

45

追加調査を必要とする 場合の研究デザイン

46

TRUMPデータ

- 移植実施
- ↓
- データ登録（現在はon site、昔は年1回の郵送調査）
- ↓
- 年次フォローアップ

47

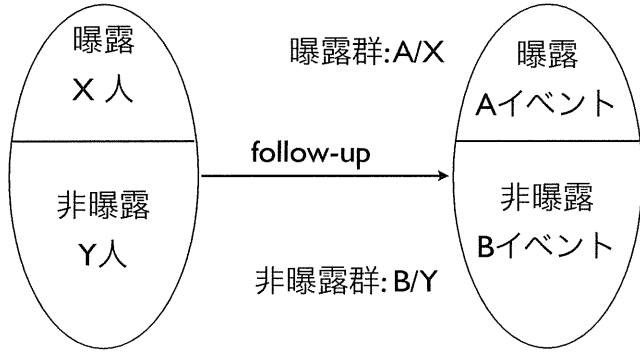
TRUMPデータ

- 基本的に適応できる研究デザインは
 - コホート研究
 - 症例対照研究（コホート内症例対象研究を勧める）
- の二つ

48

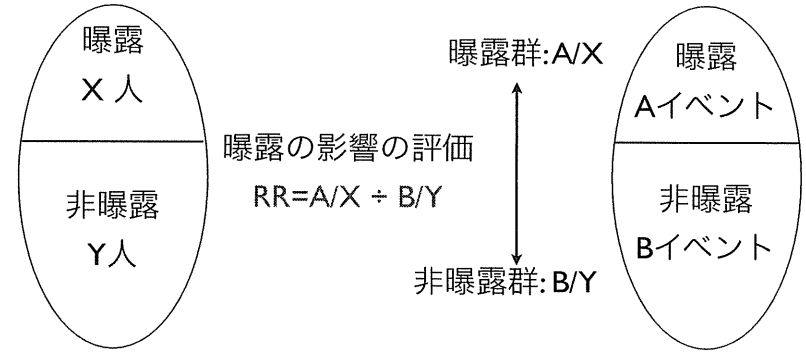
前向きコホート研究・再掲

曝露の測定を最初にする



49

前向きコホート研究・再掲

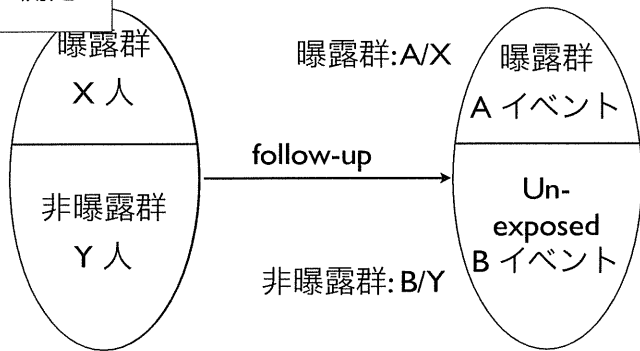


50

後向きコホート研究・再掲

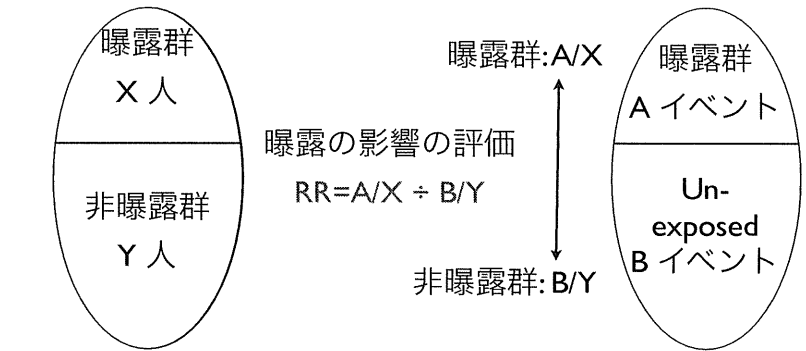
診療録レビュー等
・ 曝露の測定

既にあるリストに基づき、
「過去の曝露」を調べる



51

後向きコホート研究・再掲



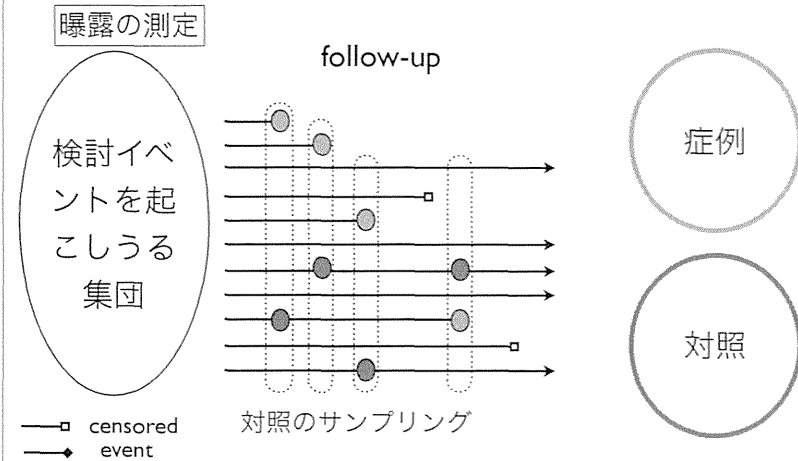
52

コホート研究

- 研究上の疑問：曝露要因のアウトカムへの影響
- アウトカムは一つに絞った方が良いが、複数のアウトカムを使い得る
- 要因曝露：もし最初のTRUMPデータに仕込まれていればそれを使える。そうでなければ全調査が必要。

53

コホート内症例対照研究



54

コホート内症例対照研究(続き)

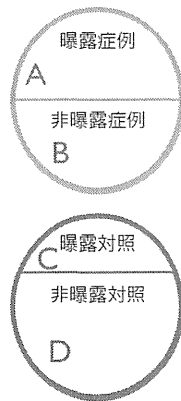
Odds in 症例: A/B

曝露の影響の評価

オッズ比
= (A/B) / (C/D)

Odds in 対照: C/D

後向きコホートでも同様の事が可能!



55

コホート内症例対照研究(続き)

- コホート研究のメリットを享受できる
 - 因果推論はコホートの枠組み
- マッチングを行われなかった対象者を欠測データとしてimputationすることも可能
- 追加調査が必要な場合に全例を調査しなくても済む
 - 実施可能性を高める
 - 但し、曝露要因がTRUMPに報告はされていないが、測定されている必要がある
- 症例・対照を定義する必要があるため、使えるイベントはただ一つだけ！
 - aGVHD、OSを両方一つの研究内で見ることが出来ない。

56

コホート内症例対照研究

- 研究上の疑問：一つの曝露のアウトカムへの影響
- アウトカムはただ一つに絞らざるを得ない。症例対照研究はアウトカムが起きた例と起きてない例を用いる
- 一つの曝露：もし最初のTRUMPデータに仕込まれていればそれを使える。そうでなければ全調査が必要
- 対照者をどう選ぶか？ STATA sttocc

57

原則

- 元データは読み込むが上書きはしない
- 全ての解析をスクリプトでやる
 - 逐次解析をしない（しても良いが、最終解析は絶対逐次解析にしない）
 - 必ずログが残る解析にする
 - 再現性が保てない解析は解析でない

59

おまけ 統計パッケージで解析する際の基本原則

- set more off (スクロールストップがかからなくなる)
- set logtype t (ログをテキスト形式にする)
- log using XXXX.txt, replace
- use xxxx.dta, clear
- save work_XXXX.dta, clear
- describe (データの場所、人数、変数とうの基本情報)
- codebook (変数の詳細な情報がでる、毎回は不要)

