

2015年度 計量生物セミナー
臨床試験における estimand と感度分析

欠測のあるデータに対する解析手法の基礎 ～(5) 事例紹介～



日本製薬工業協会医薬品評価委員会データサイエンス部会TF4
欠測のあるデータの解析チーム(IPMA 欠測チーム)

興和株式会社 棚橋昌也

日本ベーリンガーインゲルハイム株式会社 鞆飼裕之

サノフィ株式会社 渡邊大丞

武田薬品工業株式会社 黒田晋吾

2015年度 計量生物セミナー

1

取り上げる事例



1. **Brexiprazole (REXULTI®)** 調査担当者: 棚橋
■ OTSUKA PHARM CO LTD, Jul. 10, 2015
2. **Nintedanib (OFEV®)** 調査担当者: 鞆飼
■ BOEHRINGER INGELHEIM, Oct. 15, 2014
3. **Riociguat (ADEMPAS®)** 調査担当者: 渡邊
■ BAYER HLTHCARE, Oct. 8, 2013
4. **Vedolizumab (ENTYVIO®)** 調査担当者: 黒田
■ TAKEDA PHARMS USA, May. 20, 2014

X. 一般名 (商品名)
■ 申請者名, 承認日

- すべてFDAで審査, 承認されたものです.
- ここでは, 時間の都合上, 1と2をご紹介します.
■ 残り二つの事例につきましては, 配布資料をご参照ください.

2015年度 計量生物セミナー

2

はじめに



- 本発表は, 当局および公開されている論文をもとに作成しております。公開情報の範囲を超える内容についてはお答えできない事項がございます。
- 本発表における公開情報とはFDAの審査報告書および論文などを示します。
- ご理解のほど, よろしく申し上げます。

2015年度 計量生物セミナー

3

1. Brexiprazole

調査担当者: 興和株式会社 棚橋昌也



2015年度 計量生物セミナー

4

事例紹介: Brexpiprazole



➤ Brexpiprazoleの概要

- 一般名: Brexpiprazole
- 商品名: REXULTI®
- 効能・効果
 - 統合失調症
 - 大うつ病 (補助療法)
- USでの申請データパッケージにおける第3相試験
 - 統合失調症の治療 : Study 230, 231
 - 大うつ病の補助療法 : Trial 227, 228
 - ✓ 各indicationに対して、試験は2つずつ

概要



➤ いずれの試験も主解析はMMRM.

➤ 感度分析としてMNARを仮定したMI (Pattern-Mixture approach) が実施された.

- 「大うつ病の補助療法」では、shared parameter modelをはじめ、その他の感度分析も実施された。
(ただし、Statistical Reviewに詳細な報告はない。)

Indication: 統合失調症の治療 試験概要 (1/2)



➤ 二つの第3相試験 (Study 230, 231)

- 無作為化二重盲検プラセボ対照並行群間比較試験
- 観察期間は6週間
- 群構成はplaceboとBrex 3用量の4群
 - 用量の構成と割付比が異なる

Brex: Brexpiprazole

	Study 230	Study 231
群	P, 1, 2, 4 mg	P, 0.25, 2, 4 mg
割付比	3:2:3:3	2:1:2:2

- 中止割合は30~40%
 - 中止の詳細は次スライド

➤ 主要評価項目は、6週時点のPANSS Total Score変化量

- 主解析はMMRM
- 感度分析にPattern-Mixture approach

Indication: 統合失調症の治療 試験概要 (2/2)



Subject Disposition

	Study 230			
	Brexipiprazole 1 mg	Brexipiprazole 2 mg	Brexipiprazole 4 mg	Placebo
Subjects, N (%)				
Randomized	120 (100%)	186 (100%)	184 (100%)	184 (100%)
Analyzed for Efficacy	117 (97.5%)	179 (96.2%)	181 (98.4%)	180 (97.8%)
Completed	81 (69.5%)	120 (64.5%)	130 (70.7%)	118 (64.1%)
Discontinued	39 (32.5%)	67 (35.8%)	54 (29.3%)	66 (35.9%)
Lost to follow-up	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Adverse Events	11 (9.2%)	11 (5.9%)	13 (7.1%)	22 (12.0%)
Subject met withdrawal criteria	0 (0%)	0 (0%)	2 (1.1%)	1 (0.5%)
Investigator withdrew consent	2 (1.7%)	0 (0%)	0 (0%)	1 (0.5%)
Subject withdrew consent	15 (12.5%)	25 (13.4%)	23 (12.5%)	21 (11.4%)
Protocol Deviation	2 (1.7%)	1 (0.5%)	0 (0%)	0 (0%)
Lack of Efficacy	9 (7.5%)	20 (10.8%)	16 (8.7%)	21 (11.4%)

	Study 231			
	Brexipiprazole 0.25 mg	Brexipiprazole 2 mg	Brexipiprazole 4 mg	Placebo
Subjects, N (%)				
Randomized	90 (100%)	182 (100%)	180 (100%)	184 (100%)
Analyzed for Efficacy	87 (96.7%)	180 (98.9%)	178 (98.9%)	178 (96.7%)
Completed	56 (62.2%)	124 (68.1%)	121 (67.2%)	109 (59.2%)
Discontinued	34 (37.8%)	58 (31.9%)	59 (32.8%)	75 (40.8%)
Lost to follow-up	0 (0%)	0 (0%)	0 (0%)	1 (0.5%)
Adverse Events	12 (13.3%)	15 (8.2%)	17 (9.4%)	32 (17.4%)
Subject met withdrawal criteria	1 (1.1%)	0 (0%)	1 (0.6%)	0 (0%)
Investigator withdrew subject	0 (0%)	1 (0.5%)	1 (0.6%)	3 (1.6%)
Subject withdrew consent	13 (14.4%)	24 (13.2%)	31 (17.2%)	21 (11.4%)
Protocol Deviation	1 (1.1%)	1 (0.5%)	2 (1.1%)	0 (0%)
Lack of Efficacy	7 (7.8%)	17 (9.3%)	7 (3.9%)	18 (9.8%)

Indication: 統合失調症の治療
感度分析の詳細 (1/2)



➤ 感度分析はPattern-Mixture approach

- 以下の理由による中止にMNARを仮定
 - シナリオ1: Brex投与群におけるLOE LOE: Lack of Efficacy
 - シナリオ2: Brex投与群におけるLOEとAE
- その他の中止には, MARを仮定

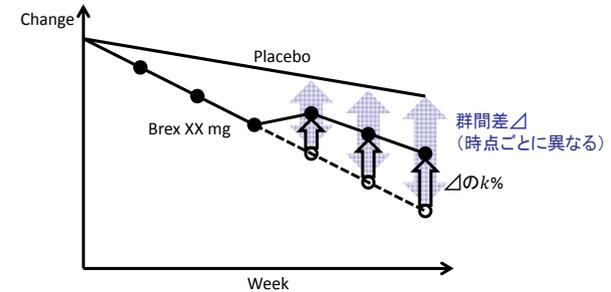
➤ MIによって欠測データを補完した後, 上記理由で中止した被験者に対しては, 主解析で推定された群間差の $k\%$ ($k=0, 10, 20, \dots, 100, \dots$) がマイナスされる。

