용자는 모든 연속형 모델에 대해 하이브리드 연속형 모델링 방법 (Hybrid continuous modeling method)과 지수정규 반응 분포 가정 (lognormal response distribution assumption, 이전에는 지수 모델 (Exponential model)에서만 사용 가능)을 사용할 수 있다.

이전 버전의 BMDS에서와 같이 사용자는 Hill, Polynomial 및 Power 모델을 제한 또는 제한 없이 실행하도록 선택할 수 있다. 선형 모델(Linear model)은 제한되지 않으며 지수 모델은 제한적으로만 실행할 수 있다.

BMDS 3.2의 새로운 기능은 연속형 종말점 모델링의 베이지안 구현의 *preview* 버전입니다. 베이지안 연속형 모델에 대한 피어리뷰는 2020년 후반으로 계획되어 있다.

[참고] 현재 EPA는 베이지안 모델링에 대한 기술 지침을 제공하지 않는다.

[참고] Preview(미리보기) 모델은 새로운 것이다. 광범위한 테스트를 거치지 않았으며 위험성 평가 목적으로는 EPA에서 공식적으로 검토 및 승인하지 않았다.



## 2) 연속형 반응데이터 입력

데이터세트 삽입 또는 가져오기에 대한 자세한 내용은 “3.분석 실행 및 분석 세부조정-2)데이터세트 추가”를 참조한다.

요약된 연속형 반응 데이터의 경우 열 머리글에는 "Dose", "N", "Mean" 및 "Std.Dev"가 기본값으로 들어가 있다.”

개별 연속형 반응 데이터의 열 머리글의 기본값은 "Dose"와 "Response"로 설정되어 있다.

### (1) adverse direction

**Figure 42.** Adverse Direction picklist for the selected dataset.

<input checked="" type="checkbox"/> Enable	DataSets	Adverse Direction
<input checked="" type="checkbox"/>	Continuous ds1	automatic
<input checked="" type="checkbox"/>	Continuous ds2	automatic
		up
		down

Adverse Direction 옵션에 대한 선택 사항은 "automatic"(기본값), "up", "down"이다.

이 옵션은 용량-반응 곡선이 위로 상승(up)하거나 하강(down)함에 따른 부작용의 증가 여부를 나타낸다. 부작용의 방향이 알려진 경우 부작용의 방향을 수동으로 선택한다.

"automatic"이 선택되면 BMDS는 관찰된 용량-반응 관계의 형태를 기반으로 부작용의 방향을 선택한다.

이 선택은 사용자 지정 BMR이 BMD를 얻기 위해 모델 결과와 함께 사용되는 방식에만 영향을 미친다.

### 3) 옵션(options)

BMDs 3 Analysis Workbook의 Main 탭에서 사용자는 하나의 "batch(일괄 처리)" 프로세스에서 여러 사용자가 선택한 모델과 여러 사용자가 선택한 데이터 세트에 적용할 여러 옵션 세트를 정의할 수 있다. Add Option Set 버튼을 선택하여 새 옵션 집합 구성을 정의한다.

Figure 43. Continuous Model options.

Option Set #	BMR Type	BMRF	Tail Probability	Confidence Level	Polynomial Restriction	Background	Applies Only to Individual Models	
							Distribution	Variance
1	Std. Dev.	1	0.01	0.95	Use dataset adverse direction		Normal	Constant
Add Option Set								

#### (1) BMD정의(defining)

Figure 44. BMD defining

Option Set #	BMR Type	BMRF	Tail Probability	Confidence Level
1	Std. Dev.	1	0.01	0.95
	Std. Dev.			
	Rel. Dev.			
	Abs. Dev.			
	Point			
	Hybrid - extra risk			
Add Option Set				

다음 옵션은 BMD 및 해당 경계의 정의와 관련된다.

- BMR(Benchmark Response) Type: BMD를 도출하는 데 사용되는 반응 수준을 결정하기 위한 선택 방법을 정의한다. 자세한 내용은 표 5를 참조한다.
- BMRF(Benchmark Response factor)는 BMR 유형에 따라 다르다. 표 5에 BMR 유형 및 BMRF와 관련된 옵션을 요약하여 나타내었다.
- Tail Probability는 부작용을 정의하기 위한 컷오프(cut-off)를 표시하며 "Hybrid extra risk(하이브리드 추가 위험)" BMR Type에만 적용된다. 예를 들어 기본값으로 0.01을 설정한다면 이는 노출이 없을 때 부작용으로 여겨지는 반응이 나타날 확률이 0.01이라고 지정했음을 나타낸다. 이것은 반응

분포의 꼬리 부분(상단 또는 하단)이 얼마나 부작용의 범위에 있는지 지정한다는 의미에서 "tail probability(꼬리 확률)"이다. 그것은 암묵적으로 정상 반응과 부작용 사이의 컷오프를 정의한다.

- 신뢰 수준은 기본적으로 0.95로 설정된다. 이 신뢰 수준은 어느 방향에서든 단측 신뢰 한계에 해당한다. 즉, 신뢰 수준이 0.95로 설정된 경우 BMDL은 BMD의 단측 95% 하한이다. BMDU는 BMD의 단측 95% 상한이다. 이 경우 BMDL에서 BMDU까지의 구간은 90% 신뢰 구간이 된다.

**Table 5. Options related to Continuous BMR Type and BMRF.**

Analysis File, Main Tab Option Name	Verbal Definition: The BMD is the dose yielding ...	Mathematical Definition	BMRF Notes
Rel. Dev.	<i>Relative Deviation:</i> ... the specified change in median response relative to the background median	$\frac{ m(BMD) - m(0) }{m(0)} = BMRF$	BMRF is the "specified change" Default value = 0.1 [10% change in median]
Abs. Dev.	<i>Absolute Deviation:</i> ... the specified change in median response	$ m(BMD) - m(0)  = BMRF$	BMRF is the "specified change" There is no default because it is very endpoint specific
Point	<i>Fixed Value:</i> ... a median equal to the specified point value	$m(BMD) = BMRF$	BMRF is the "specified point value" There is no default because it is very endpoint specific
Std. Dev.	<i>Standard Deviation:</i> ... the specified change in median relative to the control standard deviation	Normal Responses: $\frac{ m(BMD) - m(0) }{\sigma(0)} = BMRF$ Lognormal Responses: $\frac{ \ln(m(BMD)) - \ln(m(0)) }{\sigma_L(0)} = BMRF$	BMRF is the multiple of the standard deviation Default value = 1 [change in median (or log-median) equal to 1 standard deviation (or log-scale standard deviation)]
Hybrid	<i>Increased Extra Risk:</i> ... the specified extra risk, defined by the estimated distribution and background rate	If high responses are adverse: $BMRF = \frac{\Pr(X > X_0 BMD) - \Pr(X > X_0 0)}{1 - \Pr(X > X_0 0)}$ If low responses are adverse: $BMRF = \frac{\Pr(X < X_0 BMD) - \Pr(X < X_0 0)}{1 - \Pr(X < X_0 0)}$ where $X_0$ is a response value and $\Pr(X < X_0 d)$ is the probability that the response, $X$ , is less than $X_0$ at dose $d$ . For $d=0$ , the latter equals the user-specified "tail probability" and $X_0$ is then a function of that tail probability and the estimated control-group response-distribution.	BMRF is the extra risk. (Default 0.5) This option also requires specifying a "tail probability" which is the probability of extreme ("adverse") response at dose=0.

$m(x)$ 는 용량  $x$ 의 중앙값이다. 특히,  $m(BMD)$ 는 용량=BMD에서 중앙값이므로 BMD는 제시된 방정식의 해이다.  $\sigma(0)$ 은 대조군( $d=0$ )에 대한 표준편차이고  $\sigma_L(0)$ 은 대조군에 대한 로그-척도 표준편차로, 반응이 로그정규 분포로 가정될 때 사용된다.

## (2) 다항식 제한(polynomial restriction)

Figure 45. polynomial Restriction

ice	<b>Polynomial Restriction</b>	
	Use dataset adverse direction	▼
	Use dataset adverse direction	
	Non-negative	
	Non-positive	

투여 기간의 계수(coefficients)에 대한 제한은 "Use dataset adverse direction" (기본값, 감지되거나 지정된 부작용 방향에 따라 다름), "Non-negative"(>0), 또는 "Non-positive"(<0)중에 선택할 수 있다.

## (3) 분포(distribution) 및 분산(variance)

Figure 46. Distribution and Variance

ound	<b>Applies Only to Individual Models</b>	
	<b>Distribution</b>	<b>Variance</b>
	Normal ▼	Constant ▼
	Normal	
	Log-normal	

<b>Applies Only to Individual Models</b>	
<b>Distribution</b>	<b>Variance</b>
Normal ▼	Constant ▼
	Constant
	Non-Constant

기본 데이터 분포와 해당 분포의 "중심" 주변 변동성은 연결된 옵션이다. 총 세 가지 조합이 허용된다.

- 정규 분포, 등분산(기본값): 각 용량 그룹은 용량-반응 모델 매개변수와 함께 BMDS에 의해 추정되는 동일한 분산을 가진다.
- 정규 분포, 비등분산(모델링된): 각 용량 그룹은 용량 그룹의 추정 평균 값(아래 참조) 분산. 이 두 매개변수는 용량-반응 모델의 매개변수와 동시에 추정된다
- 로그 정규 분포, 일정한 변동 계수(CV, coefficient of variation): 로그 정규 분포 반응의 경우 각 용량 그룹은 동일한 CV를 가지므로 로그

척도 분산(log-scale variance)이 용량 그룹 전반에 대해 일정하다 (자natural-scale 분산은 그룹마다 다를 수 있음).

반응 분포(정규 또는 로그정규)와 관련하여 다음 사항에 유의한다.

- 로그정규 분포는 응답이 완전히 양수인 경우에만 가정할 수 있으며 양의 실수 값에만 적용할 수 있다.
- 가정된 분포에 관계없이 고려 중인 용량-반응 모델은 용량의 함수로서 반응 분포의 중앙값 변화를 나타낸다. 용량  $d$ 에서 중앙값을  $m(d)$ 로 표시하면 BMDS에 대해  $m(d) = f(d)$ 는 항상 참이다. 여기서  $f(d)$ 는 고려 중인 용량-반응 함수이다(Table 6의 list of possible function을 참고한다).
- 가정된 데이터 분포가 정규분포이면 용량  $d$ 에서의 평균  $\mu(d)$ 가 중앙값과 같다. 따라서 정규성 가정에서 용량-반응 함수를 평균 반응의 모델로 설명하고  $\mu(d) = f(d)$ 로 작성하는 것이 일반적이다. 여기서  $f(d)$ 는 Table 6에 기술된 용량-반응 기능 중 하나이다.

연속형 반응 데이터를 모델링할 때 BMDS 연속형 모델에 대한 표준 가정은 용량-반응 모델에 의해 주어진 평균과 사용자가 지정한 분산(상수 또는 평균 반응의 함수)이 있는 정규 분포(각 용량 그룹에 대해 하나씩)라는 점이다. 다른 가정은 반응이 로그정규 분포를 따른다는 것이다.

BMDS 3에서 모든 연속형 모델은 사용자가 정규 및 로그정규 반응 분포 가정 중에서 선택할 수 있다. BMDS의 이전 버전은 지수 모델(Exponential model)에 대해서만 이 선택을 허용했다.

사용자가 개별 반응 데이터에 접근한다면 해당 데이터는 분석 전에 로그 변환할 수 있지만 아래에서 설명하는 것처럼 권장되는 접근 방식은 아니다. 반응이 로그정규 분포를 따르는 것으로 의심되는 경우 권장되는 접근 방식은 기본 분포가 용량-반응 함수에 의해 정의된 중앙값과 CV상수의 추정에 따라 일정한 로그 척도 분산을 갖는 로그 정규라고 가정하고 변환되지 않은 데이터를 모델링하는 것이다.

## A. 정확하고 대략적인 MLE 솔루션

데이터를 로그정규 분포로 가정하고 데이터를 그룹별 평균 및 표준 편차로 표시하면 정확한 MLE 솔루션을 얻을 수 없다. 이 경우 결과는 "대략적인" MLE 솔루션이다. 로그 변환된 데이터의 평균( $m_L$ ) 및 표준 편차( $s_L$ )는 다음과 같이 추정한다.

$$\text{estimated log-scale sample mean (} m_L \text{):} \quad m_L = \ln(m) - \frac{s^2}{2}$$

$$\text{estimated log-scale sample standard deviation (} s_L \text{):} \quad s_L = \sqrt{\ln\left[1 + \frac{s^2}{m^2}\right]}$$

여기서  $m$ 과  $s$ 는 각각 표본 평균과 표본 표준 편차이다.

데이터가 로그 정규 분포를 따른다고 가정하고 개별 응답 데이터를 사용할 수 있는 경우 BMDS 3은 정확한 최대 우도 추정(MLE) 솔루션을 제공한다. BMDS 3.0부터는 정확한 솔루션은 개별 관측값이 입력되고 로그정규 가정이 선택될 때 구현되는 유일한 솔루션 옵션이다. BMDS 3은 근사 솔루션을 계산하기 위한 옵션을 제공하지 않는다.

사용자가 개별 관찰에서 대략적인 솔루션을 계산하려는 경우(예: 연구 목적) 다음 절차를 사용해야 한다.

1. 그룹별 표본 평균과 표본 표준 편차를 계산한다.
2. 해당 요약 통계를 기반으로 하는 분석에 대해 수행할 값을 입력한다 (그러나 여전히 분포 유형으로 로그정규를 선택함).

## B. 로그 변환된 response는 권장되지 않음

분석에 로그 변환 반응을 사용하는 것은 다음과 같은 이유로 권장되지 않는다.

- 사용자가 분석에 앞서 데이터를 로그 변환하기로 선택한 경우 BMD 및 BMDL 추정치의 해석은 신중하게 고려되어야 한다(아마도 통계 전문가와 상의하여). 로그 변환 반응을 사용할 때 데이터 해석은 자연 척도 반응 값을

사용할 때와 동일하지 않다. 실제로 모델은 자연 규모로 "다시 변환"되었을 때 표준 BMDS 모델과 일치하지 않는다.

예를 들어, 만약 로그 변환된 반응에서 파워 모델(power model)을 사용하는 경우, 사용자는 실제로 표준 BMD가 아닌 함수  $e^{(\text{background} + \text{slope} \times \text{dose})^{\text{power}}}$ 로 중앙값(natural-scale)을 암시적으로 모델링하고 있다. 모델의 특성(예: 반응의 기하급수적 증가)이 사용자가 원하는 것이 아닐 수 있다.

- 마찬가지로, BMD의 해석은 간단하게 표현되지 않는다(예: BMR이 10%의 상대 편차와 동일하게 설정된 경우 해당 상대 편차는 로그 척도로 평가되므로 BMD 또는 BMDL을 산출하지 않는다. 원래 평균 응답의 10% 변화에 해당하는 추정치).

이러한 이유로 반응 값을 로그 변환하는 것은 "best practice"로 간주되지 않으며 명시된 바와 같이 통계적 전문 지식을 뒷받침하는 경우에만 적용 및 해석해야 한다.

따라서 대부분의 경우 사용자는 변환되지 않은 값을 사용하고 데이터가 로그정규 분포를 따른다고 가정할 경우 로그정규 분포를 선택해야 한다.

#### 4) Simple design에서 연속적인 종말점을 위한 모델의 수학적 상세

이 섹션의 모델은 등지 또는 기타 복잡한 관계를 포함하지 않는 단순한 실험에서 체중 또는 효소 활성 측정과 같은 연속형 종말점과 관련이 있다. 모델은 주어진 용량에 대해 예상되는 반응의 중앙값  $m(\text{dose})$ 과 해당 중앙값 주변의 변동을 예측한다.

연속형 모델에 사용할 수 있는 옵션에 대한 이전 논의에서 알 수 있듯이 연속형 종말점의 모델링은 유사한 디자인의 이분형 종말점에 대한 모델링보다 더 많은 세부 사항을 고려해야 한다. 이 섹션에서는 BMDS에서 추정을 수행하는 방법을 결정하는 수학적 및 통계적 세부 정보를 제공한다.

## (1) 연속적인 용량-반응 모델 기능

연속형 모델의 정의는 다음 표에 완전히 지정되어 있다.  $m(\text{dose})$ 은 지정된 용량 수준에 대한 반응 중앙값이다.

**Table 6.** *The individual continuous models and their respective parameters.*

Model	Parameters	Notes
Linear and Polynomial models $m(\text{dose}) = g + \beta_1 \times \text{dose} + \beta_2 \times \text{dose}^2 + \dots + \beta_n \times \text{dose}^n$ $n$ is the degree of the polynomial, specified by user and must be a positive integer (maximum value = 21)	$g$ = control response (intercept) $\beta_0 \dots \beta_n$ : polynomial coefficients	Parameter Constraints: none User parameter restriction options: can restrict the value of the polynomial coefficients. Restricting them to be either "non-positive" or "non-negative" guarantees that the resulting function will be strictly decreasing, strictly increasing, or perfectly flat (when all the coefficients are zero). If the coefficients are unrestricted (i.e., an unrestricted form of the model is run), more complicated shapes are possible, and, particularly as the degree of the polynomial approaches the number of dose groups minus one, the polynomial will often be quite "wavy".
Linear $m(\text{dose}) = g + \beta \times \text{dose}$	$g$ = control (intercept) $\beta$ = slope	Parameter Constraints: none User parameter restriction options: none
Power $m(\text{dose}) = g + \beta \times (\text{dose})^\delta$	$g$ = control response (intercept) $\beta$ = slope $\delta$ = power	Parameter Constraints: $0 \leq \delta \leq 18$ . User parameter restriction options: $\delta$ may be further restricted to values $\geq 1$ . Note: If $\delta < 1$ , then the slope of the dose-response curve becomes infinite at the control dose. This is biologically unrealistic and can lead to numerical problems when computing confidence limits, so several authors have recommended restricting $\delta \geq 1$ .
Hill <sup>1</sup> $m(\text{dose}) = g + \frac{v \times \text{dose}^n}{k^n + \text{dose}^n}$	$g$ = control response (intercept) $k$ = dose with half-maximal change $n$ = power $v$ = maximum change	Parameter Constraints: $k > 0$ , $0 \leq n \leq 18$ . User parameter restriction options: $n$ may be further restricted to values $\geq 1$ .
Exponential <sup>1,2</sup> $\text{Exp2: } m(\text{dose}) = a \times e^{\pm b \times \text{dose}}$ $\text{Exp3: } m(\text{dose}) = a \times e^{\pm (b \times \text{dose})^d}$ $\text{Exp4: } m(\text{dose}) = a \times (c - (c - 1) \times e^{-b \times \text{dose}})$ $\text{Exp5: } m(\text{dose}) = a \times (c - (c - 1) \times e^{-(b \times \text{dose})^d})$	$a$ = control response (intercept) $b$ = slope $c$ = asymptote term $d$ = power	Parameter Constraints: $a > 0$ , $b > 0$ , $c > 1$ for responses increasing with dose, $0 < c < 1$ for responses decreasing with dose, $1 \leq n \leq 18$ . Note: The sign in " $\pm b$ " (Exp2 and Exp3 models) will change depending on the user-designated or auto-detected direction of change: + for responses increasing with dose, - for responses decreasing with dose.

