

### <日付データ>

CSV ファイルから日付データを取り込むと、Stata でのデータ形式が string(カレンダー形式)となっています。以下のコマンドにて日付データに変換する必要があります。

#### ②移植日などを日付データ(YYYY/MM/DD)に変換します

※Excel ファイルからデータを取り込んだ場合は、Stata の日付データへ自動変換されるため、このコマンドの実行は不要です。

```
gen 新変数名 = date(変数,"YMD")
```

#### ③移植年など西暦(YYYY)のみのデータが必要な場合

```
gen 新変数名 = year(変数)
```

### <新しい変数の作成/値の変更>

変数作成のコマンド

```
gen 新変数=値
```

変数の値の変更コマンド

```
replace 既存の変数=値
```

例えば、上記で作成した移植年の変数をグループ分けする場合

#### ④移植年グループの変数を作成

```
gen tyear_group = 1 if tyear<=1990
```

(tyeargroup という新変数を作成し、移植年が 1990 年以前のデータを 1 とする。)

この時点では、移植年が 1990 年以前のデータ以外は欠損値となっているため、以下の replace コマンドにて変換していきます。

#### ⑤変数の値の変更

```
replace tyear_group = 2 if tyear>=1991 & <=1995
```

同様に、変数の値を 3,4,5 に変換していきます。

### <ラベルの作成>

#### ⑥作成した変数にラベルをつけます

```
label variable 変数 "ラベルの内容"
```

#### ⑦値と対応するラベルセットを設定します

```
label define セット名 値 1 "ラベル 1" 値 2 "ラベル 2" 値 3 "ラベル 3"
```

#### ⑧ラベルセットに変数を割り付けます

```
label value 変数 セット名
```

### <データの要約>

#### ⑨⑩カテゴリ変数の要約

```
tab 変数
```

移植年グループ	Freq.	Percent	Cum.
-1990	1,581	2.92	2.92
1991-1995	6,042	11.17	14.10
1996-2000	11,854	21.92	36.02
2001-2005	17,467	32.30	68.32
2006-	17,128	31.68	100.00
Total	54,072	100.00	

```
tab 変数,nolabel
```

←ラベルなしで表示されます。

移植年グループ	Freq.	Percent	Cum.
1	1,581	2.92	2.92
2	6,042	11.17	14.10
3	11,854	21.92	36.02
4	17,467	32.30	68.32
5	17,128	31.68	100.00
Total	54,072	100.00	