- /*for table 1*/
- tab xxxx yyyy, col row chi2 exact

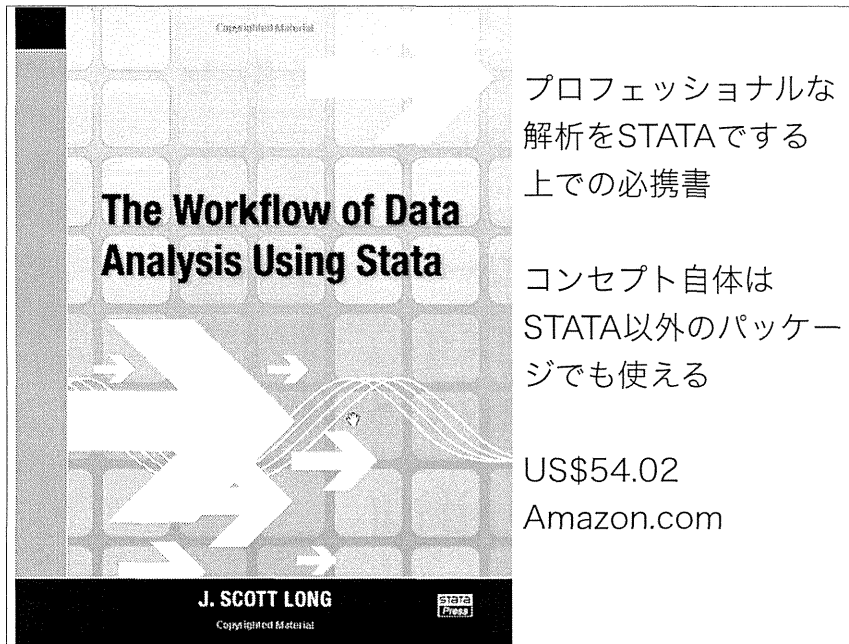
- /*for table 2*/
- stset os, f(death==1)
- stsum

- sts graph, by(sex)
- graph save xxx.gph, replace

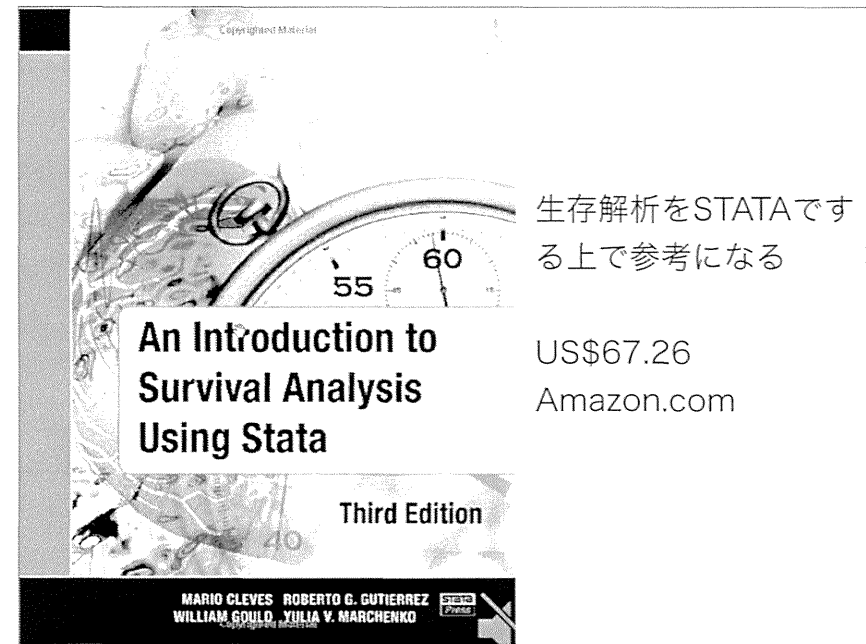
- log close

58

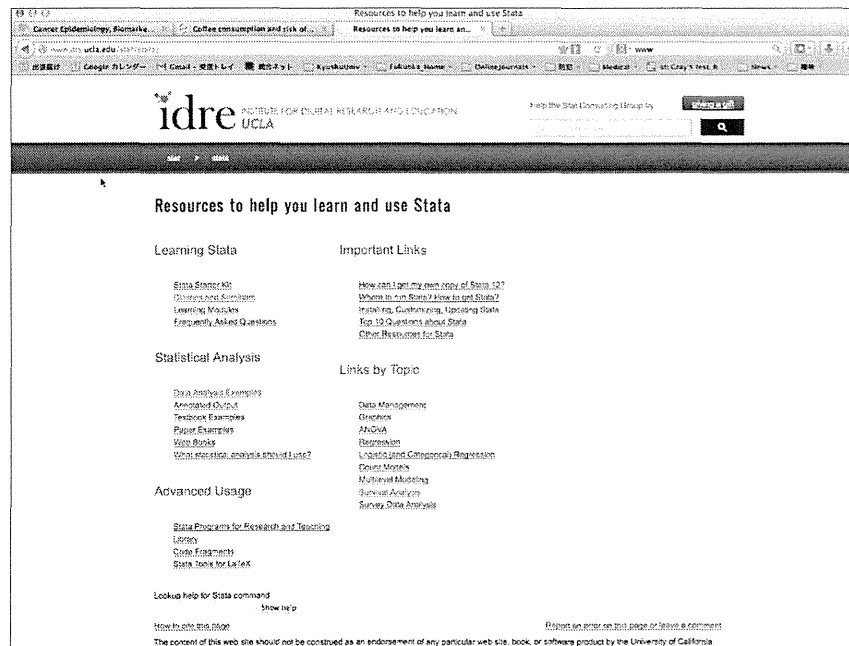
60



61



62



63

多変量解析モデル構築の具体例

多変量解析モデル構築の具体例

熱田 由子

日本造血細胞移植データセンター

多変量解析を実施する目的

- リスク因子解析
- 補正解析 (補正比較)
 - 比較可能性の保たれていない対象

Risk factor analyses

Table 1. Risk factors for second solid cancers, according to time since transplantation and type of solid cancer, among ≥ 1 year survivors of HCT

Variable	All solid cancers			Non-BCC cancers			BCC cancers		
	No. cases	RR	95% CI	No. cases	RR	95% CI	No. cases	RR	95% CI
General risk factors, ≥ 1 y survivors									
TBI	126	1.79†	1.06-3.00	93	2.29†	1.18-4.88	32	1.11	0.51-2.53
LFL no TBI	9	2.95†	1.15-7.95	4	3.94	0.92-14.77	5	2.21	0.62-7.94
Chronic GVHD*	56	1.55†	1.10-2.16	21	0.70	0.41-1.12	35	5.04†	2.30-9.00
Risk factors by time after HCT									
1-4 years after HCT									
TBI	47	1.37	0.75-2.66	37	1.63	0.76-3.81	10	0.96	0.33-3.06
LFL no TBI	2	2.09	0.25-12.40	1	1.33	0.06-11.91	1	6.52	0.12-150
Chronic GVHD*	25	1.81†	1.06-3.03	13	1.06	0.52-2.01	12	6.79†	2.44-21.73
≥ 5 y after HCT									
TBI	73	2.74†	1.17-6.82	56	6.93†	1.58-30.52	22	1.33	0.43-4.33
LFL no TBI	7	3.83†	1.24-11.57	3	11.54†	1.62-76.57	4	2.17	0.52-8.53
Chronic GVHD*	21	1.39	0.97-2.12	8	0.45†	0.20-0.91	22	4.35†	2.06-8.57

Rizzo et al. Blood 2009;113:1175

Risk factor analyses

- イベント発症のリスク因子
- 死亡であれば、予後因子
- リスクモデルの構築 (リスクスコアなど)
- ハイリスク集団の特定
- 個別集団への別治療法の開発

Adjusted analyses

Table 2. Results of multivariate analysis of outcomes in 173 recipients of cord blood and 311 recipients of bone marrow with acute myeloid leukemia, and 114 recipients of cord blood and 222 recipients of bone marrow with acute lymphoblastic leukemia

Outcome	Acute myeloid leukemia		Acute lymphoblastic leukemia	
	RR (95% CI)	P	RR (95% CI)	P
Overall survival*				
BM	1.00		1.00	
CB	1.45 (1.18-1.76)	.003	1.66 (1.21-1.57)	.001
Leukemia-free survival*				
BM	1.00		1.00	
CB	1.49 (1.28-1.74)	.001	1.32 (1.01-1.76)	.03
Relapse†				
BM	1.00		1.00	
CB	1.21 (1.09-1.34)	.001	1.42 (1.24-1.61)	<.001
TRM‡				
BM	1.00		1.00	
CB	1.47 (1.29-1.72)	.0001	1.61 (1.38-1.78)	.0001
Neutrophil recovery				
BM	1.00		1.00	
CB	0.41 (0.30-0.51)	<.001	0.37 (0.28-0.50)	<.001
Platelet recovery*				
BM	1.00		1.00	
CB	0.34 (0.27-0.44)	<.001	0.42 (0.33-0.54)	<.001
Acute GVHD§				
BM	1.00		1.00	
CB	6.30 (4.16-9.15)	<.001	0.61 (0.50-0.76)	<.001
Chronic GVHD¶				
BM	1.00		1.00	
CB	0.94 (0.63-1.42)	.79	1.02 (0.68-1.55)	.71
Chronic GVHD, extensive type††				
BM	1.00		1.00	
CB	0.56 (0.19-0.72)	.001	0.58 (0.23-1.26)	.14

