

欠測データの扱いに関する諸問題

JCOGデータセンター統計部門

水澤 純基

JCOG臨床試験セミナー中級編

2017. 10.14 (土)

注：わかりやすさを優先しているため、統計学的には正確ではない表現も含まれます

講義の目標

- 近年の欠測データ(missing data)に関するトピックを学び臨床家として結果を解釈する際に注意すべきことを身につける
 - 論文に記載されている欠測データに関する統計学的事項は何を意図したもののか
 - 統計家が何を気にしているのか
- そもそも欠測データが生じないことが最善であることを再認識する

アウトライン

- 臨床試験での欠測データ
- 近年の欠測データに関する議論
 - NASレポート、EMAガイドライン
- 欠測メカニズムの分類・欠測パターンの分類
 - MCAR、MAR、MNAR
 - monotone、non-monotone
- 欠測データを含むデータの解析法
 - 何を仮定した方法か
- 欠測データの予防とデータ収集
 - データセンター側の工夫と施設スタッフ側（医師、CRC）の工夫

臨床試験での欠測データ

欠測データとは？

Missing data are defined as values that are not available and that would be meaningful for analysis if they were observed

- もしそれが観察されていれば解析時に意味がある、観察されていない値

Little RJ, et al. N Engl J Med. 367(14):1355-1360, 2012

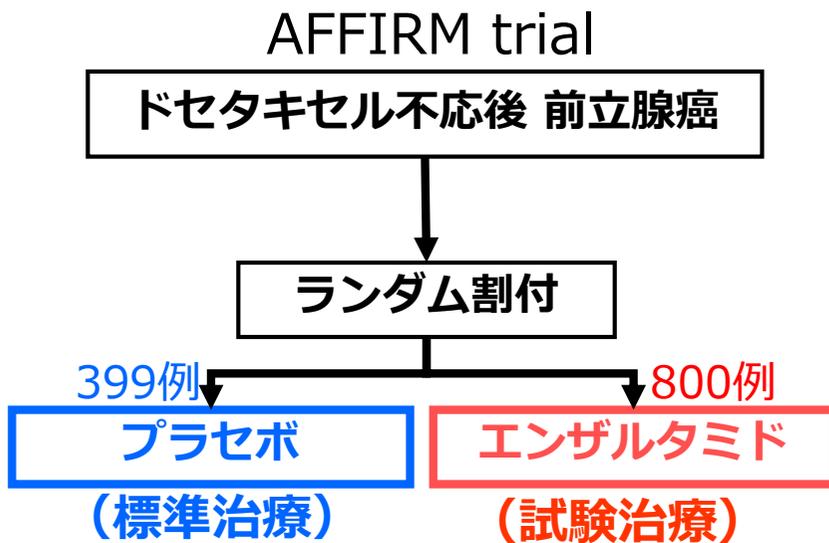
- 欠測データのうち観察されていれば意味があるもの
 - 治療中の患者のQuality of Life (QoL) スコア
 - 治療前の患者のQoLスコアや病期
- 欠測データのうち観察されていれば意味がなさそうなもの
 - 死亡した患者の死亡後のQoLスコア
 - 死亡した患者のQoLとは？
 - 多くの欠測データに関する統計学的検討では場合は議論の対象外

臨床試験の欠測データ

- 経時データのアウトカム（各対象に経時的に繰り返し測定）
 - QoLスコア
 - 患者が来院しなかった、CRCが調査用紙を渡すのを忘れた、など
- 2値データのアウトカム
 - 腫瘍縮小効果（奏効有無）
 - 一度も腫瘍縮小効果判定（CT/MRIの検査）をせずに死亡、など
- 生存時間データのアウトカム
 - イベント（死亡や再発）
 - イベントを観察しないまま追跡終了、引っ越しなどによる追跡不能、など
- 患者のベースラインデータ
 - 病期、遺伝子変異の有無、臨床検査値、など

本日の
主な焦点

臨床試験での欠測データの記載例：癌のHRQoL



Impact of enzalutamide on quality of life in men with metastatic castration-resistant prostate cancer after chemotherapy: additional analyses from the AFFIRM randomized clinical trial

D. Cella^{1*}, C. Ivanescu², S. Holmstrom³, C. N. Bui⁴, J. Spalding⁴ & K. Fizazi⁵

¹Department of Medical Social Sciences, Northwestern University Feinberg School of Medicine, Chicago, USA; ²Consulting, Quintiles, Hoofddorp; ³Global Data Science, Health Economics & Outcomes Research, Astellas Pharma Global Development, Leiden, The Netherlands; ⁴Health Economics & Clinical Outcomes Research, Astellas Pharma US, Inc., Northbrook, USA; ⁵Institut Gustave Roussy, University of Paris Sud, Villejuif, France

Received 22 August 2014; revised 3 October 2014; accepted 8 October 2014

Background: To present longitudinal changes in Functional Assessment of Cancer Therapy-Prostate (FACT-P) scores during 25-week treatment with enzalutamide or placebo in men with progressive metastatic castration-resistant prostate cancer (mCRPC) after chemotherapy in the AFFIRM trial.

Patients and methods: Patients were randomly assigned to enzalutamide 160 mg/day or placebo. FACT-P was completed before randomization, at weeks 13, 17, 21, and 25, and every 12 weeks thereafter while on study treatment. Longitudinal changes in FACT-P scores from baseline to 25 weeks were analyzed using a mixed effects model for repeated measures (MMRM), with a pattern mixture model (PMM) applied as secondary analysis to address non-ignorable missing data. Cumulative distribution function (CDF) plots were generated and different methodological approaches and models for handling missing data were applied. Due to the exploratory nature of the analyses, adjustments for multiple comparisons were not made. AFFIRM is registered with ClinicalTrials.gov, number NCT00974311.

Results: The intention-to-treat FACT-P population included 938 patients (enzalutamide, n = 674; placebo n = 264) with evaluable FACT-P assessments at baseline and >1 post-baseline assessment. After 25 weeks, the mean FACT-P total score

To estimate longitudinal changes in FACT-P scores from baseline, the primary analysis was carried out using a mixed effects model for repeated measures (MMRM) [12]. MMRM analysis uses all available data and assumes that any missing observations are missing at random. To address the possibility that missing data were not at random, a second analysis was carried out using a pattern mixture model (PMM) with placebo-based pattern imputation [13, 14]. In both models, the following covariates were included: treatment group; time; baseline ECOG score (0–1 or 2); average

