

# Kaplan-Meierプロット・Forestプロット作成の応用：

## グラフ出力範囲内・範囲外への数値出力

○魚住 龍史<sup>1</sup> 吉田 早織<sup>2</sup> 平井 隆幸<sup>2</sup> 浜田 知久馬<sup>3</sup>

<sup>1</sup>京都大学大学院医学研究科 医学統計生物情報学

<sup>2</sup>日本化薬株式会社 医薬開発本部 解析チーム

<sup>3</sup>東京理科大学 工学部 情報工学科

Advancement in both Kaplan-Meier and forest plots: quantitative results output inside or outside the graph area

Ryuji Uozumi<sup>1</sup>, Saori Yoshida<sup>2</sup>, Takayuki Hirai<sup>2</sup>, and Chikuma Hamada<sup>3</sup>

<sup>1</sup>*Department of Biomedical Statistics and Bioinformatics, Kyoto University Graduate School of Medicine*

<sup>2</sup>*Biostatistics Team, Pharmaceutical Development Division, Nippon Kayaku Co.,Ltd*

<sup>3</sup>*Department of Information and Computer Technology, Faculty of Engineering, Tokyo University of Science*

## 要旨

Kaplan-Meier プロットおよび Forest プロットは、いずれも医薬品開発でよく用いられるグラフである。本稿では、ODS GRAPHICS による機能を用いて、様々な修飾を加えた Kaplan-Meier プロットおよび Forest プロットを作成する方法を報告する。Kaplan-Meier プロットにリスク集合の大きさを修飾させて作成する場合、LIFETEST プロシジャのデフォルトではグラフ出力範囲内、すなわち横軸の上段にリスク集合の大きさが出力される。リスク集合の大きさをグラフ出力範囲外、すなわち横軸の下段に出力させる方法として、V9.4 の SGPLOT プロシジャから追加された XAXISTABLE ステートメントを用いる方法を取りあげる。次に、Forest プロットの作成を考えると、グラフの横に実際の点推定値と信頼区間も示されることが多い。V9.4 の SGPLOT プロシジャから追加された複数のステートメントを併用することで、実際の学術論文に用いられるようなクオリティーの高い Forest プロットを作成する方法を報告する。

キーワード: 生存時間解析 Kaplan-Meier プロット Forest プロット サブグループ解析 SGPLOT 修飾を加えたプロット リスク集合 XAXISTABLE YAXISTABLE X2AXIS Y2AXIS

## 1 はじめに

データのグラフ化は、データ解析の結果を可視化するために有用な手段である。SAS では、V9.2 から ODS (Output Delivery System) GRAPHICS による機能が正規版として追加され、グラフィックベースのきれいなプロットを出力することができるようになった。例えば、ODS GRAPHICS ON の状態で LIFETEST プロシジャを実行させれば、Kaplan-Meier プロットが自動的に出力される。オプションを指定すれば、様々な修飾を加

えることが可能である。また、V9.2 から追加された SG (Statistical Graphics) プロシジャである、SGPLOT プロシジャや SGPNEL プロシジャを用いれば、より見栄えの良いプロットを作成することができる。

本稿では、医薬品開発で用いられるグラフを SAS で作成することに焦点を当てる。これまで SAS によるグラフ作成に関して多くの報告が行われている。西本 (2013) は GPLOT プロシジャと SGPLOT プロシジャを比較した上で、SGPLOT プロシジャによるグラフ作成について解説した。豊泉ら (2014) によって、SGPLOT プロシジャによるデータの可視化の有用性について述べられている。さらに、魚住・浜田 (2012) は SG Annotation の機能を用いて、医薬品開発でよく用いられるグラフを作成する有用性を示した。また、より実務的な応用として、SG プロシジャ、GTL (Graph Template Language) および ODS PDF を用いた簡易解析帳票の作成方法事例や、SAS と HTML アプリケーションによる CDISC ADaM (Clinical Data Interchange Standards Consortium Analysis Data Model) 形式の解析用データセットを用いた解析帳票・グラフ簡易作成ツールの開発事例についても報告されている (高浪, 2011, 2012)。加えて、V9.4 から利用できるようになった ODS PowerPoint を用いて、SAS で作成したグラフを PowerPoint にアウトプットさせる方法も報告されている (吉田ら, 2015)。

医薬品開発で用いられるグラフのうち、Kaplan-Meier プロットは生存時間データの視覚化の手段として 20 年以上前から用いられており (大橋・浜田, 1995)、V9.2 リリース以降、SAS による Kaplan-Meier プロットの作成に関する多くの報告が行われている。長島・佐藤 (2010) は、LIFETEST プロシジャおよび GPLOT プロシジャを用いて、Kaplan-Meier プロットに付加情報を加えるマクロを開発している。魚住・浜田 (2011) は、SGPLOT プロシジャおよび SGRENDER プロシジャを用いて、Kaplan-Meier プロットを作成する手順を示している。LIFETEST プロシジャではラスター形式のグラフ出力であり、ベクター形式の出力ができない一方、SGPLOT プロシジャではベクター形式の出力も可能である点からも有用といえる (平井ら, 2015)。