광범위한 모수 값이 거의 동일한 가능성을 제공할 수 있는 경우 점근선 모수(asymptote parameter, Hill 모델 포함)가 있는 모델의 BMDL 추정값은 불안정할 수 있다. 그 지표의 문제 중 하나는 추정된 점근적 반응이 관찰된 반응의 범위를 훨씬 벗어난다는 것이다. 이러한 경향이 보이거나 의심되는 경우 사용자는 통계 전문가에게 문의해야 한다.

Power, Hill 및 Exponential 모델의 검정력 매개변수에 대한 상한은 18로 설정되었다. 이 값은 거의 모든 데이터세트 즉, 저용량에서 편평한 용량-반응에 이어 고용량에서는 매우 가파른 용량-반응을 보이는 데이터세트조차 수용해야 하는 매우 높은 곡률(curvature)을 나타내기 때문에 선택되었다.



## (2) 분산 모델

용량의 함수로서 중앙값 반응에 대한 모델 외에도 분산에 대한 모델도 정의해야 한다.

다양하다고 가정되는 반응의 경우 일반적으로 중위수 주변에서 분산 모델은 다음과 같다.

$$\sigma_i^2 = \exp\{\ln(\alpha) + \rho * \ln[m(\text{dose}_i)]\},$$

여기서  $\alpha > 0$  및  $\rho$ 는 용량-반응 함수의 매개변수와 동시에 추정된 매개변수이다 (표6 참조). 표6에서와 같이  $m(\text{dose}_i)$ 는  $i$ 번째 용량 그룹에 대한 예측 중앙값 (고려 중인 용량-반응 모델에서)이다.

사용자가 일정한 분산 모델을 지정하면 매개변수  $\rho$ 가 0으로 설정되고  $\alpha$ 만 추정된다. 그 경우,

$$\sigma_i^2 = \alpha$$

반응이 로그정규 분포라고 가정하면 모델링된 분산은 로그척도 분산이다.

$$\sigma_i^2 = \alpha$$

로그정규 데이터의 경우 BMD는 상수 로그-스케일 분산 모델(변동계수 상수와 동일)로 제한되기 때문에  $\rho$ 는 해당 방정식에 나타나지 않는다(본질적으로 로그정규 분포 반응하에 다시 0으로 설정).

위에 표시된 분산 모델의 공식화(formulation)는 일반적으로 발생하는 몇 가지 상황을 허용한다.  $\rho = 1$ 이면 분산은 중앙값에 비례한다.  $\rho = 2$ 이면 변동계수는 일정하며, 특히 생화학적 측정에 대한 일반적인 가정이며 로그정규 분포 반응의 변동계수 상수 가정을 모방한다(그러나 반응이 실제로 로그정규 분포라고 가정할 필요는 없음).

### (3) 우도 함수(likelihood function)

빈도주의 모델링 옵션의 경우 매개변수 추정값은 가능성을 최대화하는 방법으로 파생된다(즉, 최대 우도 추정값, MLE). 연속형 반응에 대한 우도 함수는 여기에서 정의된다.

용량이 있는 G 용량 그룹과

$$\text{dose}_1, \dots, \text{dose}_G$$

용량 그룹당  $N_i$  의 피험자도 있다고 가정한다. 또한  $y_{ij}$ 는  $i$ 번째 용량 그룹의  $j$ 번째 대상에 대한 측정이라고 가정한다. 로그 우도 함수의 형식은 반응이 정규 분포 또는 로그정규 분포로 가정되는지 여부에 따라 다르다.

#### A. 정규분포 반응 가정

For the assumption of Normally distributed responses, the log-likelihood function is:

$$LL = -\frac{N}{2} \ln(2\pi) - \sum_{i=1}^G \left[ \frac{N_i}{2} \ln(\sigma_i^2) + \frac{(N_i - 1)s_i^2}{2\sigma_i^2} + \frac{N_i(\bar{y}_i - m(\text{dose}_i))^2}{2\sigma_i^2} \right]$$

where

$$\bar{y}_i = \frac{\sum_{j=1}^{N_i} y_{ij}}{N_i} \text{ (the sample mean for the } i\text{th dose group),}$$

$$s_i^2 = \frac{\sum_{j=1}^{N_i} (y_{ij} - \bar{y}_i)^2}{N_i - 1} \text{ (the sample variance for the } i\text{th dose group),}$$

$$N = \sum_{i=1}^G N_i.$$

The parameters defining  $m(\text{dose}_i)$  and  $\sigma_i^2$  (see previous two subsections) are optimized to maximize the LL equation value.

#### B. 로그정규분포 반응 가정

로그 정규 분포 반응을 가정할 경우 로그 우도 함수는 다음과 같다.

$$LL = -\frac{N}{2} \ln(2\pi) - \sum_{i=1}^G \left[ N_i \bar{z}_{Li} + \frac{N_i}{2} \ln(\sigma_{Li}^2) + \frac{(N_i - 1)s_{Li}^2}{2\sigma_{Li}^2} + \frac{N_i(\bar{z}_{Li} - \ln(m(\text{dose}_i)))^2}{2\sigma_{Li}^2} \right]$$

여기서,

$\bar{z}_{Li}$  = log-scale sample mean for ith dose group, and  
 $s_{Li}^2$  = log-scale sample variance for ith dose group.

정규 분포 응답의 경우와 같이  $m(\text{dose}_i)$  및  $\sigma_{Li}^2$ 를 정의하는 매개 변수는 LL 방정식 값을 최대화하는데 최적화된다.

#### (4) AIC와 모델 비교

Akaike Information Criterion(AIC)는 동일한 데이터 세트에 대한 서로 다른 모델 적합도를 비교(동일한 피팅방법 사용, 우도를 최대화)하는 데 사용할 수 있다. AIC는 LL 값(이전 섹션 참조)과 추정된 매개변수  $p$ 의 수에 따라 달라지는 통계이다.

$$AIC = -2 \times LL + 2 \times p$$

AIC는 높은 LL 값을 달성하는 데 필요한 매개변수의 수와 관련하여 적은 수의 데이터로(parsimonious) 가능한 가장 높은 LL 값을 얻는 목표의 균형을 유지한다. AIC에 대한 방정식은 LL에 대한 음의 승수(더 크게 하고 싶은 것)와  $p$ 에 대한 양의 승수(가능한 한 작아서 여전히 "좋은 적합성"을 얻고자 함)를 가지므로 더 작은 값을 가진 모델은 다음과 같다. AIC를 기반으로 하여 다른 모델보다 AIC가 더 나은 모델로 추정된다. 이러한 방법은 정확하지 않지만 모델 선택에 유용한 지침을 제공할 수 있다.

현재 버전의 BMD5에서 "추정된 매개변수"의 수에는 경계값(bounding value)과 동일하지 않은 것으로 추정된 매개변수만 포함된다(모델 부과 제약 또는 사용자 부과 제약(Table 6)).

[참고] 이 계산 프로세스는 해당 매개변수가 적중하는 경계 값에 따라 합리적일 수도 있고 적절하지 않을 수도 있다.

예를 들어, 모델의 파워 매개변수(power parameter)가 18의 상한(upper bound)에 도달하면(즉, 동일한 것으로 추정됨) 일반적으로 해당 매개변수를 추정된 것으로 계산하고 싶지만 BMD는 그렇지 않다. 그렇게 하지 마십시오!

이러한 이유로 매개변수 경계(parameter bound)에 도달한 경우를 주의 깊게 고려하고 모델 비교 및 모델 선택과 같은 문제에 대한 영향을 고려하도록 사용자에게 알려준다.

[참고] 서로 다른 우도(즉, 정규 vs. 로그정규)가 있는 모델을 쉽게 비교할 수 있도록 LL 방정식에 표시된 모든 항을 사용하여 로그 우도를 계산한다. BMD 2.x는 보고된 LL 값이나 AIC 계산에 매개변수 독립적인 용어를 포함하지 않았다. 결과적으로 BMD 3.x 및 2.x는 동일한 데이터에서 실행될 때 LL 및 AIC에 대해 다른 값을 반환한다.

매개변수가 모델에 대한 경계(bound)에 도달하는 경우 매개변수 추정값은 제한된 의미의 최대 우도 추정값이다. 경계 매개변수에는 값이 할당되고 다른 매개변수는 할당된 값에 대한 조건부 MLE이다. 이러한 모델 결과는 AIC 측면에서 다른 모델과 엄격하게 비교할 수 없다. 이러한 경우 BMD 및 BMDL은 파워 매개변수(power parameter)의 선택에 따라 달라질 수 있다. 따라서 보고된 BMD 또는 BMDL에 의존하려는 경우 민감도 분석이 표시된다. 이것은 18의 상한에 도달한 파워 매개변수를 고려할 때 특히 중요하다.

## [주의]

표본 평균과 표본 표준 편차만 사용할 수 있고(요약된 데이터) 로그 척도 표본 평균과 표본 표준 편차는 로그 정규 분포 반응을 가정할 때만 근사화되는 상황에는 주의가 필요하다. 이러한 경우 모든 용량-반응 모델에 대해 동일한 근사치가 만들어진다. 따라서 반응이 로그 정규 분포를 따른다고 가정한 모든 실행에서 AIC 결과를 비교하는 것은 엄격히 유효하다. 그러나 정규성을 가정하여 한 세트의 결과를 얻은 결과와 로그 정규성을 가정한 한 세트를 얻은 결과의 비교는 주의해야 한다. 후자의 경우 AIC가 "유사한(similar)" 경우(여기에서 특정 지침을 제공할 수 없기 때문에 해당 용어를 느슨하게 사용) 사용자는 AIC 차이를 기반으로 모델 선택을 해서는 안 된다. AIC 차이가 "더 큰(larger)" 경우 선택하는 데 문제가 없을 수 있다. 사용된 근사값이 너무 나쁘지 않아야 하기 때문이다. 보수적인 입장은 기본 데이터가 요약된 형식으로 표시되는 경우(즉, 표본 평균과 표본 표준 편차만 사용 가능), AIC를 사용하여 정규 분포 반응을 가정하는 모델 실행과 로그 정규 반응을 가정하는 모델 실행을 비교해서는 안 된다는 것이다.

연속형 모델에 대한 BMDS 3 AIC 값이 BMDS 2.x 버전의 값과 다르더라도 모델의 기본 분포가 동일한 경우 AIC의 차이는 이전 버전의 BMDS와 동일하다. 이것은 BMDS 3.x 및 BMDS 2.x 모델 적합성이 비교되는 두 모델에 대해 동일하다고 가정한다. AIC 차이는 두 버전 간에 하나 이상의 모델 적합성이 다른 경우 동일하지 않을 수 있다(예: 하나 이상의 BMDS 3 모델이 해당 BMDS 2.x 모델보다 데이터에 대한 개선된 적합성을 제공하는 경우).

그러나 모수 분포가 다른 모델을 비교할 때 AIC 차이는 이전 BMDS 버전과 동일하지 않다. 이러한 비교의 경우 BMDS 3 소프트웨어를 사용하여 계산된 AIC가 정확하며 기본 분포에 관계없이 두 모델 간에 적절한 비교가 이루어진다.

### (5) BMDL 및 BMDU 계산

BMR 유형의 정의에 따른 BMD의 추정치는 표 5에 지정되어 있다. BMD, 즉 BMDL 및 BMDU에 대한 신뢰 범위의 유도가 여기에서 정의하였다.

BMD에 대한 신뢰 한계를 계산하는 일반적인 접근 방식은 BMDS의 모든 모델에 대해 동일하며 우도 비율의 점근적 분포(asymptotic distribution)를 기반으로 한다.

연속형 모델에 대해서는 두 가지 다른 특정 접근 방식을 따른다.

#### *파워 모델의 경우(For the Power Model)*

파워모델의 경우 기준용량 및 용량-반응 모델의 관점에서 기준 반응을 정의하는 방정식은 모델 매개변수 중 하나를 해결한다. 결과 식은 BMD가 매개변수로 명시적으로 나타나도록 모델을 다시 매개변수화하는 효과와 함께 모델 방정식으로 다시 대체된다. 그런 다음 BMD에 대한 값은 나머지 매개변수가 해당 BMD 값에 대한 조건부 우도를 최대화하도록 변경될 때 결과 로그 우도는 최대 우도 추정치보다 정확히 다음과 같이 작다.

$$\frac{\chi^2_{1,1-2\alpha}}{2}$$

#### *다항식, 힐 및 지수 모델의 경우*

#### *(For the Polynomial, Hill, and Exponential Model)*

다항식, Hill 및 지수 모델의 경우 BMD가 명시적 매개변수로 나타날 수 있도록 용량-반응 모델 함수를 명시적으로 다시 매개변수화하는 것은 비실용적 이거나 불가능하다. 이러한 모델의 경우 BMR 방정식이 비선형 제약 조건으로 사용되며 BMD의 최소값은 최대 우도 추정치에서 대수 우도와 동일하도록 결정된다.

$$\frac{\chi^2_{1,1-2\alpha}}{2}$$

때때로 모델에 대해 "BMDL computation is at best imprecise for these data(BMDL 계산은 이 데이터에 대해 정확하지 않습니다)"라는 오류 메시지가 나타날 수 있다. 이것은 필요한 수렴 수준(옵티마이저가 종료될 때까지 대상 함수의  $< 1e-3$  상대 변경)이 달성되지 않았다는 의미에서 BMDL에 대한 수렴이 "성공(successful)"하지 못했음을 나타낸다.

#### (6) Bayesian 연속모델(미리보기) 설명

[참고] 현재 EPA는 베이지안 모델링에 대한 기술 지침을 제공하지 않는다.

[참고] 미리보기 모델은 새로운 것이다. 이 기능의 사용자는 미리보기 모델이 위험 평가 목적으로 광범위한 테스트를 거치지 않았으며 EPA에서 공식적으로 검토 및 승인하지 않았음을 알려야 한다.

베이지안 관점에서 추론(inference)은 모델  $M$  및 해당 매개변수가 있는 경우 데이터 생성 메커니즘을 정의하여 진행된다.  $M$  은 반응의 확률을 결정하는 표 6에 나열된 모델 중 하나이다. 데이터 생성 메커니즘은 관찰(observation)이 해당 모델 중 하나에 의해 정의된 용량 의존 중앙값 반응과 위에서 설명한 대로 정의된 분산 항을 갖는 정규 또는 로그 정규 분포에서 얻어진 것으로 가정한다(사용자의 옵션 선택에 따라).

그런 다음 여기에서  $\mathcal{L}(D|M)$ 로 표시된 우도와 관련시킬 수 있다. 이는 모델에 대한 조건부 데이터  $D$ 의 우도임을 명시적으로 보여준다. 우도 로그의 기능적 양식은 “6. 연속형데이터의 종말점-4)Simple design에서 연속적인 종말점을 위한 모델의 수학적 상세-(3)우도 함수”에 나와 있다.

BMDS 3에서 사용된 미리보기 베이지안 연속 모델 세트는 빈도주의(MLE) 접근 방식에 사용된 모델 세트와 동일하다(Table 6). 다음에서  $\theta$ 는 이러한 모델 중 하나를 정의하는 데 필요한 매개변수의 벡터이다. 예를 들어, 검정력

모델  $\theta = (g, \beta, \delta)$ 의 경우이다. 베이지안 접근 방식에 통합된 추가 고려 사항은  $\theta$ 에 대한 사전 분포의 사양이다. 베이지안 접근 방식은 지정된 사전을 취하고 고려 중인 데이터를 사용하여 업데이트하여  $\theta$ 에 대한 "후방(posterior)" 분포를 얻는다.

BMDs는 BMD에 대한 posterior를 다음과 같이 요약한다. BMD(아마도 "중앙 추정(central estimate)"이라고 부를 수 있음)는 최대 사후(maximum a posteriori, MAP)  $\theta$  추정을 사용하여 얻은 값과 동일하다. MAP은 사후 로그 우도를 최대화하는  $\theta$ 의 값이다. 사후 밀도(posterior density) 자체는 Laplacian 근사를 사용하여 근사된다. 이 근사치는 사용자가 선택한 신뢰 수준에 해당하는 밀도의 백분위수인 BMDL 및 BMDU 값을 추정하는 데에도 사용된다.

BMDs 이분형 모델에 대한 priors는 table 7에 정의되어 있다.

### 중요 사항

매개변수에 대한 사전 설정(priors)은 조정된 용량 및 조정된 반응을 기반으로 한다. BMDs는 이 조정(scaling)을 자동으로 수행한다.

BMDs는 고려 중인 데이터셋에서 최대 용량으로 나누어 용량을 자동으로 조정한다. 즉, 용량이 0에서 1(포함)까지의 범위가 된다. BMDs는 대조군(또는 최저 용량군)의 평균 반응으로 나누어 반응을 자동으로 조정한다.

사용자는 어떤 것도 미리 조정할 필요가 없다. 프로그램에서 반환되는 매개변수 추정치 및 BMD 값은 입력 데이터 파일에 지정된 원래 용량 및 반응의 원래 척도로 다시 조정된다.