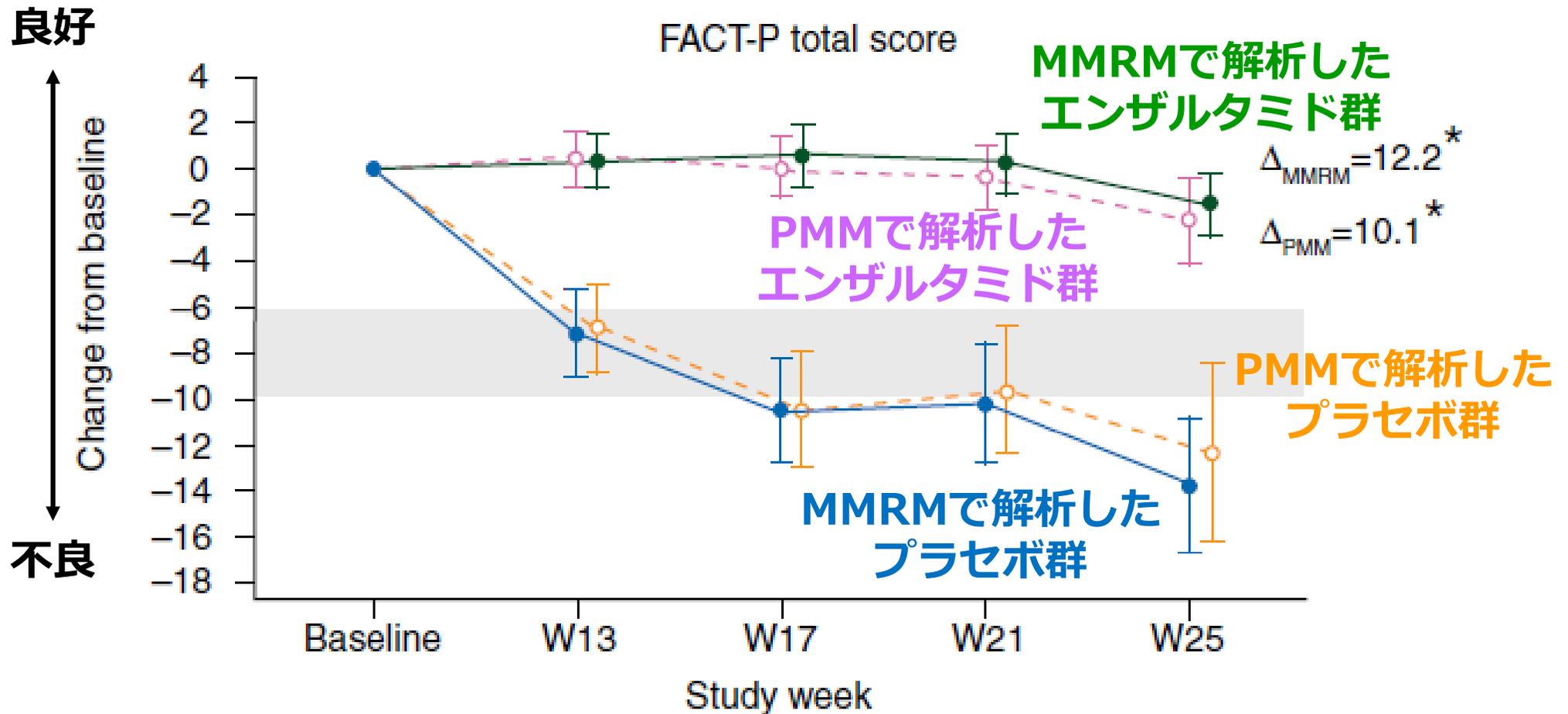
e (HRQoL)
relevant to
using use of
y regulators
re patient-

utamide, an
ent steps in
cantly pro-
men with
CRPC) after
d secondary
on HRQoL,
essment of

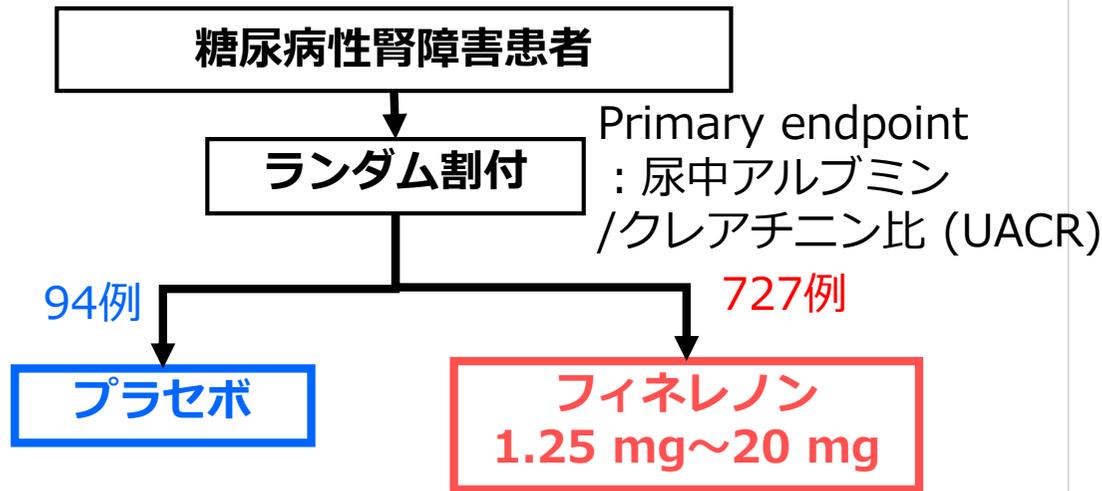
ies/by-no/4.0),
please contact

FACT-Pの解析結果

- MMRMの解析、PMMの解析共に、エンザルタミドのFACT-Pはプラセボに比べて良好であった



臨床試験での欠測データの記載例：臨床検査値



Research

Original Investigation

Effect of Finerenone on Albuminuria in Patients With Diabetic Nephropathy A Randomized Clinical Trial

George L. Bakris, MD; Rajiv Agarwal, MD; Juliana C. Chan, MD; Mark E. Cooper, MD, PhD; Ron T. Gansevoort, MD, PhD; Hermann Haller, MD, PhD; Giuseppe Remuzzi, MD; Peter Rossing, MD; Roland E. Schmieder, MD; Christina Nowack, MD; Peter Kolkhof, PhD; Amer Joseph, MBBS; Alexander Pleper, DiplStat; Nina Kimmekamp-Kirschbaum, PhD; Luis M. Rullope, MD, PhD; for the Mineralocorticoid Receptor Antagonist Tolerability Study-Diabetic Nephropathy (ARTS-DN) Study Group

IMPORTANCE Steroidal mineralocorticoid receptor antagonists, when added to a renin-angiotensin system blocker, further reduce proteinuria in patients with chronic kidney disease but may be underused because of a high risk of adverse events.

OBJECTIVE To evaluate the safety and efficacy of different oral doses of the nonsteroidal mineralocorticoid receptor antagonist finerenone, given for 90 days to patients with diabetes and high or very high albuminuria who are receiving an angiotensin-converting enzyme inhibitor or an angiotensin receptor blocker.

DESIGN, SETTING, AND PARTICIPANTS Randomized, double-blind, placebo-controlled, parallel-group study conducted at 148 sites in 23 countries. Patients were recruited from June 2012 to February 2014 and the study was completed in August 2014. Of 1501 screened and 821 received study drug.

Patients were randomly assigned to receive oral, once-daily finerenone 1.25 mg/d, n = 92; 5 mg/d, n = 100; 7.5 mg/d, n = 97; 10 mg/d, n = 98; 15 mg/d, n = 119 or matching placebo (n = 94) for 90 days.

RESULTS The primary outcome was the ratio of the urinary albumin excretion to creatinine excretion (UACR) at 90 days. Safety endpoints were changes in systolic blood pressure, heart rate, and estimated glomerular filtration rate.

At baseline, the mean age was 63 years; 78% were male. At baseline, 40.0% had an UACR ≥ 300 mg/g and 40.0% had an estimated glomerular filtration rate ≤ 30 mL/min/1.73 m². Finerenone demonstrated superior efficacy compared with placebo in reducing UACR. The placebo-corrected mean

Related article page 940
Supplemental content at jama.com

were performed. A last-observation-carried-forward (LOCF) method was applied, whereby the higher UACR value from the premature

the premature discontinuation visit), a baseline-observation-carried-forward analysis (imputing the baseline value for missing data; ie, including all missing data at day 90 v

cacy variable), a mean value imputation (imputing the value of the primary efficacy variable by the least squares mean value of the primary efficacy analysis), a random imputation (imputing the value of the primary efficacy variable by a random number from a normal distribution with least squares mean and variance [from descriptive statistics]), and a post hoc multiple imputation. The distributional model assumptions

近年の欠測データに関する議論

欠測データ：何が問題？

1. 推定値にバイアスが入る

- 状態の悪い（QoLスコアが不良）患者のQoLスコアが欠測になりやすい場合
 - 観測値だけに基づくQoLの推定値は過大評価
- 状態の良い（QoLスコアが良好）患者のQoLスコアが欠測になりやすい場合
 - 観測値だけに基づくQoLの推定値は過小評価

2. 検出力が落ちる

- 欠測値を除いた解析により有意なものが統計学的有意と言えなくなる
 - ただし、がんの第III相試験の場合、OSやPFSに基づいたサンプルサイズは、QoLの臨床的に意味のある差について十分な検出力を持つことが多いので、QoLについての検出力低下はあまり問題にならないかもしれない

Fairclough, DL. Design and analysis of quality of life studies in clinical trials. CRC press, 2010.