Atsuta et al. Blood 2009;113:1631

従属変数と独立変数

- $y = f(x_1, x_2, x_3, \dots, x_n)$
 - y: 従属変数
 - x: 独立変数
- $y = \alpha + \beta_1x_1 + \beta_2x_2 + \beta_3x_3 + \dots + \beta_nx_n$
 - y: 魅力的な職場スコア
 - x1: 通勤の便がよい
 - x2: 職員食堂の質
 - x3: 月給がよい
 - x4: 社員の教育体制
 - x5: 福利厚生が充実

変数の選択
変数の定義
連続変数のカットオフポイント

サンプル集団
想定している母集団

モデルの選択

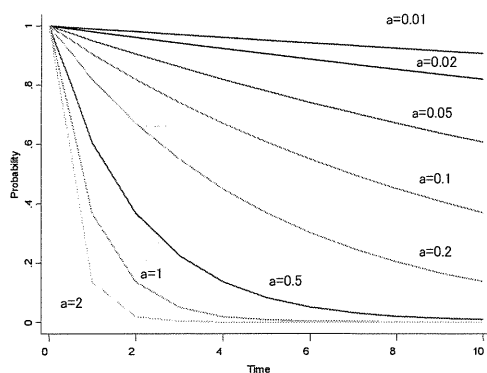
- Logistic regression model
 - $y = ax_1 + bx_2 + cx_3 + \dots$
 - 時間は考慮されない
- Cox proportional hazard model
 - Censored caseを対象に含めることが可能なモデル
 - 生存解析において、最も用いられている。
- Fine and Gray's model (competing risk regression model)
 - Competing risk eventsで用いる

Cox proportional hazard model

- モデル ← 相対的な死亡確率(ハザード比)は時間によらず一定
- ベースラインを設定して、曲線同士の比を見ている。

Factor	Relative risk	95%CI	P
Age > 40	1.30	1.13-1.49	0.003
Advanced disease	2.64	2.22-3.13	<0.0001

HR=2.0のインパクト



浜島信之 多変量解析による臨床研究 第3版

リスクスコア

- 対象
 - TRUMP サンプルデータセット
 - MAC, 16歳以上の成人
- 検討変数
 - 疾患名、疾患病期
 - 患者年齢
 - 患者性別、性別一致度
 - ドナー、幹細胞ソース

→ Do file ^

連続変数の扱い

- 連続変数の場合

Factor	Hazard ratio	P
Age	1.018617	<0.001

- Ageが1歳あがるごとにどれだけリスクが上がるか
- RRはどの年齢間でも全て同じ

順列変数も同様

連続変数の扱い

- 連続変数の場合

Factor	Hazard ratio	P
Age	1.018617	<0.001
Age > 40	1.485875	<0.001

- Cut off point
 - 過去の知見から
 - Median/percentile値を用いる
 - もっとも差が出る場所を探す (false significance になりえるので注意必要)

→ Do file ^

よく質問を受ける変数の扱い

- ドナー
 - 血縁・非血縁
- 幹細胞ソース
 - 骨髓・末梢血・臍帯血
- HLA一致度
 - 一致・不一致 (Related, UBM, UCBTにおいて一般的に用いられる基準は異なる)

考えておくべきこと

- 変数どうしの関係(相関)
 - 幹細胞ソース、臍帯血はほぼ全て非血縁
- Valueどうしを掛け合わせた場合の違い (interaction)
 - HLA不一致度がアウトカムに与える影響は、骨髓と臍帯血では異なる
- そもそも、検討したいquestionに対して、これらを一緒に検討できるか？

ドナー・幹細胞・HLA variable...

- RBM (/PB), HLA matched
- RBM (/PB), HLA mismatched
- RPB, HLA matched
- RPB, HLA mismatched
- UBM, HLA 8/8 matched
- UBM, HLA 7/8 matched
- UBM, HLA 6/8 or less matched
- UCB

→Do file←

Interaction

The interaction between main effect categories of preparative regimen and donor type (sibling, unrelated) was tested using a cutoff of 0.01 for significance. Adjusted LFS and survival probabilities were estimated through the direct adjusted survival curves estimation method.³⁸ SAS software, version 9.2 (SAS Institute, Cary, NC) was used in all analyses.

No interactions between the main effect and the covariates in the final model, including donor type ($P = .04$ for survival and $P = .11$ for LFS) and age ($P = .43$ for survival and $P = .22$ for LFS) were significant.

CY+TBI vs. ivBU+CY

blood

2013.122:3863-3870
Prepublished online September 24, 2013.
doi:10.1182/blood-2013-07-514448

Better leukemia-free and overall survival in AML in first remission following cyclophosphamide in combination with busulfan compared with TBI

Edward A. Copelan, Betty K. Hamilton, Belinda Avalos, Kwang Yoo Ahn, Brian J. Bolwell, Xiaochun Zhu, Mahmoud Ajurf, Koen van Besien, Christopher Ericsson, Jean-Yves Cahn, Luciano J. Costa, Marcos de Lima, Robert Peter Gale, Gregory A. Hale, Joerg Hattler, Maha Hamadani, Yoshihiro Inamoto, Rammurti T. Kambh, Mark R. Litzow, Alison W. Loren, David J. Marks, Eduardo Olavarria Viquei, Roy, Mitchell Sabicoff, Susan N. Savani, Matthew Seifel, Harry C. Schouten, Celestin Ustun, Edmund K. Waller, Daniel J. Weisdorf, Baldeep Vohk, Mary M. Horowitz, Mukta Avora, Jeff Szer, Jorge Cortes, Matt E. Kalaycio, Richard T. Maziarz and Václav Šabek

Interaction

ivBU+CY, related donor	ivBU+CY, unrelated donor
CY+TBI, related donor	CY+TBI, unrelated donor

Unrelated vs. related	1.3
CY+TBI vs. ivBU+CY	1.3
CY+TBI and unrelated vs. ivBU+CY and related	1.3 x 1.3

しかし、Unrelated donorからの移植では、ivBU+CYで正着不全が多く、Related donorの場合とことなり、CY+TBIのほうがNRMが低かった



“CY+TBI and unrelated vs. ivBU+CY and related”が予想外の値に

Interactionあり

患者・ドナー性別で考えたこと

- 患者性別
- 患者・ドナー性別一致度
- 患者が女性の場合と、男性の場合では、性別一致度の影響が異なるかもしれない
- 過去の報告で、F to M (特に経産婦から男性患者への移植)の際にNRMが高いというものがある
- 上記2つの変数間にinteractionがあるのではないか?

有意なinteractionを認めた場合

- Variable of interest
 - ivBU+CY研究で、Rel/UR変数とinteractionがあったら
 - Rel/URで分けた解析を実施する。同一モデルでの検討は困難
- 医学的に妥当であれば、2つの変数から、複数のvalueを持つ変数を作成

Effect of HLA disparity after UCBT

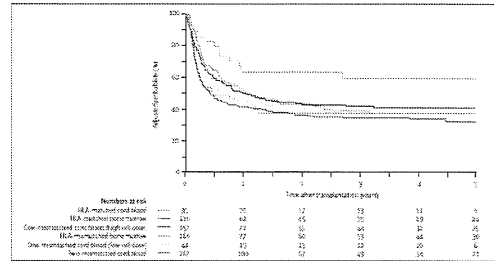
Different effects of HLA disparity on transplant outcomes after single-unit cord blood transplantation between pediatric and adult patients with leukemia

Yoshiko Aizuta,¹ Junya Kanda,² Minoko Takanashi,³ Yasunori Morishima,⁴ Shuichi Taniguchi,⁵ Satoshi Takahashi,⁶ Hiroyasu Ogawa,⁷ Kazuteru Ohashi,⁸ Yujin Ohno,⁹ Yasushi Onishi,¹⁰ Nobuyuki Aotsuka,¹¹ Tokiko Nagamura-Inoue,¹² Koji Kato,¹³ and Yoshinobu Kanda,¹⁴ on behalf of the HLA Working Group of the Japan Society for Hematopoietic Cell Transplantation