長島・佐藤 (2010) および魚住・浜田 (2011) のいずれの報告においても、Kaplan-Meier プロットのリスク集合の大きさ (number of subjects at risk) を出力させる方法について言及されている。リスク集合の大きさは、LIFETEST プロシジャのオプションを指定すれば、ODS GRAPHICS による出力で自動的に付加されるが、より工夫を凝らした出力をするための方法が報告されている。例えば、学術論文において、リスク集合の大きさは Kaplan-Meier プロットの横軸の下段に示されることが多いが、LIFETEST プロシジャのオプションで ATRISK を指定して出力すると、Kaplan-Meier プロットの横軸の上部に示される。よって、リスク集合の大きさを Kaplan-Meier プロットの下段に示す方法を考える意義があるといえる。Kaplan-Meier プロットに限らず、その他にも、例えば Forest プロットの作成においても工夫した出力を考える意義がある。Forest プロットはメタアナリシスやサブグループ解析の報告によく用いられるグラフであり、プロットに加えて、グラフの横に実際の推定値と信頼区間を示すことが多い。以上のようなことから、医薬品開発に携わる統計家および SAS プログラマにとって、グラフ出力範囲外にリスク集合の大きさや信頼区間などの情報を出力させる方法を考える必要が生じる。

本稿では、医薬品開発でよく用いられる Kaplan-Meier プロットおよび Forest プロットを作成するにあたり、グラフ出力範囲外に数値を出力させる方法として、LIFETEST プロシジャおよび SGPLOT プロシジャを用いた方法をレビューした上で、新たにできるようになった方法を報告する。

なお、本稿において示すプロットは、モノクロ印刷されても識別しやすいよう、ODS の出力としてジャーナルスタイル "STYLE = JOURNAL" を指定している。例えば、群別のグラフを示す場合、デフォルトでは群の違いが色で分けて出力されてしまう。ジャーナルスタイルの場合、モノクロ印刷で群の違いを識別できるよう、実線と破線で分けて出力される。

## 2 Kaplan-Meier プロットの作成

本節では、リスク集合の大きさの情報を修飾した Kaplan-Meier プロットを作成することを考える。Kaplan-Meier プロットを作成する対象データとして、肺癌のデータ（データセット名：VALung）を用いる。データセット VALung (Veterans Administration Lung cancer trial) は、Kalbfleisch and Prentice (2002) で使用されたデータを一部抽出したものであり ( $n = 137$ )、SAS/STAT PHREG プロシジャのマニュアル、および大橋ら (2016) 第 3 章においても用いられている。この研究の目的は、男性の進行性肺癌患者を対象としたランダム化比較試験であり、治療法として、試験治療と標準治療を比較するために行われた。評価項目は死亡までの時間 (日) (変数名: Time) で、共変量の 1 つとして組織型 (変数名: Cell) が挙げられる。組織型は 4 水準 (Cell = 'adeno' 「腺癌」, 'small' 「小細胞癌」, 'large' 「大細胞癌」, 'squamous' 「扁平上皮癌」) のカテゴリカル変数である。本節では、組織型別の Kaplan-Meier プロットを作成することを考える。

### 2.1 グラフの出力範囲内へのリスク集合の出力

LIFETEST プロシジャでは、デフォルトではリスク集合の大きさは付与されないが、PLOTS = S のオプションとして、ATRISK を指定すれば、図 1 に示すような出力が得られる。その他のオプションの指定に関しては、大橋ら (2016) を参照されたい。