数値例：欠測データによるバイアスや検出力低下

試験1	有効	無効	欠損
標準治療	20	80	10
試験治療	80	20	10

試験2	有効	無効	欠損
標準治療	20	80	50
試験治療	80	20	50

有効割合の差 (95% CI), p値

条件	試験1	試験2
① 観測データのみ(各群100例)	60.0% (48.9%-71.1%), p<0.001 $\frac{80}{100} - \frac{20}{100}$	60.0% (48.9%-71.1%), p<0.001 $\frac{80}{100} - \frac{20}{100}$
② 欠測データは観測データと同様の効果	60.0% (49.4%-70.6%), p<0.001 $\frac{88}{110} - \frac{22}{110}$	60.0% (51.0%-69.1%), p<0.001 $\frac{120}{150} - \frac{30}{150}$
③ 欠測データ全て無効とみなす	54.5% (43.5%-65.6%), p<0.001 $\frac{80}{110} - \frac{20}{110}$	40.0% (30.3%-49.7%), p<0.001 $\frac{80}{150} - \frac{20}{150}$
④ 欠測データ試験治療は無効・標準治療は全て有効とみなす	45.5% (33.7%-57.2%), p<0.001 $\frac{80}{110} - \frac{30}{110}$	6.7% (-18.0%-4.6%), P=0.248 $\frac{80}{150} - \frac{70}{150}$

欠測データについての規制当局の考え



EUROPEAN MEDICINES AGENCY
SCIENCE MEDICINES HEALTH

2 July 2010
EMA/CPMP/EWP/1776/99 Rev. 1
Committee for Medicinal Products for Human Use (CHMP)

Guideline on Missing Data in Confirmatory Clinical Trials

Discussion in the Efficacy Working Party	June 1999/ November 2000
Transmission to CPMP	January 2001
Released for consultation	January 2001
Deadline for Comments	April 2001
Discussion in the Efficacy Working Party	October 2001
Transmission to CPMP	November 2001
Adoption by CPMP	November 2001
Draft Rev. 1 Agreed by Efficacy Working Party	April 2009
Adoption by CHMP for release for consultation	23 April 2009
End of consultation (deadline for comments)	31 October 2009
Rev. 1 Agreed by Efficacy Working Party	April 2010
Adoption by CHMP	24 June 2010
Date for coming into effect	1 January 2011

This guideline replaces Points to Consider on Missing Data in Clinical Trials (CPMP/EWP/1776/99).

Keywords	<i>Baseline Observation Carried Forward (BOCF), Generalised Estimating Equations (GEE), Last observation carried forward (LOCF), Missing at random (MAR), Missing completely at random (MCAR), Missing Data, Mixed Models for Repeated Measures (MMRM), Missing not at random (MNAR), pattern mixture models.</i>
----------	---

7 Westferry Circus • Canary Wharf • London E14 4HB • United Kingdom
Telephone +44 (0)20 7418 8400 Facsimile +44 (0)20 7418 8416
E-mail info@ema.europa.eu Website www.ema.europa.eu



An agency of the European Union

© European Medicines Agency, 2011. Reproduction is authorised provided the source is acknowledged.

The Prevention and Treatment of Missing Data in Clinical Trials

Panel on Handling Missing Data in Clinical Trials

Committee on National Statistics

Division of Behavioral and Social Sciences and Education

NATIONAL RESEARCH COUNCIL
OF THE NATIONAL ACADEMIES

THE NATIONAL ACADEMIES PRESS
Washington, D.C.
www.nap.edu

Copyright © National Academy of Sciences. All rights reserved.

National Research Council (NRC)レポート

- FDAからの要請と出資を受けて作成されたレポート
 - 臨床試験の経験豊かな著名な統計家で構成された専門家集団が作成
- Phase 3のランダム化比較試験（≡検証的試験）が主な焦点
 - 試験結果に高いレベルの科学的妥当性が問われる状況
 - 他の臨床研究（Phase 2、観察研究、予後因子解析など含む）にも多くの部分は適用可
- 欠測データに関する18のrecommendationを提示
 - 特に、デザイン段階や試験実施中の欠測データの予防について重視！
 - 解析段階で何とかなるわけではないことを強調

NRCレポートの主なrecommendationの要約

1. プロトコールに試験の目的、primary outcome、どのように・いつ・誰の評価項目が測定されるのか、介入効果（何を測定したいのか[estimand]）を記載すべき デザイン・予防
2. プロトコール規定の治療が維持される患者を最大化するように試験デザインを設定すべき デザイン・予防
デザイン・予防
3. プロトコール治療中止後の患者情報も収集し解析に使用するべき デザイン・予防
4. プロトコールにプロトコール治療中止後の後治療を受けられるよう規定すべき デザイン・予防
5. プロトコール治療を受けたかどうかによらず全患者の情報を収集すべき デザイン・予防
6. プロトコールに欠測データの問題を明確に予測し、欠測データの影響を把握・制限するための対策を記載すべき デザイン・予防
7. IC文書にプロトコール治療中止後のデータも重要であることを強調すべき

上記は意訳であるため原文は以下を参照

http://www.cytel.com/hs-fs/hub/1670/file-2411099288-pdf/Pdf/MissingDataNationalAcademyof_Medicine.2010.pdf

NRCレポートの主なrecommendationの要約

8. プロトコールに欠測データの量を最小化が重要であることを記載し、どの程度の完全データがあればよいか設定すべき デザイン・予防
9. プロトコールに欠測データについての解析手法を研究者に理解するように記載する デザイン・予防 解析
10. 単一補完法は仮定が科学的に妥当でない限り、用いるべきではない 解析
11. パラメトリックモデルは仮定を説明し、モデルの適合度を検討して用いるべき 解析
12. 主たる解析法は仮定した欠測データについての仮定の下で α エラーが保たれる方法を用いることが重要 解析
13. 重みつき一般化推定方程式 (weighted generalized estimating equations) は、MARが正当化される際に広く用いられるべき 解析

上記は意識であるため原文は以下を参照

http://www.cytel.com/hs-fs/hub/1670/file-2411099288-pdf/Pdf/MissingDataNationalAcademyof_Medicine.2010.pdf

NRCレポートの主なrecommendationの要約

デザイン・予防

14. ある程度の欠測データの発生が想定される場合は欠測の理由と関係すると考えられるデータを収集すべき

解析

15. 感度解析を臨床試験の結果についての主たる報告の一部とすべき

16. FDA・NIHは一般的な欠測率や欠測データの理由について特定するために臨床試験のデータベースを作成し、異なる解析モデルがどの程度結果に影響を与えるのか研究し、今後の試験デザインに活かすべき教育など

17. FDAやスポンサー（企業）は解析担当者やレビューワーに最新の欠測データの統計解析手法を身につけるトレーニングをするべき 教育など

18. 欠測データの扱いはスポンサーの統計学的研究において優先されるべき 教育など

上記は意識であるため原文は以下を参照

http://www.cytel.com/hs-fs/hub/1670/file-2411099288-pdf/Pdf/MissingDataNationalAcademyof_Medicine.2010.pdf

NRCLレポート後の出来事

- 不適切な解析法（例えばsingle imputation）の代わりに、より適切な解析法（例えばmultiple imputation）が用いられるようになりつつある

Fielding S., et al. Quality of Life Research 25(12): 2977-2983. 2016

- 統計解析用ソフトウェアで複雑な解析に対応できるようなオプションの整備が進められるようになった

野間久史. 臨床研究における欠測データの統計解析. ICRweb講義より

- ICH E9(R1)のコンセプトペーパーでestimand（臨床試験で何の効果を推測したいのか）と感度解析の考えが追加されることが公表された

www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E9/E9__R1__Final_Concept_Paper_October_23_2014.pdf

2017/10月現在、PMDAからパブコメ中：<https://www.pmda.go.jp/int-activities/int-harmony/ich/0031.html>