- 主解析の結論が逆転する k : Tipping point
- $k=0$: MAR, $k=100$: "Back to placebo"

Indication: 統合失調症の治療
感度分析の詳細 (2/2)



➤ 感度分析のイメージ



Brex投与群でLOE(もしくはAE)で中止した被験者の中止以降のデータ(欠測データ)は, MIで補完後, 群間差 Δ の $k\%$ だけ引かれる。

Indication: 統合失調症の治療
主解析の結果



➤ 主解析の結果

Table 3. LS Mean Change from Baseline in PANSS at Week 6 (MMRM)

Study 230	Brex. 1 mg	Brex. 2 mg	Brex. 4 mg	Placebo
Number of patients	N=117	N=179	N=181	N=180
Baseline Mean (SD)	93.2 (12.7)	96.3 (12.9)	95.0 (12.4)	94.6 (12.8)
Mean Change at Week 6 (SE)	-16.9 (1.9)	-16.6 (1.5)	-20.0 (1.5)	-13.5 (1.5)
Treatment Difference	-3.4	-3.1	-6.5	-
95% Confidence Interval	(-8.1, 1.3)	(-7.2, 1.1)	(-10.6, -2.3)	-
p-value	0.16	0.15	0.0022	-
Average Effect (2mg & 4mg) versus Placebo	LS Mean Difference=-4.78, p-value=0.0093			

Study 231	Brex. 0.25 mg	Brex. 2 mg	Brex. 4 mg	Placebo
Number of patients	N=87	N=180	N=178	N=178
Baseline Mean (SD)	93.6 (11.5)	95.9 (13.7)	94.7 (12.1)	95.7 (11.5)
Mean Change at Week 6 (SE)	-14.9 (2.2)	-20.7 (1.5)	-19.7 (1.5)	-12.0 (1.6)
Treatment Difference	-2.9 (2.7)	-8.7 (2.2)	-7.6 (2.2)	-
95% Confidence Interval	(-8.3, 2.5)	(-13.1, -4.4)	(-12.0, -3.3)	-
p-value	0.29	<0.0001	0.0006	-
Average Effect (2mg & 4mg) versus Placebo	LS Mean Difference=-8.18, p-value<0.0001			

N=number of patients, SD=Standard Deviation, SE=Standard Error

Indication: 統合失調症の治療
感度分析の結果 (1/2)



➤ 感度分析の結果 (Study 230)

Table 4. Study 230 Sensitivity Analysis. LS Mean Change from Baseline in PANSS at Week 6 (Pattern Mixture Model with Multiple Imputation)

Percentage of Treatment Effect Subtracted from MAR imputed missing data	Treatment Comparison	Treatment Difference	95% Confidence Interval	P-value
Lack of Efficacy in Brex Group as MNAR シナリオ1				
K=100%	Brex.2mg vs Placebo	-2.8	(-7.3, 1.7)	0.22
	Brex.4mg vs Placebo	-5.2	(-9.7, -0.7)	0.0244
K=200%	Brex.2mg vs Placebo	-2.3	(-6.7, 2.1)	0.31
(tipping point)	Brex.4mg vs Placebo	-4.3	(-8.8, 0.1)	0.0571
Lack of Efficacy and AE in Brex Group as MNAR シナリオ2				
K=100%	Brex.2mg vs Placebo	-2.7	(-7.2, 1.9)	0.25
	Brex.4mg vs Placebo	+4.8	(-9.3, -0.2)	0.0395
K=140%	Brex.2mg vs Placebo	-2.7	(-7.2, 1.8)	0.25
(tipping point)	Brex.4mg vs Placebo	+4.4	(-8.9, 0.1)	0.0526
Dropout due to Any Reason in Brex Group as MNAR シナリオ3				
K=80%	Brex.2mg vs Placebo	-2.8	(-7.6, 2.0)	0.25
(tipping point)	Brex.4mg vs Placebo	-4.4	(-8.9, 0.1)	0.0529

シナリオ3の"Dropout due to Any ..."はレビューアが実施

Indication: 統合失調症の治療
感度分析の結果 (2/2)



➤ 感度分析の結果 (Study 231)

Table 5. Study 231 Sensitivity Analysis. LS Mean Change from Baseline in PANSS at Week 6 (Pattern Mixture Model with Multiple Imputation)

Percentage of Treatment Effect Subtracted from MAR imputed missing data	Treatment Comparison	Treatment Difference	95% Confidence Interval	P-value
Lack of Efficacy in Brex Group as MNAR シナリオ1				
K=100%	Brex.2mg vs Placebo	-9.1	(-14.0, -4.3)	0.0002
	Brex.4mg vs Placebo	-8.4	(-13.0, -3.9)	0.0003
K=500% (tipping point not reached)	Brex.2mg vs Placebo	-5.9	(-11.1, -0.8)	0.024
	Brex.4mg vs Placebo	-7.2	(-12.4, -2.0)	0.0066
Lack of Efficacy and AE in Brex Group as MNAR シナリオ2				
K=100%	Brex.2mg vs Placebo	-8.6	(-13.4, -3.9)	0.0004
	Brex.4mg vs Placebo	-7.9	(-12.5, -3.2)	0.0009
K=400% (tipping point)	Brex.2mg vs Placebo	-5.0	(-10.1, 0.2)	0.058
	Brex.4mg vs Placebo	-5.3	(-10.7, 0.2)	0.061
Dropout due to any Reason in Brex Group as MNAR シナリオ3				
K=100%	Brex.2mg vs Placebo	-7.2	(-12.3, -2.1)	0.0057
	Brex.4mg vs Placebo	-6.2	(-11.0, -1.4)	0.012
K=160% (tipping point)	Brex.2mg vs Placebo	-5.8	(-10.8, -0.90)	0.021
	Brex.4mg vs Placebo	-5.1	(-10.4, 0.2)	0.061

Indication: 統合失調症の治療
感度分析のまとめ



➤ いずれの試験も、主解析と感度分析の結果は一貫していた。

- レビューアが実施したものを除いて、tipping pointは100%を超えた。
- レビューアが実施したのも tipping pointは80%に届いた (Study 230)

レビューアが実施したものは、Brex投与群におけるすべての理由による中止にMNARを仮定するため、その他のものよりもsensitive

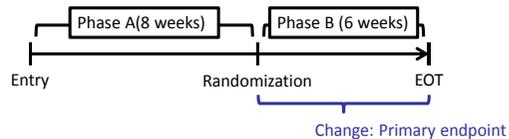
Indication: 大うつ病の補助療法
試験概要 (1/2)



➤ 二つの第3相試験 (Trial 227, 228)

- 無作為化二重盲検プラセボ対照並行群間比較試験
- 観察期間は、「既存の抗うつ治療で効果が不十分な患者」を選択するための8週間 (プラセボ単盲検, Phase A) と、有効性および安全性評価のための6週間 (二重盲検, Phase B) から成る
- 群構成はP, 1, 3 mgの3群 (Trial 227) とP, 2 mgの2群 (Trial 228)
- 中止割合は、10%未満

➤ 主要評価項目は、6週時点 (Phase B) のMADRS Total Score変化量



Indication: 大うつ病の補助療法
試験概要 (2/2)



➤ 主解析はMMRM.