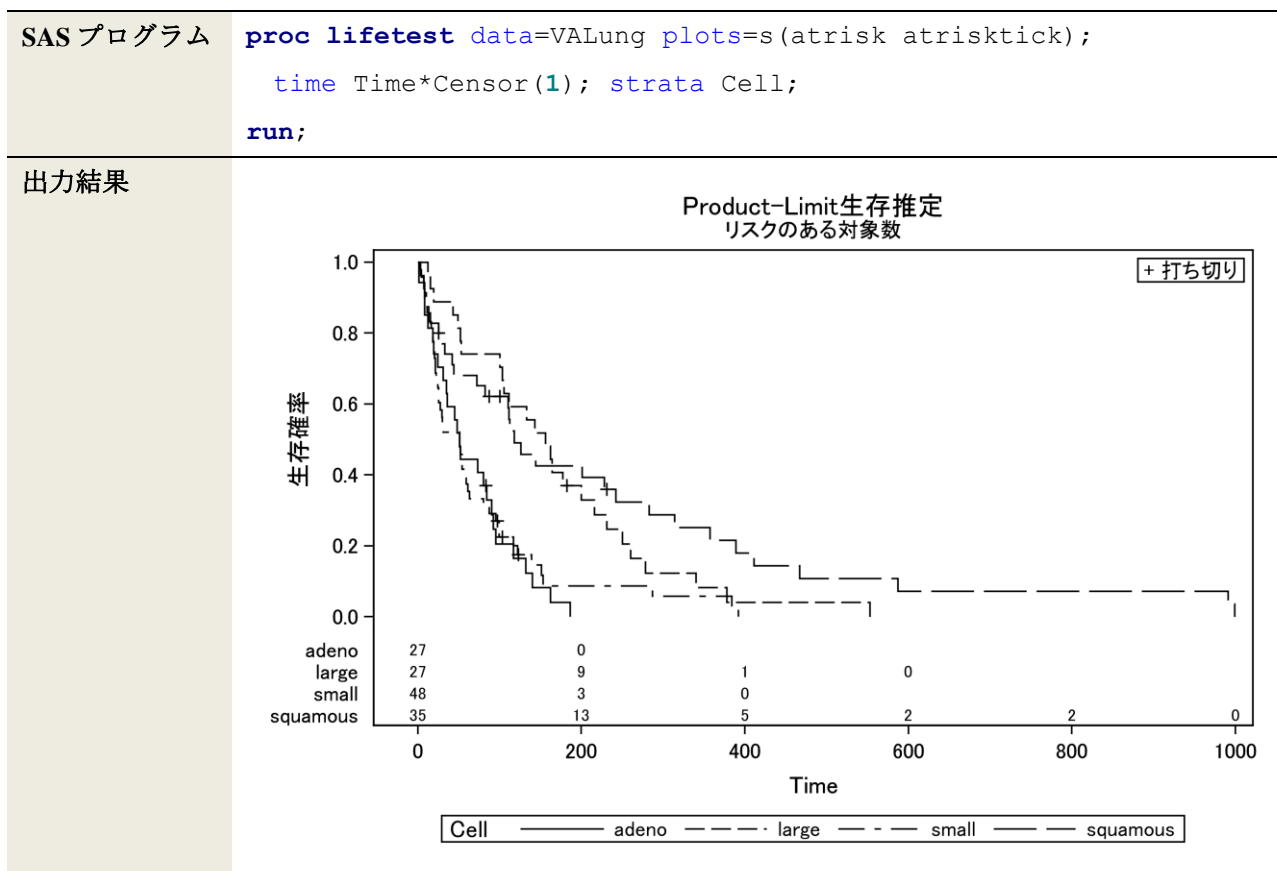


図 1 : LIFETEST プロシジャによる Kaplan-Meier プロットの作成 1

```

SAS プログラム  ods graphics;
                  ods output Survivalplot=Survivalplot;
                  proc lifetest data=VALung atrisk plots=s(atrisk);
                    time Time*Censor(1); strata Cell;
                  run;
                  *-----*;
                  *-- データハンドリング (詳細は付録A参照) --*;
                  *-----*;
                  proc sgplot data=Survivalplot2 noautolegend;
                    step x=Time y=Survival / group=StratumNum name='Survival';
                    scatter x=Time y=Censored
                          / group=StratumNum markerattrs=(symbol=plus);
                    scatter x=tAtRisk y=StratumNum / markerchar=atrisk y2axis;
                    keylegend 'Survival'
                          / location=outside noborder position=bottom;
                    yaxis offsetmin=0.20 values=(0 to 1 by 0.1) label='Survival';
                    y2axis offsetmin=0.03 offsetmax=0.85
                          min=1 max=4 label=' ' display=(noticks novalues);
                    xaxis offsetmin=0.05
                          values=(0 to &TimeMax by &TimeInterval) label='Time';
                    format StratumNum Cellf.;
                  run;