- がん臨床試験の世界では、EORTC主導でHealth-Related QoLの解析法や解釈についてのコンソーシアムが立ち上がった

Bottomley A et al., Lancet Oncol. 17:e510-14. 2016

欠測データの問題に鈍感だった（？）がん臨床試験でも今後は欠測データの問題についての理解は避けられないと思われる

欠測メカニズムと欠測パターン

なぜ欠測メカニズムの分類が必要？

- 欠測データを含むデータ解析は、欠測データのメカニズムについて特定の仮定の下で正しい推定値（バイアスのない推定値）が得られる方法
 - 欠測データのメカニズムが正しく特定できていなければ正しい結果が得られない
- NRCLレポート、EMAガイドライン、ICH E9(R1)などから、欠測データに対して様々な解析により結果の頑健性が求められる
 - 何を仮定し、何の効果を推定しようとしているのかプロトコールなどに記載するように求められつつある
 - 論文やプロトコールの統計記載で何を仮定した解析法か知っておくと、正しく結果の解釈ができる

3種類の欠測メカニズム

欠測の理由とアウトカムへの関連の程度により分類

1. Missing Completely at Random (**MCAR**)

2. Missing at Random (**MAR**)

3. Missing not at Random (**MNAR**)

1. Missing Completely at Random (MCAR)

- 観測データとも欠測データとも関係しない欠測データ
 - 得られた観測データは、欠測データを含む全てのデータからランダムに選択したものとみなせる
- 例：MCARとみなして良さそうな欠測
 - 大地震・火災が生じて患者に渡すべきQoL調査票を紛失した
 - QoL調査に用いるタブレットが偶然に故障してしまった

2. Missing at Random (MAR)

- 観測データに関係して生じたが欠測データには関係しない欠測
 - 本来得られるはずであった欠測データの値は、他の観測データのみで説明できるとみなせる
- 例：MARとみなして良さそうな欠測
 - 3時点目のQoLが欠測の患者（3時点目のQoLは欠測データ）は、観測データである2時点目のQoLが極端に悪化している患者に多い
 - 3時点目のQoLが欠損の患者は、75歳以上の男性患者に多い

3. Missing not at Random (MNAR)

- 欠測データに関係して生じた欠測
 - 多くの場合、観測データにも依存する
 - 本来得られるはずであった欠測データの値は、他の観測データだけでなく、得られるはずの欠測データもなければ説明できない
 - 欠測が生じた段階で欠測を説明できる十分なデータが得られていない場合、ほぼこれに該当すると思われる
- 例：MNARとみなして良さそうな欠測
 - 増悪した患者の増悪後のQoLスコアが欠測
 - 「増悪後のQoL値は悪そうだから取らない」と医師が考えているなら、増悪後に観察するはずだったデータに依存して欠測

3種類の欠測メカニズムのまとめ

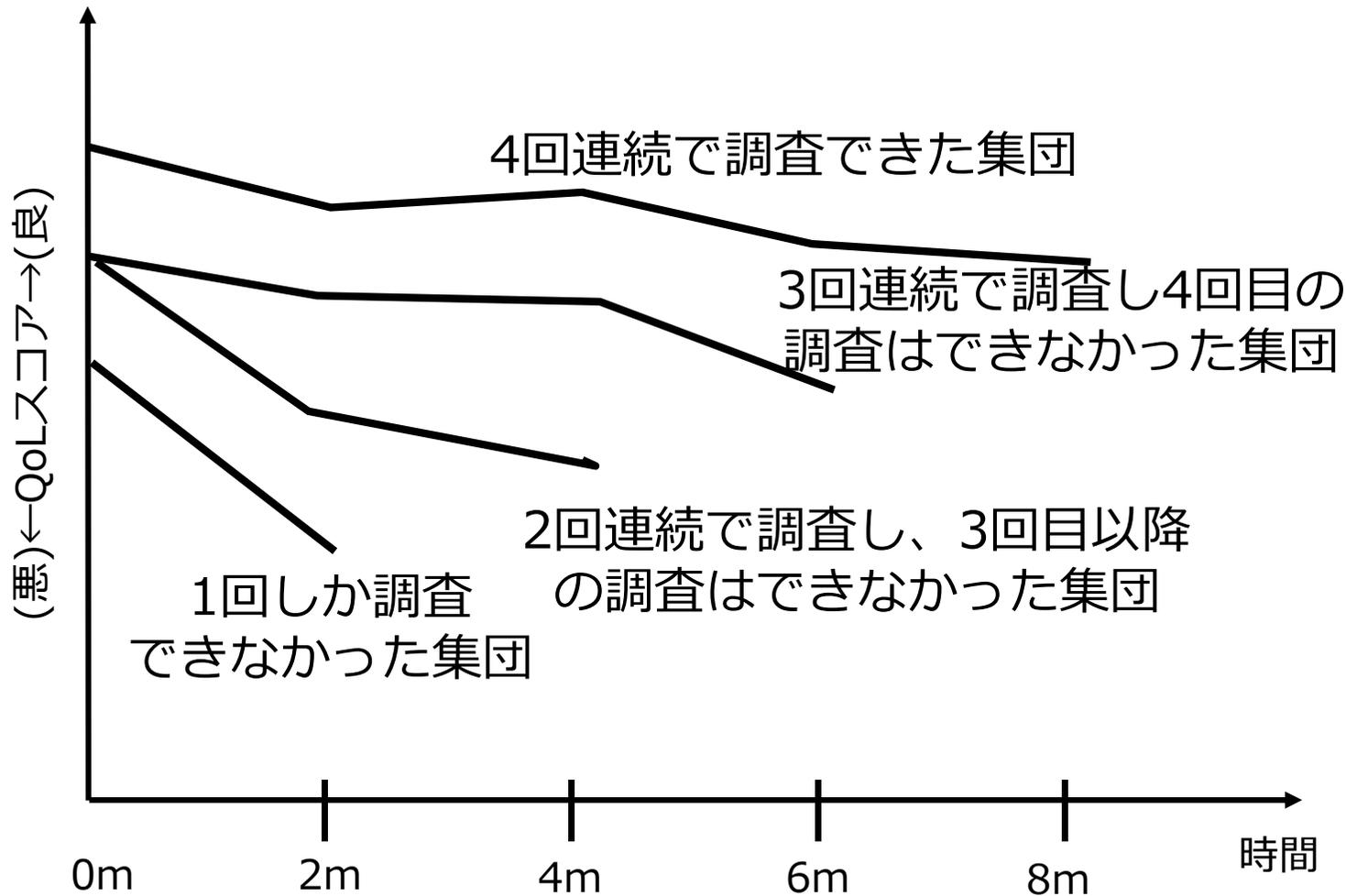
欠測メカニズム	依存する	依存しない
MCAR		観察データ 欠測データ
MAR	観察データ	欠測データ
MNAR	欠測データ (観察データ)	

- 個々の欠測に欠測メカニズムを考えることはできるが、通常、解析を行う際は集団全体に1つの欠測メカニズムを仮定した解析を行う
- 欠測データにMCARとMARが混在していると考えられる場合はMAR、MARとMNARが混在していると考えられる場合はMNARを仮定した解析を行う

どのように欠測メカニズムを区別できるか？

- 一般的に信頼性を持って区別できる手法はない
 - 出来るだけ多くの情報をCRFなどで収集して検討することが必要
 - 区別できそうなのは、MCARかMARかの違いのみ
 - 通常、MARかMNARの違いは手持ちのデータから判断できない
- 1. グラフを用いた検討
 - 患者の状態（ベースラインの年齢やPSなど）や欠測のパターンなどにより層別した集団でアウトカムの推移を図示
 - 特定の傾向があれば、少なくともMCARではないと考えられる
- 2. 検定
 - MCARであれば欠測のパターンによらずアウトカムの平均値は同じはず
 - 平均値の推移が欠測のパターンで同じかどうかを検定