¹Department of Hematopoietic Stem Cell Transplantation Data Management / Biostatistics, Nagoya University Graduate School of Medicine, Nagoya; ²Division of Hematology, Saitama Medical Center, Ichi Medical University, Saitama; ³The Japanese Red Cross Tokyo Blood Center, Tokyo; ⁴Division of Epidemiology and Prevention, Aichi Cancer Center Research Institute, Nagoya; ⁵Department of Hematology, Tomonaga Hospital, Tokyo; ⁶Department of Molecular Toxicology, The Institute of Medical Science, The University of Tokyo, Tokyo; ⁷Division of Hematology, Department of Internal Medicine, Hyogo College of Medicine, Hyogo; ⁸Hematology Division, Tokyo Metropolitan Cancer and Infectious Diseases Center, Nerima Hospital, Tokyo; ⁹Department of Internal Medicine, Atsuta City Municipal Medical Center, Midori-ku; ¹⁰Department of Hematology and Rheumatology, Tokoku University Hospital, Sendai; ¹¹Department of Hematology and Oncology, Japanese Red Cross Nara Hospital, Nara; ¹²Department of Cell Processing and Transfusion Research Hospital, The Institute of Medical Science, The University of Tokyo, and Tokyo Cord Blood Bank, Tokyo; and ¹³Department of Pediatrics, Japanese Red Cross Nagoya First Hospital, Nagoya, Japan

Effect of HLA disparity on transplant outcomes

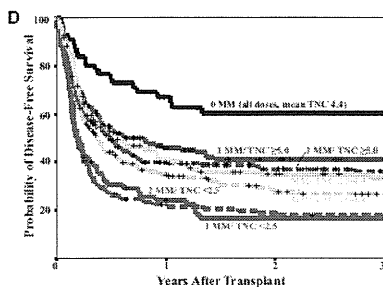
- Differences reported from analyses of mainly pediatric population
- HLA 6/6 matched UCB showed better survival



Eapen et al. Lancet 2007

Effect of HLA disparity on transplant outcomes

- HLA matching and infused total nucleated cell associated with outcome



Barker et al. Blood 2010

JSHCT HLA WG study Ped (A) and Adult (B)

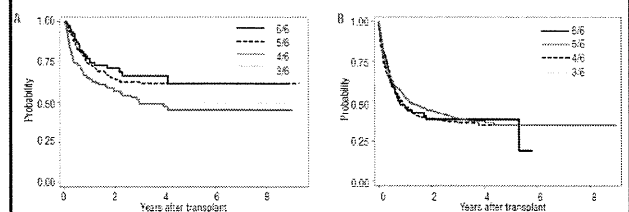


Figure 1. Unadjusted probabilities of overall survival in HLA disparity groups for pediatric (A) and adult (B) recipients with leukemia. (A) In children, the unadjusted probabilities of survival at three years post-transplant were 66% for recipients of HLA matched (6/6), 62% for one locus mismatched (5/6), 45% for two loci mismatched (4/6), and 62% for three loci mismatched (3/6) single-unit unrelated cord blood (P=0.032). (B) In adults, these probabilities were 38%, 37%, 39%, and 40% respectively (P=0.667) (6).

このstudyにおけるinteraction確認

- 背景より、成人と小児を分けて解析することを計画
- 有意なinteractionが存在したことを結果で示し、成人・小児においてHLA disparityがOSに与える影響に違いがあることを示し、上記のようなseparate analysesを行ったことが統計学的に妥当であったことを示した。

A test of the interaction between HLA disparity and age (adult vs. child) revealed that the effect of HLA disparity on overall survival differed significantly between the pediatric and adult patient groups ($P=0.009$ for HLA disparity of 0-1 mismatches vs. 2-3 mismatches).

HLA不一致座の年代別影響 (HLA WG)

bjh research paper

Impact of a single human leucocyte antigen (HLA) allele mismatch on the outcome of unrelated bone marrow transplantation over two time periods. A retrospective analysis of 3003 patients from the HLA Working Group of the Japan Society for Blood and Marrow Transplantation

Yoshitsugu Kanda,¹ Junya Kanda,¹ Yoshiko Arita,² Yoshitsugu Maeda,¹ Tatsuo Ichimura,³ Kazuteru Okada,⁴ Takahiro Inokada,⁵ Koichi Miyamura,⁶ Hiroyasu Iida,⁷ Takahiko Mori,⁸ Koji Inoue,⁹ Tetsuya Ito,¹⁰ Kazuo Kawa,¹¹ Satoshi Morita,¹² and Kazuo Morikawa^{1*}

Summary

A previous Japanese study revealed that a human leucocyte antigen (HLA) A or B allele mismatch was associated with higher overall mortality, whereas an HLA C or DRB1 allele mismatch did not affect mortality after serologically matched unrelated bone marrow transplantation (BMT). This study re-analyzed 3003 adult patients who underwent unrelated BMT from

Interactionは主たる変数の一つ

HLA (early years)		
Match	1	
A mismatch	1.16 (0.84-1.59)	0.37
B mismatch	2.47 (1.16-5.24)	0.019
C mismatch	0.96 (0.73-1.26)	0.77
DRB1 mismatch	0.83 (0.58-1.19)	0.32
HLA (late years)		
Match	1	
A mismatch	1.24 (0.95-1.62)	0.12
B mismatch	1.36 (0.78-2.35)	0.28
C mismatch	1.35 (1.15-1.59)	0.0003
DRB1 mismatch	1.45 (1.20-1.75)	0.0001
Transplantation time period		
Early period	1.00	
Late period	0.74 (0.63-0.86)	0.00016
Interactions		
Time period * A mismatch	1.07 (0.70-1.63)	0.75
Time period * B mismatch	0.55 (0.22-1.40)	0.21
Time period * C mismatch	1.41 (1.03-1.93)	0.032
Time period * DRB1 mismatch	1.74 (1.16-2.61)	0.0072

→Do file←

リスクスコアモデルの構築

Univariate analysis

- 検討因子における単変量解析
 - 因子の選出
 - 因子のカテゴリライズ方法
 - 連続変数の場合のcut-off valueの検討
 - 因子の医学的interpretation
- 検討が困難な因子
 - Missing valueの多いもの
 - Value間の偏りの著しいもの(10% >、あるいはNが著しく低いもの)

リスクスコアモデルの構築 一つの例

- 単変量解析において $P<0.1$ であるものを抽出
- ステップワイズ法を用いて変数の絞り込み($P<0.05$)を行い、最終モデルを構築する
- HRより β 値を算出し、項目ごとのスコアとする。

→Do file←

今回のリスクスコアモデル構築における問題点

- 対象、変数における検討が不十分
 - 複数回移植症例が含まれている
 - 検討変数が限られている
- リスクスコアモデルの使用目的→介入に関する変数をどう考えるか
 - 施設間の成績の差を見たい場合に補正するためのリスクスコア
 - ハイリスク集団を特定し、個別治療開発に結び付けたい

多変量解析を実施する目的

- リスク因子解析
- 補正解析(補正比較)
 - 比較可能性の保たれていない対象

Adjusted analyses

Table 2. Results of multivariate analysis of outcomes in 173 recipients of cord blood and 311 recipients of bone marrow with acute myeloid leukemia, and 114 recipients of cord blood and 222 recipients of bone marrow with acute lymphoblastic leukemia

Outcome	Acute myeloid leukemia		Acute lymphoblastic leukemia	
	RR (95% CI)	P	RR (95% CI)	P
Overall survival*				
CB	1.00		1.00	
BM	1.42 (1.06-1.91)	.020	1.66 (1.31-2.07)	.001
Leukemia-free survival†				
CB	1.00		1.00	
BM	1.49 (1.02-2.15)	.017	1.82 (1.45-2.31)	<.001
Relapses				
CB	1.00		1.00	
BM	1.41 (1.01-1.97)	.048	1.47 (1.04-2.07)	.034
TRM‡				
CB	1.00		1.00	
BM	1.47 (1.01-2.14)	.044	1.61 (1.19-2.17)	.004
Neutrophil recovery§				
CB	1.00		1.00	
BM	0.41 (0.30-0.54)	<.001	0.37 (0.29-0.46)	<.001
Platelet recovery¶				
CB	1.00		1.00	
BM	0.34 (0.27-0.44)	<.001	0.43 (0.33-0.56)	<.001
Acute GVHD**				
CB	1.00		1.00	
BM	0.40 (0.30-0.51)	.001	0.41 (0.30-0.56)	<.001
Chronic GVHD***				
CB	1.00		1.00	
BM	0.94 (0.69-1.27)	.70	1.06 (0.80-1.41)	.71
Chronic GVHD, extensive type††				
CB	1.00		1.00	
BM	0.39 (0.19-0.77)	.004	0.48 (0.28-0.82)	.005