```

## 出力結果

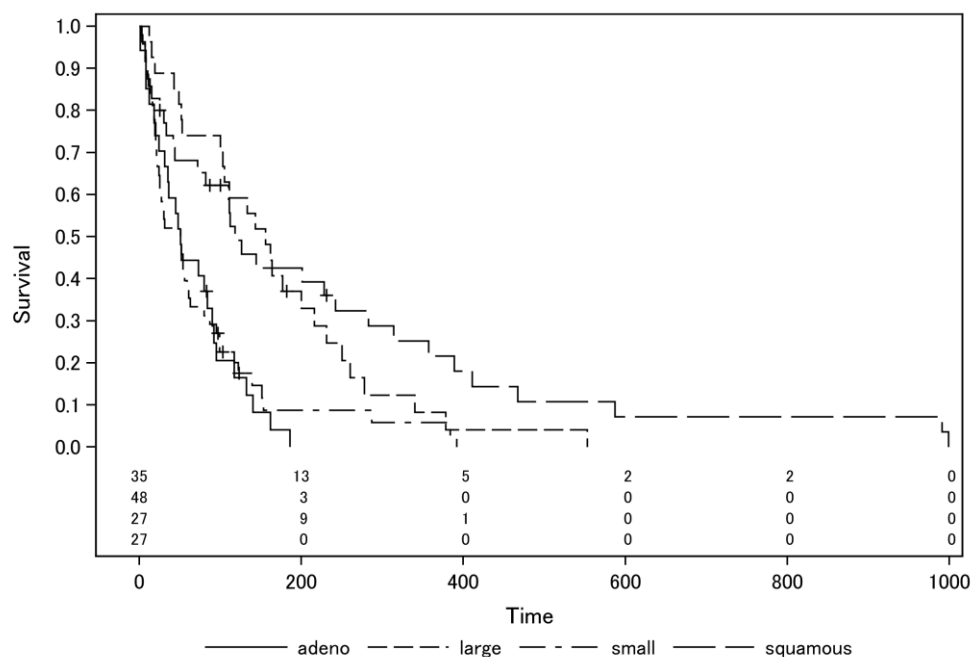


図 2 : SGPLOT プロシジャによる Kaplan-Meier プロットの作成 1

しかし、図1の出力では、リスク集合の大きさが出力されていない時点および群が存在する。データセット VALung の特徴として、例数に対するイベント割合が高いため、後半の時点ではリスク集合の大きさが0となってしまう。リスク集合が0となると、0になった最初の時点では出力されるが、それ以降の時点では0は出力されない。本稿ではリスク集合の大きさが0の場合は0に統一する出力を考える。PLOTS=Sのオプションで指定できない修飾を加える場合、生存関数の推定値を一度データセットに落として、グラフ関連のプロシジャで出力し直す必要がある。

図2はSGPLOTプロシジャを用いたプログラムおよびその出力結果を示しており、大橋ら(2016)第2章で解説されている方法に加えて、Y2AXISステートメントに加えた上でYAXISステートメントのOFFSETMIN=を指定することで、グラフの出力範囲内にリスク集合の大きさを出力させるスペースを確保している。YAXISおよびXAXISステートメントはグラフの縦軸および横軸の指定として、グラフの左側および下側の軸の設定を行っている一方、Y2AXISおよびX2AXISステートメントではグラフの右側および上側の軸の設定を行っている。OFFSETMIN=とOFFSETMAX=を指定すれば、軸の下限と上限をそれぞれ指定できる。図2のプログラムでは、YAXISステートメントでOFFSETMIN=0.20を指定し、縦軸全体の20%の値からスタートさせている。縦軸全体は生存割合1であるため、生存割合20%に相当する位置が原点として描かれている。一方、右側の縦軸に関して、Y2AXISステートメントでOFFSETMIN=0.03を指定することで、横軸とリスク集合の値が重ならないように設定している。さらに、OFFSETMAX=0.85を指定することで、全体の85%の部分を出力範囲の上限としている。なお、ODS OUTPUTによって得られたデータセットに対して、グラフで出力する時点のリスク集合の大きさが0であった場合に0が出力されるよう、データハンドリングした上で、SGPLOTプロシジャを実行させる必要がある。データハンドリングの詳細は付録Aを参照されたい。

## 2.2 グラフの出力範囲外へのリスク集合の出力

図1および図2は、いずれもグラフの出力範囲内にリスク集合の大きさを追加したKaplan-Meierプロットである。Kaplan-Meierプロットに関しては、学術論文においてリスク集合の大きさをKaplan-Meierプロットの横軸の下段に示されることが多い。そのための方法として、SGRENDERプロシジャを用いて、TEMPLATEプロシジャで事前に定義したテンプレート、特にLAYOUT LATTICEステートメントなどを活用することによって、リスク集合の大きさをKaplan-Meierプロットの下段に示すことが可能である(魚住・浜田, 2011)。しかし、テンプレートの定義には、長文のプログラムを組まなければならない。そこで、V9.4のSGPLOTプロシジャに追加されたステートメントを用いて、リスク集合の大きさをKaplan-Meierプロットの下段に示す方法を考える。

図3はSGPLOTプロシジャを用いたプログラムおよびその出力結果を示している。V9.4から利用できるXAXISTABLEステートメントを用いて、リスク集合の大きさをKaplan-Meierプロットの下段に示している。XAXISTABLEステートメントでは出力させたい値を表す変数AtRiskを指定した上、オプションとしてリスク集合の大きさを出力させる横軸の値を持つ変数tAtRiskをX=に指定し、群に該当する変数StratumNumをCLASS=に指定している。共通の横軸を用いて、Kaplan-Meierプロットの出力とリスク集合の大きさの出力を行っているため、この場合XAXISTABLEステートメントを用いている。共通の縦軸を用いる場合はYAXISTABLEステートメントを用いればよい。なお、XAXISTABLEステートメントを用いて、グラフの範囲内にリスク集合の大きさを出力させたい場合、LOCATION=INSIDEをオプションとして指定すればよい。