▶12\_gv\_outcome\_XXXXXX.do

生存解析のイベント、競合リスクのイベント(再発, 好中球回復, 血小板回復, 急性 GVHD, 慢性 GVHD)の変数の作成するプログラムです。

```
File Edit Tools View
12_gv_outcome_111205.do (Untitled.do)
1
2 *****
3 **Survival Summary**
4 **Survival Event**
5 tab survival
6 zen event_os = 1 if survival=="死亡"
7 replace event_os = 0 if survival=="生存"
8
9 *****
10 gen ldate = date(last_date, "YMD")
11 gen lday = ldate - tdate
12 gen lyear = lday / 365.25
13
14 *生存解析可能症例
15 gen sa = 1
16 replace sa = 0 if survival=="
17 replace sa = 0 if lyear==.
18 replace sa = 0 if lyear<0
19
20 count
21 tab sa, missing
22 *最終確認日
23 gen lday, detail
24
25 *****
26 **Relapse**
27 *****
28 **Relapse**
29 *****
30 **Event Relapse**
31 **Clinical/Hematological relapseを再発とする
32 **再発データの確認はエビデンスによる
33 gen relapse_manip = 1 if relapse=="寛
34 **寛解ならず移植後1日で再発したと見
35 replace relapse_manip = 1 if relapse=="寛解ならず"
36 replace relapse_manip = 1 if relapse=="寛解ならず" & dstatus=="移植後も寛解ならず"
37 replace relapse_manip = 0 if relapse=="寛
38 **Relapseを寛解の割合(移植後の病状) dstatusを参照する
39 replace relapse_manip = 0 if relapse=="寛解維持" & dstatus=="寛解維持"
40 replace relapse_manip = 1 if relapse=="寛解後再発" & dstatus=="移植後再発"
41 **再発が寛解かつ clinical/hematological relapse無しかつ寛解したもののみを再発
42 **cytogenetic/molecular relapse有り hematological relapse無しものは再発なし
43 **しかし 移植後治療を受けているものは clinical relapse有りとする
44 replace relapse_manip = 0 if relapse=="寛" & rel_clinemat_reldx=="無"
45 replace relapse_manip = 1 if rel_cyto_reldx=="寛" & rel_clinemat_reldx=="無" & addtherapy=="寛"
46 replace relapse_manip = 1 if rel_mol_reldx=="寛" & rel_clinemat_reldx=="無" & addtherapy=="寛"
47
48 gen event_relapse = 1 if relapse_manip==1
49 replace event_relapse = 2 if event_os==1 & relapse_manip==0
50 replace event_relapse = 3 if event_os==0 & relapse_manip==0 & event_relapse==.
51
52 label define event_relapse_label 1 "再発" 2 "寛解時再発" 3 "寛解"
53 label values event_relapse event_relapse_label
54
55 *****
56 **Relapse Day and Year**
57 gen rel_date = date(rel_clinemat_examdate, "YMD")
58 gen rel_day = rel_date - tdate
59 gen rel_date_old = date(old_relapse_date, "YMD")
60 /*注意!!旧レコードから移植後、旧レコードが異なることあり。用・日・年の場合は上記"日"を"年"に変更
61 gen rel_day_old = rel_date_old - tdate
62 count if event_relapse==1 & rel_day==. & rel_day_old!=.
63 replace rel_day = rel_day_old if event_relapse==1 & rel_day==. & rel_day_old!=.
64 replace rel_day = 1 if relapse=="寛解ならず"
65 replace rel_day = 1 if relapse=="寛解ならず" & dstatus=="移植後も寛解ならず"
66 replace rel_day = 1 if rel_cyto_reldx=="寛" & rel_mol_reldx=="無" & rel_clinemat_reldx=="無" & dstatus=="移植後も寛解ならず"
67 replace rel_day = 1 if rel_cyto_reldx=="無" & rel_mol_reldx=="無" & rel_clinemat_reldx=="無" & dstatus=="移植後も寛解ならず"
68 replace rel_day = 1 if rel_cyto_reldx=="寛" & rel_mol_reldx=="無" & rel_clinemat_reldx=="無" & dstatus=="移植後も寛解ならず"
69 replace rel_day = 1 if rel_cyto_reldx=="無" & rel_mol_reldx=="寛" & rel_clinemat_reldx=="無" & dstatus=="移植後も寛解ならず"
70 /*cytogenetic/molecular relapse有り hematological relapse無しかつ 再発後治療を受けている場合は再発日
71 gen rel_cyto_date = date(rel_cyto_examdate, "YMD")
72 gen rel_cyto_day = rel_cyto_date - tdate
73 gen rel_mol_date = date(rel_mol_examdate, "YMD")
74 gen rel_mol_day = rel_mol_date - tdate
75 replace rel_day = rel_cyto_day if rel_cyto_reldx=="寛" & rel_clinemat_reldx=="無" & addtherapy=="寛" & rel_cyto_day!=.
76 replace rel_day = rel_mol_day if rel_mol_reldx=="寛" & rel_clinemat_reldx=="無" & addtherapy=="寛" & rel_mol_day!=.
77
78 *****
79 **relapse解析時間変数 ci_relday, ci_relyear**
80 gen ci_relday = rel_day if event_relapse==1
81 replace ci_relday = lday if event_relapse==2 | event_relapse==3
82 count if ci_relday==. & event_relapse==1
83 count if lday==. & event_relapse==2 | lday==. & event_relapse==3
84 /*再発日に未入力が多いので注意を要する!!!
85
86 **再発日付を入力がエラーになるinputation 調査者による検討が必要
87 **再発日付が、急性GVHD発症から再発日から最終確認日の中間の生存例(死亡例でそれぞれ補充方法
88 /*この中間のコマンドを実行する場合は、"1"か"1"を"0"で取って下さい)
89 gen rel_to_last = rel_date
90 quietly sum rel_to_last if event_os==1 & diagnosis==2, detail
91 replace ci_relday = rel_to_last if event_os==1 & rel_date==. & event_os==1
92 quietly sum rel_to_last if event_os==0 & diagnosis==2, detail
93 replace ci_relday = rel_to_last if event_os==0 & rel_date==. & event_os==0
94 replace d_to_sci = if d_to_sci=0
95
96 **再発日付の再発日付の調査に最終確認日で補う(これが適切かどうかは調査者の判断が必要)
97 /*この中間のコマンドを実行する場合は、"1"か"1"を"0"で取って下さい)
98 replace ci_relday = lday if event_relapse==1 & rel_day==.
99
100 *****
101 **以下Competing risk event作成後*****
102 **Relapse-free Survival**
103 **Event rfs**
104 gen event_rfs = 1 if event_os==1
105 replace event_rfs = 1 if event_relapse==1
106 replace event_rfs = 0 if event_rfs==. & event_os==0
107
108 *****
109 **rfs day and rfs year**
110 gen rfs_day = lday
111 replace rfs_day = rel_day if event_relapse==1
112
113 gen rfs_year = rfs_day / 365.25
114
```

<生存解析に用いる変数>  
生存状況(生存/死亡)  
最終確認時までの期間  
relapse-free survival  
relapse-free survival までの期間

<競合リスクイベント解析に用いる変数>  
再発イベント  
再発日  
再発までの期間  
time variable for event  
relapse(再発あり=再発までの期間、寛解ならず、再発なし=生存期間)  
好中球 500 以上到達  
好中球 500 以上到達までの期間

time variable for event  
neut  
血小板 2 万以上到達  
血小板 2 万以上到達までの期間

time variable for event  
plt2  
血小板 5 万以上到達  
血小板 5 万以上到達までの期間

time variable for event  
plt5  
急性 GVHD 発症  
Grade2-4 の急性 GVHD

Grade2-4 の急性 GVHD 発症までの期間  
time variable for event  
agvhd24

Grade3-4 の急性 GVHD  
Grade3-4 の急性 GVHD 発症までの期間  
time variable for event  
agvhd34