Atsuta et al. Blood 2009;113:1631

Variable of interest

- 既に挙げた例においては
 - UCB vs. UBM
 - ivBU vs. CYTBI
 - HLA disparity, 6/6 vs. 5/6, 4/6など
- Tables in the paper
 - Variable of interest結果のみが表示されていることが多い(presentation)
 - 最終モデルには、他の因子が含まれている

補正因子

- 候補因子の選択方法
 - 背景でvariable of interestのvalueにおいて偏りをみとめるもの
 - 検討アウトカムに影響をあたえるもの
 - その他アウトカムへの影響因子であることが既報告であるもの
- 補正因子のselection
 - 先に示した考え方と基本的には同じ

補正因子のselection

- 変数を全て用いるフルモデル
 - 単変量解析である程度以上のインパクト(P値で評価されることが多い)があったもの全てを投入する
- 何等かの方法で、補正因子を最終モデルまでに絞り込む
 - Variable of interestは常にモデルに残した状況で他の因子をstepwiseで除く
 - Variable of interestを除いたモデルを構築し、その最終モデルでの因子で補正する

補正因子におけるmissing value

- 多変量モデルでは、complete datasetでの解析となる
- モデルを走らせる際に、必ず"N"の確認を実施する。
- 補正因子のなかでmissing valueの割合の多いものがあつた場合
 - 補正因子から外す
 - Missing valueを一つのvalueとして扱う

→Do file↵

検出力の違いによる結論の差

	HR	p value	
Stem Cell Source			
BM	1.00	67%	29%
PB	1.14	0.007	
ABO matching			
ABO matched or minor	1.00		
ABO major	0.99	0.923	
ABO missing	1.08	0.145	
Disease stage			
Standard	1.00		
Advanced	2.91	<0.0001	
Patient age			
age<40	1.00	100%	100%
Age>40	1.48	<0.0001	
	3,327例		2,255例

時間依存性変数の扱いの解説と演習

TRUMP統計セミナー2014 応用編
2014年11月30日(日) 10時~15時
@名古屋大学大幸キャンパス内

時間依存性変数の 取り扱いについて

自治医科大学附属
さいたま医療センター
諫田淳也

ランドマーク解析

観察開始後のある一定の固定した時点(ランドマーク)において、その時点でのイベント発症の有無でグループ分けして比較する方法。

注: ランドマーク時点以降に、そのイベント発症が得られても、グループの変更は行わない。

- 欠点
- ① ランドマークの設定により結果が左右される
 - ② ランドマーク以後の因子の変化は無視される。(例. 慢性GVHD発症)
 - ③ ランドマーク以前の死亡者も無視される。

前提

生存期間をあるイベントの有無で比較する場合



そのイベントの有無は生存期間の前に決定している必要がある。

観察開始後に決定したイベントの場合、不都合なことがおこる。



慢性GVHDを例に考えてみる。

ランドマーク解析の欠点

ランドマーク時点までにエンドポイントが発生した(例 死亡した)症例は解析対象から除かれる。



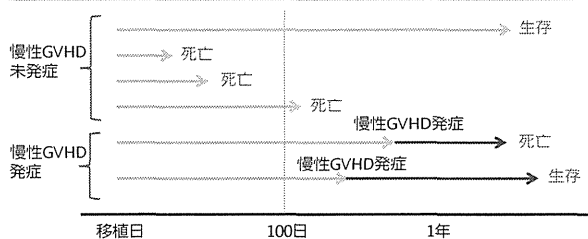
ランドマークまでの期間が長くなればなるほど、解析対象が減少し、生存率の推定が不安定となる。

ランドマーク以後にイベントが発症した症例は、イベント非発症群となる。



ランドマークまでの期間が短くなればなるほど、イベントの発症例が減少し、イベント発症群の生存率の推定が不安定となる。

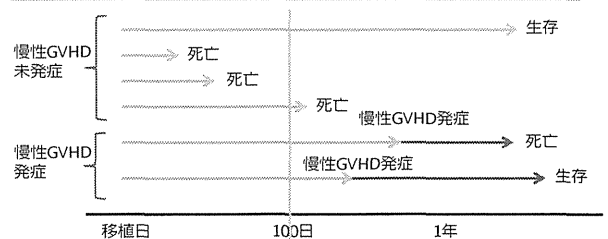
慢性GVHDの有無が生存に与える影響の解析



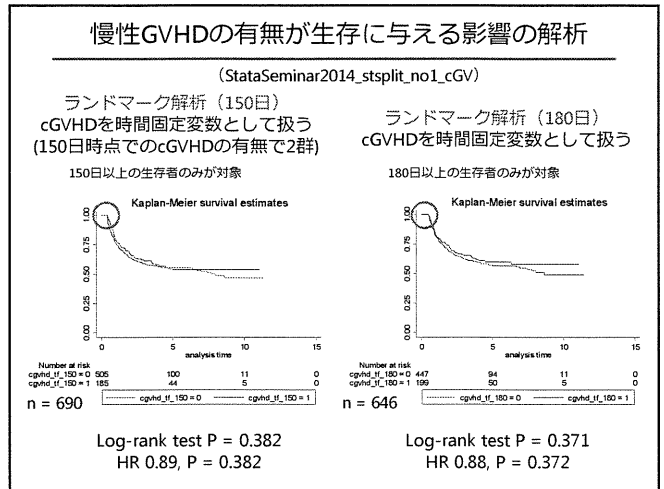
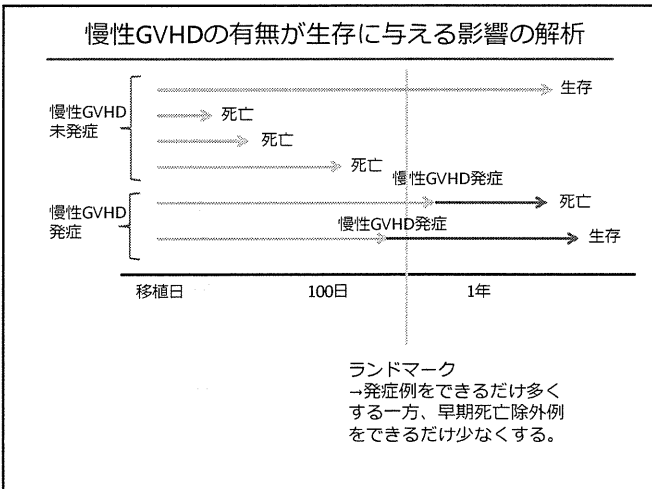
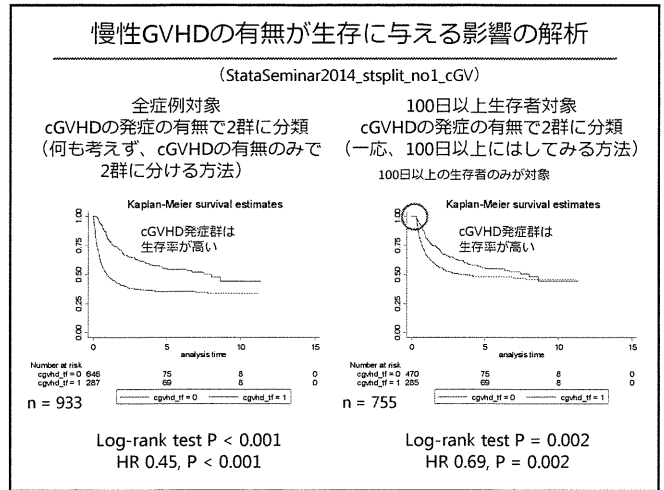
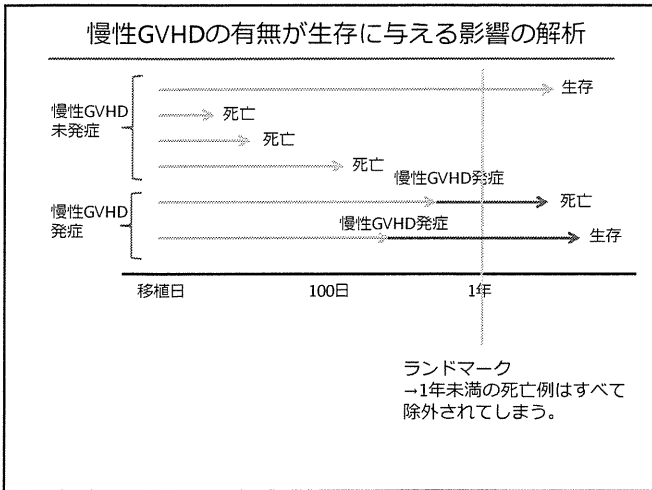
慢性GVHD発症前に死亡した症例は、慢性GVHD未発症群に含まれる。逆に言うと、慢性GVHD発症例は少なくとも100日以上生存している。

→バイアスが生じる

慢性GVHDの有無が生存に与える影響の解析



ランドマーク
→ 全員慢性GVHD未発症例にカウントされる。



慢性GVHDの解析におけるランドマークの決定方法

(StataSeminar2014_stsplit_no1_cGV)

慢性GVHD発症者の発症日中央値、75パーセンタイル、90パーセンタイルは？

centile cgvhd_day, centile(50(5)100)

Variable	Obs	Percentile	Centile	— Binom. Interp. — [95% Conf. Interval]	
cgvhd_day	451	50	104	100	111
		55	112	104	119
		60	120.4	112	132. 4077
		65	133	122	142
		70	143	133	156
		75	158	145	175. 5496
		80	181	163. 3555	197
		85	206. 4	188	226. 7877
		90	266. 6	215	338. 0268

どのポイントが妥当？

答えはないので、感度分析的にランドマーク日を動かしながら結論の正しさを確認するのが良いかもしれない。
180日や1年などがreasonable?