図3では、SGPLOT プロシジャを用いてリスク集合の大きさを Kaplan-Meier プロットの下段に示したが、実は SAS/STAT 12.1 以降の LIFETEST プロシジャでは、オプション OUTSIDE を指定することで、リスク集合の大きさを Kaplan-Meier プロットの下段に示すことが可能である (図4)。なお、OUTSIDE(*k*) と指定することで、ODS GRAPHICS で構成される全体の出力範囲のうち、リスク集合の大きさを表示させる割合を指定できる。*k* のデフォルト値は、群の数の 0.035 倍である。図4の場合、4群に対してプロットを描いているため、全体のうち 14%の範囲にリスク集合の大きさを出力させている。

```
SAS プログラム  proc sgplot data=Survivalplot2 noautolegend;
                  step x=Time y=Survival / group=StratumNum name='Survival';
                  scatter x=Time y=Censored
                        / group=StratumNum markerattrs=(symbol=plus);
                  xaxistable AtRisk / x=tAtRisk class=StratumNum;
                  keylegend 'Survival'
                        / location=outside noborder position=bottom;
                  yaxis values=(0 to 1 by 0.1) label='Survival';
                  xaxis values=(0 to &TimeMax by &TimeInterval) label='Time';
                  format StratumNum Cellf.;
run;
```

出力結果

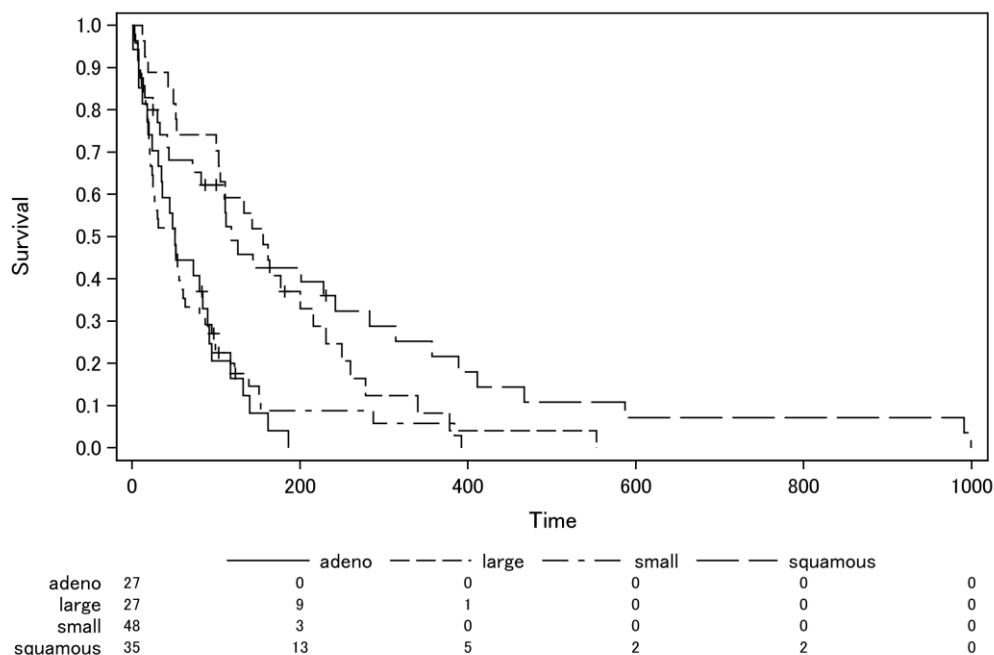


図3: SGPLOT プロシジャによる Kaplan-Meier プロットの作成 2

```
SAS プログラム proc lifetest data=VALung plots=s(atrisk atrisktick outside);  

time Time*Censor(1); strata Cell;  

run;
```

