

死亡による欠測を どう扱うか

—因果推論の応用—

京都大学大学院医学研究科
社会健康医学系専攻薬剤疫学分野
田中司朗



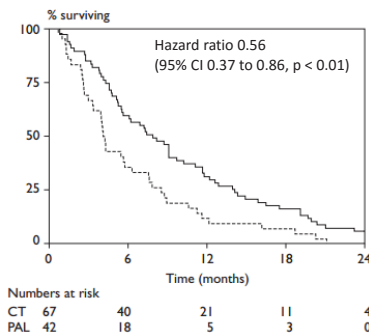
例: 非小細胞肺癌臨床試験

- 対象
 - 進行非小細胞肺癌
- ランダム化された治療
 - PAL arm
 - 緩和療法
 - CT arm
 - MMC+イホスファミド + シスプラチン
 - 最大4コース後、緩和療法
- エンドポイント
 - 全生存期間 (OS)
 - QOL (0, 3, 6, 9, 12ヶ月時点)

Item	Question
Cough	Do you have a cough?
Severe dyspnea	Do you get breathless on mild activity like dressing?
Moderate dyspnea	Do you get breathless when walking on the flat?
Mild dyspnea	Do you get breathless on stairs or walking uphill?
Haemoptysis	Have you coughed blood?
Pain	How much pain are you getting?
Appetite	Have you noticed any loss of appetite?
Anxiety	Have you been worrying?
Depression	Have you been depressed?
Dysphagia	Have you any difficulty swallowing?
Nausea	Did you feel sick during or since your last treatment? (CT arm) Have you been feeling sick? (PAL arm)
Malaise	Have you been feeling generally ill?

Cullen, et al. J Clin Oncol 1999

OSに関する生存曲線



Cullen, et al. J Clin Oncol 1999

QOLデータの欠測状況

	Q1	Q2	Q3	Q4	Q5	Total
CT arm						
Patients in study	67	67	66	53	46	
Returned questionnaires	67	63	52	45	38	265
Reasons for missing questionnaires						
Intermittent missing	0	3	1	1	0	5
Patient drop-out	0	1	13	7	8	29
PAL arm						
Patients in study	42	42	36	32	-	
Returned questionnaires	42	35	32	25	-	134
Reasons for missing questionnaires						
Intermittent missing	0	1	0	0	-	1
Patient drop-out	0	6	4	7	-	17

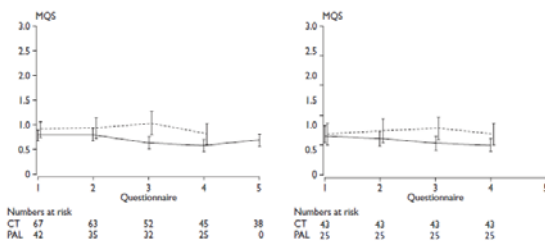
Number of missing items	Frequency (%)
0	356 (89)
1	33 (8)
2	8 (2)
3-5	2 (1)
Total	399

	CT arm	PAL arm	Total
Drop-out due to death	6	4	10
Drop-out not due to death but died within 18 weeks	9	10	19
Drop-out not due to death and died after 18 weeks	13	3	16
Drop-out last known to be alive	1	0	1
Total	29	17	46

Cullen, et al. J Clin Oncol 1999

QOLデータの経時推移

- 測定された全データの平均 (all available data)
- データが完全な患者の平均 (complete-case data)



Cullen, et al. J Clin Oncol 1999

平均の比較だとなぜだめか?