時間依存性変数 (time-dependent covariate) とは

観察開始後、時間と共に変化する独立変数のこと。
比例ハザードモデルでは、この時間依存性変数をモデルに組み込むことができる。

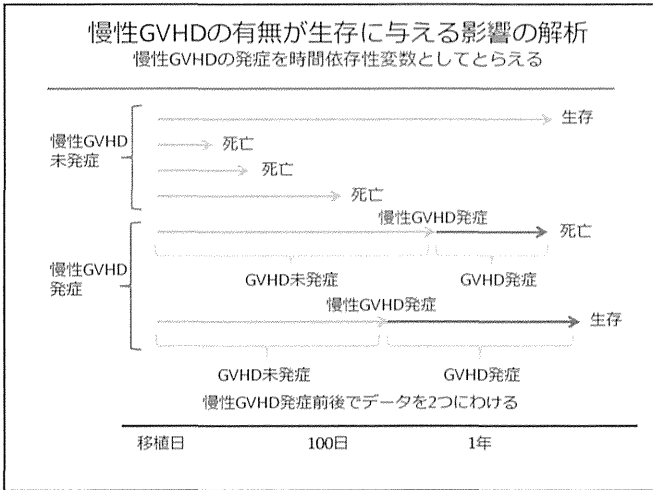
- 今回のような急性GVHD、慢性GVHDの解析
- 診断時からのアウトカムを解析する場合

例えば、診断時を起点とした、移植群と非移植群の比較。
(時間依存性変数を用いないと、移植群に、診断時から移植できるまで長期間生存したという人が含まれるというバイアスが入る)

- 移植後ステロイド投与の再発に及ぼす影響 (Imahashi, et al. ClinTransplant 2010)
- 移植後グリベック投与のcGVHD発症に及ぼす影響 (Nakasono, et al. Leukemia 2010)

観察開始時にわかっている変数は、「時間固定変数」time-fixed covariate

- 診断時年齢
- 診断時白血球数、など



慢性GVHDの有無が生存に与える影響の解析

(StataSeminar2014_stsplit_no1_cGV)

```

stsplit cgvhdsplit1 if event_cgvhdsplit1 == 1, at(0) after(time=ci_cgvhdsplit1_year)
gen cgvhdsplit2 = 0 if event_cgvhdsplit1 >= 2 & event_cgvhdsplit1 <= 4
replace cgvhdsplit2 = 0 if cgvhdsplit1 == -1
replace cgvhdsplit2 = 1 if cgvhdsplit1 == 0
    
```

ここで、発症なしが0、発症ありが1の数値が完成

```

list id event_cgvhdsplit1 _t0 _t1 in 1/20
    
```

id	event_cgvhdsplit1	cgvhdsplit2	_t0	_t1
1	1	Relapse before cGVHD	0	3 6605065
2	2	Relapse before cGVHD	0	0 04106776
3	3	cGVHD	1 31759068	8459959
4	3	cGVHD	0	31759068
5	4	cGVHD	0	15058179
6	4	cGVHD	1 15058179	1 377139
7	5	Censor	0	12867899
8	6	cGVHD	1 09582478	16700889
9	6	cGVHD	0	09582478
10	7	cGVHD	0	27104723
11	7	cGVHD	1 27104723	1 0540725
12	8	Death before cGVHD	0	04654346
13	9	Relapse before cGVHD	0	35318276
14	10	Relapse before cGVHD	0	36139631
15	11	cGVHD	1 27652293	34223136

大切なのは、データがしっかりと作成されたかどうか、確認すること。また、データの構造が変わっているため、データを上書きしてしまわないように注意すること

慢性GVHDの有無が生存に与える影響の解析

(StataSeminar2014_stsplit_no1_cGV)

```

//////*****時間依存性変数*****//////
*移植後120日でcGVHDを発症した場合
*固定時間変数
*cGVHDの値: 1 観察期間0日から最終観察日まで
*時間依存性変数
*cGVHDの値: 0→1 (120日) →1 (最終観察日まで)

/*時間依存性変数としての扱い*/
*生存に対する慢性GVHDの影響を見る場合
sort id
*id()で同一患者を示す
stset lyear, failure(event_os==1) id(id)

stsplit cgvhdsplit1 if event_cgvhdsplit1 == 1, at(0) after(time=ci_cgvhdsplit1_year)
gen cgvhdsplit2 = 0 if event_cgvhdsplit1 >= 2 & event_cgvhdsplit1 <= 4
replace cgvhdsplit2 = 0 if cgvhdsplit1 == -1
replace cgvhdsplit2 = 1 if cgvhdsplit1 == 0
    
```

慢性GVHDの有無が生存に与える影響の解析

(StataSeminar2014_stsplit_no1_cGV)

100日以上生存者対象
cGVHDを時間依存変数として扱う

100日以上生存者対象
cGVHDの発症の有無で2群に分類

```

st graph if lday >= 100 & lday < ., by(cgvhdsplit2) risktable
    
```

HR 0.86, P = 0.210

Log-rank test P = 0.002
HR 0.69, P = 0.002

慢性GVHDの有無が生存に与える影響の解析

(StataSeminar2014_stsplit_no1_cGV)

```

stset lyear, failure(event_os==1) id(id)
stsplit cgvhdsplit1 if event_cgvhdsplit1 == 1, at(0) after(time=ci_cgvhdsplit1_year)
list id event_cgvhdsplit1 _t0 _t1 in 1/20
    
```

id	event_cgvhdsplit1	cgvhdsplit1	_t0	_t1
1	1	Relapse before cGVHD	0	3 6605065
2	2	Relapse before cGVHD	0	0 04106776
3	3	cGVHD	1 31759068	8459959
4	3	cGVHD	0	31759068
5	4	cGVHD	0	15058179
6	4	cGVHD	-1	0 15058179
7	5	Censor	0	12867899
8	6	cGVHD	0	09582478
9	6	cGVHD	-1	0 09582478
10	7	cGVHD	-1	0 27104723
11	7	cGVHD	0	27104723
12	8	Death before cGVHD	0	04654346
13	9	Relapse before cGVHD	0	35318276
14	10	Relapse before cGVHD	0	36139631
15	11	cGVHD	0	27652293
16	11	cGVHD	-1	0 27652293
17	12	Death before cGVHD	0	12594114
18	13	Relapse before cGVHD	0	9 6372347
19	14	Censor	0	6 9678302
20	15	cGVHD	0	1 3716632

慢性GVHD発症年に関しては発症年(ci_cgvhdsplit1_year)で情報が2つに分けられている。cgvhdsplit1がその変数。-1と0で分類される。_t0がGVHD発症年。このままだとcgvhdsplit1の変数は使えないので、もう少し調整 (次スライド)

慢性GVHDの有無が生存に与える影響の解析

(StataSeminar2014_stsplit_no1_cGV)

100日以上生存者対象
cGVHDを時間依存変数として扱う

ランダム解析 (180日)
cGVHDを時間固定変数として扱う

```

st graph if lday >= 100 & lday < ., by(cgvhdsplit100) risktable
    
```

HR 0.86, P = 0.210

Log-rank test P = 0.371
HR 0.88, P = 0.372

慢性GVHDの有無が生存に与える影響の解析

(StataSeminar2014_stsplit_no1_cGV)

Cox解析

単変量解析

stcox cgvhdsplit2 if lday >= 100 & lday < .