## 出力結果

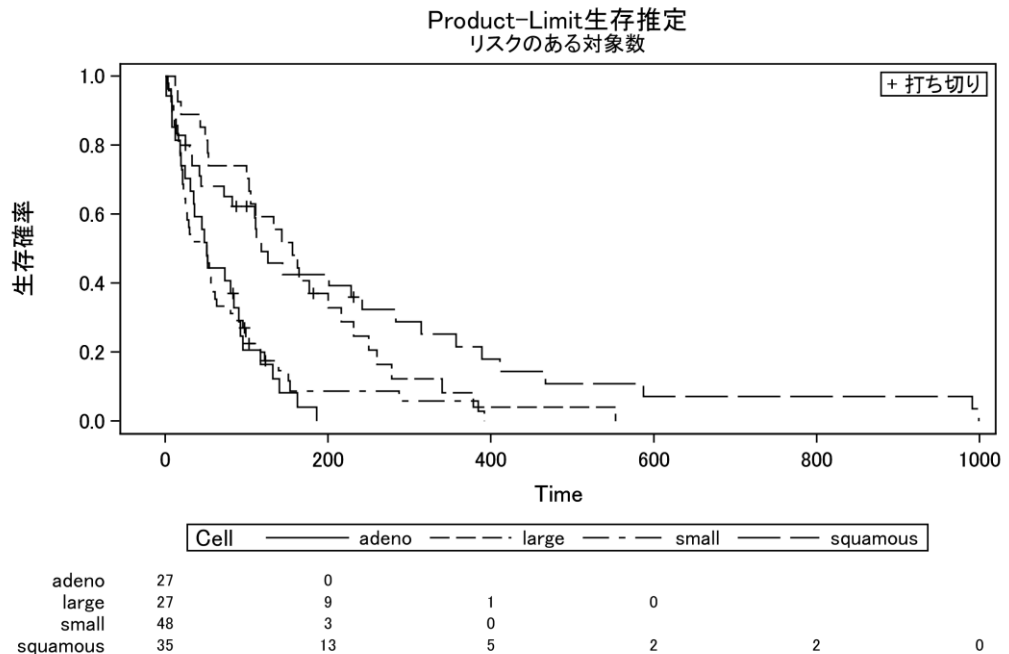


図 4: LIFETEST プロシジャによる Kaplan-Meier プロットの作成 2

## 3 Forest プロットの作成

医薬品開発では様々なプロットを用いてデータの可視化が行われるが、メタアナリシスやサブグループ解析の報告においては Forest プロットがよく用いられる。メタアナリシスであれば研究間の異質性を視覚的に確認するために、サブグループ解析であればサブグループ間でどのような傾向がみられるか確認するために用いられる。

Forest プロットにおいても、Kaplan-Meier プロットに対するリスク集合の大きさのように、実際の推定値およびその信頼区間などの修飾を加えることが多い。近年報告された臨床論文におけるサブグループ解析の結果を例に挙げると、Borghaei *et al.* (2015) はハザード比とその 95% 信頼区間に加え、各サブグループの例数も修飾して出力している。Motzer *et al.* (2015) は各群のイベント割合を修飾しており、点推定値のプロットの大きさについてもそれぞれのサブグループの例数に依存させて変化させている。Turner *et al.* (2015) は各サブグループにおける交互作用の p 値についても修飾させている。

SAS による Forest プロットの報告として、堀田 (2013) は、GPLOT プロシジャとその Annotation の機能を用いて、連続量データに対して Forest プロットを作成する方法を取りあげている。本稿では、Forest プロットに修飾する内容として、点推定値のプロットの大きさを例数に依存して変化させた上、各サブグループにおける例数、イベント割合、点推定値と信頼区間、交互作用の p 値を修飾して出力することとする。前節で用いたデータセット VALung における治療法 (Therapy2 = 1 「試験治療」、2 「標準治療」) の違いのサブグループ解析として、年齢カテゴリ (AgeC = 1 「< 65 歳」、2 「≥ 65 歳」)、組織型 (CellC = 1 「腺癌」、2 「大細胞

癌」, 3「小細胞癌」, 4「扁平上皮癌」), 既往歴の有無 (Prior = 0「無」, 1「有」) をサブグループとした, 比例ハザードモデルによる解析を実施することを考える. 生存時間データに対するサブグループ解析であるため, 点推定値としてハザード比およびその 95%信頼区間を示すこととする. なお, データセット VALung に含まれている文字変数を数値変数として定義した変数が Therapy2, AgeC (Age から IF ステートメントで作成), CellC である.

表 1: Forest プロット作成時のデータセットに格納されている変数の概要

変数名	変数のタイプ	内容
Subgroup	文字	サブグループ名およびサブグループ内のカテゴリ名 [プロットに修飾]
Intpval	数値	交互作用の p 値 [プロットに修飾]
Type	数値	サブグループ名とサブグループ内のカテゴリ名の判別
N	数値	サブグループにおける例数
Grp1	文字	試験治療群におけるイベント数/例数 (%) [プロットに修飾]
Grp2	文字	標準治療群におけるイベント数/例数 (%) [プロットに修飾]
HazardRatio	数値	ハザード比
HRLowerCL	数値	ハザード比の 95%信頼下限
HRUpperCL	数値	ハザード比の 95%信頼上限
HR	文字	ハザード比 (95%信頼区間) [プロットに修飾]
Indent	数値	サブグループ名とサブグループ内のカテゴリ名の出力の 違いを表すインデントを設定
Forestval	数値	プロットの縦軸
Ref	数値	プロット行のうち, 色を変える行に対応する縦軸の値

SGPLOT プロシジャで Forest プロットを作成する場合, PHREG プロシジャを用いた比例ハザードモデルによる解析などを実施し, 表 1 の変数が含まれるデータセットを準備する. なお, PHREG プロシジャによるサブグループ解析から SGPLOT プロシジャで使用するデータセット作成までのプログラムは付録 B, 表 1 のデータセットのアウトプットは付録 C を参照されたい. 表 1 のデータを用いて, SGPLOT プロシジャで Forest プロットを作成するプログラムおよび出力結果を図 5 に示す. 図 5 では, BUBBLE ステートメントでハザード比の点推定値の大きさを●印で示し, 各サブグループにおける例数によって大きさを変更している. ただし, メタアナリシスで用いる Forest プロットの場合, 点推定値の大きさは各研究の例数の大きさとは限らず, 各研究における点推定値の分散の逆数とする場合もある (丹後, 2002). また, 横軸はハザード比であるので, XAXIS ステートメントで TYPE = LOG を指定することで対数スケールとしている. ただし, 縦軸の一番下においてサブグループでない Overall の結果が示されており, 縦軸に指定している変数 Forestval の値の大きさ