グラフを用いた検討の例



1. すぐに欠測になる患者集団はベースラインの時点からQoLスコアが不良の患者集団

2. 欠測前の最後の観測データは、それ以前の観測データよりも悪化する

↓
少なくとも観測データに依存しているのでMCARではなくMAR

欠測パターン

○：観測、-：欠測

Terminal Dropout (Monotone パターン)			
○	○	○	○
○	○	○	-
○	○	-	-
○	-	-	-

Intermittent (Non-monotone パターン)			
○	○	-	○
○	-	○	○
-	○	-	○
○	-	-	○

Mixed パターン (Non-monotone パターン)			
○	-	○	-
-	○	○	-
-	○	-	-
-	-	○	-

- 例：
- 患者拒否
 - 増悪後の調査なし
 - (死亡)
- 例：
- 地震で来院困難
 - AEへの対処のため未実施
- 例：
- 左記、2パターンの組み合わせ

欠測パターンにより解析手法や欠測データを補完する手法が
変わり得るため解析段階では確認が必要
(本講義では詳細は省略)

欠測データを含むデータの解析法

欠測データを含んだ解析法の分類

1. 欠測データを無視した解析法
 - A) Complete case (CC) 解析
2. 欠測データに何かしらの値を補完 (imputation) した方法
 - A) Single imputation (単一補完法)
 - a. Last observation carried forward (LOCF) 法
 - B) Multiple imputation (多重補完法)
3. 欠測データのメカニズムと観測データを統計モデルで組み合わせた解析法
 - A) Inverse Probability weighting (IPW) 法
 - B) Mixed model for repeated measure (MMRM) 法
 - C) Pattern Mixture Model (PMM) 法

Complete case (CC) 解析

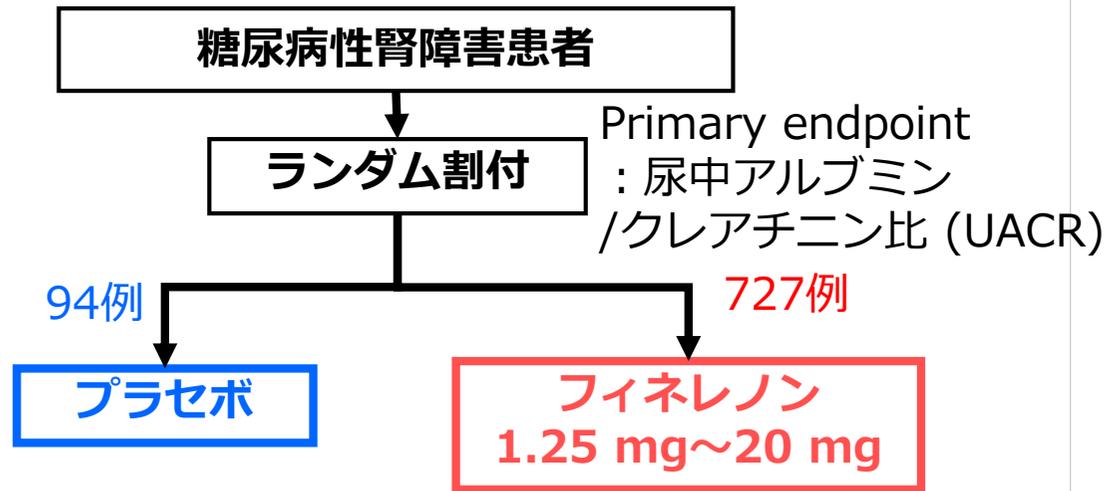
- すべての測定値が観察された患者のみ解析対象
 - 1時点でも欠測データがあれば除外
 - 下図の場合に、最初から25例・50例だったかのように解析する
- 欠測データについての仮定
 - MCAR (MCARが成り立てばバイアスはなし)
- 問題点
 - 患者数が減少することにより検出力低下
 - そもそもMCARが妥当という状況はほぼあり得ない



欠測データを含んだ解析法の分類

1. 欠測データを無視した解析法
 - A) Complete case (CC) 解析
2. 欠測データに何かしらの値を補完 (imputation) した方法
 - A) Single imputation (単一補完法)
 - a. Last observation carried forward (LOCF) 法
 - B) Multiple imputation (多重補完法)
3. 欠測データのメカニズムと観測データを統計モデルで組み合わせた解析法
 - A) Inverse Probability weighting (IPW) 法
 - B) Mixed model for repeated measure (MMRM) 法
 - C) Pattern Mixture Model (PMM) 法

再掲：臨床試験での欠測データの記載例



Research

Original Investigation

Effect of Finerenone on Albuminuria in Patients With Diabetic Nephropathy A Randomized Clinical Trial

George L. Bakris, MD; Rajiv Agarwal, MD; Juliana C. Chan, MD; Mark E. Cooper, MD, PhD; Ron T. Gansevoort, MD, PhD; Hermann Haller, MD, PhD; Giuseppe Remuzzi, MD; Peter Rossing, MD; Roland E. Schmieder, MD; Christina Nowack, MD; Peter Kolkhof, PhD; Amer Joseph, MBBS; Alexander Pleper, DiplStat; Nina Kimmeskamp-Kirschbaum, PhD; Luis M. Rullope, MD, PhD; for the Mineralocorticoid Receptor Antagonist Tolerability Study-Diabetic Nephropathy (ARTS-DN) Study Group

IMPORTANCE Steroidal mineralocorticoid receptor antagonists, when added to a renin-angiotensin system blocker, further reduce proteinuria in patients with chronic kidney disease but may be underused because of a high risk of adverse events.

OBJECTIVE To evaluate the safety and efficacy of different oral doses of the nonsteroidal mineralocorticoid receptor antagonist finerenone, given for 90 days to patients with diabetes and high or very high albuminuria who are receiving an angiotensin-converting enzyme inhibitor or an angiotensin receptor blocker.

DESIGN, SETTING, AND PARTICIPANTS Randomized, double-blind, placebo-controlled, parallel-group study conducted at 148 sites in 23 countries. Patients were recruited from June 2012 to February 2014 and the study was completed in August 2014. Of 1501 screened and 821 received study drug.

Patients were randomly assigned to receive oral, once-daily finerenone 1 mg/d, n = 92; 5 mg/d, n = 100; 7.5 mg/d, n = 97; 10 mg/d, n = 98; 15 mg/d, n = 119 or matching placebo (n = 94) for 90 days.

SURES The primary outcome was the ratio of the urinary albumin excretion at baseline. Safety endpoints were changes in estimated glomerular filtration rate.

At baseline, 78% were male. At baseline, 40.0% had an estimated glomerular filtration rate of 73 mL/min or lower. Finerenone demonstrated superior efficacy compared with placebo (p < .001 for all comparisons).

Related article page 940
Supplemental content at jama.com

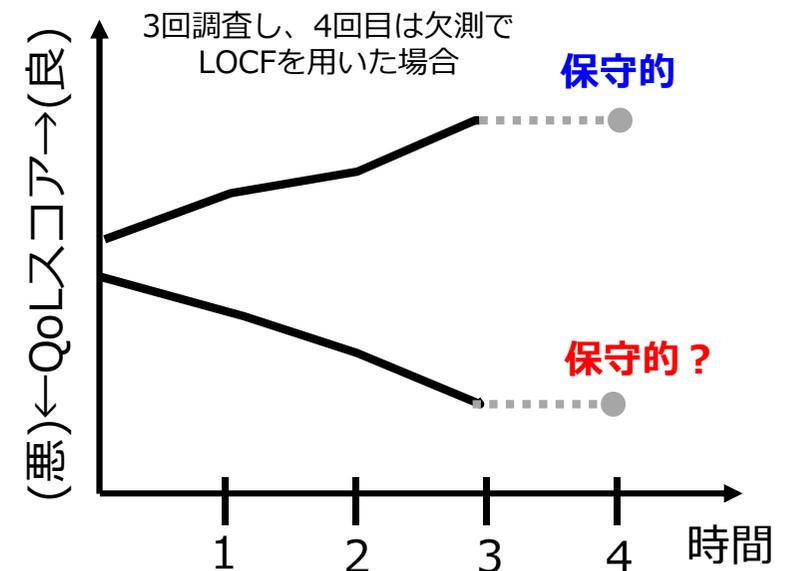
were performed. A last-observation-carried-forward (LOCF) method was applied, whereby the higher UACR value from the premature discontinuation visit), a baseline-observation-carried-forward analysis (imputing the baseline value for missing data; ie, including all missing data at day 90 v of the primary efficacy variable), a mean value imputation (imputing the value of the primary efficacy variable by the least squares mean and variance [from descriptive statistics]), and a post hoc multiple imputation. The distributional model assumptions