**Table 7.** The individual Bayesian continuous models (preview) and their parameter priors.

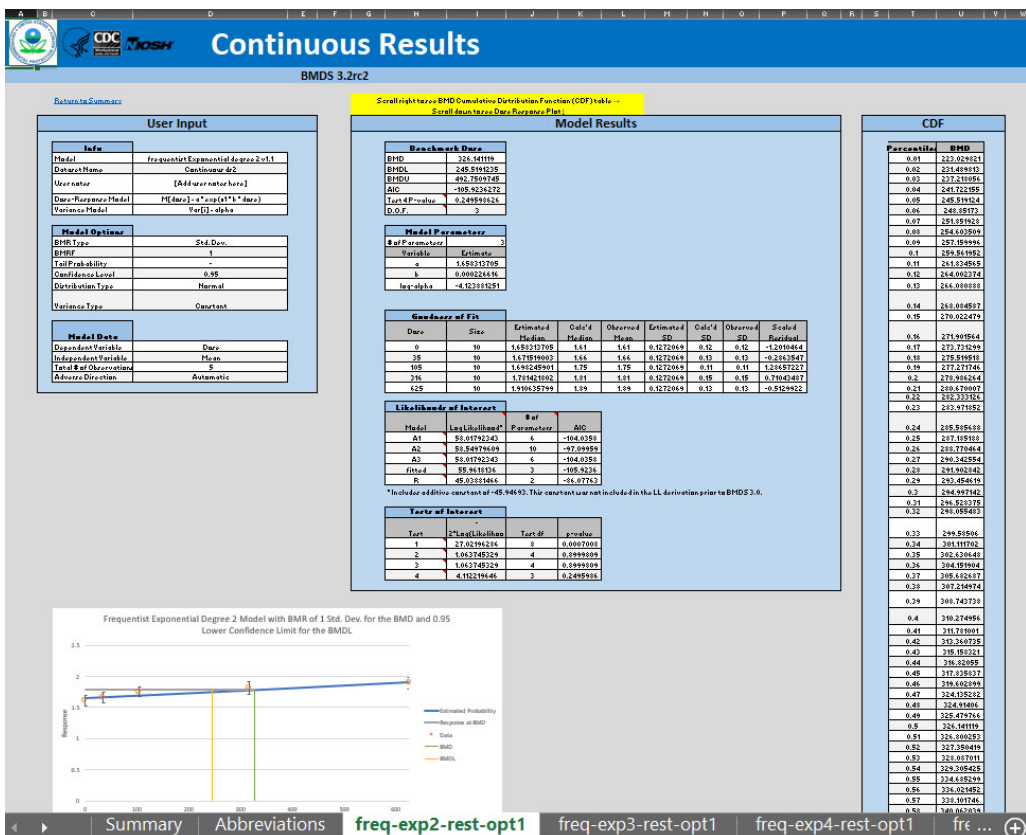
Model	Constraints	Priors
Polynomial $p(\text{dose}) = g + \beta_1 \text{dose} + \beta_2 \text{dose}^2 + \dots + \beta_n \text{dose}^n$	$0 < g < 1e6$ $-1e-4 < \beta_i < 1e4$	$g \sim \text{Lognormal}(0,1)$ $\beta_i \sim \text{Normal}(0,2)$
Linear $p(\text{dose}) = g + \beta_1 \text{dose}$	$0 < g < 1e6$ $-1e-4 < \beta_1 < 1e4$	$g \sim \text{Lognormal}(0,1)$ $\beta_1 \sim \text{Normal}(0,2)$
Power $p(\text{dose}) = g + \beta \times (\text{dose})^\delta$	$0 < g < 1e6$ $-1e-4 < \beta < 1e4$ $0 < \delta < 40$	$g \sim \text{Lognormal}(0,1)$ $\beta \sim \text{Normal}(0,1)$ $\delta \sim \text{Lognormal}(0.405465, 0.5)$
Hill $p(\text{dose}) = g + \frac{v \times \text{dose}^n}{k^n + \text{dose}^n}$	$0 < g < 18$ $-18 < v < 18$ $0 < k < 18$ $0 < n < 18$	$g \sim \text{Lognormal}(0,1)$ $v \sim \text{Normal}(1,2)$ $k \sim \text{Lognormal}(-0.69315, 1)$ $n \sim \text{Lognormal}(0.405465, 0.2501)$
Exponential $p(\text{dose}) = a \times e^{\pm b \times \text{dose}}$ $p(\text{dose}) = a \times e^{\pm (b \times \text{dose})^d}$ $p(\text{dose}) = a \times (c - (c - 1) \times e^{-b \times \text{dose}})$ $p(\text{dose}) = a \times (c - (c - 1) \times e^{-(b \times \text{dose})^d})$	$0 < a < 1e6$ $0 < b < 18$ $-20 < \ln(c) < 20$ $0 < d < 18$	$a \sim \text{Lognormal}(0,1)$ $b \sim \text{Lognormal}(0,1)$ $\ln(c) \sim \text{Normal}(0,1)$ $d \sim \text{Lognormal}(0, 0.2501)$

모든 모델에서  $\rho$  및  $\alpha$ 는 분산모델의 매개변수이다.  $\sigma_i^2 = \exp\{\ln(\alpha) + \rho * \ln[m(\text{dose}_i)]\}$   
 모든 모델에서  $\ln(\alpha)$ 에 대한 prior는  $\text{Normal}(0,1)$ 이고 -18에서 18 사이의 값으로 제한한다. 매개변수  $\rho$ 는 분산이 일정하지 않은 것으로 정의된 경우에만 위에서 지정한 prior가 할당된다. 이 경우  $\ln(\rho)$ 는  $\text{Normal}(0, 0.2501)$ 이고  $\rho$ 는 0과 18사이의 값으로 제한된다.  $\text{Normal}(x, y)$ 는 평균  $x$ , 표준편차  $y$ 를 가지는 정규분포를 나타낸다.  $\text{Lognormal}(w, z)$ 은 로그 척도 평균이  $w$ 이고 로그 척도 표준 편차가  $z$ 인 로그 정규 분포를 의미한다.

[참고] 미리보기 베이지안 연속형 모델은 새로운 기능이다. 사용자는 이 기능이 위험 평가 목적으로 광범위한 테스트를 거치지 않았으며 EPA에서 공식적으로 검토 및 승인하지 않았음을 알려야 한다.

## 5) Frequentist 연속형 모델에 대한 특이적인 출력

Figure 47. Sample Results Workbook tab for a frequentist continuous model run.



## (1) Fit table의 좋은 점

**Figure 48.** Sample Goodness of Fit table, with Normal assumption.

Goodness of Fit		Estimated Median	Calc'd Median	Observed Mean	Estimated SD	Calc'd SD	Observed SD	Scaled Residual
Dose	Size							
0	10	1.658313705	1.61	1.61	0.12720687	0.12	0.12	-1.201046387
35	10	1.671519003	1.66	1.66	0.12720687	0.13	0.13	-0.286354713
105	10	1.698245901	1.75	1.75	0.12720687	0.11	0.11	1.286572266
316	10	1.781421802	1.81	1.81	0.12720687	0.15	0.15	0.710434871
625	10	1.910635799	1.89	1.89	0.12720687	0.13	0.13	-0.512992166

**Figure 49.** Sample Goodness of Fit table, with Lognormal assumption.

Goodness of Fit		Estimated Median	Calc'd Median	Observed Mean	Estimated GSD	Calc'd GSD	Observed SD	Scaled Residual
Dose	Size							
0	10	1.648797808	1.6055465	1.61	1.12433068	1.077271	0.12	-0.109122205
35	10	1.663398431	1.65493293	1.66	1.12433068	1.081332	0.13	-0.009558383
105	10	1.692599677	1.74655307	1.75	1.12433068	1.064809	0.11	0.161443391
316	10	1.780620575	1.80381636	1.81	1.12433068	1.08625	0.15	0.082632182
625	10	1.909523218	1.8855449	1.89	1.12433068	1.071117	0.13	-0.054910745

적합도(Goodness of Fit) 표는 입력된 관찰(또는 "계산") 데이터와 관련된 모델 예측을 각 용량 그룹에 대해 한 행으로 표시한다. 일반적으로 사람들은 가능한 입력 데이터와 모델예측이 일치하기를 바란다. 참고로 이 표에서

- 추정 중앙값(estimated median) = 모형 예측 중앙값(model predicted median, 정규 분포 데이터의 경우 평균과 같음)
- 계산된 중앙값(Calc'd) = 관찰 데이터에서 계산된 중앙값. 정규 분포 데이터의 경우 이는 표본 평균과 같다. 로그 정규 분포 반응의 경우 중앙값은  $\exp(Z_L)$ 로 계산되며, 여기서 Z는 그 척도 평균이며, “(6)연속형 데이터의 종말점-(3) 옵션-다)분포 및 분산”에 표시된 대로 요약된 반응 데이터에서 필요한 경우 추정된다.
- 관찰된 평균 = 표본 평균
- 추정된 [G]SD = 모형에 의해 추정된 표준 편차(또는 로그 정규 분포 데이터의 경우 기하 표준 편차)

- Calc'd [G]SD = 관찰 데이터에서 계산된 표준 편차(또는 로그 정규 분포 데이터의 경우 기하 표준 편차). 정규 분포 데이터의 경우 이는 반응의 표본 표준 편차와 같다. 로그 정규 분포 데이터의 경우 이것은  $\exp(s_L)$ 과 같다. 여기서  $s_L$ 은 로그 척도 표준 편차이며, “(6)연속형 데이터의 종말점-(3)옵션-다)분포 및 분산”에 표시된 대로 요약된 반응 데이터에서 필요한 경우 추정된다.
- 관찰된 SD = 표본 표준 편차
- Scaled Residual = 정규 반응의 경우 다음과 같다.

$$(\text{Calc'd Median} - \text{Estimated Median})/(\text{Estimated SD}/\sqrt{N_i})$$

로그 정규 반응의 경우 척도화된 잔차(residual)는 다음과 같다.

$$(\ln(\text{Calc'd Median}) - \ln(\text{Estimated Median}))/(\ln(\text{Estimated GSD})/\sqrt{N_i})$$

조정된 잔차 값은 관측값과 모델 예측 간의 "보정된(calibrated)" 차이이다. 반응 분포에 대한 가정과 관계없이 정규 분포에 해당하는 척도로 계산된다. 또한, 계산을 위한 분모는 모델 예측에 대한 불확실성(평균의 표준 오차)의 정도를 추정한다. 따라서, 절대값에서 약 2보다 큰 보정된 잔차 값은 적어도 국소적으로 더 낮은 적합성을 나타낼 수 있는 예측된 값과 관찰된 값 사이의 불일치를 나타낸다.

## (2) Interest table의 우도(likelihood)

Likelihoods of Interest 표에는 고려 중인 용량 반응 모델("fitted")을 포함하여 5개 모델에 대한 로그 우도, 매개변수 수 및 AIC가 표시된다. BMD5는 모델 매개변수를 추정하고 궁극적으로 위험성 평가 데이터를 기반으로 추천하기 위해 우도 이론을 사용한다. 최대 우도는 모델 매개변수를 추정하는 프로세스이다. 우도 함수는 고려 중인 모델의 형태와 데이터가 주어지면 가능한 만큼 커진다(최대화). 즉, 매개변수 값은 주제 모델(예: 다항식 또는 거듭제곱)이 모델의 매개변수 구조의 제약 조건이 주어지면 데이터에 가장 적합한 적합성을 얻도록 "선택"된다.

Figure 50. Likelihoods of Interest

Likelihoods of Interest			
Model	Log Likelihood*	# of Parameters	AIC
A1	58.01792343	6	-104.03585
A2	58.54979609	10	-97.099592
A3	58.01792343	6	-104.03585
fitted	55.9618136	3	-105.92363
R	45.03881466	2	-86.077629

\* Includes additive constant of -45.94693. This constant was not included in the LL derivation prior to BMDs 3.0.

이전에 언급했듯이 각 모델에 대한 매개변수의 수는 추정을 위해 설정된 경계 중 하나에 값이 있는 매개변수를 제외한다(사용자가 지정한 경계 또는 모델과 관련된 고유한 제약, table 6 참조).

5개의 로그 우도 모델은 적합성 검정을 포함하여 점근적으로 카이 제곱되는 가설 검정에 사용할 수 있다. 이러한 각 로그 우도 값은 사용자가 데이터 분석에서 고려할 수 있는 모델에 해당한다. 5가지 모델은 다음 표에 요약되어 있다.

Table 8. Likelihood values and models for continuous endpoints.

Model	Description
A1: "Full" Constant Variance Model	$Y_{ij} = \mu_i + e_{ij},$ $Var\{e_{ij}\} = \sigma^2$
A2: "Fullest" Model	$Y_{ij} = \mu_i + e_{ij},$ $Var\{e_{ij}\} = \sigma_i^2$
A3: "Full" Model with variance structure specified by the user	$Y_{ij} = \mu_i + e_{ij},$ $Var\{e_{ij}\} = \alpha \times \mu_i^p$
R: "Reduced" Model	$Y_i = \mu + e_i,$ $Var\{e_i\} = \sigma^2$
Fitted Model	The user-specified model

모델 A1은 관찰된 용량 그룹에 대한 개별적이고 독립적인 평균을 추정하지만 (이 점에서 "full" 또는 "saturated"됨) 해당 그룹에 대해 일정한 분산을 가정한다.

모델 A2는 관찰된 선량 그룹(모델 A1에서와 같이)에 대해 개별적이고 독립적인 평균을 추정하고 해당 그룹에 대한 개별적이고 독립적인 분산도 추정한다는 점에서 "fullest(가장 완전한)" 모델이다. 평균 간에 또는 용량 그룹 간의 분산 간에는 기능적 관계가 가정되지 않는다. 이 모델은 종종 "포화" 모델이라고 한다(용량 그룹만큼 많은 평균 및 분산 매개변수가 있음). 이 모델에 대해 얻은 로그 우도는 고려 중인 데이터에 대해 얻을 수 있는 최대값이다.

모델 A3은 모델 A2와 유사하며 분산 매개변수에 대해서만 다를 수 있다. 모델 A2는 관찰된 용량 그룹(예: A1)에 대해 개별적이고 독립적인 평균을 추정한다. 사용자가 피팅된 모델에 대해 일정한 분산을 지정하면 모델 A3도 이를 가정하고 모델 A1과 동일하게 된다. 사용자가 피팅된 모델에 대해 일정하지 않은 분산을 가정하면 모델 A3도 분산에 대해 동일한 기능 형식을 가정한다.

Reduced 모델(R)은 용량 수준에 대한 평균 또는 분산의 차이가 없음을 의미하는 모델이다. 즉, 모든 용량 수준에서 해당 평균을 중심으로 동일한 분산을 갖는 일정한 평균 반응 수준을 가정한다.

마지막 모델인 fitted 모델은 사용자 지정 모델(예: 거듭제곱 또는 다항식 등)이다. 사용자는 특정 모델이 데이터를 잘 설명할 수 있다고 믿을 만한 이유가 있을 수 있으므로 이를 사용하여 BMD 및 BMDL을 계산할 수 있다.

## (3) Interest table의 test(tests of Fit)

Figure 51. Tests of Interest

Tests of Interest			
Test	-2*Log(Likelihood Ratio)	Test df	p-value
1	27.02196286	8	0.00070084
2	1.063745329	4	0.89998092
3	1.063745329	4	0.89998092
4	4.112219646	3	0.24959863

Interest test 표는 interest 우도 표의 로그 우도를 기반으로 하는 4 개의 테스트 결과를 보여준다. 검정과 관련된 p-값은 우도 비율의 점근적 속성을 기반으로 한다.

너무 기술적이지 않으면 우도 비율은 BMD5 출력에 많은 값이 제공되는 두 개의 우도 값의 비율일 뿐이다. 통계이론은 표본크기가 커지고 용량군의 수가 많아 질수록  $-2*\ln(\text{우도비, likelihood ratio})$ 가 카이제곱 확률변수로 수렴함을 증명한다. 이러한 값은 차례로 모델 적합성(model fit)에 대한 추론을 수행하기 위한 대략적인 확률을 얻는 데 사용할 수 있다. 카이-제곱 표는 거의 모든 통계 참고서에서 찾을 수 있다.

사용자가 두 모델 A와 B의 적합도를 테스트하려고 한다고 가정하자. 이 테스트를 위해 모델 A가 모델 B에 "내포되어(nested within)" 있다고 추정했다. 즉, 단순화된 모델이 모델 A가 되는 방식으로 모델 B를 단순화할 수 있다(모델 B의 일부 매개변수 제한을 통해). 이는 모델 A가 더 적은 가변 매개변수를 갖는다는 것을 의미한다. 예를 들어, linear(선형) 모델은 파워 매개변수(power parameter)가 1로 제한되어 있기 때문에 파워 모델에 비해 선형 모델이 "단순한(simpler)" 또는 "중첩된(nested)" 모델이라고 가정한다.

[참고] 더 많은 수의 매개변수가 있는 모델은 항상 이 비율의 분모에 있다.

$L(X)$ 가 모델  $X$ 의 우도를 나타낸다고 가정하자. 이 이론을 이용하여,  $-2 \times \ln\left\{\frac{L(A)}{L(B)}\right\}$  카이제곱 확률변수에 접근한다. 이는 비율의 로그(log of ratio)가 로그 간의 차이 또는 아래식과 같다는 사실을 이용하여 단순화할 수 있다.

$$-2 \times \ln\left\{\frac{L(A)}{L(B)}\right\} = -2 \times (\ln\{L(A)\} - \ln\{L(B)\}) = 2 \times \ln\{L(B)\} - 2 \times \ln\{L(A)\}$$

Likelihoods of Interest 테이블의 값은 실제로 위에서 논의한 바와 같이  $\ln\{L(B)\}$  및  $\ln\{L(A)\}$  로그 우도이므로 이 우도 비율 계산은 단지 빼기 문제가 된다. 그런 다음 이 값은 지정된 수의 자유도를 가진 카이-제곱 확률 변수와 비교할 수 있다.

Likelihoods of Interest 표와 관련하여 언급했듯이 각 로그 우도 값에는 연관된 수의 매개변수가 있다. 카이-제곱 검정 통계량의 자유도 수는 단순히 두 모델 매개변수 개수 간의 차이이다. 위의 미니 예제에서 모델 A에 5개의 적합된 매개변수가 있고 모델 B에 8개의 매개변수가 있다고 가정하면, 이 경우 비교할 카이-제곱 값은  $8 - 5 = 3$ 의 자유도를 갖는 카이-제곱이 된다.

A와 B의 예시에서는 정확히 무엇을 테스트하는가? 가설을 세운다면 다음과 같다.

$H_0$ : A는 B만큼 데이터를 모델링한다.

$H_1$ : B는 A보다 데이터를 더 잘 모델링한다.

이를 염두에 두고 자유도 3을 기준으로

$2 \times \log\{L(B)\} - 2 \times \log\{L(A)\} = 4.89$  라고 가정한다. 또한 기각 기준(rejection criteria)이 .05 미만의 카이제곱 확률이라고 가정한다. 카이-제곱 표에서 4.89는 0.10에서 .25 사이의 p-값을 갖는다. 이 경우  $H_0$ 는 기각되지 않으며 Model A를 사용하여 데이터를 모델링하는 것이 적절할 것이다. BMDs는



자동으로 사용자에게 대해 "table look-up(테이블 조회)"를 수행하고 위에서 기술된 자유도를 가지는 계산된 로그 우도 비와 관련된 p-값을 제공한다.

Tests of Interest 표는 네 가지 기본 테스트를 제공한다. 테스트 결과에 대한 EPA의 해석(즉, EPA가 선택한 p-값과 관련하여)을 포함하는 테스트 결과의 요약된 해석을 표시하기 위해 액세스할 수 있는 "hover box(호버 상자)"가 이러한 각 테스트와 연결되어 있다. 그러나 계산된 p-값은 사용자가 원하는 거부 기준을 자유롭게 사용할 수 있도록 표시된다. 연속형 모델에 대해 제공되는 네 가지 기본 테스트 각각은 아래에서 자세히 설명한다.