を逆順にしなければならない。このプロットを PowerPoint などに貼り付けて、ハザード比とその 95%信頼区間やサブグループ名などの修飾を加えて、図のファイルとして用いることが多いのではないだろうか。V9.4 から ODS PowerPoint も利用できるようになり (吉田ら, 2015), 手動で PowerPoint に貼り付けなくても、プログラムを実行させることで PowerPoint に出力できるような環境が整ってきているといえる。

しかし、上記のような手順でなく、SAS ですべて操作して修飾を加えたプロットを作成したい場合、プロットの場合はどのような方法で行えばよいだろうか。

```
SAS プログラム proc sgplot data=forestdata noautolegend;
  refline ref / axis=y lineattrs=(thickness=30 color=cxf0f0f0);
  scatter y=forestval x=HazardRatio
    / errorbarattrs=(thickness=1 color=blue) markerattrs=(size=0)
      xerrorupper=HRUpperCL xerrorlower=HRLowerCL;
  bubble y=forestval x=HazardRatio size=N
    / bradiusmax=4.2 bradiusmin=0.2
      fillattrs=(color=blue) lineattrs=(color=blue);
  refline 1 / axis=x
      lineattrs=(pattern=shortdash) transparency=0.5;
  xaxis type=log minor min=0.1 max=5
      display=(nolabel) valueattrs=(size=7);
  yaxis display=none;
run;
```

#### 出力結果

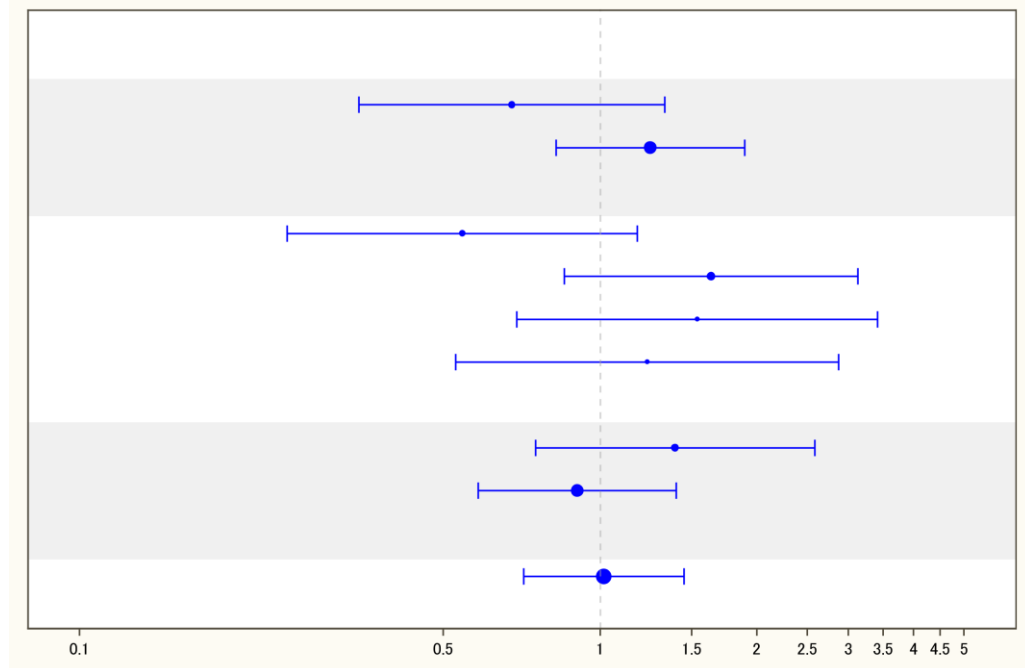


図 5 : SGPLOT プロシジャによる Forest プロットの作成 1

Forest プロットに修飾を加えるためには、前節のように Y2AXIS ステートメントを用いた記述を図 5 のプログラムに加えれば作成できそうである。図 5 は V9.3 の SGPLOT プロシジャで作成可能であるが、V9.4 においてさらに多くのステートメントおよびオプションが追加されている。

図 6 は、V9.4 から利用できるようになった SGPLOT プロシジャの機能を用いて、修飾を加えた Forest プロットを作成するプログラムである。DATTRMAP=ATTRMAP として指定しているデータセットの詳細は、付録 D を参照されたい。図 5 までのプログラムに比べて長文になった上、出力結果も他のプロットよりも大きく示すことを意図して、プログラムと出力結果を別々に示している。