単一補完法の例

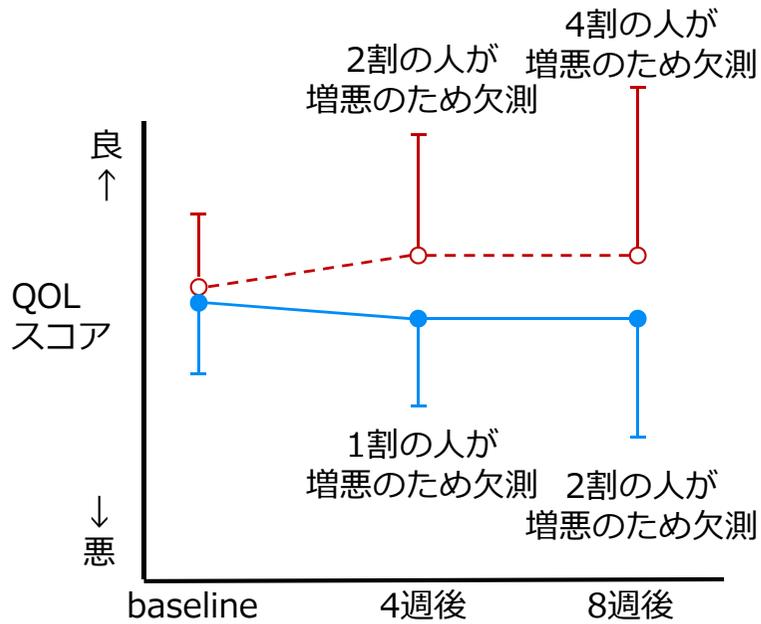
多重補完法

Last observation carried forward (LOCF) 法

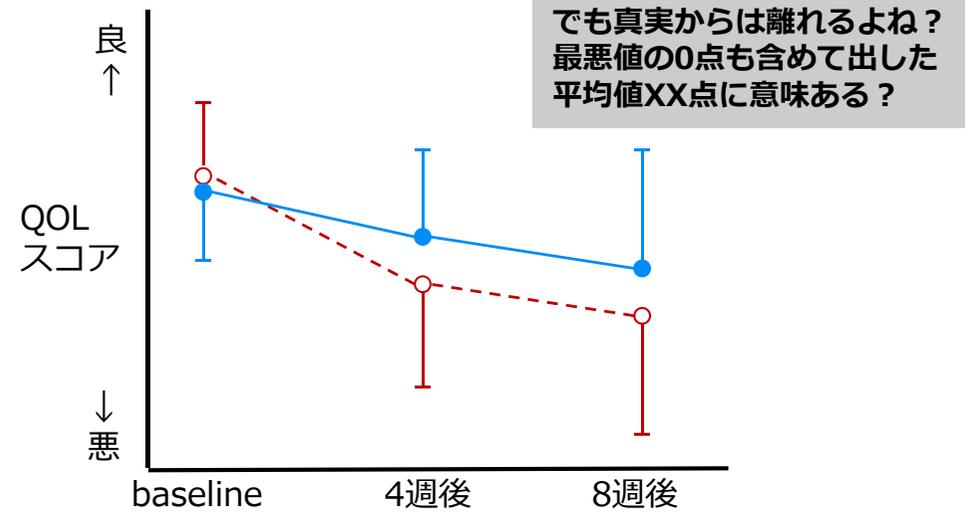
- 欠測データを1つの値で補完する方法 (=単一補完法) のうち、欠測データをその患者の最後に観察された値で補完する方法
 - 補完後はすべての患者がすべてのデータが測定された状態になるので、CC解析に準じた解析を行う
 - 保守的な方法として用いられることが多かった
- 欠測データについての仮定
 - 欠測データは最後に観察された値に等しい
- 問題点
 - 試験治療に有利なバイアスが入ることがある
 - 仮定が妥当という状況は考えにくい
 - 見かけ上データのばらつきが小さくなり、通常の解析では α エラー増加



JCOGの標準：2値に落とし込みQoL非改善と扱う

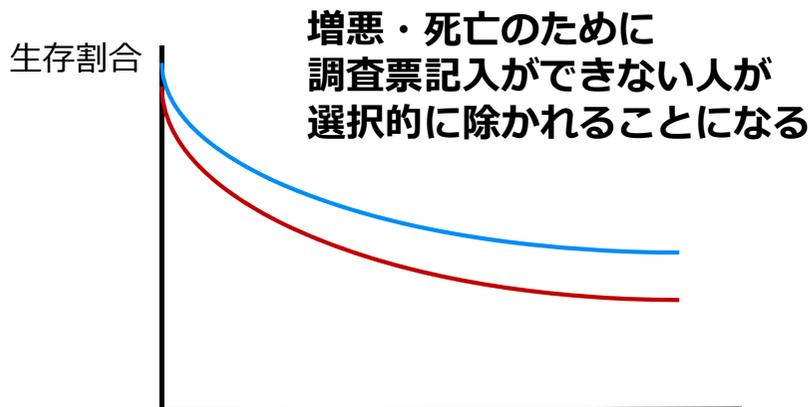


欠測の場合は最悪値（0点）を補填

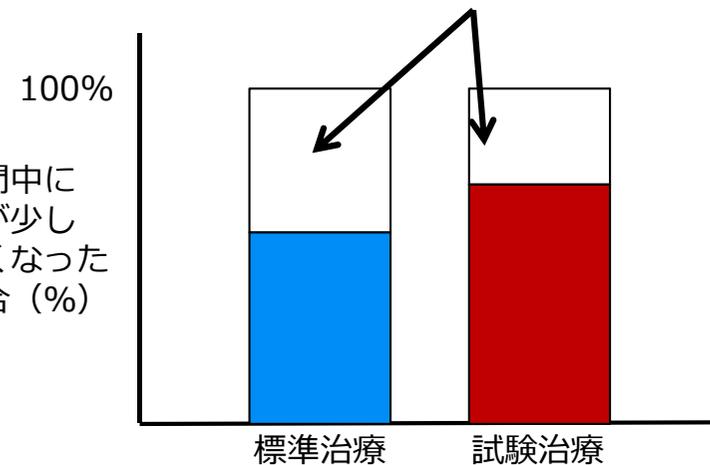


ASCOの
教育セッション
では・・・

増悪・死亡のために調査票記入ができない人は
こっちに入る（解析から除かれない）→ 過大評価にはならない



観察期間中に
スコアが少し
でもよくなった
人の割合 (%)



スコアがよくなった人が試験治療に多かった・・・
でもやっぱり臨床的な意味はよくわからない

2.欠測データに値を補完した方法

B) Multiple imputation (多重補完法)

- 欠測データに値を補完した完全なデータセットを複数作成し、個々のデータセットで求めた結果を統合して1つの治療効果の推定値や信頼区間を推定する方法
 - 補完後はすべてのデータが測定された患者ばかりになるので、CC解析に準じた解析を行う
- 欠測データについての仮定
 - MAR (MARの仮定が正しければバイアスなく推定できる)
- Single imputation (単一補完法) 法と比べた利点
 - MARが正しいとすれば妥当な解析 (バイアスが入らない)
 - 補完値がばらつくことを考慮し、 α エラーも制御できる
- NRCレポートでも推奨されている方法の1つ

方法の詳細はICRweb講義：野間久史、「臨床研究における欠測データの統計解析」を参照ください

欠測データを含んだ解析法の分類

1. 欠測データを無視した解析法
 - A) Complete case (CC) 解析
2. 欠測データに何かしらの値を補完 (imputation) した方法
 - A) Single imputation (単一補完法)
 - a. Last observation carried forward (LOCF) 法
 - B) Multiple imputation (多重補完法)
3. 欠測データのメカニズムと観測データを統計モデルで組み合わせた解析法
 - A) Inverse Probability weighting (IPW) 法
 - B) Mixed model for repeated measure (MMRM) 法
 - C) Pattern Mixture Model (PMM) 法