**Test 1(A2 vs R): 반응과 분산이 용량 수준 간에 다르지 않다는 귀무 가설을 검정한다. 이 검정이 귀무 가설을 기각하지 못하면 용량-반응이 없을 수 있다.**

이 테스트는 모델 R(simpler 모델)과 모델 A2를 비교한다. 모델 R은 모든 평균 매개변수를 서로 동일하게 제한하고 모든 분산 매개변수를 서로 동일하게 제한함으로써 A2에서 R을 얻을 수 있기 때문에 더 간단한 A2(또는 A2 내에 중첩됨)이다. 이 테스트가 귀무 가설을 기각하지 못하면 단순한 모델(R)이 포화(saturated) 모델보다 훨씬 나쁘지 않다는 추론이 있기 때문에 용량-반응이 없을 수 있다. 테스트의 기본 p-값(출력의 Test of Interest 섹션에 보고됨)은 0.05이다. 0.05 미만의 p-값은 반응 또는 용량 수준별 분산에 차이가 있다는 의미이며 데이터를 모델링하기 위한 결론을 뒷받침한다. 0.05보다 큰 p-값은 데이터가 용량-반응 모델링에 적합하지 않을 수 있음을 나타낸다.

**Test 2(A1 vs A2): 분산이 동일하다는 귀무 가설을 검정한다. 이 검정이 귀무 가설을 기각하지 못하면 더 간단한 상수 분산 모델이 적절할 수 있다.**

이 검정은 A1(simpler model)과 모델 A2를 비교한다. 모델 A1은 모든 분산 매개변수가 서로 동일하도록 제한하여 A2에서 A1을 얻을 수 있기 때문에 더 간단한 A2(또는 A2 내에 중첩됨)이다. 이 검정이 귀무 가설을 기각하면 상수 분산 가정이 맞지 않으며 데이터를 적절하게 나타내기 위해 모형화된 분산이 필요하다는 추론이다. 테스트의 기본 p-값(출력의 Tests of Interest 섹션에서

보고됨)은 0.05이다. 0.05보다 작은 p-값은 사용자가 비등분산 모델 실행을 고려해야 함을 나타낸다. 0.05보다 큰 p-값은 일정한 분산 가정이 용량-반응 모델링에 적합할 수 있음을 나타낸다.

**Test 3(A3 vs A2): 분산이 적절하게 모형화되었다는 귀무 가설을 검정한다.** 이 검정이 귀무 가설을 기각하지 못하면 분산이 적절하게 모형화되었음을 유추할 수 있다.

여기서 테스트는 사용자 지정 분산 모델이 적합한지 확인하는 테스트이다. 사용자 지정 분산 모델이 "일정한 분산(constant variance)"인 경우 모델 A1과 A3은 동일하다. 이 테스트는 동일한 해석으로 테스트 2와 동일하다. 사용자 지정 분산 모델이 일정하지 않은 경우( $\sigma^2 = \alpha \times \mu_i^{\rho}$ ), 이 테스트는 해당 방정식이 용량 그룹 간의 분산을 설명하는 데 적절한지 여부를 결정한다. 모델 A3은 일정하지 않은 분산 방정식에 맞게 분산을 제한하여 얻은 모델 A2의 더 간단한 버전이다. 테스트의 기본 p-값(출력의 Tests of Interest 섹션에서 보고됨)은 0.05이다. 0.05보다 작은 p-값은 사용자가 다른 분산 모델을 고려하기를 원할 수 있음을 나타낸다. 0.05보다 큰 p-값은 용량-반응 모델링에 대해 모델링된 분산의 사용을 지원한다.

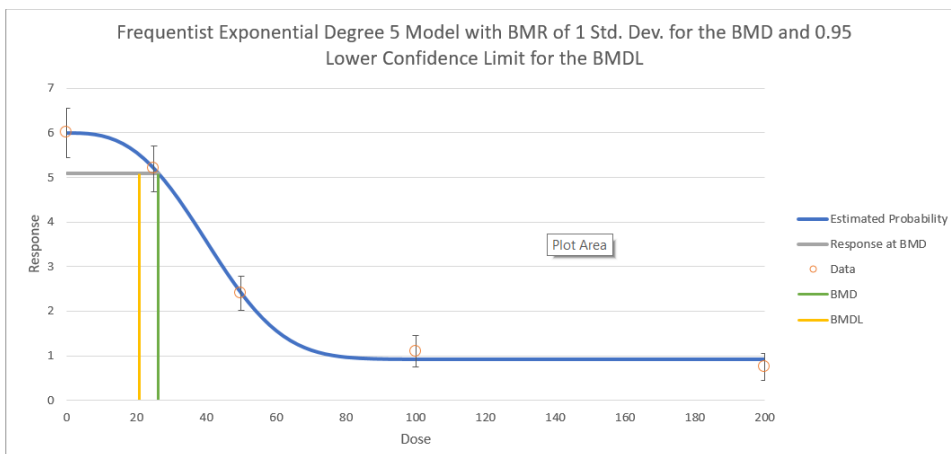
**Test 4(Fitted vs A3): 평균에 대한 모형이 데이터를 fit한다는 귀무 가설을 검정한다.** 이 검정이 귀무 가설을 기각하지 못하면 사용자는 선택한 모델을 지지한다.

이 검정은 적합 모형(Fitted Model)을 모형 A3과 비교한다. 적합 모형은 고려 중인 용량-반응 함수에 의해 설명될 수단(A3에서 제한 없음)을 제한하여 얻을 수 있기 때문에 모형 A3(또는 모형 A3 내에 중첩됨)보다 더 간단하다. 이 검정이 귀무 가설을 기각하지 못하면 적합 모형이 평균의 용량 관련 변화를 설명하기에 더 적합하다는 결론이 난다(분산 모델의 형태에 대한 조건부, 분산 모델의 형태는 fitted model 및 모델 A3과 동일함). 귀무 가설을 기각하지 못하는 것은 고려 중인 용량-반응 함수의 형태에 대한 수단의 제한이 적절하다는 추론과

관련이 있다. 검정의 기본 p-값(출력의 Tests of Interest 섹션에서 보고됨)은 0.1이다. 0.1보다 p-값이 작으면 사용자가 다른 모델을 시도하기를 원할 수 있다(즉, 적합 모형의 적합도가 충분하지 않음). 0.1보다 큰 p-값은 fitted model이 용량-반응 모형화에 적합함을 나타낸다.

#### (4) Plot과 error bar 계산

**Figure 52.** Frequentist results plot for continuous data.



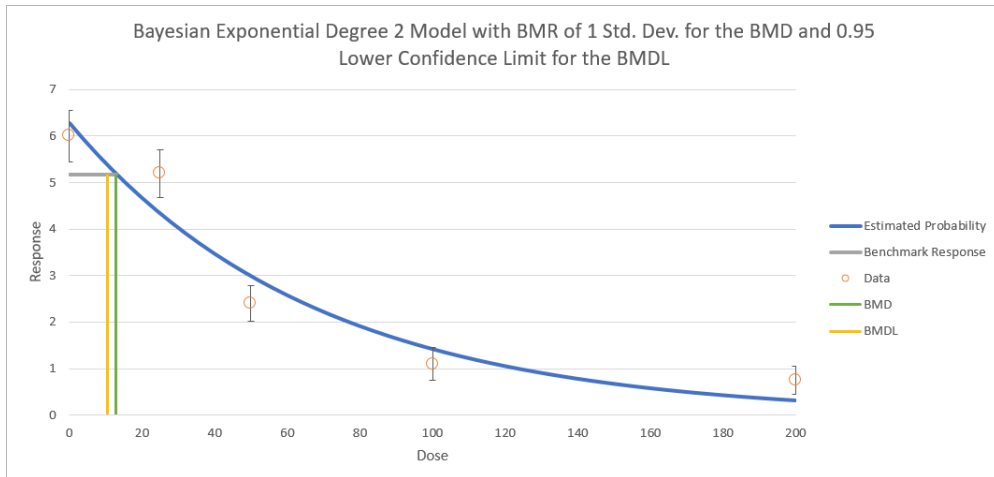
그래픽 출력, 즉 플롯(plot)은 모델링 결과를 시각적으로 표현한 것이다. 모든 모델에 공통적인 플롯은 “5)모델 유형의 공통 출력물-(5)그래픽 출력”에서 언급했기 때문에 여기에서는 연속형 모델에 특정한 한 가지 추가 세부 사항, 즉 오차 막대 계산을 설명하였다.

- Plotting routine은 각 그룹에 대한 평균의 표준 오차(SEM, standard error of the mean)을 계산한다. 루틴은 그룹 특이적으로 관찰된 분산 (obs standard deviation squared, obs 표준 편차 제곱)을 그룹별 샘플 크기로 나눈다.
- 루틴은 SEM에 그룹별 샘플 크기에 적합한 Student-T 백분위수(각각 하한 (lower) 및 상한(upper)에 대해 2.5번째 백분위수 또는 97.5번째 백분위

수)를 곱한다(즉, 해당 샘플보다 1 작은 자유도를 가짐). 루틴은 오차 막대의 하단 및 상단을 정의하기 위해 관찰된 평균에 결과물을 추가한다.

## 6) Bayesian 연속형 모델(미리보기)에 대한 특이적인 출력

Figure 53. Bayesian results plot for continuous data.



두 베이저안 모델 간의 차이를 비교하기 위해 정규화되지 않은 로그 사후 확률(LPP, Log Posterior Probability)이 제공되며, 이를 통해 베이즈 인수(BF, Bayes factor)를 계산하여 두 모델을 비교할 수 있다. BF는 두 LPP 간의 지수 차이와 같다. 예를 들어, Log-Logistic 모델(모델 A)( $LPP_A$  산출)을 다단계 2차 모델(모델 B,  $LPP_B$  포함)과 비교하기 위해 BF를 아래와 같이 추정한다.

$$BF = \exp(LPP_A - LPP_B)$$

이 계산은 두 모델이 선형적으로(priori) 동일한 확률을 갖는다고 가정한다. 이 값은 한 모델이 다른 모델보다 더 정확하고 베이저안 가설 검정에 사용되는 사후 확률(posterior odds)로 해석된다. 위의 예에서 BF가 2.5인 경우 Log-logistic 모델이 다단계 모델보다 사후 가능성(posteriori)이 2.5배 더 높다고 해석할 수 있다. 이러한 값이 적절한 확률로 정규화되면 모델 평균화에 제공된 사후 모델 확률과 동일하다(다시, 동일한 모델 확률을 선형적으로 가정). 표9에는 Jeffreys 등이 제시한 Bayes Factors의 일반적인 해석을 나타내었다.

**Table 9.** Bayes factors for continuous models.

Bayes Factor	Strength of Evidence for $H_A$
< 1	negative (supports $H_B$ )
1 to 3.2	not worth mentioning
3.2 to 10	substantial
10 to 31.6	strong
31.6 to 100	very strong
100	decisive

BMDs 3.0의 경우 모든 LPP 및 해당 사후 모델 확률은 Laplace 근사를 이용하여 계산한다. 이 값은 일반적으로 사용되는 BIC(Bayesian Information Criterion)와 다르며 BIC를 독점적으로 사용하는 다른 모델 평균화 접근 방식을 기반으로 두 값을 혼동해서는 안 된다. BIC에서 추정된 사후 확률의 오류는  $O(1)$  추정기(estimators)이다. Laplace 근사치를 사용하여 추정된 사후 확률 (posterior probability)의 오류는  $O(n^{-1})$ 이다. 즉, 데이터가 증가함에 따라 후자의 근사치는 실제 사후 모델 확률로 가고 전자는 BIC를 사용하여 실제 값이 되지 않을 수 있다.

## 7. 이분형 자료의 종말점

BMDS에는 관측치가 서로 독립적인 이분법적 종말점에 대한 모델이 포함되어 있다. 이 모델에서 용량-반응 모델은 실험 단위(예: 독성 시험에서 쥐 또는 마우스)가 주어진 용량에서 부작용을 가질 확률을 정의한다. 부작용이 있는 실제 동물 수는 이항 분포(binominal distribution)된 것으로 가정한다.

이러한 데이터세트의 구체적인 예로는 성체 동물을 다양한 농도의 독성 물질에 노출시킨 다음 간 독성의 존재를 평가하는 연구 등이 있다. 이분형 압 모델 및 여러 종양 종점에 대한 모델 예측 조합은 “8. 다중 종양 분석”을 참조한다.

이 섹션에서는 다음 주제에 대한 세부 정보를 제공한다.

- BMDS 3에서 이분형 모델 구현
- 이분형 모델 데이터 입력
- 이분형 모델 옵션
- 이분형 모델 출력
- 특정 모델 매개변수의 값을 제한하는 옵션
- 이분형 반응 모델링에 대한 베이지안 접근 방식, 사전 정의 방법 및 모델 평균화 방법 지정

### 1) 이분형 반응 모델

BMDS 3은 BMDS의 이전 버전과 각 모델의 베이지안 버전에서 사용할 수 있는 전통적인 빈도주의적 이분형 반응 모델과 베이지안 모델 평균화 기능을 제공한다.

Figure 54. BMDS 3 dichotomous models, as they appear in the Analysis Workbook.

	MLE		Alternatives		
	Frequentist Restricted	Frequentist Unrestricted	Bayesian	Bayesian Model Average	
Model Name	Enable <input type="checkbox"/>	Enable <input type="checkbox"/>	Enable <input type="checkbox"/>	Enable <input type="checkbox"/>	Prior Weights
Dichotomous Hill	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	0.0000%
Gamma	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	0.0000%
Logistic	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	0.0000%
Log-Logistic	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	0.0000%
Log-Probit	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	0.0000%
Multistage	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	0.0000%
Probit	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	0.0000%
Quantal Linear	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	0.0000%
Weibull	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	0.0000%
				Total Weight	0.000%

대부분의 빈도주의 모델은 일부를 제한하거나 제한없이 실행할 수 있다. 초기 실행에 대한 EPA 기본 권장 사항은 Dichotomous Hill, Gamma, Log-Logistic, Multistage 및 Weibull 모델을 제한하고 Log-Probit 모델을 제한 해제하는 것이다. Logistic, Probit 및 Quantal Linear 모델에는 제한 옵션이 없다.

사용자가 제한된 버전의 모델을 선택할 때의 효과는 table 10을 참조한다 ("User parameter restriction options"). 일반적으로 제한은 용량이 0인 군에서 용량-반응 곡선의 기울기가 무한대가 되는 것을 방지한다. 이는 생물학적으로 비현실적인 것으로 여겨지며 신뢰 한계(confidence limit)를 계산할 때 수치적 문제로 이어질 수 있으므로 적절히 매개변수를 제한할 것을 권장하고 있다.

## 2) 옵션

BMDS 3 Analysis Workbook의 Main 탭에서 사용자는 단일 "batch(일괄 처리)" 프로세스에서 여러 사용자가 선택한 모델과 여러 사용자가 선택한 데이터 세트에 적용할 여러 Option Sets를 정의할 수 있다. Add Option Set 버튼을 선택하여 새로운 Option Set 구성을 정의한다.

Figure 55. Dichotomous Model options.

Option Set #	Risk Type	BMR	Confidence Level	Background
1	Extra Risk ▼	0.1	0.95	Estimated ▼
Add Option Set				

### (1) 위험 유형(risk type)

Figure 56. Risk type

Option Set #	Risk Type
1	Extra Risk ▼
	Extra Risk
	Added Risk

위험 유형에 대한 선택은 "Extra Risk"(기본값) 또는 "Added Risk"이다.

Added risk는 용량이 있을 때 반응하는 총 실험 단위의 추가 비율 또는 용량  $d$ ,  $P(d)$ 에서 예상되는 반응 확률에서 노출이 없을 때의 예상 반응 확률  $P(0)$ 을 뺀 것이다.

$$\text{Added risk at dose } d = P(d) - P(0)$$

Extra risk는 added risk를 노출 없이 반응하지 않을 동물의 예측 비율로 나눈 값이다( $1 - P(0)$ ).

$$\text{Extra risk at dose } d = \frac{P(d) - P(0)}{1 - P(0)}$$



## (2) BMR

Figure 57. BMR

<b>BMR</b>
0.1

BMR은 BMD가 추정되는 위험 값(사용자가 지정한 extra 또는 added)이다. BMR은 0에서 1 사이의 값이어야 한다(1은 포함하지 않음).  $P(0) > 0$ 인 경우 BMR 값이  $1 - P(0)$ 보다 크면 위험 유형이 added risk일 때 오류가 발생한다. 달성할 수 있는 최대 added risk가  $1 - P(0)$ 이기 때문이다. 실제로는 사람들은 보통 0.01에서 약 0.10 사이의 BMR 값에 관심이 있기 때문에 일반적으로는 문제가 되지 않는다.

### BMR 및 그래프

그래픽 모델 출력에 표시되는 BMR과 관련된 반응은  $P(0) = 0$ 일 때만 BMR과 동일하다.

실제 반응 값을 얻으려면 위에서 논의한 added 또는 extra risk에 대한 방정식에서  $P(d)$ 에 대해 풀어야 하기 때문이다.

그래픽 모델 출력에 표시되는 BMD를 도출하는 데 사용되는 반응 수준을 나타내는 수평 막대는 백그라운드 반응인  $P(0)$ , 0과 같다.

$P(0)$ 이 0과 같지 않으면 위에서 설명한 Extra Risk 방정식을 사용하여 실제 반응 수준을 계산할 수 있다.

## (3) 신뢰수준(confidence level)

Figure 58. Confidence level

<b>Confidence Level</b>
0.95

신뢰 수준은 0과 1 사이의 분수이다. EPA는 0.95를 권장한다.

신뢰 수준 값은 0에서 1 사이의 값이어야 한다(1은 포함하지 않음).  $x$ 의 신뢰 수준에 대해 BMDS는 BMDL 및 BMDU 추정치를 출력하며, 각각은 수준  $x$ 에서 단측 신뢰 범위이다. 예를 들어, 사용자가 신뢰 수준을 0.95(기본값)로 설정하면 BMDL은 BMD 추정치에 대한 95% 단측 신뢰 하한이다. BMDU는 BMD 추정치에 대한 95% 단측 신뢰 상한이다. 이 예에서 BMDL에서 BMDU까지의 범위는 90% 신뢰 구간을 구성합니다(해당 구간 외부의 각 꼬리에서 5%).

#### (4) 백그라운드

Figure 59. Background

사용자는 백그라운드 매개변수를 추정할지(기본값) 또는 다른 매개변수를 피팅 및 추정하기 전에 0으로 설정할지 지정할 수 있다.

[참고] 용량이 0인 그룹에 반응이 0보다 많은 경우 사용자는 백그라운드를 0으로 설정하지 않아야 한다. 이것은 논리적으로 불가능한 상황, 즉 해당 용량군의 반응 확률이 0이지만 실제로는 일부 반응이 있는 상황을 나타낸다.

### 3) Simple designs에서 이분형 종말점을 위한 모델의 수학적 상세

BMDS에는 다음 표에 정의한 것과 같이 이분형 종말점에 대한 9가지 모델이 포함되어 있다.