➤ 感度分析で、Pattern Mixture Model based on MIに加えて、shared parameter model, random coefficient pattern mixture modelが行われた。

- その他にも、OC-ANCOVA, LOCF-ANCOVA

Indication: 大うつ病の補助療法
主解析の結果



➤ 主解析の結果

試験中に randomization criteria が変更された。
新しい criteria に合致した集団 ("per protocol set") を対象に
解析すると LS mean difference = -1.93, p=0.008

Table 5. Trial 227; Primary Analysis Result of the Primary Endpoint

Variable	1mg Brex+ADT N=225	3mg Brex+ADT N=226	Placebo+ADT N=218
MADRS Total Score, MMRM			
Mean (SD) End of Phase A	26.69 (5.61)	26.31 (5.24)	26.23 (5.27)
LS Mean (SE) Change At Week 14	-7.65 (0.50)	-7.98 (0.51)	-6.45 (0.51)
LS Mean Difference (95% CI)	-1.19 (-2.58, 0.20)	-1.52 (-2.92, -0.13)	-
P-value (多重性は調整していない)	0.0925	0.0327	-

多重性を考慮すると N.S.
(Hochbergの方法)

Table 12. Trial 228; Key Efficacy Results - MMRM

Endpoint Parameter	2 mg/day Brex +ADT N=187	Placebo +ADT N=191
Primary Efficacy Endpoint		
MADRS Total Score, MMRM by Randomized Treatment Group (Efficacy Sample - Primary Analysis)		
Mean (SD) end of Phase A (Week 8)	26.61 (5.79)	27.14 (5.60)
LS mean change (SE) at end of Phase B (Week 14)	-8.27 (0.61)	-5.15 (0.60)
LS mean difference ^a (95% CI)	-3.12 (-4.70, -1.54)	-
P-value ^b	0.0001	-

Indication: 大うつ病の補助療法
感度分析の結果 (1/2)



➤ 「感度分析の結果は主解析と同様の結果であった」と記載があるのみ。

- 中止割合が小さく(10%未満)で欠測値による問題はわずかなものであった。

Indication: 大うつ病の補助療法
感度分析の結果 (2/2)



➤ Subject Disposition

Table 3. Trial 227; Subject Disposition during Phase B (Randomized Sample)

	Brex 1mg (N=226)		Brex 3mg (N=230)		PLACEBO (N=221)		TOTAL (N=677)	
NUMBER OF SUBJECTS:	n	(%)	n	(%)	n	(%)	n	(%)
RANDOMIZED	226	(100.0)	230	(100.0)	221	(100.0)	677	(100.0)
COMPLETED	216	(95.6)	210	(91.3)	208	(94.1)	634	(93.6)
DISCONTINUED	10	(4.4)	20	(8.7)	13	(5.9)	43	(6.4)

Table 10. Trial 228; Subject Disposition during Phase B (Randomized Sample)

	Brex 2mg (N=188)		PLACEBO (N=191)		TOTAL (N=379)	
NUMBER OF SUBJECTS:	n	(%)	n	(%)	n	(%)
RANDOMIZED	188	(100.0)	191	(100.0)	379	(100.0)
COMPLETED	174	(92.6)	178	(93.2)	352	(92.9)
DISCONTINUED	14	(7.4)	13	(6.8)	27	(7.1)

参考



➤ 大塚製薬ニュースリリース

<https://www.otsuka.co.jp/company/release/detail.php?id=2997>

➤ Statistical Review

http://www.accessdata.fda.gov/drugsatfda_docs/n da/2015/205422Orig1Orig2s000StatR.pdf

2. Nintedanib

調査担当者: 日本ペーリンガーインゲルハイム株式会社 鶴飼裕之

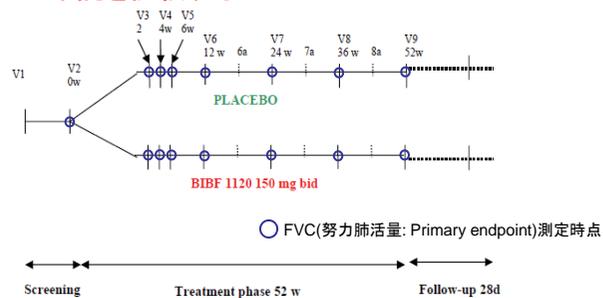
事例紹介: Nintedanib

➤ Nintedanibの概要

- 一般名: Nintedanib
- 商品名: OFEV®
- 効能・効果
 - 特発性肺繊維症
- 申請データパッケージにおける第3相試験(Sister trial)
 - 1199.32(INPULSE 1): 52週間の並行群間比較試験
 - 1199.34(INPULSE 2): 52週間の並行群間比較試験

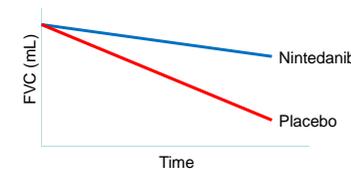
INPULSE 試験概要: 試験デザイン

- 二重盲検ランダム化並行群間プラセボ対照比較試験
- 欠測データの数を少なくするために投与中止後でも52週までの来院を依頼するプロトコル



試験概要: 主要評価項目と解析

- 主要評価項目: FVCの年間減少率(mL/year)
- 主解析: ランダム係数回帰モデル
 - 投与中止後のデータがある場合には主解析にそのデータを含める
- 仮説の設定
 - H_0 : プラセボ群に比べNintedanib群のベースラインから52週までのFVCの年間減少率に差がない
 - H_1 : プラセボ群に比べNintedanib群のほうがベースラインから52週までのFVCの年間減少率が小さい(片側検定)



試験概要:ランダム係数回帰モデル

製薬協

$$Y_{itk} = (\alpha + a_i) + (\gamma + \beta_s T_k + g_i)t + \beta_g \text{Gender}_i + \beta_a \text{Age}_i + \beta_h \text{Height}_i + \epsilon_{it}$$

$H_0: \beta_s = 0$

$H_1: \beta_s > 0$

i : 被験者 k : 治療群 t : 時間(連続量)

Y : 応答変数 $T_k=0$: プラセボ群, $T_k=1$: Nintedanib群

β_s : Nintedanib群の効果 α, γ : 切片および傾き a_i, g_i : 被験者*i*の傾きと切片

$\beta_g, \beta_a, \beta_h$: 被験者個別の背景係数

被験者*i*の

Gender_i : 性別(男性を基準とする), Age_i : ベースライン時の年齢, Height_i : ベースライン時の身長

ϵ_{it} : 被験者*i*時間*t*におけるランダム誤差 $\sim \text{i.i.d } N(0, \sigma_\epsilon^2)$

補足:

被験者内誤差はランダム係数モデルの切片と傾きに従う

分散共分散構造に無構造(UN)を用いる。収束しない場合には他の相関構造を用いる

観測されたデータに基づき解析を行う

25

FDAとのやりとり(1/2)

製薬協

タイミング	FDAのコメント
EOP2 meeting (2010年12月)	主解析についてMARを仮定しているが、治験中止は治療の影響を受けていると考えるので、MARの仮定について懸念している。感度分析で欠測データの影響を評価してほしい。

*EOP2: End of phase 2 meeting



BI: 感度分析にPattern mixture modelおよびMultiple Imputation (MI)を用いるとプロトコルに記載

*当時の感度分析に関する詳細な方法については公開情報に記載がない

BI: Boehringer Ingelheim

2015年度 計量生物セミナー

26

FDAとのやりとり(2/2)

製薬協

タイミング	FDAのコメント
TYPE C meeting (2013年8月)	BIが提案したPattern mixture modelに対してコメントを述べた。

*TYPE C meeting: Type A(e.g. clinical hold), Type B(e.g. EOP2)以外の話題について議論するmeeting

*詳細な情報については公開情報に記載がない



BI: FDAからのコメントを受けて、被験者を4つのパターンに分け、補完方法を変更した感度分析を行うこととした

2015年度 計量生物セミナー

27

感度分析:欠測データの取り扱い(1/3)

製薬協

- On-treatmentのデータ(治験薬投与中止前までのデータ)のみを使用する。
 - IPF患者に対する薬剤の効果をもっとも反映している
 - 主要モデルを使用
- 主要モデルに肺移植後データを含める
- 利用可能なデータによって被験者を異なるパターンに分ける
 - 中止後のデータも解析に使用する
 - 詳細は次スライド

2015年度 計量生物セミナー

28

感度分析: 欠測データの取り扱い(2/3)



被験者を4つのパターンに分ける

52週のデータがある被験者

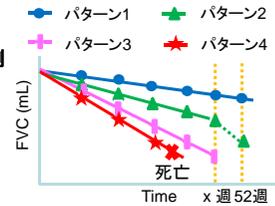
パターン1: 完了例

パターン2: 52週までに投与中止した症例

52週のデータがない被験者

パターン3: 生存例

パターン4: 死亡例



解析モデルは主解析と同じ

- パターン 1, 2は観測データを使用する
- パターン 3, 4の52週時のデータのみをMIによって補完する
 - 52週前のデータについては補完しない
 - 詳細は次スライド

感度分析: 欠測データの取り扱い(3/3)



$$Y_{itk} = (\alpha + a_i) + (\gamma + \beta_s T_k + g_i)t + \beta_g \text{Gender}_i + \beta_a \text{Age}_i + \beta_h \text{Height}_i + \epsilon_{it}$$

*補完モデルは主解析モデルに近いものを用いた(未記載)

感度分析	52週時の欠測データの補完方法 (パターン3)	52週時の欠測データの補完方法 (パターン4)
感度分析c1	パターン2の各群の傾きとSEに基づいてMI(reg)で補完	パターン2のプラセボ群の傾きとSEに基づいてMI(reg)で補完
感度分析c2	パターン2のプラセボ群の傾きとSEに基づいてMI(reg)で補完	ただし、生存している被験者よりもより厳しくするように設定する。 (($-\infty, \hat{\beta}$])の範囲で切断分布 $f(\beta)/2$ を用いる)
感度分析c3	主解析のプラセボ群の傾きとSEに基づいてMI(reg)で補完	主解析のプラセボ群の傾きとSEに基づいてMI(reg)で補完。ただし生存している被験者よりもより厳しい設定をする (($-\infty, \hat{\beta}$])の範囲で切断分布 $f(\beta)/2$ を用いる)

解析手順: 感度分析c1の場合 (1/2)



*52週時よりも以前のデータは補完しない

- 各群の傾きおよびSEを求めるため、パターン2の患者を対象にMMRM(主解析に近いモデルを使用)を実施する
 - $\hat{\beta}_D, \hat{\sigma}_D$ をNintedanib群の傾きとSEとする
 - $\hat{\beta}_P, \hat{\sigma}_P$ をプラセボ群の傾きとSEとする
- $\hat{\beta}_D, \hat{\beta}_P$ をstep1で得られた傾きを用いて $\beta_D \sim N(\hat{\beta}_D, \hat{\sigma}_D^2)$, $\beta_P \sim N(\hat{\beta}_P, \hat{\sigma}_P^2)$ に従う分布とする
- パターン3の52週時のデータをMI(reg, 1000回)によって補完する、i: i番目の被験者, j: j番目の補完
 - 補完したFVC week 52 = 被験者iの最後のFVC値 + $\hat{\beta}_{D \text{ or } P ij}$ (被験者iの最後のFVCの日付から予定されていた52週時の日付までの期間[日])

解析手順: 感度分析c1の場合 (2/2)



- パターン4の52週時のデータをMIによって補完する
 - Step 2で定義した切断分布 $f(\beta_P)/2$ を用いて範囲を $(-\infty, \hat{\beta}_P]$ に制限した、プラセボ群のランダムな傾きを使用する
 - Step3と同じ手順で行う
- 補完したデータセットに対して主解析と同じ方法で補完したデータセットごと(by_IMPUTATION_)を行う
- Step 5の推定値をProc MIANALYZEによって解析する

試験結果:脱落割合

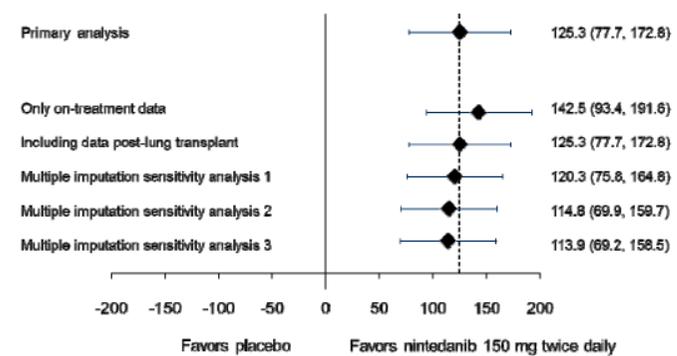
製薬協

	1199.32(N=515)		1199.34(N=551)	
	Nintedanib (n=309)	Placebo (n=206)	Nintedanib (n=331)	Placebo (n=220)
投与例数	309	204	329	219
投与完了例数	231(75%)	168(82%)	251(76%)	175(80%)
投与中止例数	78(25%)	36(18%)	78(24%)	44(20%)
早期投与中止例の理由(抜粋)				
有害事象	65(21%)	24(12%)	62(19%)	35(16%)
治験継続拒否	9(3%)	7(3%)	11(3%)	6(3%)
試験完了例数	260(84%)	174(85%)	272(83%)	179(82%)
試験中止例数	49(16%)	30(15%)	57(17%)	40(18%)
試験中止例の理由(抜粋)				
有害事象	25(8%)	15(7%)	42(13%)	30(14%)
治験継続拒否	23(7%)	12(6%)	9(3%)	7(3%)