再掲：臨床試験での欠測データの記載例

Annals of Oncology 26: 179–185, 2015
doi:10.1093/annonc/mdu510
Published online 30 October 2014

Impact of enzalutamide on quality of life in men with metastatic castration-resistant prostate cancer after chemotherapy: additional analyses from the AFFIRM randomized clinical trial

D. Cella^{1*}, C. Ivanescu², S. Holmstrom³, C. N. Bui⁴, J. Spalding⁴ & K. Fizazi⁵

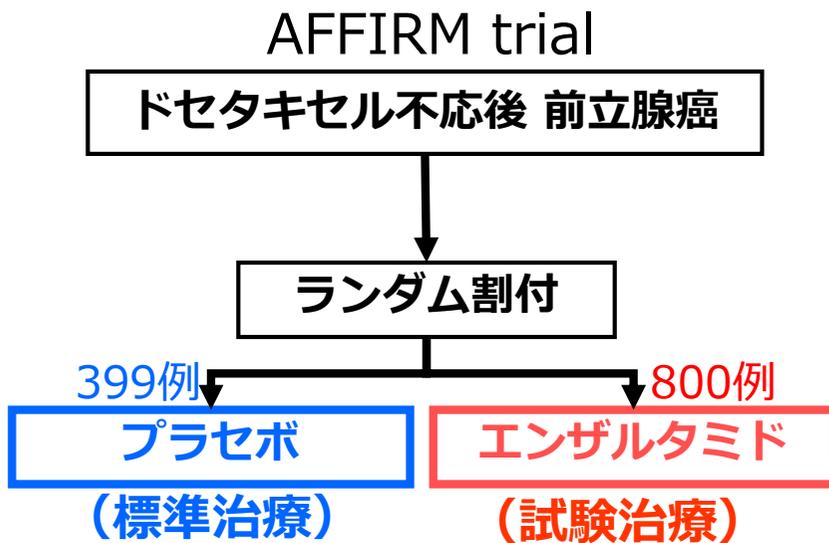
¹Department of Medical Social Sciences, Northwestern University Feinberg School of Medicine, Chicago, USA; ²Consulting, Quintiles, Hoofddorp; ³Global Data Science, Health Economics & Outcomes Research, Astellas Pharma Global Development, Leiden, The Netherlands; ⁴Health Economics & Clinical Outcomes Research, Astellas Pharma US, Inc., Northbrook, USA; ⁵Institut Gustave Roussy, University of Paris Sud, Villejuif, France

Received 22 August 2014; revised 3 October 2014; accepted 8 October 2014

Background: To present longitudinal changes in Functional Assessment of Cancer Therapy-Prostate (FACT-P) scores during 25-week treatment with enzalutamide or placebo in men with progressive metastatic castration-resistant prostate cancer (mCRPC) after chemotherapy in the AFFIRM trial.

Patients and methods: Patients were randomly assigned to enzalutamide 160 mg/day or placebo. FACT-P was completed before randomization, at weeks 13, 17, 21, and 25, and every 12 weeks thereafter while on study treatment. Longitudinal changes in FACT-P scores from baseline to 25 weeks were analyzed using a mixed effects model for repeated measures (MMRM), with a pattern mixture model (PMM) applied as secondary analysis to address non-ignorable missing data. Cumulative distribution function (CDF) plots were generated and different methodological approaches and models for handling missing data were applied. Due to the exploratory nature of the analyses, adjustments for multiple comparisons were not made. AFFIRM is registered with ClinicalTrials.gov, number NCT00974311.

Results: The intention-to-treat FACT-P population included 938 patients (enzalutamide, n = 674; placebo n = 264) with evaluable FACT-P assessments at baseline and >1 post-baseline assessment. After 25 weeks, the mean FACT-P total score



To estimate longitudinal changes in FACT-P scores from baseline, the primary analysis was carried out using a mixed effects model for repeated measures (MMRM) [12]. MMRM analysis uses all available data and assumes that any missing observations are missing at random. To address the possibility that missing data were not at random, a second analysis was carried out using a pattern mixture model (PMM) with placebo-based pattern imputation [13, 14]. In both models, the following covariates were included: treatment group; time; baseline ECOG score (0–1 or 2); average

e (HRQoL)
relevant to
using use of
y regulators
are patient-

utamide, an
ent steps in
cantly pro-
men with
CRPC) after
d secondary
on HRQoL,
essment of

100/ty-no/4.0),
please contact

A) Inverse Probability Weighting (IPW)法

- 観察されるデータの観察確率の逆数で重み付けした解析
 - 観察確率は観察時点より前のアウトカムや共変量から推定
 - ある時点の観察確率が0.25なら、観察されたデータの重みを4とする
 - 下図の場合、実際に観察されたのは25/50人でも共に100人ずつのデータが観察されたかのように解析する
- 欠測データについての仮定
 - MAR (MARの仮定が正しければバイアスなく推定できる)

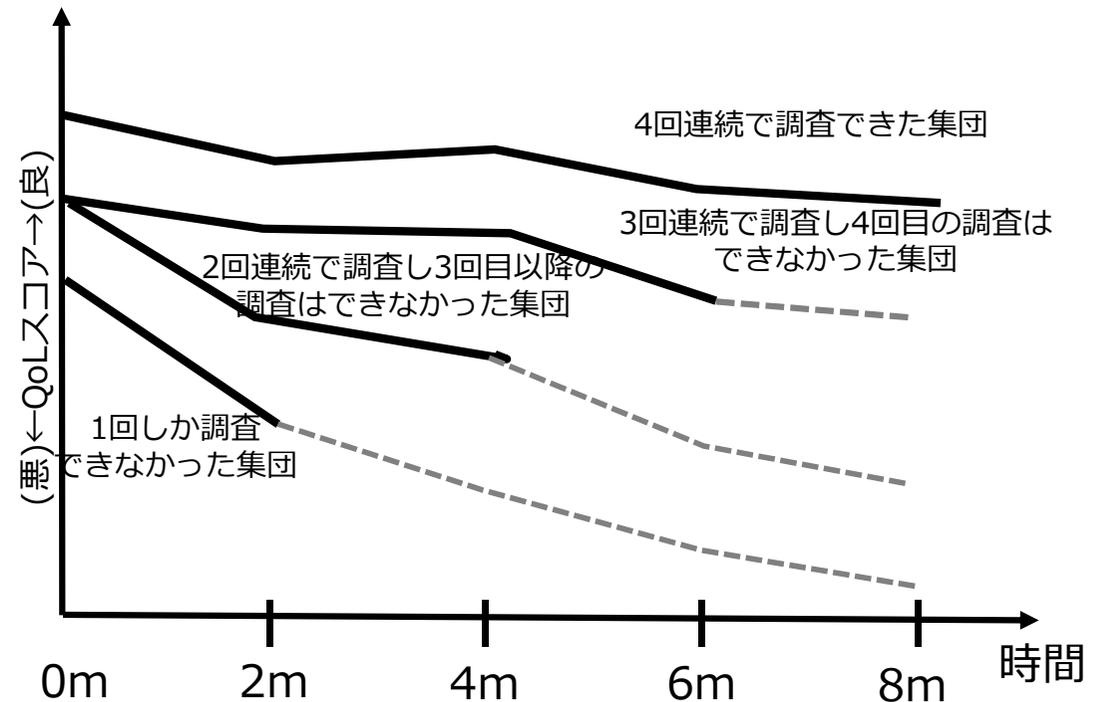


B) Mixed model for repeated measure (MMRM) 法

- 経時測定データとして解析する場合、アウトカムに特定の分布と回帰モデルを仮定した解析
 - 欠測メカニズムの分布を無視しても、観測データに基づき統計モデルを設定すれば妥当な解析になるという前提の解析法
- 欠測データについての仮定
 - MAR (MARの仮定が正しければバイアスなく推定できる)
- 利点
 - 計画段階でモデルの詳細を決めやすいので主たる解析に使いやすい
 - 通常の統計解析ソフトウェアで比較的容易に解析できる