**Table 10.** *The individual dichotomous models and their respective parameters.*

Model	Parameters	Notes
Multistage $p(\text{dose}) = g + (1 - g)(1 - \exp[-\sum_{j=1}^n \beta_j \text{dose}^j])$	$g$ = background $\beta_j$ = dose coefficients	Parameter Constraints: $0 \leq g < 1$ $n \leq 23$ User parameter restriction options: can restrict all $\beta$ coefficients to $\geq 0$ . Doing so will guarantee that the multistage model will be either perfectly flat or always increasing. Per EPA Guidance, when the Multistage model is used for cancer analyses (e.g., in Multi-tumor analyses) all $\beta$ coefficients are restricted to be non-negative.
Weibull (and Quantal Linear) $p(\text{dose}) = g + (1 - g)(1 - \exp[-\beta \text{dose}^\alpha])$	$g$ = background $\alpha$ = power $\beta$ = slope	Parameter Constraints: $0 \leq g < 1$ $0 < \alpha \leq 18$ $0 < \beta$ User parameter restriction options: $1 \leq \alpha$ The <b>Quantal Linear</b> model results from setting $\alpha = 1$ .
Gamma $p(\text{dose}) = g + \frac{1-g}{\Gamma(\alpha)} \int_0^{\beta \text{dose}} t^{\alpha-1} \exp(-t) dt$	$g$ = background $\alpha$ = power $\beta$ = slope	Parameter Constraints: $0 \leq g < 1$ $0.2 < \alpha \leq 18$ $0 < \beta$ User parameter restriction options: $1 \leq \alpha$ Note that for the unrestricted Gamma model, $\alpha > 0.2$ for numerical reasons.
Logistic $p(\text{dose}) = \frac{1}{1 + \exp[-\alpha - \beta(\text{dose})]}$	$\alpha$ = intercept $\beta$ = slope	Parameter Constraints: $0 < \beta$ User parameter restriction options: none.

### (1) 우도 함수(Likelyhood function)

빈도주의적 모델링 접근 방식의 경우 이분형 모델은 최대 우도법을 사용하여 fitting한다. 이 섹션에서는 이분형 모델을 fitting하는 데 사용되는 우도 함수에 대해 설명한다.

G 용량군이 있는 데이터세트가 있다고 가정하고 아래처럼 나타내었다.

$$\text{dose}_1, \text{dose}_2, \dots, \text{dose}_G$$

각 용량군의 총 관찰 수 및 반응 수는 각각 다음과 같이 나타낸다.

$$N_1, N_2, \dots, N_G \text{ 및 } n_1, n_2, \dots, n_G$$

$n_i$ 의 분포는 확률이 있는 이항(binomial)으로 추정된다.

$$p_i = p(\text{dose}_i; \theta), i = 1, 2, \dots, G$$

여기서  $\theta$ 는 용량-반응 모델 매개변수의 벡터이다(각 모델의 매개변수의 목록 표를 참조). 그 후 로그 우도 함수 LL은 아래와 같이 쓰일 수 있다.

$$LL = \sum_{i=1}^G LL_i(N_i, n_i, \text{dose}_i; \theta)$$

여기서

$$LL_i(N_i, n_i, \text{dose}_i; \theta) = n_i \ln(p_i) + (N_i - n_i) \ln(1 - p_i), i = 1, 2, \dots, G$$

이 표현식은 매개변수 벡터 값과 무관한 상수항을 무시하므로 해당 매개변수의 추정에 영향을 미치지 않는다.

위의 표에서 일부 모델의 파워 매개변수의 상한과 다른 일부 모델의 기울기 매개변수(slope parameter)는 18로 설정되었다. 이 값은 매우 높은 곡률(curvature)을 나타내기 때문에 선택되었으며, 저용량에서 매우(혹은 절대적으로) 균일한 용량-반응을 보인 다음 고용량에서 매우 가파른 용량-반응을 나타내는 데이터 세트를 포함하여 거의 모든 데이터 세트를 수용할 수 있어야 한다.

이러한 매개변수 값이 18과 같다고 리포트되거나 문제의 추정치가 "Bounded"로 리포트되는 경우 ("7. 이분형 자료의 종말점-4)Frequentist 이분형 모델에 대한 특이적인 출력-(2)일탈표의 분석"에서 이분형 모델 실행의 출력에 대한 설명 참조), 매개변수 추정치는 문제의 매개변수에 값이 할당되었고 다른 매개변수는 할당된 값에 대한 조건부 MLE라는 제한된 의미로만 가능한 최대 우도 추정치가 된다. 이러한 모델 결과는 AIC 측면에서 다른 모델과 엄격하게 비교할 수 없다. 이러한 경우 BMD 및 BMDL은 파워 매개변수의 선택에 따라 달라질 수 있다. 따라서 보고된 BMD 또는 BMDL에 의존하려는 경우 민감도 분석(sensitivity analysis)이 표시된다.

## (2) AIC와 model 비교

Akaike Information Criterion(AIC)은 동일한 데이터세트에 대한 서로

다른 모델 적합도를 비교하는 데 사용할 수 있다. AIC는 LL의 값과 추정된 매개변수의 수  $p$ 에 따라 달라지는 통계이다.

$$AIC = -2 \times LL + 2 \times p$$

이분형 종말점에 대한 AIC는 이전 섹션에서 정의된 LL이 해당 용어를 무시하기 때문에 매개변수 독립적인 용어를 무시한다. 이것은 그 값이 가정된 기본 데이터 분포(정규 또는 로그 정규)에 의존하기 때문에 매개변수 독립 항이 무시되지 않는 연속형 종말점의 경우와 다르다. 이분형 종말점의 경우 반응의 수에 대해 가정된 분포(이항 분포)가 하나밖에 없으므로 매개변수 독립 항은 효과가 없어 무시할 수 있다.

AIC는 높은 LL 값을 달성하는 데 필요한 매개변수의 수와 관련하여 적은 수의 데이터로(parsimonious) 가능한 가장 높은 LL 값을 얻는 목표의 균형을 맞춘다. AIC에 대한 방정식은 LL에 대한 음의 승수(더 크게 하고 싶은 것)와  $p$ 에 대한 양의 승수(가능한 한 작아서 여전히 "좋은 적합성"을 얻고자 함)를 가지므로 더 작은 값을 가진 모델은 다음과 같다. AIC를 기반으로 하여 다른 모델보다 AIC가 더 나은 모델로 추정된다. 이러한 방법은 정확하지 않지만 모델 선택에 유용한 지침을 제공할 수 있다.

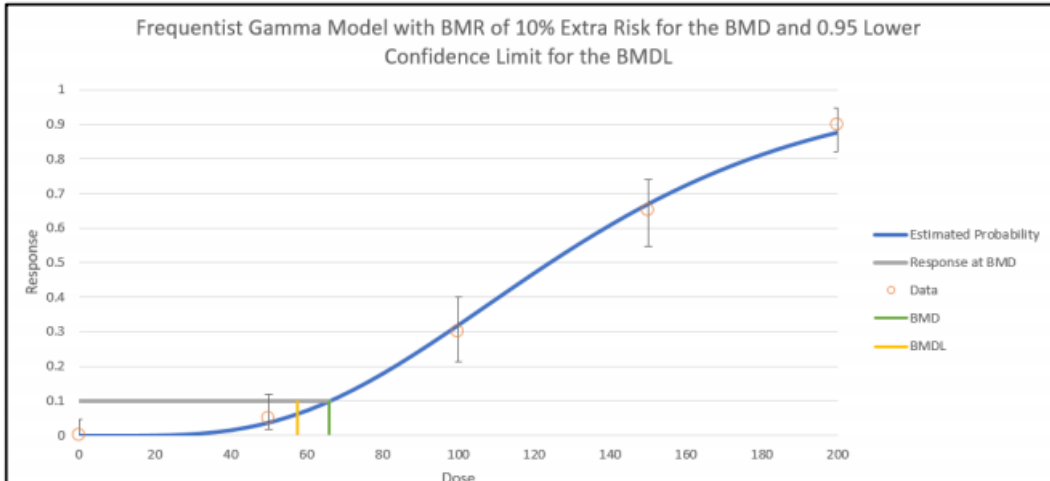
현재 버전의 BMDS에서 "추정된 매개변수(estimated parameter)"의 수에는 모델이 부과한 제약 조건이나 사용자가 부과한 제약 조건에서 경계 값과 동일하지 않은 것으로 추정된 매개변수만 포함된다. 자세한 내용은 table 10을 참조한다.

**[중요 사항]** 이 계산 프로세스는 해당 매개변수가 적중하는 경계값(boundary value)에 따라 합리적일 수도 있고 아닐 수도 있다. 예를 들어, 모델의 파워 매개변수가 18의 상한에 도달하면 일반적으로 해당 매개변수가 추정된 것으로 계산하고 싶지만 BMDS는 그렇지 않다. 그렇게하지 마십시오!

이러한 이유로 사용자는 매개변수 경계에 도달한 경우를 주의 깊게 확인하고 모델 비교 및 모델 선택 등의 문제에 대한 영향을 고려하도록 한다.

### (3) Plot과 error bar 계산

Figure 60. Dichotomous endpoint plot.



그래픽 출력, 즉 플롯은 모델링 결과의 시각적 묘사방법이다. 일반적으로 플롯은 “5. 모델 유형의 공통 출력물-5)그래픽 출력”에서 논의되었기 때문에 여기서는 이분형 모델에 특정한 한 가지 추가 세부 사항, 즉 오차 막대 계산에 대해 설명하였다.

이분형 데이터의 플롯에 표시된 오차 막대는 Wilson interval의 변형 (modification)(점수 통계 기반)을 사용하지만 continuity correction 방법에서 파생되었다. 비율  $p_i$ 를 찾는 계산식은 아래와 같다.

$$\frac{|p - p_i| - \frac{1}{2n}}{\sqrt{\frac{p_i \times (1 - p_i)}{n}}} = z$$

이 식에서

- $p$ 는 관찰된 비율이다.
- $n$ 은 해당 그룹의 총 수이다.
- $z = Z_{1-\frac{\alpha}{2}}$ 는  $1 - \frac{\alpha}{2}$ 에서 평가된 역 표준 정규 누적 분포 함수이다.

이는 다음의 하한 및 상한에 대한 방정식으로 이어진다.

$$LB = \frac{(2np+z^2-1)-z\sqrt{z^2-(2+\frac{1}{n})+4p(nq+1)}}{2(n+z^2)}$$

$$UB = \frac{(2np+z^2+1)+z\sqrt{z^2+(2-\frac{1}{n})+4p(nq-1)}}{2(n+z^2)}$$

여기서  $q = 1-p$ 이다.

BMDs 플롯에 표시된 오차 막대는  $\alpha = 0.05$ 를 사용하므로 관찰된 비율에 대한 95% 신뢰 구간을 나타낸다(모델과 무관).

#### (4) BMD 계산

BMD는 고려 중인 모델의 매개변수 함수로 계산한다(table 10). Table 11은 모든 이분형 모델에 대한 BMD에 대한 솔루션을 지정한다.

**Table 11.** Calculation of the BMD for the individual dichotomous models<sup>a</sup>.

Model	BMD Calculation
Multistage	There is no general analytic form for the BMD in terms of the BMR and the estimated model parameters for the multistage model. Instead, the BMD is the root of the equation $\beta_1 BMD + \dots + \beta_n BMD^n + \ln(1-A) = 0$ , where $A = \begin{cases} \text{BMR extra risk} \\ \frac{\text{BMR}}{1-g} \text{ added risk} \end{cases}$
Weibull	$BMD = \begin{cases} \left[ \frac{-\ln(1-BMR)}{\beta} \right]^{\frac{1}{\alpha}} \text{ extra risk} \\ \left[ \frac{-\ln(1-\frac{BMR}{1-g})}{\beta} \right]^{\frac{1}{\alpha}} \text{ added risk} \end{cases}$

Model	BMD Calculation
Gamma	Let $G(x; \alpha) = \frac{1}{\Gamma(\alpha)} \int_0^x t^{\alpha-1} e^{-t} dt$ be the incomplete Gamma function and $G^{-1}(\cdot; \alpha)$ be its inverse function. Then $\text{BMD} = \begin{cases} \frac{G^{-1}(\text{BMR}; \alpha)}{\beta} & \text{extra risk} \\ \frac{G^{-1}\left(\frac{\text{BMR}}{1-g}; \alpha\right)}{\beta} & \text{added risk} \end{cases}$
Logistic	$\text{BMD} = \frac{\ln\left(\frac{1-Z}{1+Z+e^{-\alpha}}\right)}{\beta} \text{ where } Z = \begin{cases} \text{BMR extra risk} \\ \text{BMR} \times \frac{1+e^{-\alpha}}{e^{-\alpha}} & \text{added risk} \end{cases}$
Log-Logistic	$\ln(\text{BMD}) = \begin{cases} \frac{\ln\left(\frac{\text{BMR}}{1-\text{BMR}}\right) - \alpha}{\beta} & \text{extra risk} \\ \frac{\ln\left(\frac{\text{BMR}}{1-g-\text{BMR}}\right) - \alpha}{\beta} & \text{added risk} \end{cases}$
Probit	$\text{BMD} = \begin{cases} \frac{\Phi^{-1}(\text{BMR}[1-\Phi(\alpha)] + \Phi(\alpha)) - \alpha}{\beta} & \text{extra risk} \\ \frac{\Phi^{-1}(\text{BMR} + \Phi(\alpha)) - \alpha}{\beta} & \text{added risk} \end{cases}$
Log-Probit	$\ln(\text{BMD}) = \begin{cases} \frac{\Phi^{-1}(\text{BMR}) - \alpha}{\beta} & \text{extra risk} \\ \frac{\Phi^{-1}\left(\frac{\text{BMR}}{1-g}\right) - \alpha}{\beta} & \text{added risk} \end{cases}$
Dichotomous Hill	Added risk: $\text{BMD} = e^{\frac{-\alpha - \log\left(-\frac{\text{BMR} - v + gv}{\text{BM}}\right)}{\beta}}$ Extra risk: $\text{BMD} = e^{\frac{-\alpha - \log\left(-\frac{\text{BMR} - v + gv - \text{BMR}gv}{\text{BMR}(-1+gv)}\right)}{\beta}}$

## (5) BMDL과 BMDU 계산

BMDS는 현재 BMD 관행에 따라 단측 신뢰 구간을 계산한다. BMD(여기서는 BMDL 및 BMDU라고 함)에 대한 신뢰 한계를 계산하는 일반적인 접근 방식은 BMDS의 모든 모델에 대해 동일하며 우도 비율의 점근적 분포(asymptotic distribution)를 기반으로 한다. 이 모델에서는 두 가지 다른 특정 접근 방식을 따른다.

다단계 모델(multistage model)의 경우 BMD가 명시적 매개변수로 나타날



수 있도록 용량-반응 모델 기능을 명시적으로 다시 매개변수화하는 것은 비효율적이다. 이러한 모델의 경우 BMR 방정식이 비선형 제약 조건으로 사용되며 BMD의 최소값은 최대 우도 추정치에서 지수 우도와 동일하게 결정된다.

$$\frac{\chi^2_{1,1-2\alpha}}{2}$$

나머지 모델의 경우 BMD와 용량-반응 모델(table 9) 측면에서 BMR을 정의하는 방정식은 모델 매개변수 중 하나에 대해 해결된다. 결과 식은 BMD가 매개변수로 명시적으로 나타나도록 모델을 다시 매개변수화하는 효과와 함께 모델 방정식으로 다시 대체된다. 그런 다음 나머지 매개변수가 우도를 최대화하기 위해 변경될 때 결과 로그 가능성이 최대 가능성 추정치보다 정확히 다음과 같이 작은 BMD 값이 발견된다.

$$\frac{\chi^2_{1,1-2\alpha}}{2}$$

모든 경우에 추가 제약 조건은 BMDL이 BMD보다 작고 BMDU가 BMD보다 크도록 지정한다.

#### (6) Bayesian 이분형 모델 설명

[참고] 현재 EPA는 베이지안 모델링 또는 베이지안 모델 평균화에 대한 기술 지침을 제공하지 않는다.

베이지안의 관점에서 추론은 모델  $M$  및 해당 매개변수가 있는 경우 데이터 생성 메커니즘을 정의하여 진행된다.  $M$ 은 반응 확률을 결정하는 table 10에 나열된 모델 중 하나이다. 데이터 생성 메커니즘은 관찰이 이항 샘플링에서 얻은 것으로 가정하며, 해당 모델 중 하나에 의해 정의된 용량 의존적 반응 확률(해당 모델의 특정 매개변수 값 포함)을 갖는다.

그런 다음 여기에서  $\mathcal{A}(D|M)$ 로 표시된 우도와 관련시킬 수 있다. 이는 모델에 대한 조건부 데이터  $D$ 의 우도임을 명시적으로 보여준다. 우도로그의 기능적 형식은 “7. 이분형 자료의 종말점-3) Simple designs에서 이분형 종말점을 위한 모델의 수학적 상세-(1)우도 함수”에 나와 있다.

BMDs 3에 사용된 베이지안 이분형 모델 세트는 빈도주의(MLE) 접근법에 사용된 모델 세트와 동일하다(table 10). 다음에서  $\theta$ 는 이러한 모델 중 하나를 정의하는 데 필요한 매개변수의 벡터이다. 예를 들어 Weibull 모델의 경우  $\theta = (g, \alpha, \beta)$ 입니다. 베이지안 접근 방식에 통합된 추가 고려 사항은  $\theta$ 에 대한 사전 분포의 사양이다. 베이지안 접근 방식은 지정된 사전을 취하고 고려 중인 데이터를 사용하여 업데이트하여  $\theta$ 에 대한 "후방" 분포를 얻는다.

베이지안 관점에서  $\theta$ 의 함수도 사후 밀도를 갖는다. 따라서 table 11에 요약된 방정식을 사용하여 BMD에 대한 사후 분포를 도출할 수 있다.

**[중요 사항]** 매개변수에 대한 사전 설정(prior)은 조정된 용량 및 조정된 반응을 기반으로 한다. BMDs는 이 조정(scaling)을 자동으로 수행한다.

BMDs는 고려 중인 데이터세트의 최대 용량으로 나누어 용량을 자동으로 조정한다. 즉, 고려 중인 용량이 0에서 1(1 포함)까지이다. BMDs는 대조군(또는 최저 용량) 그룹의 평균 반응으로 나누어 반응을 자동으로 조정한다.

사용자는 사전에 어떤 것도 확장할 필요가 없다. 이는 프로그램에서 반환된 매개변수 추정치 및 BMD 값이 입력 데이터 파일에 지정된 원래 용량 및 반응의 원래 척도로 다시 조정되었음을 의미한다.

BMDs는 BMD에 대한 후방을 다음과 같이 요약한다. BMD(아마도 "중앙 추정치"라고 부를 수 있음)는 최대 MAP(사후추정)  $\theta$  추정치를 사용하여 얻은 값(table 11 참조)과 같다. MAP은 사후 로그 가능성을 최대화하는  $\theta$ 의 값이다.

사후 밀도 자체는 Laplacian 근사를 사용하여 근사된다. 이 근사치는 사용자가 선택한 신뢰 수준에 해당하는 해당 밀도의 백분위수인 BMDL 및 BMDU 값을 추정하는 데에도 사용된다.