慢性 GVHD 発症  
慢性 GVHD 発症までの期間  
time variable for event  
cgvhd

relapse-free survival  
relapse-free survival までの期間

## ▶21\_characteristics\_XXXXXX.do

データを要約する際に便利な検定コマンドの解説も入っています。

```

25  ****Characteristics****
26
27  /**データの表示：表形式**/
28  *tab 変数1 変数2,missing
29  *missingを入れると欠損データをカウントすることができます。
30
31  /**データの表示：中央値、パーセンタイル値など**/
32  *sum 変数,detail
33  *総数、平均値、標準偏差、中央値、パーセンタイル値、Min、Max、分散、歪度、尖度を算出することができます。
34
35  /**検定：Mann-Whitney検定**/
36  *ranksum 変数,by(group)
37
38  /**検定：chi-squared検定**/
39  *tab 変数1 変数2,chi2
40
41  /**検定：Fisherの正確検定**/
42  *tab 変数1 変数2,exact
43
44  /**検定：t検定**/
45  *ttest 変数,by(group)
46
47  /**検定：1群のt検定**/
48  *ttset 変数=数値
49
50  /**検定：paired t-test**/
51  *ttset 変数1=変数2
52
53  /**検定：two-sample t-test**/
54  *ttset 変数1=変数2,unpair
55
56
57  *年齢
58  tab pt_age_group2
59
60  *性別、年齢グループ
61  tab sex pt_age_group2,missing
62  ranksum pt_age_group2,by (sex)
63
64  *性別マッチ、年齢グループ
65  tab sex_mismatch1 pt_age_group2,missing
66  ranksum pt_age_group2,by (sex_mismatch1)
67
68  tab sex_mismatch2 pt_age_group2,missing
69
70  tab diagnosis pt_age_group2,missing
71
72
73
74  tab diagnosis pt_age_group2,missing
75  tab stage_aml pt_age_group2,missing
76  tab abo_match pt_age_group2,missing
77

```

### <検定>

- カテゴリ変数の要約値
- 連続変数の要約値
- ノンパラメトリック検定
- χ<sup>2</sup>乗検定
- Fisherの正確検定
- t検定

▶31\_univariate\_OS\_XXXXXX.do  
単変量解析の解説です。

<pre> 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80                 </pre>	<p>①</p> <p>②</p> <p>③</p> <p>④</p> <p>⑤-1</p> <p>⑤-2</p>	<p>&lt;解説内容&gt; Kaplan-meierの生存関数 生存曲線の書き方</p> <p>Long-rank 検定 Wilcoxon 検定</p> <p>群分けでの解析方法 様々な設定でのサブグループでの解析結果の抽出</p>
--	---	--

①はデータが生存時間データであることをStataに伝えるコマンド書式です。  
コマンド failure は、エンドポイントを示します。  
ここで例えるコマンド書式の場合では、stset lyear, failure(event\_os==1)  
値 1 はエンドポイントで、その他の値は打ち切りを示すと設定されます。

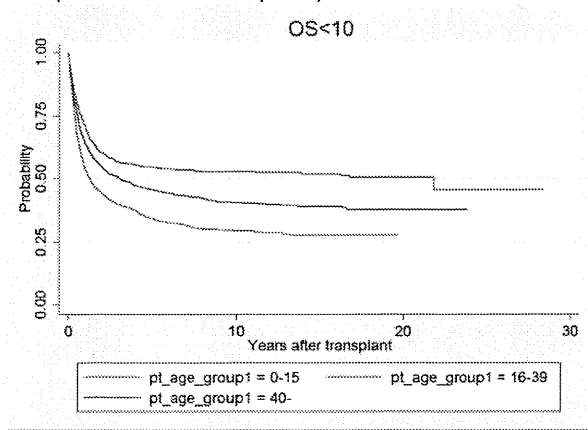
②Kaplan-Meier 曲線を描きます。

**sts graph if (条件), by(群分けグループ変数) title(グラフのタイトル) ytitle(y 軸のタイトル) xtitle(x 軸のタイトル) オプション**

様々な設定でのサブグループで抽出する際には、by( )グループ変数名を入力します。

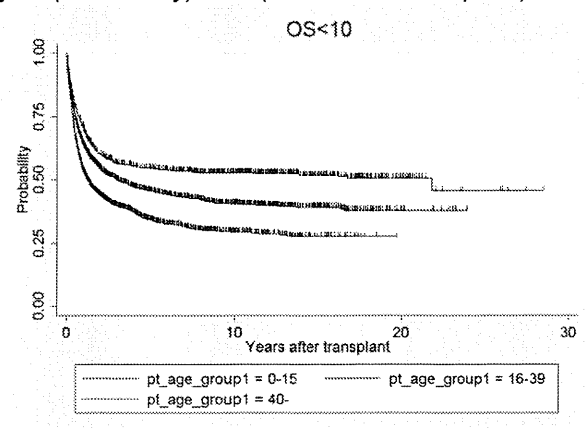
例：AML、血縁移植で年齢別の解析

`sts graph if diagnosis==1 & t_cd>=2 & t_cd<=3, by(pt_age_group1) title(OS<10) ytitle(Probability) xtitle(Years after transplant)`



打ち切りイベントが発生した時点に縦線をつける場合には、オプション箇所に `censored(single)` を加えます。

`sts graph if diagnosis==1 & t_cd>=2 & t_cd<=3, by(pt_age_group1) censored(single) title(OS<10) ytitle(Probability) xtitle(Years after transplant)`