C) Pattern Mixture Model (PMM) 法

- 集団には患者の特徴による複数のパターン（分布）が混在していると仮定し、統計モデルにパターンごとに別々のアウトカムの分布を当てはめる解析
 - 右下図の、欠測となった破線部分に仮定をおいて、破線部分の推移を推定する
- 欠測データについての仮定
 - MAR
 - MNAR
 - どちらを仮定した場合も可能
- MNARを仮定した感度解析として用いられることが多い

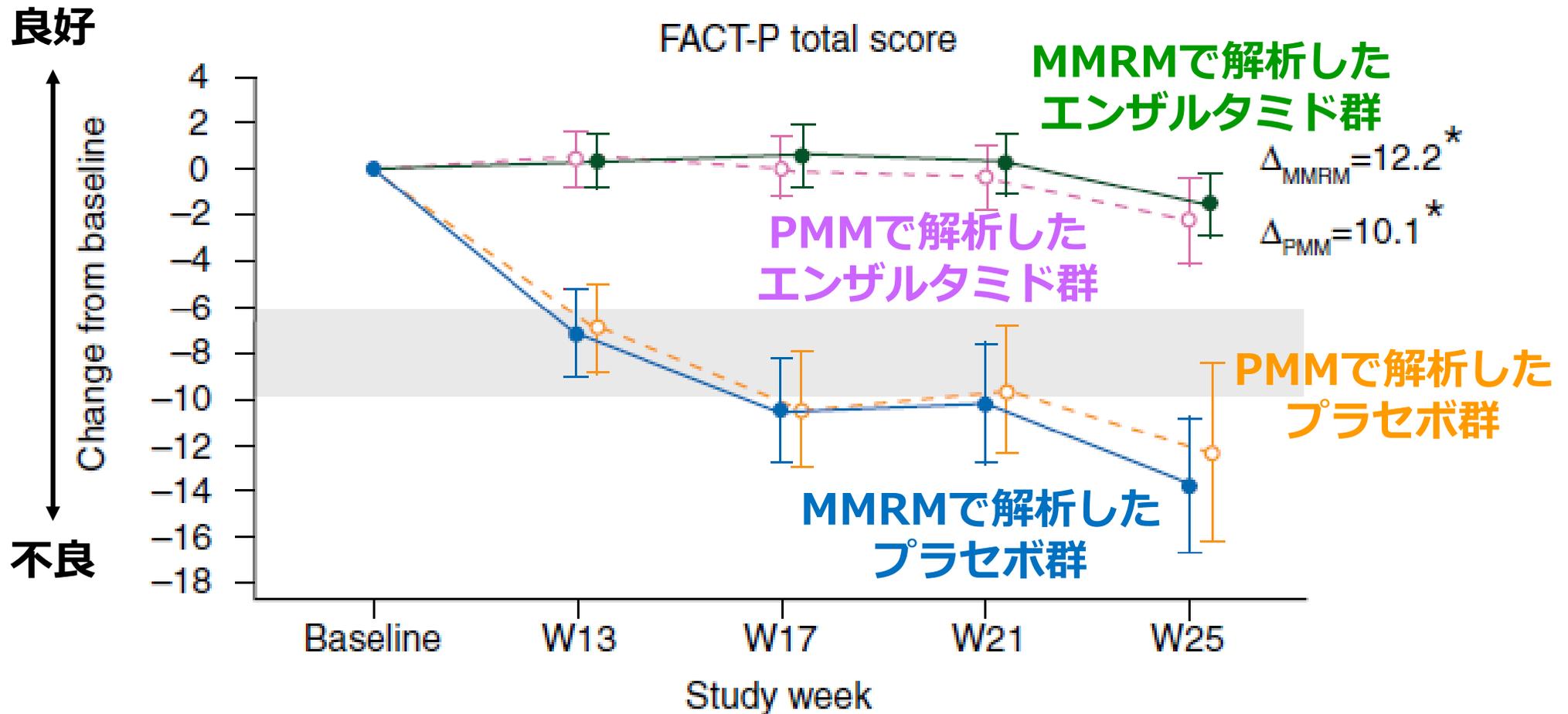


結局、どれを用いれば良い？

- 何を推定したいのか（estimand）による [NASレポートより]
誤解を恐れずに、すごく単純化すると・・・
 - 最初に割り付けた治療を行った場合の効果（ITTベースの「**有用性**」の推定）
 - 最初に割り付けた治療を遵守した場合の効果（PPSベースの「**有効性**」の推定）
 - estimandを決める → 欠測メカニズムを仮定する → 解析法を決める という流れ
 - 本講義のスコープから外れるので詳細は割愛
- 少なくとも、CC解析、単一補完法は用いるべきでない [NASレポートより]
- 異なる欠測メカニズムを仮定した感度解析を実施し、結果の頑健性を確認することが重要
 - 感度解析：一般的には異なる仮定（手法や対象）で主たる結果が変わらないこと（robustness：頑健性）を確認するための解析のこと
 - 今回の場合、MARを仮定し**MMRM**、MARを仮定し**MI**、MNARを仮定し**PMM**など

再掲：FACT-Pの解析結果

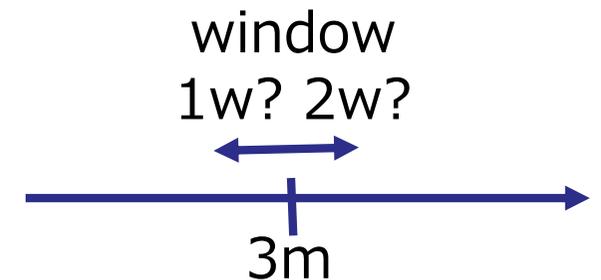
- MARを仮定したMMRMの解析、MNARを仮定したPMMの解析共に、エンザルタミドのFACT-Pはプラセボに比べて良好であった



欠測データの予防とデータ収集

比較的容易に実施できなそうなものを抜粋

1. 評価回数を最小限にする
 - 被験者や研究者の負担を減らせば回収率が高まる
2. 不要なデータは最小限にする
 - CRFなどの収集項目数を減らせば回収率が高まる
3. 記入しやすい・わかりやすいCRFデザイン
4. 来院なしで測定可能なアウトカムの設定
5. 十分な期間のwindowの設定
 - Windowから外れた測定を欠測とみなさずに済む
6. 回収率の悪い施設・研究者の除外
 - モニタリングで把握し、アクションを起こす
7. 施設スタッフへのトレーニング・意義の理解
 - 患者へのICの仕方の工夫、スタッフのモチベーション↑による効果



まとめ

- 欠測データは統計学的方法論の中でも最もホットな話題の1つであり、臨床研究における取り扱いの整理が進みつつある
- 論文やプロトコールの記載を読むときは、どの欠測メカニズムを想定した解析法かを念頭に
- NRCLレポートのkey messageが大切
 1. 欠損データが生じても解析で何とかなるわけではない
 2. 欠損データの予防が最も大切、あらゆる手段を講じて欠測を予防

参考文献

より詳細に学びたい方は以下の文献および、その中で引用されている文献をご覧ください

- 日本製薬工業協会.臨床試験の欠測データの取り扱いに関する最近の展開と今後の課題について- NASレポート, EMAガイドライン, estimandと解析方法の概説-.
http://www.jpma.or.jp/medicine/shinyaku/tiken/allotment/pdf/data_140704_1.pdf
- 日本製薬工業協会.臨床試験のestimand を理解するための因果推論.
<http://www.jpma.or.jp/medicine/shinyaku/tiken/allotment/pdf/estimand.pdf>
- 日本製薬工業協会.欠測のある連続量経時データに対する統計手法について.
<http://www.jpma.or.jp/medicine/shinyaku/tiken/allotment/pdf/statistics01.pdf>