33

試験結果(INPULSIS-1):企業側の解析

製薬協



2015年度 計量生物セミナー

34

試験結果(INPULSIS-1):FDA Reviewerの解析

製薬協

- FDAのreviewerによる解析も実施された
 - ただし、Statistical reviewには結果のみ記載されており、詳細な実施方法等は記載されていない

治療群	N	LS Mean	LS Mean Diff	95%CI	P値
FDAによる感度分析: Treated set (*pMIを用いたANCOVA)					
Nintedanib	309	-101	88	(45,131)	<0.001
Placebo	204	-189			
FDAによる感度分析: Treated set (もっとも低いRankで補完したRANK ANCOVA)					
Nintedanib	309	--	--	--	<0.001
Placebo	204	--			

*Placebo群の完了例のデータ

2015年度 計量生物セミナー

35

FDA Reviewerの結論

製薬協

- 申請者は主解析におけるFVCの欠測データについて、治療中止後に線形的に肺機能が減少する仮定を置き、データを補完しなかった。各患者の中止前の肺機能の悪化後の利用可能なデータに基づいて傾きを推定し、保守的にFVCの年間減少率を推定していることから、主解析は受け入れ可能とReviewerは判断した。
- 申請者は感度分析によって欠測データに対する影響を調べていた
- Reviewerが実施した感度分析の結果とも一貫性がとれていた

2015年度 計量生物セミナー

36

まとめ



➤ 主解析:ランダム化係数モデル

- 投与中止後のデータも解析に使用する

■ 感度分析

- On-treatmentのデータのみを使用
- モデルに肺移植後データを含める
- 被験者を4つのパターンに分ける
 - ✓52週時のデータがある:完了例,投与中止例
 - ✓52週時のデータがない:生存例,死亡例

参考文献



➤ FDA審査報告書

- <http://www.fda.gov/Drugs/DevelopmentApprovalProcess/DrugInnovation/ucm429247.htm>

➤ Richeldi et al. Efficacy and Safety of Nintedanib in Idiopathic Pulmonary Fibrosis. N Engl J Med. 2014; 370:2071-2082

- <http://www.nejm.org/doi/full/10.1056/NEJMoa1402584>

3. Riociguat

調査担当者: サノフィ株式会社 渡邊大丞



事例紹介: Riociguat



➤ Riociguatの概要

- 一般名: Riociguat
- 商品名: Adempas
- 効能・効果
 - 慢性血栓塞栓性肺高血圧症 (CTEPH)
 - 肺動脈性高血圧 (PAH)
- 申請データパッケージにおける第3相試験
 - 11348 (CHEST-1): [CTEPH]16週間の平行群間比較試験
 - 12934 (PATENT-1): [PAH]12週間の平行群間比較試験
 - ✓ 本資料で扱う試験

PATENT-1試験概要:試験デザイン

製薬協

- 国際共同プラセボ対照二重盲検比較試験
 - 用量調節期終了時の用量を主治療期に投与する。
- 中止後の規定最終来院日までの**継続はしていない**。



2015年度 計量生物セミナー

41

試験概要:主要評価項目と解析

製薬協

- 主要評価項目: 6分間歩行距離 (6MWD) のベースラインからの変化量
- 主解析: 共分散分析 (ANCOVA)
 - 欠測データの補完ルール
 - 治験完了又は中止した被験者: **第16週までの最終測定値**
 - Effectiveness (*Last observed value*)? Efficacy (*LOCF strategy*)?
 - 死亡・臨床的悪化で中止来院がない被験者、又は中止来院時の測定値がない被験者: **最悪値 (0m)**
 - 完了例又は中止例とは異なる取り扱い
 - 層別Wilcoxon検定
 - 正規性検定結果が有意であったため、事前に規定していた層別Wilcoxon検定を正式な有意差検定結果とした。

2015年度 計量生物セミナー

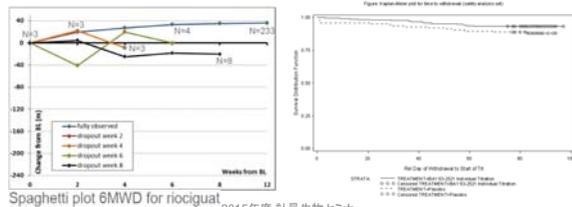
42

感度分析(1/4)

データの取り扱い及び中止に関する種々の解析

製薬協

- 主解析から変更/追加されたルール
 - 中止Visitを次の予定Visitとする。予定外Visitは解析から除く。
 - 死亡・臨床的悪化で中止来院がない被験者、又は中止来院時の測定値がない被験者に対する最悪値 (0m) の補完は行わない。
- 感度分析前に実施した解析
 - 中止理由ごとの割合、 Kaplan-Meierプロット
 - 中止時期別の6MWDの平均変化量のプロット (Spaghetti plot)



2015年度 計量生物セミナー

43

感度分析(2/4)

欠測メカニズムごと解析内容

製薬協

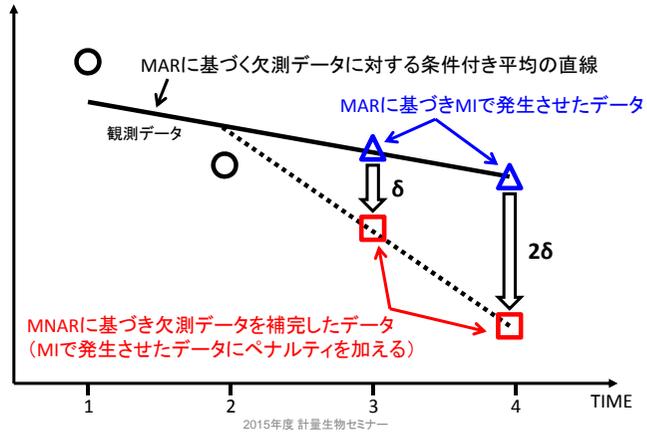
- MARに基づく解析: MMRM
 - 全ての被験者が治験完了したと仮定した場合: **Efficacyの評価**
- MNARに基づく解析
 - Delta method (Pattern mixture model) [Carpenter & Kenward (2008)]
 - 盲検化レビューの結果、脱落例は完了例に比べて6MWDの変化量が低値であった。
 - 特にプラセボ群ではそれが顕著であった。
 - 6MWDが脱落後の来院ごとに一定量低下すると仮定する。
 - MARに基づき Multiple imputation (MI) を用いて欠測データを補完し、シナリオ (Fixed penalty, Decreasing slope) ごとに設定されたペナルティを加える。
 - 補完したデータで共分散分析を実施し、結果を併合する。
 - 補完回数: 50回