*時間固定変数を用いて、多変量解析を行い、cGVHDのoutcomeに対する影響を補正することも可能

*例えば患者年齢 pt_age, 患者性別 pt_sex, 診断 diagnosis, 病期 stage3 (CR1/CR2 vs. others)により補正する場合、

多変量解析

stcox cgvhdsplit2 pt_age pt_sex diagnosis stage3 if lday >= 100 & lday < .

急性GVHDの有無が生存に与える影響の解析

(StataSeminar2014_stsplit_no2_aGV)

Grade 2-4 aGVHDが生存に与える影響

1. ハザード比の算出(G2-4 vs. G0-1 aGVHD)

(単変量解析・多変量解析←年齢・性別・疾患・ステージで補正)

2. 図示(KM curve)

方法は1と2の方法を用いる

1. ランドマーク法(移植後60日をランドマーク日とする。)を用いる。

2. 時間依存性共変量を用いる。

慢性GVHDの有無が生存に与える影響の解析
図示の方法

- ① ランドマーク法
 - ② セミランドマーク法?
 - ③ 時間依存性変数を用いる
(stsplitを用いて時間依存性変数前後でデータを分けた後にsts graphを用いて図を書くのみ)
- ①で図示されることが最も多い
②、③ともに使用される頻度は低い。
(②はまれ)

急性GVHDの有無が生存に与える影響の解析

(StataSeminar2014_stsplit_no2_aGV)

```
/*Landmark解析 day60*/
gen agvhd_tf_60 = 1 if event_agvhd24==1 & agvhd_day<=60
replace agvhd_tf_60 = 0 if event_agvhd24==1 & agvhd_day>60
replace agvhd_tf_60 = 0 if event_agvhd24>=2 & event_agvhd24<=4
tab agvhd_tf_60 if lday>=60 & lday< .
```

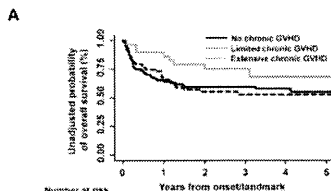
```
/*時間依存性変数としての扱い*/
sort id
stset lday, failure(event_os==1) id(id)
```

```
stsplit agvhd24_split1 if event_agvhd24==1, at(0) after(time=ci_agvhd24)
gen agvhd24_split2 = 0 if event_agvhd24>=2 & event_agvhd24<=4
replace agvhd24_split2 = 0 if agvhd24_split1==1
replace agvhd24_split2 = 1 if agvhd24_split1==0
```

セミランドマーク法

cGVHD発症者: cGVHD発症日をランドマーク日

cGVHD未発症者: cGVHD発症者におけるcGVHD発症日の中央値をランドマーク日とする。



Baron F, et al. J Clin Oncol. 2005;23:1993-2003.
Kanda J, et al. Blood. 2012;119:2141-8.
Ishida T, et al. BBMT. 2013;19:1731-1739.

ハザード比の算出(G2-4 vs. G0-1 aGVHD)

Landmark解析 (StataSeminar2014_stsplit_no2_aGV)

stcox agvhd_tf_60 if lday >= 60 & lday < .

_t	Haz. Ratio	Std. Err.	z	P> z	[95% Conf. Interval]
agvhd_tf_60	1.188112	.1196894	1.71	0.087	.9752323 1.447461

stcox agvhd_tf_60 pt_age pt_sex diagnosis stage3 if lday >= 60 & lday < .

_t	Haz. Ratio	Std. Err.	z	P> z	[95% Conf. Interval]
agvhd_tf_60	1.064986	.1101031	0.61	0.543	.8696466 1.304202
pt_age	1.019576	.0030412	3.50	0.000	1.004633 1.015554
pt_sex	1.166921	.1184836	1.52	0.128	.9563447 1.423864
diagnosis	1.09167	.1279445	0.75	0.454	.867821 1.373577
stage3	3.142105	.3365135	10.69	0.000	2.547174 3.875991

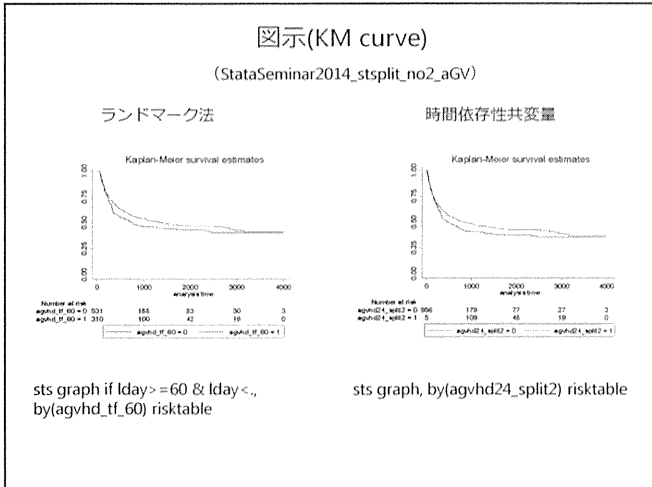
時間依存性変数

stcox agvhd24_split2

_t	Haz. Ratio	Std. Err.	z	P> z	[95% Conf. Interval]
agvhd24_split2	1.225535	.110784	2.25	0.024	1.02655 1.463091

stcox agvhd24_split2 pt_age pt_sex diagnosis stage3

_t	Haz. Ratio	Std. Err.	z	P> z	[95% Conf. Interval]
agvhd24_split2	1.088805	.102175	1.01	0.312	.9155358 1.31828
pt_age	1.013578	.0027209	5.02	0.000	1.008259 1.018925
pt_sex	1.130503	.1016011	1.36	0.172	.9479208 1.348252
diagnosis	1.102185	.1156202	0.93	0.354	.8973522 1.353775
stage3	3.057825	.2933677	11.65	0.000	2.53366 3.690429



急性GVHDと慢性GVHDを同時にモデルに入れる

今まで、慢性GVHDおよび急性GVHDに関して、それぞれ、ランドマーク解析あるいは時間依存性変数を用いて、解析を行った。

では、急性GVHD/慢性GVHDを同時に解析するにはどうすれば良いだろうか。

- ① 急性GVHD、慢性GVHDをそれぞれ時間依存性変数として取り扱う。*対象者は、全症例? 100日以上(無病)生存者?
- ② 100日以上(無病)生存者に限定した上で、急性GVHDは時間固定変数として取り扱い、慢性GVHDのみを時間依存性変数として取り扱う。

→②の場合、急性GVHDが与える早期(100日以内)の影響を無視することになる。急性GVHDの影響は全症例で行うべき?

*急性GVHD発症は慢性GVHD発症の最も強力な予測因子であるが、これら2つを同時にモデルに含めることは妥当?

GVHDの各グレードが生存に与える影響の解析

では、Grade 1 aGVHD、Grade 2 aGVHD、Grade 3-4 aGVHDが全生存率に与える影響に関してはどのように解析すべきだろうか。

Grade 1 aGVHD、Grade 2 aGVHD、Grade 3-4 aGVHDそれぞれの発症日がわかっているならば、Grade 0→Grade 1→Grade 2→Grade 3-4と移行させればよい。

ただ、実際は、Grade 1→Grade 2→Grade 3-4への移行期間は非常に短いので、Grade 0→Grade 1
Grade 0→Grade 2
Grade 0→Grade 3-4
として扱っても (agvhd24_split2をG1=1, G2=2, G3-4=3とする) 結果に大きな差はないであろう。

しかし、Limited cGVHDとextensive cGVHDでは影響があるかもしれない。発症日のデータがそろえば、none→limited
cGVHD→extensive cGVHDとわけて再解析することが必要?

他の論文ではどうなっているだろうか。

Graft-versus-Host Disease Induced Graft-versus-Leukemia Effect: Greater Impact on Relapse and Disease-Free Survival after Reduced Intensity Conditioning

Daniel Weisdorf,^{1,2} Manje Zhang,^{2,3} Mukta Anas,^{1,4} Mary M. Horewitz,^{2,3} Douglas Rizzo,^{1,2} Mary Eapen^{2,1}
CIBMTRの研究

The first analysis included all patients (N=5791), and aGVHD and cGVHD were treated as time-dependent covariates.

The second analysis was a landmark analysis that included patients who survived at least 1-year disease-free after their transplantations (N=2369).

For the landmark analysis of 1-year disease-free survivors, aGVHD and/or cGVHD status was evaluated at 1 year by which time all aGVHD and 95% of cGVHD had developed.