③グラフの保存

**translate @Graph (グラフ名).emf, trans(Graph2emf)**

graph 名にはスペース、<>等は使用できません  
emf の形式にてデータ(.dta ファイル)保存先のフォルダに保存されます

④生存解析結果を表示

個々のエンドポイントまたは打ち切りポイントにおける survival function, standard error, 95%CI を表示することができます。

**sts list,by(グループ変数) at (時間時点リスト)**

例：`sts list if diagnosis==1 & t_cd>=2 & t_cd<=3, by(pt_age_group1) at(1 3 5)`

この例の場合は、エンドポイントまたは打ち切り観察時間 1 年、3 年、5 年時点での生存関数の結果を示します。

コマンドを実行すると下記の結果が表示されます。

```
failure_d: event_os == 1
analysis time_t: lyear
```

Time	Beg. Total	Fail	Survivor Function	Std. Error	[95% Conf. Int.]	
-----						
0-15						
1	670	274	0.7118	0.0147	0.6819	0.7395
3	485	134	0.5644	0.0163	0.5319	0.5956
5	429	14	0.5475	0.0164	0.5147	0.5790
16-39						
1	1280	736	0.6432	0.0106	0.6221	0.6635
3	837	264	0.5017	0.0113	0.4794	0.5236
5	636	71	0.4565	0.0115	0.4338	0.4788
40-						
1	1055	948	0.5401	0.0110	0.5183	0.5614
3	591	254	0.3960	0.0112	0.3739	0.4179
5	378	70	0.3420	0.0114	0.3197	0.3644
-----						

左から 1 列目の time は、エンドポイントもしくは打ち切りを観察した時間を示します。

2 列目の Beg.total は、ある時点の開始時の観察者の人数を示します。

3 列目の Fail は、この期間に観察された failure の数を示します。

4 列目は Survival Function(生存率)を示します。

5 列目は Standard Error を示します。

6-7 列目は 95%CI を示します。

at(時間時点リスト)のオプションをつけずに実行すると、打ち切りまたはイベントが生じたエンドポイント(この例の場合は死亡)のすべての時点での生存率を計算します。

## ⑤生存解析における検定

### 1.Log-rank 検定

```
sts test グループ変数,オプション
```

オプションで設定変更を行わない限り Log-rank 検定を行います。  
Wilcoxon 検定を行う場合には、オプションで指定してください。

### 2.Wilcoxon 検定

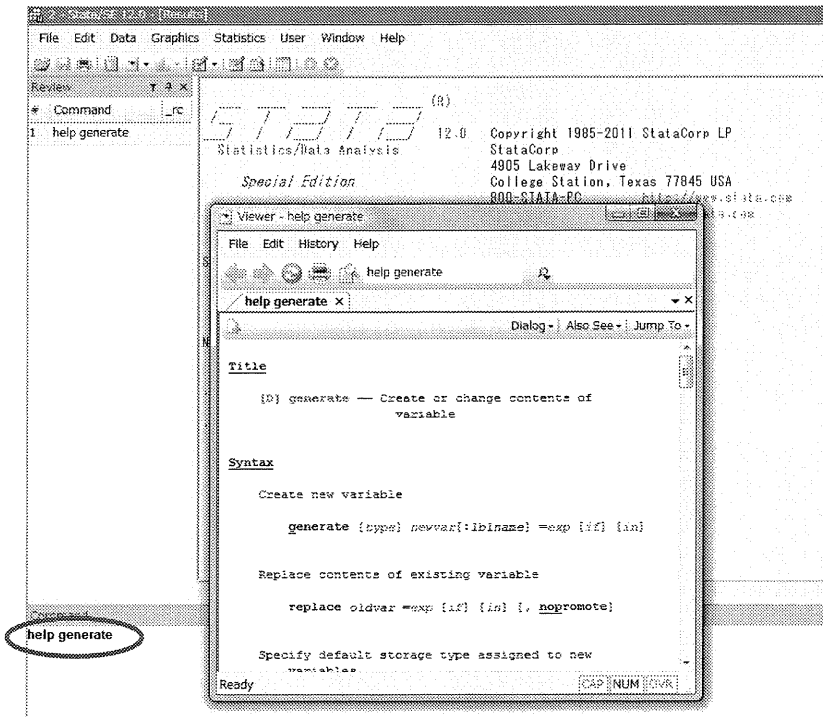
```
sts test グループ変数,wilcoxon
```

# ヘルプの使い方

## 1. コマンドについて調べる

① コマンドウインドウに調べたいコマンドを入力する。

**help** 探すコマンド  
**search** 探すコマンド または キーワード



マニュアル作成

血液疾患臨床研究サポートセンター  
 名古屋大学医学部造血細胞情報管理・生物統計学

倉田美穂  
 熱田由子

## 《模擬プロトコール》

Stata Seminar 2012

Sample protocol

研究課題：移植アウトカムに対するドナーとレシピエントの血液型メジャーミスマッチの影響

### 1.0 目的：

ドナーとレシピエントの血液型メジャー不一致が、全生存、移植関連死亡を含む移植アウトカムに与える影響について検討する。

### 2.0 利用データ

- ・日本造血細胞移植学会 : 成人血縁者間移植
- ・日本小児血液学会 : 小児血縁者間移植
- ・骨髄移植推進財団 : 非血縁者間骨髄移植
- ・日本さい帯血バンクネットワーク : 非血縁者間臍帯血移植

### 3.0 データセット

Stata Seminar 2012用サンプルデータ (TRUMP 2011固定データセットからの無作為抽出データ)

- ・ StataSeminar2012

### 4.0 解析対象：

選択規準：

- ・対象疾患：急性白血病  
(AML, ALL)
- ・患者年齢：年齢制限なし
- ・移植時期：2000年1月から2009年12月までに移植を受けた症例
- ・移植回数：第1回目の同種移植症例のみ
- ・ドナー・レシピエントの血液型一致度が判明しているもの

除外規準：

- ・複数臍帯血移植
- ・UR-PBSCT, BMT+PBSCTを除く (極めて少数のため)
- ・幹細胞種類が不明、あるいはその他
- ・ドナーが不明のもの
- ・GVHD予防なしのもの
- ・生存状況あるいは生存時間が不明のもの (主要評価項目の解析が不可)