- 比較可能性の問題
 - 治療群間で欠測割合に差があると、QOLの平均に差があったとしても、治療の効果なのか、欠測の影響なのか分からない
 - 5/67 vs 1/42
 - 29/67 vs 17/42
- 過大評価の問題
 - 一般に、状態の悪い患者に欠測が生じやすい
 - そのためQOLの平均には過大評価のバイアスが生じる

	Q1	Q2	Q3	Q4	Q5	Total
CT arm						
Patients in study	67	67	66	53	46	
Returned questionnaires	67	63	52	45	38	265
Reasons for missing questionnaires						
Intermittent missing	0	3	1	1	0	5
Patient drop-out	0	1	13	7	8	29
PAL arm						
Patients in study	42	42	36	32	-	
Returned questionnaires	42	35	32	25	-	134
Reasons for missing questionnaires						
Intermittent missing	0	1	0	0	-	1
Patient drop-out	0	6	4	7	-	17

THE NEW ENGLAND JOURNAL OF MEDICINE

SPECIAL REPORT

The Prevention and Treatment of Missing Data in Clinical Trials

Roderick J. Little, Ph.D., Ralph D'Agostino, Ph.D., Michael L. Cohen, Ph.D., Kay Dickerson, Ph.D., Scott S. Emerson, M.D., Ph.D., John T. Farrar, M.D., Ph.D., Constantine Frangakis, Ph.D., Joseph W. Hogan, Sc.D., Geert Molenaar, Ph.D., Susan A. Murphy, Ph.D., James D. Neaton, Ph.D., Andrea Rotnitzky, Ph.D., Daniel Scharfstein, Sc.D., Weichung J. Shih, Ph.D., Jay P. Siegel, M.D., and Hal Stern, Ph.D.

There is **no universal method for handling missing data** in a clinical trial, since each trial has its own set of design and measurement characteristics. The range of approaches to modeling and inference is extremely broad, and no single method or class of methods is suitable for all situations. ... In general, the panel favored estimating-equation methods and **methods that are based on a statistical model for the data.**

Little, et al. NEJM 2012

欠測データの問題に関する規制

- 2008年 FDAが生物統計家のパネルを招集しガイドラインの基となる報告書作成を依頼
- 2010年 National Academy of Science (NAS) 報告書
- 2010年 Guideline on Missing Data in Confirmatory Clinical Trials (EMA)
- 2012年 Little, et al. NEJM 2012
Little, et al. Stat Med 2012
O'Neill and Temple. Clin Pharmacol Ther 2012
- 2013年 ICH E9 Revision 1 (推定対象・感度解析)

欠測データ解析でよく用いられる統計手法

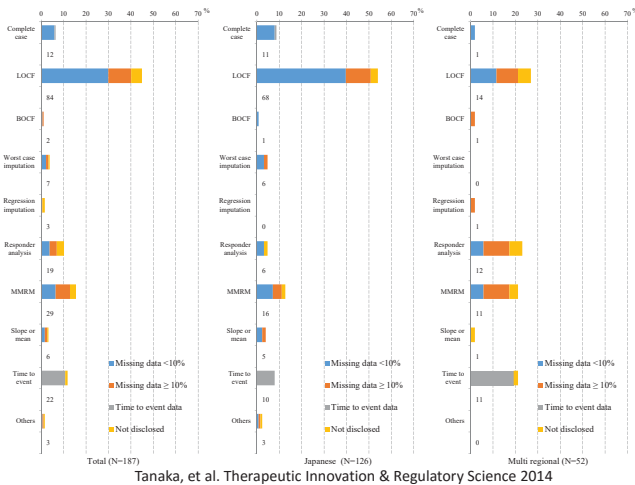
- Last observation carried forward (LOCF)
 - 欠測データは最終測定値により補完
- Worst case imputation
 - 欠測データは最悪値により補完
- Responder analysis
 - エンドポイントを、改善/非改善の2値データと定義
 - 欠測を生じた患者は非改善に分類
- Mixed model repeated measure (MMRM)
 - 測定された全データに統計モデルを当てはめる方法
 - 個人内に反復測定があるため、個人内相関が生じる
 - 個人内相関を考慮するために、個人を表す効果（変量効果）を含める

なぜMMRMが好まれるのか？

- 欠測メカニズムの分類
 - Missing completely at random (MCAR)
 - $p(M|Y^{obs}, Y^{mis})=p(M)$
 - 完全にデータに依存しない
 - Missing at random (MAR)
 - $p(M|Y^{obs}, Y^{mis})=p(M|Y^{obs})$
 - 欠測データに依存しない
 - Missing not at random (MNAR)
 - 欠測確率は、 Y^{obs} で条件付けても、欠測データに依存する
- MMRMは、MCARまたはMARの状況下では、バイアスがないことが知られている