Biol Blood Marrow Transplant 18:1727-1733, 2012

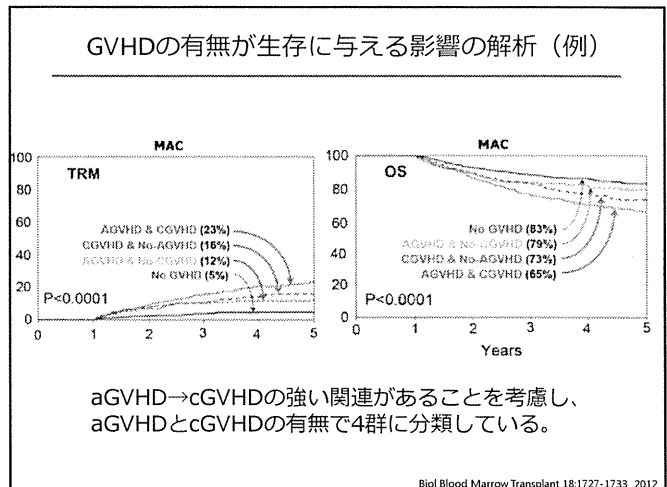
GVHDの有無が再発に与える影響

GVHDの有無が再発に与える影響を中心に解析するためには、どういったpopulationを選択するか。

非寛解期の移植で、移植後寛解ならず、の症例をどのように扱うか。GVHDの有無が再発に与える影響を解析するには、寛解期移植に限定するのが本来は望ましい。

スクリプトでは、移植後寛解ならずは、0.1日での再発と定義されている。GVHD未発症でかなりの症例が再発してしまう。これが影響を及ぼすかどうか。早期再発例を除外すべき?

また、ATLのように、再発が必ずしも再発による死亡につながる場合は(再発後もGVHD発症→GVL効果で再度寛解に至る可能性がある場合)、再発関連死亡/治療関連死亡をエンドポイントとして解析するののも一つの方法。



GVHDの有無が生存に与える影響の解析 (例)

Table 2. Influence of GVHD on Outcome after HCT: Time-Dependent Multivariate Analysis of All Patients

Survival	MAC				
No GVHD		1.00			<.0001*
aGVHD only		1.71 (1.55-1.89)			<.0001
cGVHD only		0.99 (0.86-1.15)			.90
aGVHD + cGVHD		1.37 (1.21-1.57)			<.0001

The analysis with all patients allowed examination of any early effect of GVHD within 1 year.

Table 3. Influence of GVHD on Outcome after HCT: Landmark Multivariate Analysis among 1-year Disease-Free Survivors

Survival	MAC				
Disease-free survivors at 1 year ^b (N = 1739)					<.0001†
No GVHD	511	1.00			
aGVHD only	198	1.24 (0.95-1.62)	.27		
cGVHD only	524	1.74 (1.31-2.30)	.0001		
aGVHD + cGVHD	506	2.20 (1.68-2.87)	<.0001		

The landmark analysis allowed us to examine any late effect of GVHD on those who were disease-free for at least 1 year.

Biol Blood Marrow Transplant 18:1727-1733, 2012

GVHDの有無が生存に与える影響の解析

(StataSeminar2014_stsplit_no3_aGvcGV)

全症例を対象として解析
stcox agvhd24_split2 cgvhd_split2

_t	Haz. Ratio	Std. Err.	z	P> z	[95% Conf. Interval]
agvhd24_split2	1.266234	.1198288	2.49	0.013	1.051868 1.524286
cgvhd_split2	.7645812	.0891293	-2.30	0.021	.6084105 .9608387

100日生存症例を対象として解析
stcox agvhd24_split2 cgvhd_split2 if lday>100

_t	Haz. Ratio	Std. Err.	z	P> z	[95% Conf. Interval]
agvhd24_split2	1.224218	.1393952	1.78	0.076	.9793473 1.530314
cgvhd_split2	.8203551	.0986042	-1.65	0.099	.6481717 1.038278

どちらも有意ではなくなりました。症例数の問題？

急性GVHD、慢性GVHDをそれぞれ時間依存性変数として取り扱うには。

(StataSeminar2014_stsplit_no3_aGvcGV)

```
gen ci_agvhd24_year = ci_agvhd24/365.25
(まずはaGVHD/cGVHD発症日ともにyearにあわせる。)
```

```
stsplit agvhd24_split1 if event_agvhd24==1, at(0) after(time=ci_agvhd24_year)
gen agvhd24_split2 = 0 if event_agvhd24>=2 & event_agvhd24<=4
replace agvhd24_split2 = 0 if agvhd24_split1==1
replace agvhd24_split2 = 1 if agvhd24_split1==0
```

(cGVHDに関しても、単純にstsplitをもう一度行えばよい)

```
stsplit cgvhd_split1 if event_cgvhd==1, at(0) after(time=ci_cgvhd_year)
gen cgvhd_split2 = 0 if event_cgvhd>=2 & event_cgvhd<=4
replace cgvhd_split2 = 0 if cgvhd_split1==1
replace cgvhd_split2 = 1 if cgvhd_split1==0
```

GVHDの有無が生存に与える影響の解析

(StataSeminar2014_stsplit_no3_aGvcGV)

全症例を対象として解析
stcox agvhd24_split2 cgvhd_split2
急性GVHDがOSに及ぼす影響はconsistentであったが慢性GVHDに関しては、非常に不安定

_t	Haz. Ratio	Std. Err.	z	P> z	[95% Conf. Interval]
agvhd24_split2	1.266234	.1198288	2.49	0.013	1.051868 1.524286
cgvhd_split2	.7645812	.0891293	-2.30	0.021	.6084105 .9608387

100日生存症例を対象として解析
stcox agvhd24_split2 cgvhd_split2 if lday>100

_t	Haz. Ratio	Std. Err.	z	P> z	[95% Conf. Interval]
agvhd24_split2	1.224218	.1393952	1.78	0.076	.9793473 1.530314
cgvhd_split2	.8203551	.0986042	-1.65	0.099	.6481717 1.038278

100日無病生存症例を対象として解析 (DFSやrelapseも解析することを予定して)
stcox agvhd24_split2 cgvhd_split2 if rfs_day>100 cGVHDのHRはほぼ1.00となった

_t	Haz. Ratio	Std. Err.	z	P> z	[95% Conf. Interval]
agvhd24_split2	1.27181	.1614126	1.89	0.058	.9917255 1.630996
cgvhd_split2	1.00662	.1309259	0.05	0.960	.7801072 1.298904

GVHDの有無が生存に与える影響の解析

(StataSeminar2014_stsplit_no3_aGvcGV)

全症例を対象として解析
stcox agvhd24_split2 cgvhd_split2

_t	Haz. Ratio	Std. Err.	z	P> z	[95% Conf. Interval]
agvhd24_split2	1.266234	.1198288	2.49	0.013	1.051868 1.524286
cgvhd_split2	.7645812	.0891293	-2.30	0.021	.6084105 .9608387

2-4 aGVHDの発症は致死率の増加と有意な関連がある一方、cGVHDの発症は、致死率の低下と関連がある。

cGVHD発症はGVL効果と関連する？

GVHDの有無が治療関連死亡に与える影響の解析
競合因子を考慮する方法

(StataSeminar2014_stsplit_TRMComp_aGvcGV)

```
sort id
stset ci_rel_year, failure(event_relapse==2) id(id)
```

```
gen ci_agvhd24_year = ci_agvhd24/365.25
```

```
stsplit agvhd24_split1 if event_agvhd24==1, at(0) after(time=ci_agvhd24_year)
gen agvhd24_split2 = 0 if event_agvhd24>=2 & event_agvhd24<=4
replace agvhd24_split2 = 0 if agvhd24_split1==1
replace agvhd24_split2 = 1 if agvhd24_split1==0
```

```
stsplit cgvhd_split1 if event_cgvhd==1, at(0) after(time=ci_cgvhd_year)
gen cgvhd_split2 = 0 if event_cgvhd>=2 & event_cgvhd<=4
replace cgvhd_split2 = 0 if cgvhd_split1==1
replace cgvhd_split2 = 1 if cgvhd_split1==0
```

```
stcrreg agvhd24_split2 cgvhd_split2 if rfs_day>100, compete(event_relapse==1)
```

通常通り、stcrregを用いて、競合リスクを指定すれば良い。ただ、競合リスクのイベント発症後は、competing risk regression modelではそのsubjectは“at risk”として扱うが、時間依存性変数は、競合因子のイベント発症時の値に固定されるため、時間依存性変数が複数の値をとるような時はバイアスがかかる可能性がある。