### 5.0 アウトカム：

Primary endpoint

- ・全生存

Secondary endpoints

- ・非再発死亡 (*non-relapse mortality, NRM*)
- ・再発
- ・好中球回復



- grade 2 to 4 acute GVHD

### 解析に用いる変数：

Main effects：

- matched or minor mismatched vs. major mismatched or major-minor mismatched

Recipient-related：

- 年齢：0-15, 16-39, 40-
- 原疾患：AML, ALL
- 移植時病期：
  - 白血病のみ
    - Standard (AML：CR1, CR2, ALL：1CR, CML：1CP, MDS：RA/RARS)
    - Advanced (all others)
- 移植前PS：0-1 vs.  $\geq 2$

Donor-related：

- ドナーとの関係：血縁 vs. 非血縁
- 幹細胞種類：BM vs. PBSC vs. CB
- HLA適合：HLA well matched (including identical) vs. HLA partially matched vs. mismatched
- (こちらを採用) ドナー-細胞種類 combination：
  - Rel-BM
  - Rel-PB
  - UR-BM
  - UR-CB

Transplant-related：

- 移植前処置：TBI照射なし vs. TBI照射あり
- 移植前処置：骨髄破壊的 (MAC) vs. 骨髄非破壊的 (RIC)
- GVHD予防：CsA based vs. Tac based

## 6.0 研究デザイン

本邦において1回目の血縁者間末梢血幹細胞・骨髄移植、非血縁者間骨髄移植、非血縁者間臍帯血移植を施行された成人の造血器悪性疾患症例を対象として、ドナーとレシピエントの血液型の組み合わせが、全生存、非再発死亡、好中球回復、急性GVHD、あるいは再発に及ぼす影響を検討する。

血液型の組み合わせでは、ドナー・レシピエント間に関して、マッチに対しての他の組み合わせのものの全生存に対するリスクを単変量、および多変量解析で実施し、最終解析における群分け方法を決定する。

ドナー・移植種類によるサブグループ解析を追加し、血液型の影響がどのような移植において明らかであるか検討する。

## 7.0 統計解析

### 7.0.1 主評価項目の解析

#### (1) 全生存

Main effect 群分け変数において、移植日 (day0) を起算日とし、原因を問わない死亡をイベントとしてKaplan-Meier法を用いて全生存曲線の描出および全生存率の算出を行う。群間比較はlogrank testで検定する。同アウトカムに関して8.0.4で示す方法

での多変量解析による補正解析を実施する。

#### 7.0.2 副次的評価項目の解析

##### (1) 非再発死亡 (Non-relapse mortality)

Main effect 群分け変数において、移植日 (day0) を起算日とし、寛解時死亡(非再発死亡)をイベントとし、再発を競合リスクイベントとして Cumulative incidence 法を用いて再発率を算出し、Cumulative incidence curve を描出する。群間比較は logrank test および Gray's test で検定する。同アウトカムに関して 8.0.4 で示す方法での多変量解析による補正解析を実施する。

##### (2) 再発

Main effect 群分け変数において、移植日 (day0) を起算日とし、再発をイベントとし、寛解時死亡を競合リスクイベントとして Cumulative incidence 法を用いて再発率を算出し、Cumulative incidence curve を描出する。群間比較は logrank test および Gray's test で検定する。同アウトカムに関して 8.0.4 で示す方法での多変量解析による補正解析を実施する。

##### (3) 好中球回復

Main effect 群分け変数において、移植日 (day0) を起算日とし、好中球回復 (3 ポイント以上好中球  $500/\mu\text{l}$  以上を測定できた場合に好中球回復とし、最初の検査日を回復日とする) をイベントとし、回復前死亡を競合リスクイベントとして Cumulative incidence 法を用いて好中球回復率を算出し、Cumulative incidence curve を描出する。群間比較は logrank test および Gray's test で検定する。同アウトカムに関して 8.0.4 で示す方法での多変量解析による補正解析を実施する。

##### (4) Grade 2 to 4 急性 GVHD

Main effect 群分け変数において、移植日 (day0) を起算日とし、II 度以上の急性 GVHD をイベントとし、II 度以上の急性 GVHD 発症前死亡と II 度以上の急性 GVHD 発症前再発を競合リスクイベントとして Cumulative incidence 法を用いて好中球回復率を算出し、Cumulative incidence curve を描出する。群間比較は logrank test および Gray's test で検定する。同アウトカムに関して 8.0.4 で示す方法での多変量解析による補正解析を実施する。

#### 7.0.3 サブグループ解析

以下の移植種別における主評価項目および副次的評価項目のサブグループ解析を行う。

- ・ 血縁者間移植
- ・ 非血縁者間骨髄移植
- ・ 非血縁者間臍帯血移植

#### 7.0.4 多変量解析

以下の因子で補正し、再発率、全生存率 (OS) および移植関連死亡 (TRM) の多変量解析を行う。

- ・ 年齢 (0-15, 16-39, 40-)
- ・ 患者性別 (F/M)
- ・ 患者・ドナー性別一致度 (match, M to F, F to M)
- ・ 原疾患 (AML, ALL)
- ・ 移植時病期 (standard, advanced)
- ・ ドナー--細胞種類 combination :

Rel-BM

Rel-PB

UR-BM

UR-CB

- ・ 移植前処置 (MAC / RIC)
- ・ GVHD 予防 (CyA-based, Tac-based)