死亡による欠測をどう扱うか？

- MMRMの問題点
 - 死亡による欠測は、ランダムな欠測（missing completely at randomまたはmissing at random）ではない（それでもなおMMRMを用いることもある）
 - つまり、MMRMの前提条件を満たさない
- Worst value imputation
 - 死亡による欠測データは効用値0により補完
- Quality adjusted life year (QALY)
 - 全生存期間とQOLを一つの指標として扱う
- Principal stratification
 - 「どの治療を受けても生存するものの集団」を推定して、その集団におけるQOLへの治療効果を推定

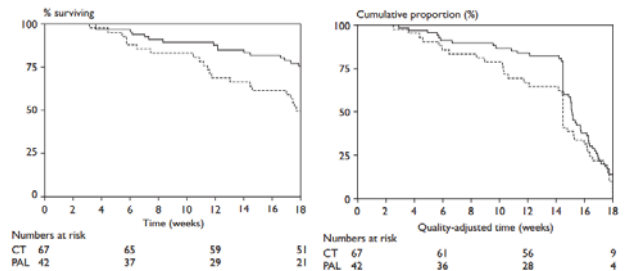


Tanaka, et al. Therapeutic Innovation & Regulatory Science 2014

Quality adjusted life year

- 全生存期間とQOLの複合エンドポイント
- 打ち切りがなければ、個々の患者の生存期間を分割し、分割ごとに効用値で重みを付けることで、QALYを計算することができる
- QALYが計算できれば通常の統計手法を適用できる

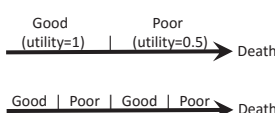
18ヶ月までのOS（左）とQALY（右）に関する生存曲線



Cullen, et al. J Clin Oncol 1999

QALYを扱うための統計手法

- 打ち切りがあると個々の患者のQALYは計算できない
- その場合でも対処法はいくつかある
 - 全生存確率と効用値の平均を組み合わせる方法
 - 全生存期間を効用値が変化する時点で分割し、各々の期間の平均を求める方法

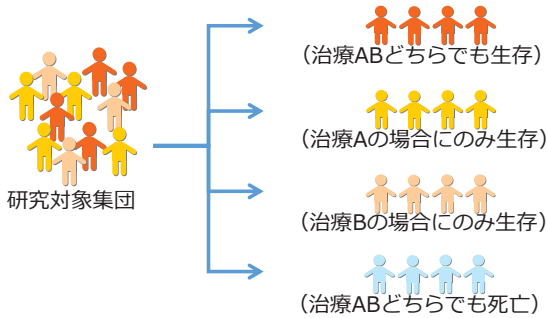


Health state	CT arm: restricted mean survival time (weeks)	PAL arm: restricted mean survival time (weeks)
GOOD1	3.37	4.22
POOR1	8.10	7.96
GOOD2	3.16	2.15
POOR2	1.66	0.19

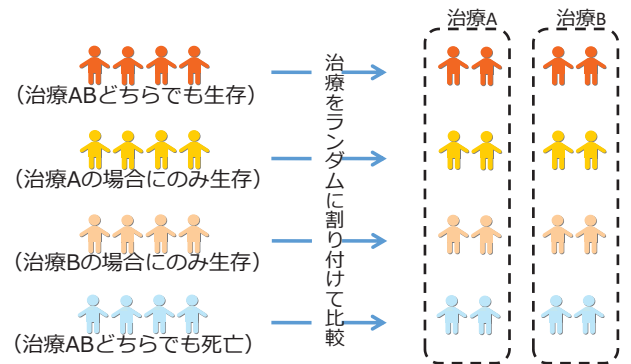
Principal stratificationの発想

- 死亡した後のQOLには意味がないのだから、「生存したものの集団においてQOLが治療により改善したかどうか」を調べたらどうだろうか
- しかし、データが完全な患者の平均（complete-case data）には、以下のような問題があった
 - 治療群間で欠測割合に差があると、QOLの平均に影響があったとしても、治療の効果なのか、欠測の影響なのか分からない
- どのようにすれば「生存したものの集団」をうまく定義できるだろうか