2015年度 計量生物セミナー

44

感度分析(3/4)

Delta method [Carpenter & Kenward (2008)]

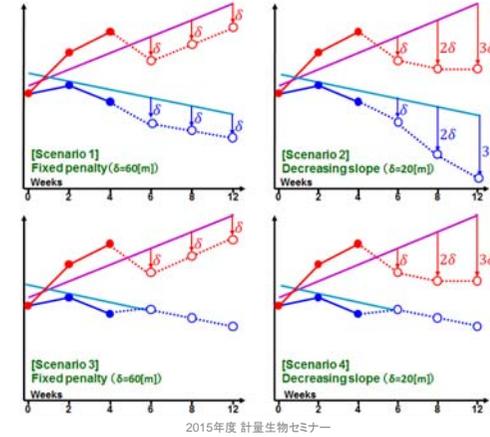
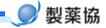


2015年度 計量生物セミナー

45

感度分析(4/4)

PATENT-1試験で仮定したシナリオ



2015年度 計量生物セミナー

46

解析結果(1/1)

脱落割合



	Placebo (n=126)	Riociguat 2.5 mg (n=254)
投与例数	126	254
投与完了例数	111 (88%)	237 (93%)
投与中止例数	15 (12%)	17 (7%)
投与中止例の理由(抜粋)		
有害事象	7 (12%)	8 (3%)
死亡	2 (2%)	0

2015年度 計量生物セミナー

47

解析結果(1/2)

プラセボ群とRiociguat 2.5mg群との群間差



解析方法	群間差	95%CI
主要解析		
ANCOVA	35.78	(20.06, 51.51)
感度分析		
MMRM	30.02	(16.11, 43.94)
Scenario 1 [Fixed penalty:-60m for both groups]	33.10	(18.49, 47.71)
Scenario 2 [Decreasing slope:-20m per visit for both groups]	35.03	(20.49, 49.57)
Scenario 3 [Fixed penalty:-60m only for Riociguat]	26.31	(12.32, 40.30)
Scenario 4 [Decreasing slope:-20m per visit only for Riociguat]	27.16	(13.36, 40.96)

2015年度 計量生物セミナー

48

FDA Comments



- 肺動脈性高血圧患者対象の試験は一般的に多くの欠測データを伴う傾向があるが、PATENT-1試験の欠測データは少数であった。
- スポンサーは欠測データに対して異なる仮定のもと多くの感度分析を実施した。
- 感度分析の結果は主要解析の結果と同様に適度に大きな効果を示した。
=>欠測データに対する問題は特になかった。

参考文献



- Statistical Review, Drug Approval Package, Adempas (riociguat) tablets. Food and Drug Administration website
 - http://www.accessdata.fda.gov/drugsatfda_docs/nda/2013/204819Orig1s000StatR.pdf
- Briefing Information for the August 6, 2013 Meeting of the Cardiovascular and Renal Drugs Advisory Committee. Food and Drug Administration website
 - <http://www.fda.gov/downloads/AdvisoryCommittees/CommitteesMeetingMaterials/Drugs/CardiovascularandRenalDrugsAdvisoryCommittee/UCM363541.pdf>
- Briefing Information for the August 6, 2013 Meeting of the Cardiovascular and Renal Drugs Advisory Committee. Food and Drug Administration website
 - <http://www.fda.gov/downloads/AdvisoryCommittees/CommitteesMeetingMaterials/Drugs/CardiovascularandRenalDrugsAdvisoryCommittee/UCM363543.pdf>
- Missing Data Sensitivity Analysis of a Continuous Endpoint - An Example from a Recent Submission. Arno Fritsch. Clinical Statistics Europe, Bayer
 - <http://www.amstat.org/chapters/newjersey/Documents/events/2011to2015/2014/Bayer/2%20NJ-ASA-Bayer-Workshop-Fritsch.pdf>
- Arno Fritsch. (2014) Missing Data Sensitivity Analysis of a Continuous Endpoint - An Example from a Recent Submission. November 21, 2014 - ASA NJ Chapter / Bayer Workshop, Whippany
 - <http://www.amstat.org/chapters/newjersey/Documents/events/2011to2015/2014/Bayer/2%20NJ-ASA-Bayer-Workshop-Fritsch.pdf>
- Carpenter, J. R., Kenward, M. G. (2008). Missing data in clinical trials - a practical guide. National Health Service Coordinating Centre for Research Methodology: Birmingham.
 - http://missingdata.lshtm.ac.uk/downloads/rm04_jh17_mk.pdf
- Ghofrani H-A et al. (2013) Riociguat for the treatment of pulmonary arterial hypertension. N Engl J Med.
 - <http://www.nejm.org/doi/pdf/10.1056/NEJMoa1209655>

4. Vedolizumab

調査担当者: 武田薬品工業株式会社 黒田晋吾



事例紹介: Vedolizumab



- Vedolizumabの概要
 - 一般名: Vedolizumab
 - 商品名: Entyvio
 - 効能・効果
 - 中等症又は重症のクローン病 (CD)
 - 中等症又は重症の潰瘍性大腸炎 (UC)
 - USでの申請データパッケージにおける第3相試験
 - C13007: CDの導入療法及び維持療法
 - C13006: UCの導入療法及び維持療法
 - C13011: CDの導入療法
 - C13008: UC及びCDの長期安全性

本資料で扱う試験

事例紹介: Vedolizumab

背景

	C13007試験		C13006試験	
	導入療法期	維持療法期	導入療法期	維持療法期
主要評価項目の評価時点	6週	52週	6週	52週
試験中止例数の割合	9%	52%	6%	44%

- 導入療法期、維持療法期ともにFDAのreviewerから主解析とは異なる補完方法を用いた感度分析を求められ、申請者が追加解析を実施
- 本事例紹介では、FDAのstatistical reviewにおける、**維持療法期**の欠測データの取扱いに関する議論を紹介

事例紹介: Vedolizumab

- **C13007試験 (CD)**
- **C13006試験 (UC)**

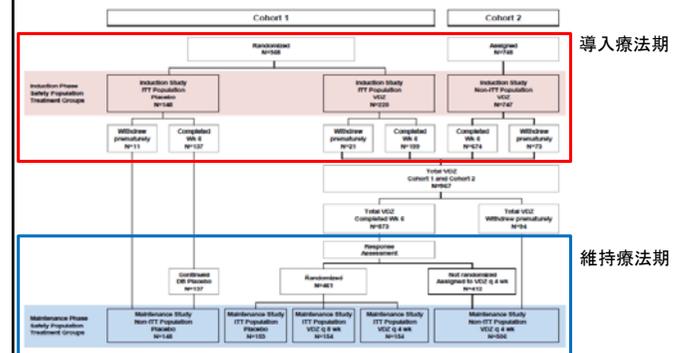
事例紹介: Vedolizumab C13007 (CD、維持療法期)