全生存は、Cox 比例ハザードモデルを用い、競合リスクアウトカム（非再発死亡、再発、好中球回復、急性 GVHD）は、競合リスク回帰モデルを用いる。多変量解析のモデルに加える補正因子の選択方法は、いくつかあるが、セミナーでは以下の方法①および③-1 を実践する。

- ① 変数をすべて用いるフルモデルで補正する。  
：文献や医学的見地を参照し、主要評価項目全生存に対するリスク因子を用いる。
- ② 群間において比較(検定)を行い、単変量解析において、有意にあるいは **marginal** に ( $P < 0.1$ ) 検討するエンドポイントに影響を与えた因子をモデルに加える。
- ③ 選択した因子をモデルに加える。
  - ③-1：全ての検討する因子を含めたフルモデルからスタートし、血液型の変数は常に残した上で、**backward stepwise** にて P 値が高い変数から順にフルモデルから除き血液型の変数以外は有意な因子のみを残した最終モデルとする。
  - ③-2：血液型を除いた変数のみを用いて、**forward or backward stepwise** にて有意な因子のみを残したモデルを最初に作成する。絞り込まれた因子と目的の因子をモデルに加えて最終モデルとする。

## 8.0 背景表

《変数表》

	Variables	Value	Value label	コメント
移植日	tdate	移植日付		
移植年	tyear	移植年		
診断から移植までの日数	dx_to_sct			
診断から移植までの日数グループ	dx_to_sct_group	0	<=90	
		1	90 to 180	
		2	>180	
移植年グループ	tyear_group	1	-1990	
		2	1991-1995	
		3	1996-2000	
		4	2001-2005	
		5	2006-	
幹細胞種類	source	1	BM	
		2	PB	
		3	BM+PB	
		4	CB	
		5	Other	
ドナー種類	donor_cd	1	Auto	
		2	Syn	
		3	Rel	
		4	UR	
ドナー血縁	rel	1	Sib	
		2	Other_rel	
		9	Missing	
auto_allo	auto_allo	1	Auto	
		2	Allo	
移植種類	sct_type	1	Auto	
		2	Rel-BM	Syngeneicは含めていない
		3	Rel-PB	Syngeneicは含めていない
		4	UR-BM	
		5	UR-PB	
		6	UR-CB	
移植時患者年齢	pt_age	患者移植時年齢		
移植時患者年齢グループ	pt_age_group1	1	0-15	
		2	16-39	
		3	40-	
年齢グループ2	pt_age_group2	1	0-9	
		2	10-19	
		3	20-29	
		4	30-39	
		5	40-49	
		6	50-59	
		7	60-	
年齢50歳以上	age_50	0	0-49	
		1	50-	
child/adult	adult	1	0-15	
		2	16-	
患者性別	pt_sex	0	Female	
		1	Male	
ドナー性別	donor_sex	0	Female	
		1	Male	
性別マッチ/ミスマッチ	sex_mismatch1	0	Match	
		1	Mismatch	
	sex_mismatch2	0	Match	
		1	M to F	
		2	F to M	

	Variables	Value	Value label	コメント
ECOG PS	ps24	0	PS 0-1	
		1	PS 2-	
患者ABO	pt_abo	O		文字列、TRUMPオリジナル項目
		A		
		B		
		AB		
ドナーABO	do_abo	O		文字列、TRUMPオリジナル項目
		A		
		B		
		AB		
ABOメジャーマッチ/ミスマッチ	abo_majormis	0		
		1		
ABOマイナーマッチ/ミスマッチ	abo_minormis	0		
		1		
ABOミスマッチ	abo_mismatch	0	Matched	
		1	Minor mismatch	
		2	Major mismatch	
		3	Major-minor mismatch	

	Variables	Value	Value label	コメント
疾患名	diagnosis2	1	AML	
		2	ALL	
		3	ATL	
		4	CML	
		5	MDS	
		6	Other leukemia	その他白血病
		7	MPN	CMML、JMMLを含む
		8	NHL	
		9	HL	
		10	CLL	
		11	Other lymphoma	type missingは含めていない
		12	MM	
		13	PCL	
		14	POEMS	
		15	Amyloidosis	
		16	Plasma Cell Neoplasmas	
		17	AA	
		18	PRCA	
		19	PNH	
		20	Congenital Bone Marrow Failure	
		21	Other Bone Marrow Failure	
		22	Solid Tumor	
		23	EB related Disorders	
		24	HPS or LCH	
		25	Autoimmune Disease	
		26	Inborn Metabolism Errors	
		27	Primary Immunodeficiency	
		28	Others	
		99	not written	
	diagnosis	1	AML	JSHCT報告書作成時の解析variable
		2	ALL	
		3	ATL	
		4	CML	
		5	MDS	
		6	ot leukemia	その他白血病
		7	MPD	CMML、JMMLを含む
		8	NHL	
		9	HL	
		10	Other lymphoma / type missing	
		11	MM/PCD	
		12	AA	
		13	PRCA	
		14	PNH	
		15	Cong. 造血障害	先天性造血障害
		16	ST	固形腫瘍
		17	Others	
99	not written			
疾患大分類	diag_cd1	1	Leukemia	
		2	Lymphoma	
		3	MM	
		4	ST	
		5	AA/PRCA/PNH	
		6	Other	
固形癌疾患名	diag_st	1	Breast cancer	
		2	Ovarian cancer	
		3	Lung cancer	
		10	Germ cell tumor	
		11	Neuroblastoma	
		12	Rhabdomyosarcoma	
		13	Central nerve tumors	
		14	Hepatoblastoma	
15	Wilms tumor			
16	Ewing's sarcoma			
17	Osteosarcoma			
AML分類	diag_aml	0	M0	
		1	M1	
		2	M2	
		3	M3	