BMDS 이분법 모델에 대한 우선순위는 다음 표에 정의되어 있다.

**Table 12. Bayesian dichotomous models and their respective parameter priors.**

Model	Constraints	Priors	Notes
Multistage $p(\text{dose}) = g + (1 - g)(1 - \exp[-\sum_{i=1}^N \beta_i \text{dose}^i])$	$0 \leq g < 1$ $\beta_i > 0$ $N \geq 2$	$\text{logit}(g) \sim \text{Normal}(0, 2)$ $\beta_1 \sim \text{Lognormal}(0, 0.25)$ $\beta_i \sim \text{Lognormal}(0.1, i \geq 2)$	The prior over the $\beta_1$ parameter reflects the belief that the linear term should be strictly positive if the quadratic term is positive in the two-hit model of carcinogenesis. The difference in priors between Multistage and Quantal Linear models is by design. The objective is to emphasize the higher-order terms in each model. For model averaging purposes, $N = 2$ .
Weibull $p(\text{dose}) = g + (1 - g)(1 - \exp[-\beta \text{dose}^\alpha])$	$0 \leq g < 1$ $\alpha > 0$ $\beta > 0$	$\text{logit}(g) \sim \text{Normal}(0, 2)$ $\alpha \sim \text{Lognormal}(\ln(2), \sqrt{0.18})$ $\beta \sim \text{Lognormal}(0.1)$	The prior for $\alpha$ entails that there is only a 0.05 prior probability the power parameter will be less than 1. This allows for models that are supra-linear; however, it requires a large amount of data for the $\alpha$ parameter to go much below 1.
Quantal linear $p(\text{dose}) = g + (1 - g)(1 - \exp[-\beta \text{dose}])$	$0 \leq g < 1$ $\beta > 0$	$\text{logit}(g) \sim \text{Normal}(0, 2)$ $\beta \sim \text{Lognormal}(0.1)$	The difference in priors between Quantal Linear and the following Multistage model is by design. The objective is to emphasize the higher-order terms in each model. Quantal Linear is <b>not</b> the same as Multistage-1. This is important for model averaging purposes.
Gamma $p(\text{dose}) = g + \frac{1-g}{\Gamma(\alpha)} \int_0^{\beta \text{dose}} t^{\alpha-1} \exp(-t) dt$	$0 \leq g < 1$ $\alpha > 0.2$ $\beta > 0$	$\text{logit}(g) \sim \text{Normal}(0, 2)$ $\alpha \sim \text{Lognormal}(\ln(2), \sqrt{0.18})$ $\beta \sim \text{Lognormal}(0.1)$	The prior for $\alpha$ entails that there is only a 0.05 prior probability the power parameter will be less than 1. This allows for models that are supra-linear; however, it requires a large amount of data for the $\alpha$ parameter to go much below 1. The $\alpha$ parameter is also constrained to be greater than 0.2, for numerical reasons.
Logistic $p(\text{dose}) = \frac{1}{1 + \exp[-\alpha - \beta \text{dose}]}$	$-20 < \alpha < 20$ $\beta > 0$	$\alpha \sim \text{Normal}(0, 2)$ $\beta \sim \text{Lognormal}(0.1, 1)$	
Log-Logistic $p(\text{dose}) = g + \frac{1-g}{1 + \exp[-\alpha - \beta \ln(\text{dose})]}$	$0 \leq g < 1$ $-40 < \alpha < 40$ $\beta > 0$	$\text{logit}(g) \sim \text{Normal}(0, 2)$ $\alpha \sim \text{Normal}(0, 1)$ $\beta \sim \text{Lognormal}(\ln(2), 0.5)$	
Probit $p(\text{dose}) = \Phi(\alpha + \beta \text{dose})$	$-8 < \alpha < 8$ $\beta > 0$	$\alpha \sim \text{Normal}(0, 2)$ $\beta \sim \text{Lognormal}(0.1, 1)$	$\Phi(\cdot)$ is the standard normal cumulative density function

Model	Constraints	Priors	Notes
Log-Probit $p(\text{dose}) = g + (1 - g)\Phi[\alpha + \beta \ln(\text{dose})]$	$0 \leq g < 1$ $-8 < \alpha < 8$ $\beta > 0$	$\text{logit}(g) \sim \text{Normal}(0, 2)$ $\alpha \sim \text{Normal}(0, 1)$ $\beta \sim \text{Lognormal}(\ln(2), 0.5)$	$\Phi(\cdot)$ is the standard normal cumulative density function
Dichotomous Hill $p(\text{dose}) = g + \frac{v(1-g)}{1 + \exp[-a - b \ln(\text{dose})]}$	$0 \leq g < 1$ $-40 < a < 40$ $b > 0$ $0 < v \leq 1$	$\text{logit}(g) \sim \text{Normal}(-1.2)$ $a \sim \text{Normal}(-3, 3.3)$ $b \sim \text{Lognormal}(\ln(2), 0.5)$ $\text{logit}(v) \sim \text{Normal}(0, 3)$	

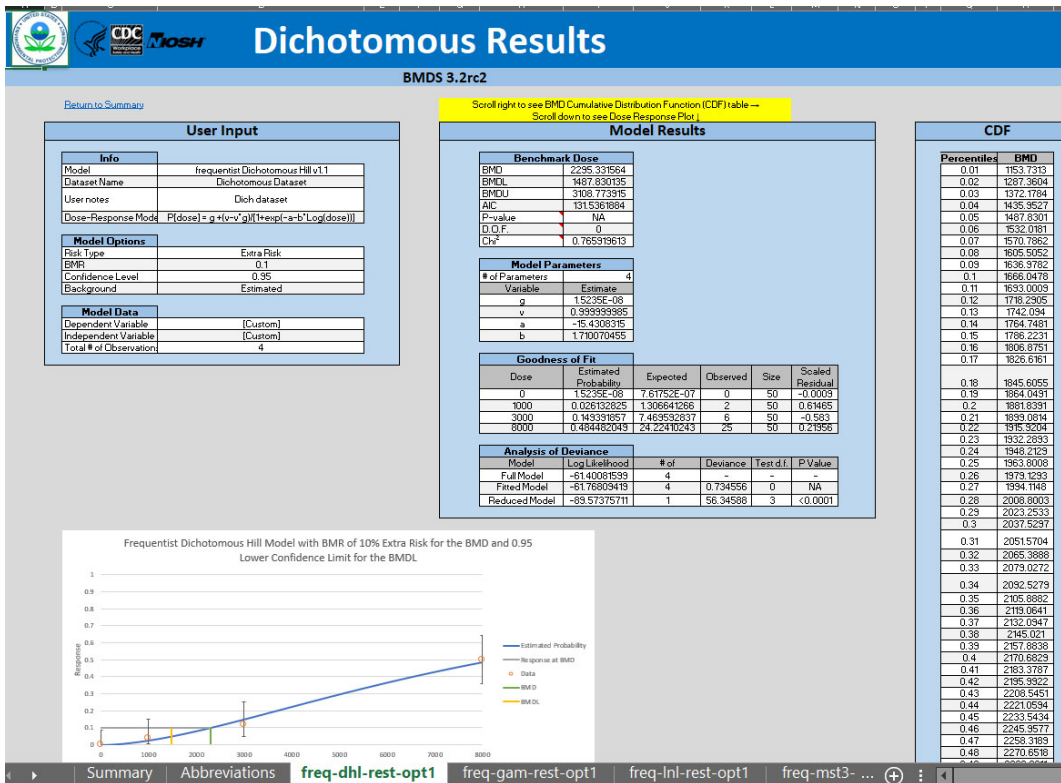
Notes:  $\text{logit}(g) = \ln\left(\frac{g}{1-g}\right)$ .  $\text{Normal}(x, y)$  denotes a Normal distribution with mean  $x$  and standard deviation  $y$ .  $\text{Lognormal}(w, z)$  denotes a lognormal distribution with log-scale mean  $w$  and log-scale standard deviation  $z$ .

데이터세트의 관측치가 증가함에 따라 베이지안 접근법과 빈도주의 접근법에서 얻은 매개변수와 BMD 사이의 양적 차이가 적어야 한다.

데이터가 적으면 priors가 베이지안 추정에 영향을 미친다. 데이터가 "hockey-stick" 용량-반응 관계를 제시하거나 강력한 초선형 행동을 시사할 때 영향이 가장 두드러질 수 있다. 이러한 경우에 베이지안 접근 방식에 대해 위에서 지정한 priors 설정은 기울기 변화가 더 점진적인 더 부드러운 용량-반응 관계를 얻기 위해 매개변수 추정치를 "shrink back(축소)"하는 경향이 있다.

#### 4) Frequentist 이분형 모델의 특이적인 출력

Figure 61. Sample Results Workbook tab for a dichotomous model run.



## (1) Fit table의 좋은점

Figure 62. Goodness of Fit

Goodness of Fit					
Dose	Estimated Probability	Expected	Observed	Size	Scaled Residual
0	1.5235E-08	7.61752E-07	0	50	-0.000873
1000	0.026132825	1.306641266	2	50	0.6146529
3000	0.149391857	7.469592837	6	50	-0.58302
8000	0.484482049	24.22410243	25	50	0.2195627

개별 모델 결과의 이 표는 데이터 목록("Observed" 및 "Size")과 모델 추정 반응 확률 및 해당 예상 응답자 수를 보여준다. 이것은 사용자가 BMD 표("5. 모델 유형의 공통 출력물-2)BMD 평가와 key fit 통계" 참조) 및 이탈 분석 테이블("7.이분형 자료의 종말점-4)Frequentist 이분형 모델에 대한 특이적인 출력-(1)Fit table의 좋은점" 참조)에 보고된 전체 적합도 통계와 함께 모델의 적절성을 판단하기에 좋다. 모델이 잘 맞는다면 관찰되고 예상되는 응답자 수가 상대적으로 비슷해야 한다.

표 끝에 인쇄된 조정된 잔차값(scaled residual value)은 다음과 같이 정의된다.

$$\frac{(\text{Obs} - \text{Expected})}{SE}$$

여기서 "Expected(예상)"은 모델에서 예측된 응답자의 수이고 SE는 해당 예측된 수에서 추정된 표준 오차이다. 이 모델의 경우 추정된 표준 오차는 다음과 같다.

$$\sqrt{[n \times p \times (1 - p)]}$$

n은 표본 크기(이 표에서 "size")이고, p는 모델 예측 응답 확률이다(이 표의 "Estimated probability(추정 확률)").

데이터에 대한 모델의 순응도(adherence)는 임의의 용량군, 특히 대조군 또는 BMD에 가장 가까운 용량을 가진 군에 대한 척도 잔차 값이 2보다 크거나 -2보다 작은 경우 문제가 될 수 있다.

## (2) 이탈 표(deviance table)의 분석

**Figure 63.** Analysis of Deviance

Analysis of Deviance					
Model	Log Likelihood	# of Parameters	Deviance	Test d.f.	P Value
Full Model	-61.40081599	4	-	-	-
Fitted Model	-61.76809418	3	0.73455639	1	0.3914103
Reduced Model	-89.57375711	1	56.3458822	3	<0.0001

이탈도 분석(analysis of deviance) 표에는 세 가지 로그 우도 값이 표시된다.

- 첫 번째는 "full model(전체 모형)"이다. Full model은 각 용량 그룹에 대해 개별적이고 독립적인 반응 확률을 가정한다. 이러한 확률을 제한하는 용량-반응 함수는 없다. 표시된 로그 우도는 주어진 데이터세트에 대해 달성할 수 있는 최대값이다. Full model에 대한 매개변수의 수는 용량군의 수와 동일하다(각각 고유한 독립적 확률 매개변수가 있음).
- 두 번째 로그 우도는 "fitted model(적합된 모형)"에 대한 것이다. 고려 중인 모델에 대해 얻을 수 있는 최대 로그 우도 값이다. 매개변수 추정값 표에 표시된 값과 동일한 매개변수 세트가 있는 모델에 해당한다. 매개변수의 수는 "Bounded"로 보고되지 않은 해당 표의 매개변수 수와 같다.
- 마지막 로그 우도값은 "reduced model(축소된 모형)"에 대한 것이다. 모든 용량군에 동일한 반응 확률이 적용된다고 가정할 경우 얻을 수 있는 최대 로그 우도이다. Reduced model, 즉 가정된 일정한 응답 확률에 대한 매개변수는 1개뿐이다.

이 세 가지 모델 각각에는 편차, 자유도(검정 d.f.) 및 P-값의 세 가지 값이 연결되어 있다. 이탈도는 fitted 또는 reduced 모형과 전체 모형 로그 우도값

간의 차이의 두 배이다. 이 편차는 또 다른 적합도 메트릭(goodness-of-fit metric)이다. 편차가 작으면 "smaller(더 작은)" 모형(즉, fitted 또는 reduced 모형)이 전체 모형과 거의 마찬가지로 데이터를 설명한다. 편차는 "검정 d.f"로 지정된 자유도를 갖는 대략 카이제곱 확률 변수이다. 이는 비교되는 두 모델에 대한 매개변수 수의 차이이다. "P-값"은 적합도 차이의 중요성을 평가하기 위해 이 카이제곱 근사값의 사용을 반영한다. 편차가 클수록 더 작은 p-값에 대응하므로 작은 p-값은 더 작은 모델이 전체 모델에 적합하지 않다는 것을 나타낸다. 사용자는 모델 적합도가 적절한지 테스트하기 위해 기각 수준(0.05가 일반적임)을 선택할 수 있다.

Fitted 모형의 경우 데이터에 대한 모델의 fitting에 대한 또 다른 측정이다. Reduced 모형의 경우 해당 모델을 기각하지 못하면(사용자가 선택한 기각 수준보다 큰 p-값) 사용자는 반응 확률에 용량 관련 영향이 없다고 추론할 수 있다.

### (3) 제한된 다단계 모델에만 추가

**Figure 64.** Slope Factor (last row in table) appears only on Restricted Multistage Model results.

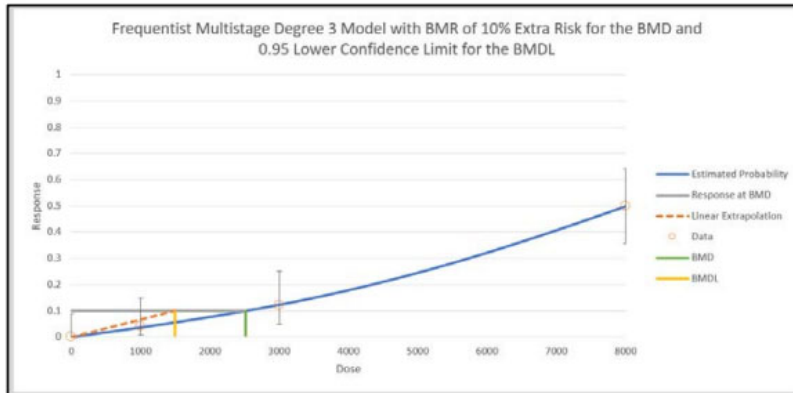
Benchmark Dose	
BMD	2512.859821
BMDL	1492.16541
BMDU	3755.154706
AIC	128.8193162
P-value	0.893543819
D.O.F.	1
Chi <sup>2</sup>	0.017908162
Slope Factor	6.70167E-05

고려 중인 모델이 다단계 모델이고 "제한된" 실행된 경우 몇 가지 추가 평가 도구가 제공된다. 이것은 암 데이터 모델링을 위한 EPA 표준 설정이지만 다단계 모델이 이러한 방식으로 실행될 때 이분형 종말점에 대해 이러한 추가 결과가 표시된다.

제한된 다단계 모델에 대한 BMD 표에는 BMDL(10)에서 extra risk (0.1)과 백그라운드(일반적으로 0 용량)에서 extra risk(0) 사이의 선형 기울기로 EPA에서 정의한 기울기 계수 추정값이 포함된다.

제한된 다단계 모델 플롯에는 이 선형 기울기를 나타내는 점선도 포함되어 있다.

**Figure 65.** The dashed line for the Multistage model plot representing linear slope.

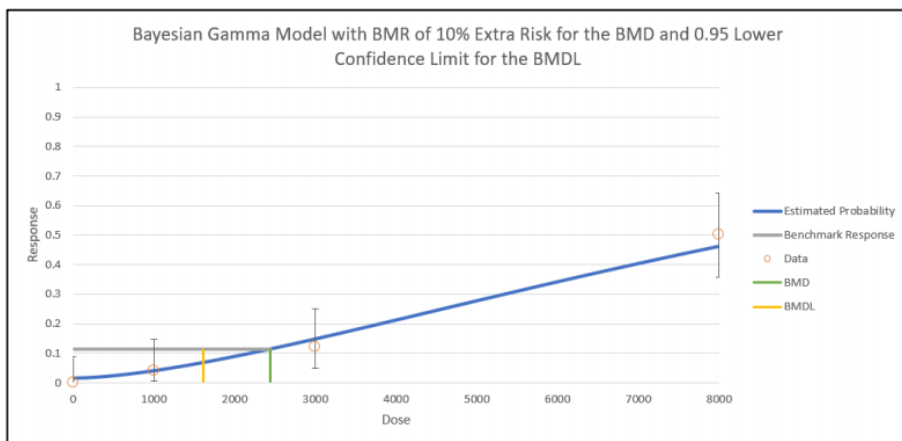


용량 단위가 mg/kg-day인 경우 이는 IRIS에서 정의한 OSF(oral slope factor, 구강 기울기 계수)와 같다. 용량 단위가  $\mu\text{g}/\text{m}^3$ 인 경우 이는 IRIS에서 정의한 흡입 단위 위험도(IUR)와 같다. 자세한 내용은 통합 위험 정보 시스템(IRIS) 웹 페이지에 대한 기본 정보의 "IRIS Toxicity Values" 섹션을 참조한다.



## 5) Bayesian 이분형 모델에 대한 특이적인 출력

Figure 66. Sample Bayesian dichotomous results plot.



두 베이지안 모델 간의 차이를 비교하기 위해 정규화되지 않은 로그 사후 확률(LPP)이 제공되며, 이를 통해 베이즈 인수(BF)를 계산하여 두 모델을 비교할 수 있다. BF는 두 LPP 간의 지수 차이와 같다. 예를 들어, 로그-로지스틱 모델 (모델 A)( $LPP_A$  산출)과 다단계 2차 모델(모델 B,  $LPP_B$  포함)을 비교하려면 BF를 다음과 같이 추정한다.

$$BF = \exp(LPP_A - LPP_B)$$

이 계산은 두 모델이 *prior*에서 동일한 확률을 갖는다고 가정한다. 이 값은 한 모델이 다른 모델보다 더 정확하고 베이지안 가설 검정에 사용되는 사후 확률(posterior odds)로 해석된다. 위의 예에서 Bayes Factor가 2.5인 경우 Log-logistic 모델이 다단계 모델보다 posterior가 2.5배 더 높다고 해석할 수 있다. 이러한 값이 적절한 확률로 정규화되면 모델 평균화에 제공된 posterior 모델 확률과 동일하다(다시, 동일한 모델 확률을 사전에 가정). 표13에는 Jeffreys 등이 제시한 Bayes Factors의 일반적인 해석을 나타내었다.