試験デザイン

- 目的: 中等症又は重症の**CD患者**に対するvedolizumabの維持療法における有効性及び安全性を検討する
- 対象: 中等症又は重症のCD患者で、vedolizumabを投与され導入療法期の6週時点において改善が見られた被験者
- 主要評価項目: 52週時点での寛解割合
 - CDAIスコアが150以下のときに寛解と定義
- 投与と群: プラセボ群、Vedolizumab Q4W群、Q8W群
- 主解析: CMHカイ2乗検定 (Vedolizumab Q4W vs placebo, Vedolizumab Q8W vs placebo)
 - 試験中止例は「寛解なし」として取扱った

事例紹介: Vedolizumab C13007 (CD)

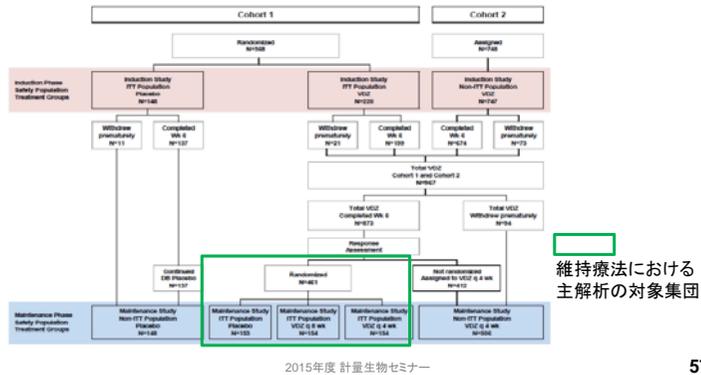
フローチャート



事例紹介: Vedolizumab C13007 (CD)

製薬協

➤ フローチャート



57

事例紹介: Vedolizumab C13007 (CD、維持療法期)

製薬協

➤ 被験者の内訳

	プラセボ群	vedolizumab Q8W群	vedolizumab Q4W群
無作為化例数	153	154	154
試験中止例数(%)	89 (58)	81 (53)	72 (47)
有害事象	15 (10)	12 (8)	9 (6)
試験実施計画書からの逸脱	1 (<1)	2 (1)	3 (2)
有効性の欠如	64 (42)	58 (38)	48 (31)
依頼者による試験の中止	0	0	0
同意撤回	7 (5)	6 (4)	9 (6)
追跡不能	1 (<1)	3 (2)	2 (1)
その他	1 (<1)	0	1 (<1)

2015年度 計量生物セミナー

58

事例紹介: Vedolizumab C13007 (CD、維持療法期)

製薬協

➤ 感度分析

■ 欠測データに対処するため、FDAからの要求に従って以下の感度分析を実施した

- observed case (※この試験ではcomplete caseと一致)
 - ✓ 主要評価項目が欠測の場合は解析から当該被験者を除外
- worst case (※statistical reviewで②の結果の記載はない)
 - ✓ ① 主要評価項目が欠測の場合は「寛解なし」として取扱う (主解析と同様の補完方法)
 - ✓ ② 主要評価項目が欠測の場合は、プラセボ群では「寛解あり」、vedolizumab群では「寛解なし」として取扱う

□ LOCF analysis

□ multiple imputation

✓ PROC MIを用いて補完回数を10回として実施

2015年度 計量生物セミナー

59

事例紹介: Vedolizumab C13007 (CD、維持療法期)

製薬協

➤ 感度分析

Table 28 Sensitivity Analyses – Clinical Remission at Week 52

Study C13007				
Analysis	Placebo	VDZ q8w	Difference	P-value
Primary	33/153 (21.6%)	60/154 (39.0%)	17.4%	0.0007
Observed Case	33/63 (52.4%)	59/72 (81.9%)	30.5%	<0.0001
Per Protocol	33/147 (22.4%)	56/149 (37.6%)	15.0%	0.0042
LOCF	45/153 (29.4%)	67/154 (43.5)	14.1%	0.0090
Multiple Imputation	66/153 (43.1%)	119/154 (77.3%)	34.2%	<0.0001

Analysis	Placebo	VDZ q4w	Difference	P-value
Primary	33/153 (21.6%)	56/154 (36.4%)	14.7%	0.0042
Observed Case	33/63 (52.4%)	56/81 (69.1%)	17.2%	0.0350
Per Protocol	33/147 (22.4%)	55/144 (38.2%)	15.9%	0.0029
LOCF	45/153 (29.4%)	71/154 (46.1)	16.6%	0.0023
Multiple Imputation	66/153 (43.1%)	103/154 (66.9%)	23.7%	<0.0001

2015年度 計量生物セミナー

60

事例紹介: Vedolizumab C13007 (CD、維持療法期)



➤ 感度分析

- reviewerによる結果の解釈
 - 52週時点における欠測データの数が投与群間で異なっており、missing at randomの仮定は適切でない可能性が考えられることから、multiple imputationによる解析結果は過大推定している傾向があるかもしれない
 - 52週時点においてプラセボ群では58%、vedolizumab Q8W群では53%、vedolizumab Q4W群では47%のデータが欠測であり、得られた群間差は信頼できる結果ではないかもしれない
 - 欠測データの割合はCDを対象とした同様のデザインの他試験と同様ではあるが、**欠測データが多いことにより試験結果の解釈を困難としている**

2015年度 計量生物セミナー

61

事例紹介: Vedolizumab C13007 (CD、維持療法期)



➤ FDAのstatistical team leaderの見解

- 申請者の主解析は正当化される
 - 本試験での中止割合は他試験から想定される中止割合と同程度
 - 中止理由のほとんどは有効性の欠如又は有害事象
 - 投与群間で欠測データの割合が同程度の場合、欠測データを「寛解なし」として取扱うことによって治療効果を過大推定することはない
- FDAの要求により主解析とは異なる補完方法を用いて感度分析を実施した結果、すべての解析においてプラセボ群と比較してvedolizumab群で良好な結果であった

2015年度 計量生物セミナー

62

事例紹介: Vedolizumab C13007 (CD、維持療法期)



➤ 結論

- 維持療法期の主要評価項目について、プラセボと比較してvedolizumabの統計的に有意なベネフィットが示された

2015年度 計量生物セミナー

63

事例紹介: Vedolizumab



- C13007試験 (CD)
- C13006試験 (UC)

2015年度 計量生物セミナー

64

事例紹介: Vedolizumab C13006(UC、維持療法期)

製薬協

試験デザイン

- 目的: 中等症又は重症のUC患者に対するvedolizumabの維持療法における有効性及び安全性を検討する
- 対象: 中等症又は重症のUC患者で、vedolizumabを投与され導入療法期の6週時点において改善が見られた被験者
- 主要評価項目: 52週時点での寛解割合
 - 完全Mayoスコアが2以下でかつすべてのサブスコアが1以下のときに寛解と定義
- 投与群: プラセボ群、Vedolizumab Q4W群、Q8W群
- 主解析: CMHカイ2乗検定 (Vedolizumab Q4W vs placebo, Vedolizumab Q8W vs placebo)
 - 試験中止例は「寛解なし」として取扱った