	Variables	Value	Value label	コメント
		4	M4	
		5	M5	
		6	M6	
		7	M7	
		8	with MD	多系統の形態異常を伴うAML
		9	other AML	
		10	AML UK	
chromosome_ph	chromosome_ph	1		ALL分類(Ph-ALL)
MDS分類	diag_mds	1	RA	
		2	RARS	
		3	RAEB	
		4	RAEBt	
MDS分類2	diag_mds2	1	RA/RARS	RA,RARS
		2	RAEB/RAEBt	RAEB,RAEBt
MPD分類	diag_mpd	1	JMML	
		2	CMMoL	
		3	MF	
NHL分類	diag_nhl	1	DLBCL	びまん性大細胞型B細胞性リンパ球
		2	Follicular	濾胞性リンパ球
		3	Mantle	マントル細胞リンパ腫
		4	Lymphoblastic	B前駆細胞腫瘍,T前駆細胞腫瘍
先天性造血障害	diag_cmf	1	Fanconi Anemia(AA)	Fanconi貧血(造血障害のみ)
		2	Diamond-Blackfan	Diamond-Blackfan貧血
		3	Congenital Neutropenia	先天性好中球減少症
先天性代謝異常症	diag_met	1	MPS	ムコ多糖症
		2	ALD	副腎白質ジストロフィー
原発性免疫不全症	diag_id	1	SCID	重症複合免疫不全症(SCID)
		2	WAS	Wiskott Aldrich症候群
		3	CGD	慢性肉芽腫症



	Variables	Value	Value label	コメント
病期	disease_status	1	CR1	AML or ALL
		2	CR2	AML or ALL
		3	CR3-	AML or ALL
		4	NR	AML or ALL、初発状態(未治療)を含む
		5	CP1	CML
		6	CP2	CML
		7	CP3-	CML
		8	AP	CML
		9	BC	CML
		10	RA	MDS
		11	RAEB	MDS
		12	RAEBt	MDS
		13	RARS	MDS
AML移植時病期	stage_aml	1	1CR	
		2	2CR	
		3	>=3CR	
		4	NCR	初発状態(未治療)を含む
		9	Missing	
ALL移植時病期	stage_all	1	1CR	
		2	2CR	
		3	>=3CR	
		4	NCR	初発状態(未治療)を含む
		9	Missing	
CML移植時病期	stage_cml	1	1CP	
		2	2CP	
		3	>=3CP	
		4	AP	
		5	BC	
Lymphoma移植時病期	stage_lym	1	1CR	HL,NHL共通
		2	2CR	
		3	>=3CR	
		4	1PR	
		5	2PR	
		6	NIR	
MM移植時病期	stage_mm	1	CR/PR/MR	
		2	NC/PD	
固形癌移植時病期	stage_st	1	CR	
		2	NCR	寛解導入不能例,寛解導入例:CR以外
疾患リスク分類1	stage1	0	Standard	AML:1CR/2CR,ALL:1CR,CML:1CP,MDS:RA/RARS
		1	Advanced	all others
疾患リスク分類2	stage2	0	Low risk	AML/ALL:CR1,CR2、CML:CP1,CP2、MDS
		1	High risk	AML/ALL:CR3-,NR、CML:CP3-,AP,BC
	stage3	0	Standard	AML/ALL CR1,CR2、CML CP1,CP2、AP、MDS RA,RARS
		1	Advanced	AML/ALL CR3-,NR、CML CP3-,BC、MDS RAEB, RAEBt(as AML)

	Variables	Value	Value label	コメント
MAST/RIST	ric_mac	1	RIC	
		0	MAC	
TBI	tbi	0		TBI照射なし
		1		TBI照射あり
Full TBI	tbi_full	1		TBI有りで8Gyを超え、15.75Gy以下。これを超える場合は入力ミスの可能性が高いとした
TBI regimen	tbi_regimen	0		
		1		TBI有りで線量が8Gy以下のもの、および線量が不明でRICと記載があったものを除いたもの(線量不明でRICと記載のないものは1に含まれる)
移植前処置	cond	1	CY+TBI±	
		2	Other TBI regimen	
		3	BY+CY±	
		4	Other non TBI	TBIが空欄の場合、化学療法剤2剤以上が“有”で1-3, 5-6を満たさないもの。MMのMelphala単剤およびAAのCY単剤はここに含む
		5	FL±TBI±	
		6	Other RIST	
移植前処置2	cond2	9	Missing	
		1	FL+BU±	RISTの細分類
		2	FL+CY±	
		3	FL+Mel±	
GVHD予防	gvhd_pro	4	Other RIST	
		0	無し	
		1	CyA+MTX	
		2	CyA±	
		3	Tac+MTX	
		4	Tac±	
CSAあるいはTACの使用	gvhd_pro2	5	Other	
		9	Missing	
体内T細胞除去	tcd_invivo	0		
		1		Campath、ATG、ALGのいずれかが有り (conditioning or gvhd prophylaxis)