Principal stratificationの発想

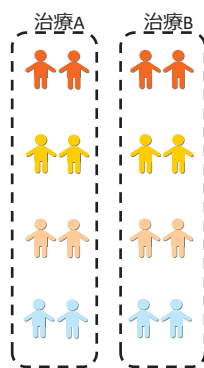


Principal stratificationの発想



Principal stratificationの発想

- 生存するかどうかは、治療の影響を受けるため、実際に観察された「生存したものの集団」内で治療群間を比較すると、バイアスがある（事後層別）
- 一方で、このように定義した集団は、ランダム化とは独立なので、バイアスが生じない



Principal stratificationの数値例

- TreatmentとControlの二つの治療を比較するランダム化臨床試験を考える
- 対象者は、四つの層に分類できると仮定する
 - Live-Live : 20%
 - Live-Death : 40%
 - Death-Live : 20%
 - Death-Death : 20%

% population	Principal stratum	Treatment		Control		Treatment effect on QOL
		$S_i(T)$	$Y_i(T)$	$S_i(C)$	$Y_i(C)$	
20	LL	L	900	L	700	200
40	LD	L	600	D	*	*
20	DL	D	*	L	800	*
20	DD	D	*	D	*	*

Principal stratificationの数値例

- 表には、仮にTreatmentに割り付けられた場合とControlに割り付けられた場合の、生存状態 (S) とQOLの期待値 (Y) が示されている
- QOLは、S=Liveの場合にのみ観察される

% population	Principal stratum	Treatment		Control		Treatment effect on QOL
		$S_i(T)$	$Y_i(T)$	$S_i(C)$	$Y_i(C)$	
20	LL	L	900	L	700	200
40	LD	L	600	D	*	*
20	DL	D	*	L	800	*
20	DD	D	*	D	*	*

Principal stratificationの数値例

- 例えば“Live-Death”の集団だと
 - Treatment: S=Live, Y=600
 - Control: S=Death, Y=*
- この集団の占める割合は40%
- つまり、TreatmentはControlに比べて生存患者を20%増やす効果がある

% population	Principal stratum	Treatment		Control		Treatment effect on QOL
		$S_i(T)$	$Y_i(T)$	$S_i(C)$	$Y_i(C)$	
20	LL	L	900	L	700	200
40	LD	L	600	D	*	*
20	DL	D	*	L	800	*
20	DD	D	*	D	*	*

Principal stratificationの数値例

- 関心のある治療効果は、“Treatment”の場合のQOLの平均と“Control”の場合のQOLの平均の差である
- “Live-Death”, “Death-Live”, “Death-Death”の集団では、一部のQOLが定義されない (*で表記)
- この表では、“Live-Live”の集団ではQOLが200ポイント改善していることが分かる

% population	Principal stratum	Treatment		Control		Treatment effect on QOL
		$S_i(T)$	$Y_i(T)$	$S_i(C)$	$Y_i(C)$	
20	LL	L	900	L	700	200
40	LD	L	600	D	*	*
20	DL	D	*	L	800	*
20	DD	D	*	D	*	*

まとめ: 統計手法の比較

- Mixed model repeated measure (MMRM)
 - 測定された全データに統計モデルを当てはめる方法
 - 死亡による欠測は、ランダムな欠測 (missing completely at random または missing at random) ではない
- Worst value imputation
 - 死亡による欠測データは効用値0により補完
- Quality adjusted life year (QALY)
 - 全生存期間とQOLを一つの指標として扱う
- Principal stratification
 - QOLへの死亡を調整した治療効果を推定

まとめ: 統計手法の比較

- QOLなどの経時データ解析ではMMRMが標準的
- しかし, 死亡による欠測があり, その頻度が治療群間で異なる場合, 「バイアスがある」と言われると反論できない
- QALYを用いた解析では, この問題は生じない
- しかし, 延命効果とQOL改善効果を区別することはできない
- QOL改善効果を推定するための統計手法として, worst value imputationとprincipal stratificationがある