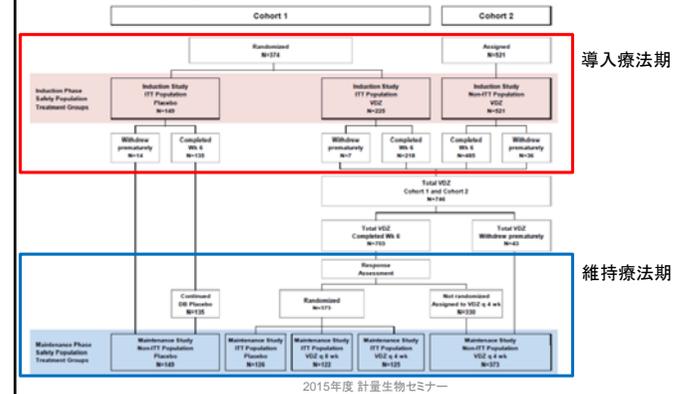
2015年度 計量生物セミナー

65

事例紹介: Vedolizumab C13006(UC)

製薬協

フローチャート



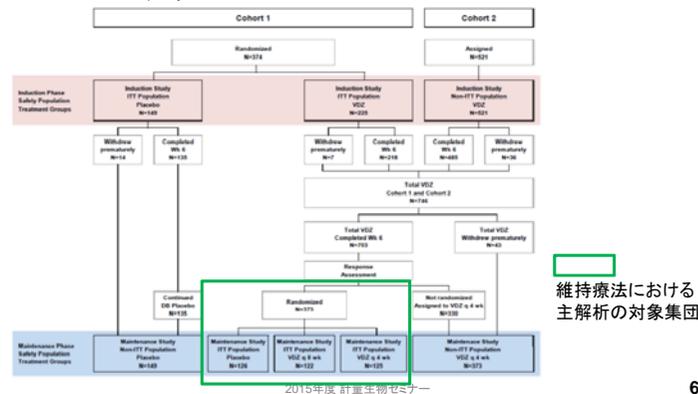
2015年度 計量生物セミナー

66

事例紹介: Vedolizumab C13006(UC)

製薬協

フローチャート



2015年度 計量生物セミナー

67

事例紹介: Vedolizumab C13006(UC、維持療法期)

製薬協

被験者の内訳

	プラセボ群	vedolizumab Q8W群	vedolizumab Q4W群
無作為化例数	126	122	125
試験中止例数(%)	78 (62)	45 (37)	41 (33)
有害事象	15 (12)	7 (6)	6 (5)
試験実施計画書からの逸脱	0	0	0
有効性の欠如	61 (48)	31 (25)	33 (26)
依頼者による試験の中止	0	0	0
同意撤回	2 (2)	5 (4)	2 (2)
追跡不能	0	2 (2)	0
その他	0	0	0

2015年度 計量生物セミナー

68

事例紹介: Vedolizumab
C13006 (UC、維持療法期)



➤ 感度分析

■ 欠測データに対処するため、FDAからの要求に従って以下の感度分析を実施した

□ observed case (※この試験ではcomplete caseと一致)

✓ 主要評価項目が欠測の場合は解析から当該被験者を除外

□ worst case

✓ ① 主要評価項目が欠測の場合は「寛解なし」として取扱う (主解析と同様の補完方法)

✓ ② 主要評価項目が欠測の場合は、プラセボ群では「寛解あり」、vedolizumab群では「寛解なし」として取扱う

□ LOCF analysis

□ multiple imputation

✓ PROC MIを用いて補完回数を10回として実施

事例紹介: Vedolizumab
C13006 (UC、維持療法期)



➤ 感度分析

Table 27 Clinical Remission at Week 52 – Sensitivity Analysis

Analysis Set	Placebo	VDZ 300 mg Q8W	Difference (VDZQ8w – placebo)	p-value	VDZ 300 mg Q4W	Difference (VDZQ4W – placebo)	p-value
Primary	20/126 (15.9%)	51/122 (41.8%)	26.1%	< 0.0001	56/125 (44.8%)	29.1%	< 0.0001
Per Protocol	20/121 (16.5%)	49/117 (41.9%)	25.7%	< 0.0001	55/121 (45.5%)	29.1%	< 0.0001
Observed Case	20/48 (41.7%)	51/77 (66.2%)	24.1%	0.0079	56/83 (67.5%)	25.0%	0.0042
ITT with Revised eDiary Requirements ^a	20/126 (15.9%)	48/122 (39.3%)	23.6%	< 0.0001	56/125 (44.8%)	29.1%	< 0.0001
LOCF ^b	27/126 (21.4%)	57/122 (46.7%)	25.4%	< 0.0001	58/125 (46.4%)	25.2%	< 0.0001
Worst Case ^c	98/126 (77.8%)	51/122 (41.8%)	-36.0%	< 0.0001	56/125 (44.8%)	-32.9%	< 0.0001
Multiple Imputation	67/126 (53.2%)	91/122 (74.6%)	21.1%	0.0004	88/125 (70.4%)	17.1%	0.0039

a: 日誌データの取扱いに関する感度分析
c: worst case ②の結果

事例紹介: Vedolizumab
C13006 (UC、維持療法期)



➤ 感度分析

■ reviewerによる結果の解釈

□ 感度分析の結果、worst case②以外においてはプラセボ群と比較してvedolizumab群で良好な結果であった

□ 52週時点において、プラセボ群で60%以上、vedolizumab群で30%以上が欠測しており、主解析では欠測データは「寛解なし」として補完しているため、得られた群間差が過大推定されている可能性はあるが、中止理由の多くは有効性の欠如又は有害事象であった

事例紹介: Vedolizumab
C13006 (UC、維持療法期)



➤ FDAのstatistical team leaderの見解

■ 申請者の主解析は擁護可能

□ reviewerが指摘する通り、52週時点において、プラセボ群で60%以上、vedolizumab群で30%以上が欠測しており、欠測データは「寛解なし」として補完しているため、得られた群間差が過大推定されている可能性はある

□ しかし、中止理由の多くは有効性の欠如又は有害事象

□ 本試験での中止割合は他試験から想定される中止割合と同程度

■ FDAの要求により主解析とは異なる補完方法を用いて感度分析を実施した結果、すべての解析においてプラセボ群と比較してvedolizumab群で良好な結果であった

事例紹介: Vedolizumab C13006(UC、維持療法期)



➤ 結論

- 維持療法期の主要評価項目について、プラセボと比較してvedolizumabの統計的に有意なベネフィットが示された

事例紹介: Vedolizumab



➤ 参考文献

- Label
http://www.accessdata.fda.gov/drugsatfda_docs/label/2014/125476s000lbl.pdf
- Statistical Review
http://www.accessdata.fda.gov/drugsatfda_docs/nda/2014/125476Orig1s000StatR.pdf