**Table 13.** Bayes factors for dichotomous models.

Bayes Factor	Strength of Evidence for $H_A$
< 1	negative (supports $H_0$ )
1 to 3.2	not worth mentioning
3.2 to 10	substantial
10 to 31.6	strong
31.6 to 100	very strong
100	decisive

BMDs 3.0의 경우 모든 LPP 및 해당 사후 모델 확률은 Laplace approximation을 사용하여 계산한다. 이 값은 일반적으로 사용되는 BIC(Bayesian Information Criterion)와 다르며 BIC를 독점적으로 사용하는 다른 모델 평균화 접근 방식을 기반으로 두 값을 혼동해서는 안된다. BIC에서 추정된 사후 확률의 오류는  $O(1)$  추정기(estimators)이다. Laplace approximation을 사용하여 추정된 사후 확률의 오류는  $O(n^{-1})$  이다. 즉, 데이터가 증가함에 따라 후자의 근사값은 실제 사후 모델 확률로 가고 전자는 BIC를 사용하여 실제 값이 되지 않을 수 있다.

## 8. 다중 종양 분석(Multiple tumor analysis)

Figure 67. Multi-tumor (MS\_Combo) analysis.

The screenshot shows the 'BMDs 3.2' software interface with the 'Define Analysis' window open. The window contains the following fields and controls:

- Analysis Name:** MT tumors
- Select Output Directory:** C:\Users\mbrown06\Environmental Protection Agency (EPA)\EMVL Projects -
- Analysis Description:** Comparing 3 datasets
- Select Model Type:** Dichotomous - Multi-tumor (MS\_Combo)
- Buttons:** Load Analysis, Save Analysis, Run Analysis

Below these fields are two tables:

Enable	Datasets	Degree	Background
<input checked="" type="checkbox"/>	tumor 1	[auto-select] [E estimated]	
<input checked="" type="checkbox"/>	tumor2	[auto-select] [2ans]	
<input checked="" type="checkbox"/>	tumor3	[1] [E estimated]	

Option Set #	Risk Type	BMR	Confidence Level
1	Added Risk	0.1	0.95
2	Extra Risk	0.1	0.95

Below the second table is an 'Add Option Set' button.

### 1) 이분형 자료: multi-tumor models과 옵션

모델링 옵션 'Dichotomous -Multi-tumor (MS\_Combo)'는 이전에 논의되지 않은 이분형 모델링의 특수 응용 프로그램이다. 두 개 이상의 독립적이고 이분형 반응의 결합 효과를 평가하는 데 관심이 있는 사용자에게 편의를 제공한다. 용량-반응 관계에 맞게 다단계 모델만 실행한다는 것에 특화되어 있다. 문제의 종말점 중 하나 이상에 반응할 위험에 대한 BMD 추정값(및 관련 범위)을 반환한다. 구성 데이터세트가 다양한 위치에서 발생하는 종양인 암 데이터 분석에 가장 자주 사용되므로 그렇게 명명되었다.

BMDs의 이전 버전에서와 같이 BMDs 3는 사용자가 EPA의 다종양(MS\_Combo) 모델을 실행하여 여러 종양 유형 중 하나 이상을 경험할 특정 added 또는 extra risk와 관련된 BMD, BMDL 및 BMDU를 결정할 수 있도록 한다. 그러나 BMDs의 이전 버전과 달리 BMDs 3는 사용자 가 수동으로 선택하거나 BMDs가 각 데이터세트에 적용할 다단계 모델의 정도를 "자동 선택"하도록

허용하는 옵션을 제공한다. 자동 선택 프로세스는 다른 모델링 시나리오에 대해 EPA에서 설명한 모델 선택 프로세스와 다른 암 데이터세트 분석을 위한 다단계 모델 정도를 선택하기 위한 최신 EPA 기술 지침을 따른다.

## 2) 추정(assumptions)

여러 종양에 대한 분석은 다음과 같은 가정과 결과를 가지고 있다.

1. 종양은 통계적으로 서로 독립적이다.

[참고] 종양 유형이 독립적이지 않다는 것을 나타내는 상당한 생물학적 증거가 없는 한(모델 매개변수 값에 대한 조건부) 독립성을 기반으로 한 접근 방식이 적절한 것으로 간주된다.

2. 다단계 모델은 각 종양에 대해 개별적으로 적절한 모델이다. 그러나 개별 종양에 적합한 개별 다단계 모델이 동일한 다항식 차수를 가질 필요는 없다.
3. 사용자가 분석 중인 하나 이상의 종양에 걸릴 위험을 추정하는 데 관심이 있다. 결과는 사용자 정의 BMR 수준과 관련된 BMD 및 BMDL을 나타내며, 여기서 BMD 및 BMDL은 "combination"(하나 이상의 종양 발생)에서 BMR과 동일한 extra risk를 제공하는 것으로 추정되는 용량의 최대 우도 및 하한 추정치이다.

EPA 암 지침에 따라 Multiple Tumor Analysis는 항상 제한된 형태의 다단계 모델을 실행한다.

BMDs 3을 사용하면 사용자가 각 종양 데이터세트에 대해 다단계 모델의 적절한 다항식 정도를 BMDs "Auto-Select"할 수 있다. "Auto-Select" 기능을 사용하면 BMDs는 다단계 모델의 모든 관련 형식을 실행하고 종양 분석을 위한 현재 EPA 다단계 모델 선택 기준을 기반으로 사용할 다항식 정도를 선택한다. 이것은 BMDs 3.0의 기본 옵션이지만 사용자는 각 데이터세트에 대해 다항식 차수를 수동으로 설정하도록 선택할 수도 있다. 어떤 경우든, 모델링 매개변수에

대한 다항식 및 기타 선택의 정도가 분석 중인 데이터 세트에 대해 원하는 대로 적절한지 확인하는 것은 궁극적으로 사용자의 책임이다.

### 3) Multi-tumor(MS\_Combo) 모델 설명

[참고] MS\_Combo를 사용하기 전에 암 모델링을 위한 다단계 모델의 적절한 단계 선택에 대한 기술 지침을 참조하는 것이 좋다. 기술 지침에는 BMDs MS\_Combo 프로그램의 가정 및 적용에 대한 배경 정보가 포함되어 있다.

BMDs에서 MS\_Combo 프로그램의 목적은 개별 종양의 용량-반응이 다단계 암 모델로 모델링되었을 때 사용자가 종양 조합(하나 이상의 종양이 발생할 정의된 위험에 해당)에 대한 BMDs와 BMDL을 계산할 수 있도록 하기 위함이다.

MS\_Combo 실행의 출력은 각 개별 종양(해당 종양에 대한 BMD 및 BMDL 포함)을 비롯하여 특정한 종양 반응의 조합에 대한 결합된 로그 우도, BMD 및 BMDL을 맞춘 결과를 표시한다.

실제로 사용자는 각 종양을 개별적으로 조사하고 각 개별 종양에 가장 적합한 다단계 암 모델의 정도를 결정해야 한다. 그 결정에는 적합성, AIC 등의 모든 일반적인 고려 사항이 포함된다.

다단계-암 모델의 특정 형태가 관심 있는 각 종양에 대해 선택되면(해당 종양 전체에 걸쳐 동일한 정도가 필요하지 않음) 사용자는 MS\_Combo 실행에서 해당 선택을 지정해야 한다.

[참고] 다음 설명은 종양이 서로 독립적인 것으로 가정되는 경우에만 유효하다(용량에 대해 조건부).

다단계 모델의 형식 때문에 결합된 위험에 대한 MLE 추정치는 개별 종양 다단계 모델 적합에 대해 얻은 매개변수 값의 함수이다. 실제로 결합 확률 함수는 다단계 모델 형식을 갖는다.

$$P(d) = 1 - e^{-(\beta_0 + \beta_1 d + \beta_2 d^2 + \dots)}$$

결합 확률 함수( $\beta_0, \beta_1, \dots$ )의 항은 다음과 같이 지정된다.

$$\beta_0 = \sum \beta_{0i}$$

$$\beta_1 = \sum \beta_{1i}$$

$$\beta_2 = \sum \beta_{2i}$$

합계가  $i = 1, \dots, t$  이상인 경우

$t$ 는 고려 중인 종양의 수이고,

$\beta_{xj}$ 는 종양  $j$ 에 대한  $x$ 번째 매개변수( $0, 1, \dots$ )이다.

$\beta_{xj}$  값은 개별 종양에 대해 수행된 다단계 암 실험에서 직접 사용할 수 있지만 MS\_Combo는 사용자를 위해 계산을 수행하여 개별 용어의 합계를 완료하고 결합된 매개변수 값과 사용자 지정 BMR을 기반으로 BMD를 계산한다.

Profile-likelihood(프로필 우도) 접근 방식은 BMDL을 도출하는 데 사용된다.

1. MLE 솔루션과 관련된 BMD 및 로그 우도(log-likelihood)가 주어지면 사용자가 지정한 신뢰 수준(예: 95%)을 기반으로 목표 가능성이 정의된다.
2. 그 목표 가능성은 사용자가 지정한 신뢰 수준(실제로는 신뢰 수준과 관련된 알파에 2를 곱한 값)에 해당하는 카이 제곱(자유도 1)의 백분위수로부터 파생된다.
3. 해당 백분위수를 2로 나누고 최대 로그 우도에서 뺀다.

4. 그러한 도출은 자유도가 1인 우도 비율(likelihood ratio) 검정을 기반으로 한다. BMDL을 추정하는 것은 결합되는 종양의 수에 관계없이 자유도 1을 잃는 것에 해당한다는 것을 보여준다.
5. 결합 반응(관심 종양 중 하나 이상)에 대한 BMDL은 다음 두 조건이 충족되는 최소 용량 D로 정의된다.
  - i. D를 사용하는 결합된 로그 우도와 해당 매개변수가 목표 우도보다 크거나 같도록 매개변수 세트가 있으며,
  - ii. 해당 매개변수 세트의 경우 D에서의 위험은 사용자가 지정한 BMR과 같다.

결합된 로그 우도는 종양별  $\beta$ 값을 사용하여 얻은 개별 종양의 적합도(fit, 개별 로그 우도의 합) 함수이다. 따라서 결합된 다단계-암 모델의 매개변수에 대한 검색은 개별 로그 우도가 원하는 범위(목표 이상) 내에서 결합 가능성까지 합산되는 방식으로 개별 종양  $\beta$ 값을 변경한다.

그러나 두 번째 제약 조건을 충족하기 위해 개별 종양 매개변수의 합(위에서 결합 확률 함수의 매개변수로 표시됨)을 사용하여 제안된 BMDL, D에 대한 위험을 평가한다.

개별 종양은 다단계 암 모델의 동일한 정도를 사용하여 모델링할 필요가 없다. 개별 종양에 포함되지 않은 모든 용어는 위에 표시된 합계에서 0으로 가정된다(BMDL 최적화 동안 0으로 유지됨).

#### 4) Multi-tumor 데이터 입력

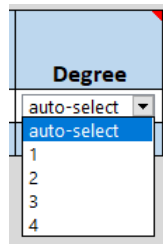
Figure 68. Dataset options for multi-tumor data.

<input checked="" type="checkbox"/> Enable	DataSets	Degree	Background
<input checked="" type="checkbox"/>	tumor 1	auto-select	Estimated
<input checked="" type="checkbox"/>	tumor2	auto-select	Zero
<input checked="" type="checkbox"/>	tumor3	1	Estimated

사용자는 데이터 세트당 (다항식) Degree(차수) 및 Background의 두 가지 옵션을 설정할 수 있습니다.

### (1) Polynomial degree(다항식 차수) 세팅

**Figure 69.** Setting Polynomial Degree



BMDs가 적절한 다항식 차수를 자동 선택할 것인지 또는 사용자 지정할 것 인지를 선택한다.

BMDs는 Logic 탭에 있는 의사결정 로직 (decision logic) 및 설정 (setting)을 기반으로 모델 차수를 권장한다. 자세한 내용은 “10. BMDs 권장 사항 및 의사결정 로직”을 참조한다.

- 사용자가 BMDs의 auto-select를 선택하면 Technical Guidance on choosing the appropriate stage of a multistage model for cancer modeling(암 모델링을 위한 다단계 모델의 적절한 단계 선택에 대한 기술 지침)에 따라 가장 적합한 모델이 선택된다. 해당 기준에 따라 모델을 선택할 수 없으면 MS\_Combo 결과에서 모델이 제거된다.
- MS\_Combo 결정 로직은 다음 기준에 대해 Logic 탭의 사용자 정의 검정 임계값(user-defined test thresholds)을 사용한다.
  - A. Goodness of fit p-test(cancer)
  - B. Ratio of BMD/BMDL(Cautious)
  - C. Abs(Residual of Interest) too large



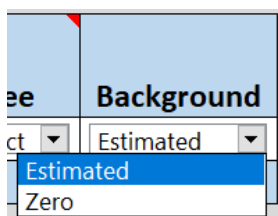
D. Abs(Residual at control) too large

E. 결과가 Logic 탭의 "Ratio of BMD/BMDL (Caution)"에서 설정한 "Test Threshold" 값을 충족하지 않는 경우 BMDS는 "BMD/BMDL ration > [value]"이라는 팝업 메시지를 사용자에게 표시한다. 이런 경우에는 통계학자와의 상담을 고려하는 것이 좋다. 그러나 그 모델은 MS\_Combo 결과에서 삭제되지 않는다.

- 사용자가 주어진 모델에 대해 차수를 지정하면 해당 차수는 모델 적합성에 관계없이 사용된다.

## (2) 백그라운드

Figure 70. Background



선택 사항은 Estimated(기본값) 또는 Zero이다. 이는 일반적으로 고려 중인 화학 물질에 대한 노출이 없을 때 반응 확률이 절대적으로 0이라는 강력한 증거가 존재하지 않는 한 Estimated를 선택해야 한다.

[참고] 대조군에 반응이 있는 경우 백그라운드를 0으로 설정하지 않는다.

## 5) 옵션

Figure 71. Options

Option Set #	Risk Type	BMR	Confidence Level
1	Extra Risk	0.1	0.95

Add Option Set

다종양 옵션은 이분형의 옵션과 동일하다.

### (1) 위험 유형(risk type)

Figure 72. Risk Type

Option Set #	Risk Type
1	Extra Risk

Extra Risk  
Added Risk

선택 사항은 "Extra Risk"(기본값) 및 "Added Risk"이다.

Added risk는 용량이 있는 경우 반응하는 총 동물의 추가 비율 또는 용량d, P(d)에서 예상되는 반응 확률에서 노출이 없을 때의 예상 반응 확률을 뺀 것이다.

$$P(0).I.e.added\ risk = P(d) - P(0)$$

Extra risk는 추가 위험을 노출이 없을 때 반응하지 않을 동물의 예측 비율로 나눈 값이다.

$$1 - P(0).I.e.,extra\ risk = \frac{P(d) - P(0)}{1 - P(0)}$$

## (2) BMR

Figure 73. BMR

<b>BMR</b>
0.1

BMR은 BMD가 추정되는 위험 값(사용자가 지정한 extra 또는 added)이다. BMR은 0에서 1 사이의 값이어야 한다(1은 미포함).  $P(0) > 0$ 인 경우 BMR 값이  $1 - P(0)$ 보다 크면 위험 유형이 added risk일 때 오류가 발생한다. 달성할 수 있는 최대 added risk가  $1 - P(0)$ 이기 때문이다. 실제로는 사람들은 일반적으로 0.01에서 약 0.10 사이의 BMR 값에 관심이 있기 때문에 문제가 되지 않는다.

## (3) 신뢰수준(confidence level)

Figure 74. Confidence level

<b>Confidence Level</b>
0.95

신뢰 수준은 0과 1 사이의 실수이다. EPA는 0.95를 권장한다.

## 6) 분석수행 및 결과 보기

"Run Analysis"를 선택하면 다중 종양 결과의 별도 Results Workbook이 생성된다. Workbook에는 개별적으로 고려된 각 개별 종양에 대한 결과(선택한 데이터 세트별 옵션 사용)와 사용자가 지정한 위험 유형, BMR 및 신뢰 수준에 대한 결합된 종양 확률에 대한 BMD 및 BMDL의 해당 추정치가 포함된다.

개별 다단계 모델 실행에 대한 플롯은 개별 모델 결과 탭에 표시된다. Multistage polynomial degree(다단계 다항식 차수)를 선택하기 위해 "Auto-Select" 기능을 사용한 경우 사용자는 결과 모델 핏(fit)이 원하는 용량-

반응 영역에 적절한지 확인해야 한다. 사용자가 다른 다단계 다항식 차수를 적용하려면 "Auto-Select" 대신 지정된 차수를 사용하여 분석을 다시 실행할 수 있다.

Results Workbook 표에 대한 자세한 내용은 “7. 이분형 자료의 종말점-4) 빈도주의적 이분형 모델에 대한 특이적인 출력”을 참조한다.

## 7) 종양분석의 문제해결책(troubleshooting)

하나 이상의 종양이 시험했던 최고 용량보다 3배이상 큰 BMD를 갖는 것으로 추정되는 경우(해당 종양에 대해) 다중 종양 분석은 중간 지점(intermediate point), 즉 해당 종양에 대한 피팅(fitting)이 완료된 후 중단되고 해당 BMD의 크기가 결정된다. 해당 종양 아래에 나열된 종양은 분석되지 않으며 조합도 완료되지 않는다.

다른 종양의 BMD가 최고 용량의 3배 미만이라면 해당 종양이 종양 조합에 대한 BMD 추정에 실질적으로 추가되지 않는 경우가 있을 수 있다. 문제의 종양에 대한 반응의 크기가 높은 수준의 노출 시의 BMR 수준에 도달하지 않았기 때문에 분석 대상 종양 중 하나 이상이 발생할 위험에 대한 기여도가 다른 종양에 비해 작기 때문일 것이다. 사용자는 문제의 종양이 포함되지 않은 조합을 시도할 수 있다.

## 9. 고려해야 할 점

### 1) 연속형 모델의 AIC

서로 다른 우도(즉, 정규 대 로그 정규)가 있는 모델을 쉽게 비교할 수 있도록 정규 및 로그 정규 분포에 대한 로그 우도는 모든 정규화 상수를 사용하여 계산된다. 그 결과 이전 BMDS 버전에 제공된 것과 다른 수치적 AIC 값이 생성된다.

연속형 모델에 대한 BMDS 3 AIC 값이 BMDS 2.x 버전의 값과 다르더라도 모델의 기본 분포가 동일한 경우 AIC의 차이는 이전 버전의 BMDS와 동일하다. 이것은 BMDS 3 및 BMDS 2.x 모델 핏이 비교되는 두 모델에 대해 동일하다고 가정한다. AIC 차이는 두 버전 간에 하나 이상의 모델 핏이 다른 경우 동일하지 않을 수 있다(예: 하나 이상의 BMDS 3 모델이 해당 BMDS 2.x 모델보다 데이터에 대한 개선된 핏을 제공하는 경우).