	Variables	Value	Value label	コメント
幹細胞種類	source1	1		骨髄(BMT)
		2		末梢血(PBSCT)
		3		臍帯血(CBSCT)
		4		骨髄+末梢血(BMT+PBSCT)
		5		複数臍帯血(MultiCBSCT)
		6		その他
		9		Null
移植日	txdt	移植日付		
移植年	txdt_y	移植年		
HLA_血清型_患者A1	serora1			
HLA_血清型_患者A2	serora2			
HLA_血清型_ドナーA1	seroda1			
HLA_血清型_ドナーA2	seroda2			
HLA_血清型_患者B1	serorb1			
HLA_血清型_患者B2	serorb2			
HLA_血清型_ドナーB1	serodb1			
HLA_血清型_ドナーB2	serodb2			
HLA_血清型_患者C1	serorc1			
HLA_血清型_患者C2	serorc2			
HLA_血清型_ドナーC1	serodc1			
HLA_血清型_ドナーC2	serodc2			
HLA_血清型_患者DR1	serordr1			
HLA_血清型_患者DR2	serordr2			
HLA_血清型_ドナーDR1	seroddr1			
HLA_血清型_ドナーDR2	seroddr2			
HLA_遺伝子型_患者A1	genora1			
HLA_遺伝子型_患者A2	genora2			
HLA_遺伝子型_ドナーA1	genoda1			
HLA_遺伝子型_ドナーA2	genoda2			
HLA_遺伝子型_患者B1	genorb1			
HLA_遺伝子型_患者B2	genorb2			
HLA_遺伝子型_ドナーB1	genodb1			
HLA_遺伝子型_ドナーB2	genodb2			
HLA_遺伝子型_患者DRB1_1	genordr1			
HLA_遺伝子型_患者DRB1_2	genordr2			
HLA_遺伝子型_ドナーDRB1_1	genoddr1			
HLA_遺伝子型_ドナーDRB1_2	genoddr2			
HLA_遺伝子型_患者C1	genorc1			
HLA_遺伝子型_患者C2	genorc2			
HLA_遺伝子型_ドナーC1	genodc1			
HLA_遺伝子型_ドナーC2	genodc2			
HLA抗原マッチング				
血清型 A GVH方向ミスマッチ 座数	serogvhmisa	0		0座ミスマッチ
		1		1座ミスマッチ
		2		2座ミスマッチ
血清型 A HVG方向ミスマッチ 座数	serohvgmisa	0		0座ミスマッチ
		1		1座ミスマッチ
		2		2座ミスマッチ
血清型 B GVH方向ミスマッチ 座数	serogvhmisb	0		0座ミスマッチ
		1		1座ミスマッチ
		2		2座ミスマッチ
血清型 B HVG方向ミスマッチ 座数	serohvgmisb	0		0座ミスマッチ
		1		1座ミスマッチ
		2		2座ミスマッチ
血清型 C GVH方向ミスマッチ 座数	serogvhmisc	0		0座ミスマッチ
		1		1座ミスマッチ
		2		2座ミスマッチ
血清型 C HVG方向ミスマッチ 座数	serohvgmisc	0		0座ミスマッチ
		1		1座ミスマッチ
		2		2座ミスマッチ
血清型 DR GVH方向ミスマッチ 座数	serogvhmisdr	0		0座ミスマッチ
		1		1座ミスマッチ
		2		2座ミスマッチ

	Variables	Value	Value label	コメント
血清型 DR HVG方向ミスマッチ 座数	serohvgmisdr	0		0座ミスマッチ
		1		1座ミスマッチ
		2		2座ミスマッチ
血清型 A,B,DR GVH方向ミスマッチ 座数	serogvhmis6abdr	0		0座ミスマッチ
		1		1座ミスマッチ
		2		2座ミスマッチ
		3		3座ミスマッチ
		4		4座ミスマッチ
		5		5座ミスマッチ
血清型 A,B,DR HVG方向ミスマッチ 座数	serohvgmis6abdr	0		0座ミスマッチ
		1		1座ミスマッチ
		2		2座ミスマッチ
		3		3座ミスマッチ
		4		4座ミスマッチ
		5		5座ミスマッチ
血清型 A,B,C, DR GVH方向ミスマッチ 座数	serogvhmis8abcdr	0		0座ミスマッチ
		1		1座ミスマッチ
		2		2座ミスマッチ
		3		3座ミスマッチ
		4		4座ミスマッチ
		5		5座ミスマッチ
		6		6座ミスマッチ
		7		7座ミスマッチ
血清型 A,B,C, DR HVG方向ミスマッチ 座数	serohvgmis8abcdr	0		0座ミスマッチ
		1		1座ミスマッチ
		2		2座ミスマッチ
		3		3座ミスマッチ
		4		4座ミスマッチ
		5		5座ミスマッチ
		6		6座ミスマッチ
		7		7座ミスマッチ
血清型 Aミスマッチ座数	seromisa	0		serogvhmisaとserohvgmisaのうち大きい方
		1		
		2		
血清型 Bミスマッチ座数	seromisb	0		serogvhmisbとserohvgmisbのうち大きい方
		1		
		2		
血清型 Cミスマッチ座数	seromisc	0		serogvhmiscとserohvgmiscのうち大きい方
		1		
		2		
血清型 DRミスマッチ座数	seromisdr	0		serogvhmisdrとserohvgmisdrのうち大きい方
		1		
		2		
血清型 A,B,DRミスマッチ 座数	seromis6abdr	0		0座ミスマッチ
		1		1座ミスマッチ
		2		2座ミスマッチ
		3		3座ミスマッチ
		4		4座ミスマッチ
		5		5座ミスマッチ
血清型 A,B,DRミスマッチ 座数	seromis6abdr2	0		0座ミスマッチ
		1		1座ミスマッチ
		2		2座ミスマッチ
		3		3座ミスマッチ