그러나 모수 분포가 다른 모델을 비교할 때 AIC 차이는 이전 BMDS 버전과 동일하지 않다. 이러한 비교의 경우 BMDS 3 소프트웨어를 사용하여 계산된 AIC가 정확하며 기본 분포에 관계없이 두 모델 간에 적절한 비교가 이루어진다.

## [주의]

로그 정규 분포 반응을 가정할 때에는 로그 척도 모수(log-scale parameters)가 근사화되는 표본 평균과 표본 표준 편차(요약된 데이터)만 사용할 수 있는 상황에 대한 주의가 필요하다. 이러한 경우 로그 정규 로그 우도(lognormal log-likelihood)에 대한 정규화 상수(normalizing constant)는 근사값일 뿐이다. 로그 정규 분포 가정 하에서 모든 모델 적합에 대해 동일한 정규화 상수이므로 해당 가정을 사용하는 모델 간의 비교가 유효하다.

그러나 정규성을 가정하여 한 세트의 결과를 얻은 결과와 로그 정규성을 가정한 한 세트를 얻은 결과의 비교는 주의해야 한다. AIC가 "similar(유사한)" 경우(여기에서 특정 지침을 제공할 수 없기 때문에 해당 용어를 느슨하게 사용) AIC 차이를 기반으로 모델 선택을 해서는 안 된다. AIC 차이가 "larger(더 큰)" 경우에는 모델 선택에 문제가 없을 것이다. 이는 사용된 근사값(approximation)이 너무 나쁘지 않아야 하기 때문이다.

보수적인 입장은 기본 데이터가 요약된 형식으로 표시되는 경우(즉, 표본 평균과 표본 표준 편차만 사용할 수 있는 경우) 정규 분포를 가정하는 모델과 대수 정규 분포를 가정하는 모델을 비교할 때 AIC를 사용해서는 안 된다는 것이다.

[참고] 이분형 모델의 AIC 값은 BMDS 2.x에서 BMDS 3.x까지 동일해야 한다.

## 2) 음수의 평균값(negative means)을 이용한 연속형 반응 데이터

평균이 음수인 데이터는 상수 분산 모델(constant variance model)로만 모델링해야 한다.

변환된 데이터를 모델링할 때 사용자가 음수 데이터를 모델링해야 하는 경우가 있다. 이 경우 사용되는 변환(transformation)은 분산 안정화 변환(variance-stabilizing transformation)이어야 일정-분산 모델(constant-variance model)이 적합하다.

표준 편차 기반 BMR을 사용하여 BMD 계산을 정의하는 경우 모든 관측치(또는 평균)에 상수를 추가하여 값(평균)을 양수로 만들 수 있다. 이렇게 하면 관측값의 표준 편차가 바뀌지 않으며 사용자가 분산을 모델링할 수 있다.

## 3) 동일한 종말점을 위해 두 개의 데이터세트를 결합한 검정

BMDs에는 공변량 값(covariate value)(예: 종 또는 성별과 같은 등급 변수(class variance) 간)에 걸친 용량 반응 유사성에 대한 공식적인 검정은 포함되어 있지 않다. EPA의 범주형 회귀 소프트웨어인 CatReg에는 이러한 기능이 있다.

그러나 사용자가 결합을 고려하고 있는 두 실험에 대한 용량-반응 데이터가 있는 경우(예: 한 종 내의 두 성별 또는 두 종 등) BMDs에서 다음 절차를 사용할 수 있다.

Step 1. 두 데이터세트에 적용할 단일 모델을 선택한다.

Step 2. 두 데이터세트를 개별적으로 모델링한다. 각 실행에 대해 다음을 기록한다.

A. 각 데이터세트에 대한 최대 로그 우도. 합산된 로그 우도를 얻기 위해 두 개의 로그 우도(각 데이터세트에서 하나씩)를 추가한다.

- B. 각 데이터세트에 대한 제약 없는 매개변수의 수. 합산된 비제약 매개변수를 얻으려면 각 실행에서 해당 숫자를 추가한다.

Step 3. 두 실험의 데이터를 결합하고 함께 모델링하며 다음을 기록한다.

- A. 결합된 데이터세트에 대한 최대 로그 우도. 이것은 결합된 로그 우도가 된다. Fitting된 모델 로그 우도는 Analysis of Deviance(이분형 종말점) 또는 Likelihoods of Interest(연속형 종말점) 표에 기록된다.
- B. 결합된 데이터세트에 대한 제약 없는 매개변수의 수. 이것은 결합되지 않은 매개변수가 된다.

Step 4. 합산된 로그 우도에서 결합된 로그 우도를 뺀다. 그런 다음 차이에 2를 곱한다.

Step 5. Step 4의 값을 카이제곱 분포와 비교한다. 해당 카이 제곱 분포의 자유도는 합산된 비제약 모수(unconstrained parameters, Step 2)와 결합된 비제약 모수(Step 3) 간의 차이가 된다. Step 4의 값이 문제의 카이 제곱 분포의 꼬리에 있는 경우(즉, 95번째 백분위수보다 큼), 두 세트가 동일한 용량-반응 관계를 갖는다는 귀무 가설을 기각한다. 귀무 가설이 기각되면 두 데이터 세트를 결합하는 것이 적절하지 않다고 추론한다.



## 10. BMDS 권장 및 결정 로직

BMDS Wizard 응용 프로그램에 익숙한 사용자는 BMDS 3이 2012 EPA Benchmark Dose Technical Guidance와 일치하는 모델 선택과 관련하여 모델링 결과를 분석하고 자동 권장 사항을 만드는 유사한 접근 방식을 사용한다는 점에 주목할 것이다.

이들 기준은 Figure 75와 같이 BMDS 3 Analysis Workbook의 Logic 탭에서 변경할 수 있다. Decision logic을 켜거나 끌 수 있으며 특정 기준은 다양한 데이터세트 유형별로 활성화 또는 비활성화할 수 있다. 로직은 분석하려는 데이터 유형(연속형, 이분형, 등지형)에 따라 다르다.

Figure 75. BMDS 3 Logic tab with EPA default recommendation decision logic.

**Decision-Logic**

Recommend model in Viable Bin?	TRUE
Recommend model in Questionable Bin?	FALSE
BMDL range deemed "sufficiently close" to use lowest AIC instead of lowest BMDL in viable models:	3

**Model Recommendation/Bin Placement Logic**

Test Description	Test On/Off			Test Threshold (where appropriate)	Bin Placement if Test is Failed	Notes to Show
	Continuous	Dichotomous	Nested			
BMD calculated	On	On	On	NA	Unusable Bin	BMD not estimated
BMDL calculated	On	On	On	NA	Unusable Bin	BMDL not estimated
BMDU calculated	Off	Off	Off	NA	No Bin Change (Warning)	BMDU not estimated
AIC calculated	On	On	On	NA	Unusable Bin	AIC not estimated
Constant Variance	On	On	On	0.05	Questionable Bin	Constant variance test failed (Test 2 p-value < 0.05)
Non-Constant Variance	On	On	On	0.05	Questionable Bin	Non-constant variance test failed (Test 3 p-value < 0.05)
Goodness of fit p-test	On	On	On	0.1	Questionable Bin	Goodness of fit p-value < 0.1
Goodness of fit p-test (cancer)	On	On	On	0.05	Questionable Bin	Goodness of fit p-value < 0.05
Ratio of BMD/BMDL (serious)	On	On	On	20	Questionable Bin	BMD/BMDL ratio > 20
Ratio of BMD/BMDL (caution)	On	On	On	3	No Bin Change (Warning)	BMD/BMDL ratio > 3
Abs(Residual of interest) too large	On	On	On	2	Questionable Bin	Residual for Dose Group Near BMD  > 2
BMD higher than highest dose	On	On	On	1	No Bin Change (Warning)	BMD higher than maximum dose
BMDL higher than highest dose	On	On	On	1	No Bin Change (Warning)	BMDL higher than maximum dose
BMD lower than lowest dose (warning)	On	On	On	3	No Bin Change (Warning)	BMD 3x lower than lowest non-zero dose
BMDL lower than lowest dose (warning)	On	On	On	3	No Bin Change (Warning)	BMDL 3x lower than lowest non-zero dose
BMD lower than lowest dose (serious)	On	On	On	10	Questionable Bin	BMD 10x lower than lowest non-zero dose
BMDL lower than lowest dose (serious)	On	On	On	10	Questionable Bin	BMDL 10x lower than lowest non-zero dose
Abs(Residual at control) too large	On	On	On	2	No Bin Change (Warning)	Residual at control  > 2
Poor control dose std. dev.	On	On	On	1.5	No Bin Change (Warning)	Modeled control response std. dev. > 1.5 (actual response std. dev. d.f. = 0, saturated model (Goodness of fit test cannot be calculated))
D.O.F. equals 0	On	On	On	NA	Questionable Bin	

위에서 설명한 대로 사용자가 입력한 Decision logic을 기반으로 BMDS는 "recommended(권장)" 모델을 선택하려고 시도한다. 사용자는 궁극적으로는 모델을 선택해야 하며 BMDS 자동 결정에 동의하지 않을 수 있다.

BMDs 3는 Results Workbook 요약표 및 Word Report File 표의 "BMDs Recommendation" 및 "BMDs Recommendation Notes" 열에 제안 문구를 자동으로 생성한다.

Results Workbook에서 일부 형식의 수정(예: 행 높이, 열 너비, 플롯의 크기, 디자인 및 위치)는 허용되지만 텍스트 및 숫자 결과는 수정할 수 없다. 그러나 Word Report 파일은 광범위하게 수정할 수 있으며 사용자는 이러한 유연성을 활용하여 테이블 헤더와 모델이 선택된 이유에 대해 수정할 수 있다.

BMDs 3은 각 모델을 다음 3개의 빈(bin, 저장소) 중 하나에 넣는다.

- Viable(실행 가능) - 최고 품질의 모델, 사용자 정의 로직을 기반으로 심각한 결함이 발견되지 않았지만 경고가 포함될 수 있음
- Questionable(문제 있음) - 사용자 정의 decision logic을 기반으로 하는 모델에 몇 가지 심각한 결함이 있음
- Unusable(사용할 수 없음) - BMD 또는 BMDL과 같은 필수 출력이 추정되지 않음

Bin 배치를 결정하는 요인(검정)의 기본 설정은 EPA Benchmark dose Technical guidance(U.S. EPA, 2012)와 일치하며 "Reset to Default Logic (기본 로직으로 재설정)"을 선택하여 재설정할 수 있다.

검정 실패를 유발하여 bin 배치에 영향을 줄 수 있는 기본 설정을 아래와 같이 나타내었으나, EPA BMD 지침에서는 명시적으로 제공하지 않는다. 이는 일반적인 EPA 관행에 따라 정리하였으며 따라서 사용자 재량에 더 개방적이다.

- BMDL range default fail: > 3-fold  
(BMDL 범위 기본값 설정 실패: 3배 초과)
- constant and non-constant variance  $p$ -value<sup>10</sup> default fail: < 0.05  
(일정 및 일정하지 않은 분산  $p$ -값<sup>10</sup> 기본값 설정 실패: 0.05미만)

- Ratio of BMD/BMDL(serious) default fail: > 20  
(BMD/BMDL(심각한) 기본 실패 비율: > 20)
- BMD lower than lowest dose(serious) default fail: > 10  
(최저 복용량 보다 낮은 (심각한) BMD 기본값 실패: 10초과)
- BMDL lower than lowest dose(serious) default fail: > 10  
(최저 복용량 보다 낮은 (심각한) BMDL 기본값 실패: 10초과)

Option Set이 동일한(BMR 유형, BMRF 등) 모든 모델이 3개의 bin 중 하나에 배치된 후, 모델은 2012 EPA Benchmark Dose Technical Guidance(U.S. EPA, 2012)에 정의된 BMDL 또는 AIC criteria에 근거하여 “Viable” bin에서 추천된다.

## 1) Decision logic의 변경(changing)

BMDS는 2012년 EPA 지침(U.S. EPA, 2012년)과 이전 섹션에서 설명한 추가 기준(additional criteria)을 사용하여 가장 적합한(fitting) BMDS 모델을 자동으로 추천하려고 시도한다. 이러한 기준은 Logic worksheet에서 변경할 수 있다. Decision logic을 켜거나 끌 수 있으며 다양한 데이터세트 유형에 대해 특정 기준을 활성화하거나 비활성화할 수 있다.

사용자가 입력한 decision logic을 기반으로 BMDS는 가장 적합한 모델(best-fitting model)로 "recommended"할 모델을 선택하려고 시도한다.

BMDS에서 기본 설정되어 있는 로직에 대한 모든 변경은 result 또는 report에 기록되어야 한다.

로직 표의 일부 회색 셀은 선택할 수 없으므로 편집할 수 없다. 또한 "Bin Placement if Test is Failed(테스트가 실패한 경우 bin 배치)" 및 "Notes to Show(표시할 메모)" 열의 값은 편집할 수 없다.

로직 설정은 Setting 및 Results Workbook과 함께 저장되므로 이전에 실행한 분석을 로드하면 사용자 지정된 로직세팅이 복원된다.

### ***Enabling/Disabling Model Tests(모델 검정의 활성화/비활성화)***

특정 모델 검정을 켜거나 끄려면 셀의 아래쪽 삼각형을 선택하고 "on(켜기)" 또는 "off(끄기)"를 선택한다.

**Figure 76.** *Toggling model tests on/off in the Logic table.*

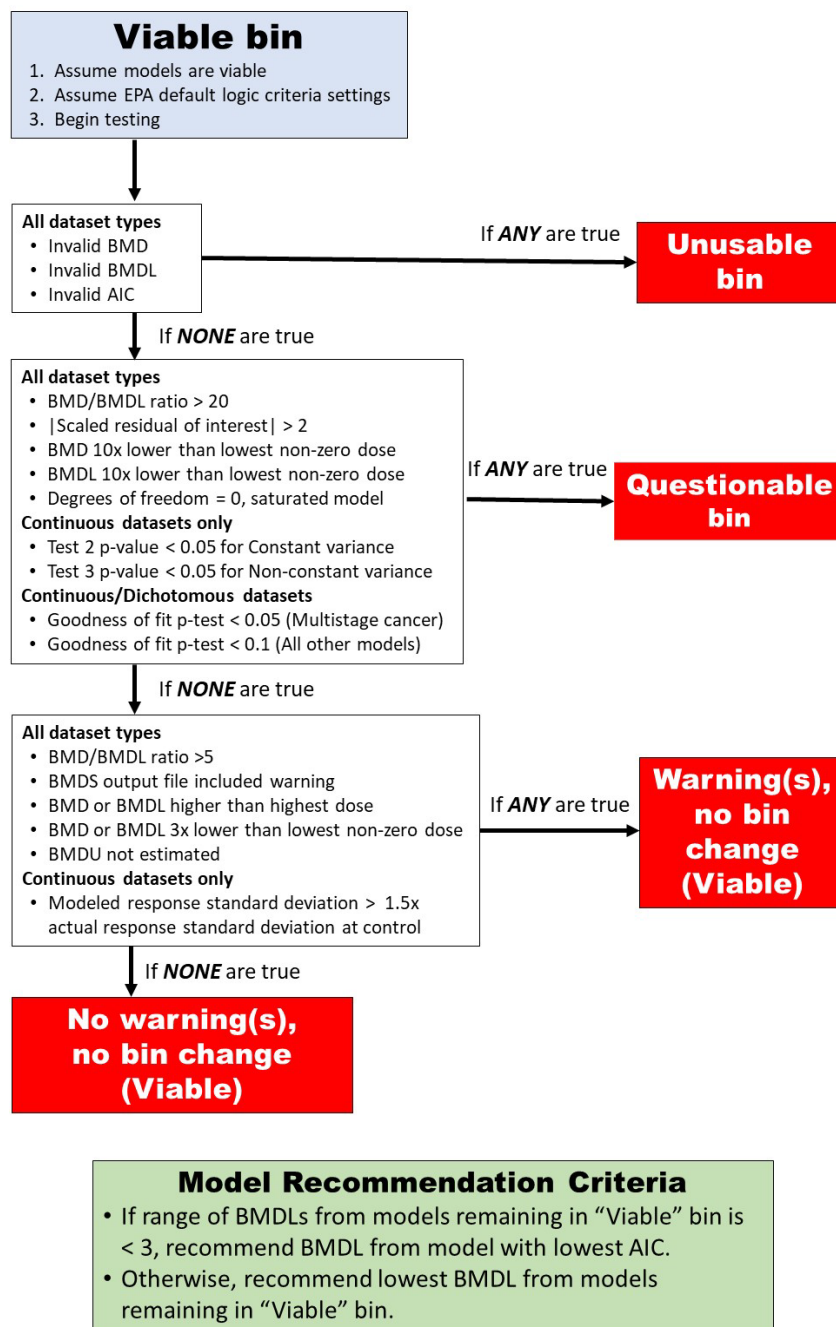
Test On/Off		
Continuous	Dichotomous	Nested
On	On	On
On	On	On
Off	Off	Off
On	On	On

### ***To Edit Test Threshold Values (검정 임계값을 편집하려면)***

셀을 선택하고 더블클릭하여 커서가 표시된다. 원하는 대로 셀 값을 편집한다. 예를 들어 사용자는 이분형 모델 분석에 대한 임계값을 변경한 다음 연속형 분석에 대한 원래 값을 복원할 수 있다.

[참고] Decision logic을 변경할 때 criteria 선택이 합당한지 확인하려면 숙련된 사용자 또는 통계 전문가와 상의해야 한다.

**Figure 77.** Flowchart of BMDS 3 model recommendation logic using EPA default logic assumptions.



### 연구진

연구기관 : 산업안전보건연구원

연구책임자 : 이미주 (연구위원, 흡입독성시험1부)

### 연구기간

2022. 01. 01. ~ 2022. 11. 30.

본 연구보고서의 내용은 연구책임자의 개인적 견해이며,  
우리 연구원의 공식견해와 다를 수도 있음을 알려드립니다.

**산업안전보건연구원장**

**흡입독성시험 결과 활용도 제고를 위한 Benchmark dose 도입 연구  
(2022-산업안전보건연구원-858)**

**발행일** : 2022년 11월 30일

**발행인** : 산업안전보건연구원 원장 김은아

**연구책임자** : 산업안전보건연구원 연구위원 이미주

**발행처** : 안전보건공단 산업안전보건연구원

**주소** : (44429) 울산광역시 중구 종가로 400

**전화** : 042-869-8533

**팩스** : 042-869-8696

**Homepage** : <http://oshri.kosha.or.kr>

**I S B N** : 979-11-92782-59-1

**공공안심글꼴** : 무료글꼴, 한국출판인회의, Kopub바탕체/돋움체