---

```
proc sgplot data=forestdata noautolegend
    dattrmap=attrmap nowall noborder nocycleattrs;
    styleattrs axisextent=data;
    refline ref / axis=y lineattrs=(thickness=13 color=cxf0f0f0);
    scatter y=forestval x=HazardRatio
        / errorbarattrs=(thickness=1 color=blue)
            xerrorupper=HRUpperCL xerrorlower=HRLowerCL markerattrs=(size=0);
/* highlow y=forestval low=HRLowerCL high=HRUpperCL / lineattrs=(color=blue); */
    bubble y=forestval x=HazardRatio size=N
        / bradiusmax=4.2 bradiusmin=0.2
            fillattrs=(color=blue) lineattrs=(color=blue);
    scatter y=forestval x=HazardRatio / markerattrs=(size=0) x2axis;
    refline 1 / axis=x lineattrs=(pattern=shortdash) transparency=0.5;
    text x=x1 y=forestval text=text / position=center contributeoffsets=none;
    yaxistable subgroup / location=inside position=left
        textgroup=valtype labelattrs=(size=8) labeljustify=left
        titlejustify=left textgroupid=text indentweight=indent;
    yaxistable grp1 grp2 / location=inside position=left
        labelattrs=(size=8) valueattrs=(size=7)
        valuehalign=center valuejustify=right;
    yaxistable HR intpval / location=inside position=right
        labelattrs=(size=8) valueattrs=(size=7) valuehalign=center;
    yaxis reverse display=none offsetmin=0.1
        colorbands=odd colorbandsattrs=(transparency=1);
    xaxis type=log values=(0.2 0.5 2 5) minor display=(nolabel) valueattrs=(size=7);
    x2axis label='Hazard Ratio (95% CI)'
        display=(noline noticks novalues) labelattrs=(size=8);
run;
```

---

図 6: SGPLOT プロシジャによる Forest プロットの作成 1 (V9.4 から実行可能)

表 2 : Forest プロット作成時にデータセットに格納されている変数概要

ステートメント	役割	バージョンに関する注意
<b>STYLEATTRS</b>	<ul style="list-style-type: none"> <li>X 軸の範囲をデータに合わせて作成</li> </ul>	V9.4 以降で指定可能 (一部 TS1M3 以降でないと指定できないオプションあり)
<b>REFLINE</b>	<ul style="list-style-type: none"> <li>色を変えるサブグループにおいて色の縦方向の幅を指定</li> <li>ハザード比 = 1 に対する参照線を出力</li> </ul>	-
<b>SCATTER</b>	<ul style="list-style-type: none"> <li>点推定値とその信頼区間のグラフを作成 (HIGHLOW ステートメントでも可能)</li> <li>グラフの上部における Hazard Ratio (95% CI) のラベルを修飾するために指定 (X2AXIS ステートメントとあわせて指定)</li> </ul>	-
<b>BUBBLE</b>	<ul style="list-style-type: none"> <li>点推定値を表す●印を出力 (サブグループにおける例数に依存して大きさを変化)</li> </ul>	-
<b>TEXT</b>	<ul style="list-style-type: none"> <li>グラフの下部 (&lt;-- Test Better や Standard Better --&gt;) の修飾</li> </ul>	V9.4 以降で指定可能
<b>YAXISTABLE</b>	<ul style="list-style-type: none"> <li>左側のサブグループ名の装飾</li> <li>左側の各群のイベント割合の装飾</li> <li>右側の点推定値とその信頼区間および交互作用の p 値の装飾</li> </ul>	V9.4 以降で指定可能 (一部 TS1M3 以降でないと指定できないオプションあり)
<b>YAXIS</b>	<ul style="list-style-type: none"> <li>縦軸が変数 Forestval の値の大きさと逆順で出力</li> <li>左側のサブグループ名の装飾</li> </ul>	一部オプションは V9.4 以降で指定可能
<b>XAXIS</b>	<ul style="list-style-type: none"> <li>横軸を対数スケールで出力</li> </ul>	-
<b>X2AXIS</b>	<ul style="list-style-type: none"> <li>グラフの上部における Hazard Ratio (95% CI) のラベルを修飾する際の細かい設定</li> </ul>	-

図 6 で指定している SGPLOT プロシジャのステートメントについて、表 2 にそれぞれの役割と V9.4 でのみ実行可能な注意点について示している。図 7 は、図 6 のプログラムの実行結果である。図 5 と異なり、縦軸において指定している値の大きさと逆順で出力されている。これは YAXIS ステートメントで REVERSE をオプション指定しているためである。また、参考までに、BUBBLE ステートメントの代わりに、HIGHLOW ステートメントを指定した上で作成したプロットを図 8 に示している。図 8 では、信頼下限・上限において縦方向のヒゲを装飾していない。Motzer *et al.* (2015) による報告で示したハザード比の信頼区間のプロットは、図 8 に近いといえる。図 7 および図 8 を見ると、学術論文に投稿する際の図として文句のないクオリティのプロットを作成できたといえる。前節の Kaplan-Meier プロットの作成では XAXISTABLE ステートメントを利用したが、図 7 および図 8 では YAXISTABLE ステートメントを利用している。YAXISTABLE ステート

メントのオプションとして、LOCATION = INSIDE と指定しているのので、グラフの出力範囲内への修飾を行っていることになる。図 6 の SGPLOT プロシジャでは、YAXIS ステートメントにおいて DISPLAY = NONE を指定しているが、この記述を削除すれば、サブグループの項目名の左側に縦軸が出力され、グラフの出力範囲内への修飾であることを確認できる。

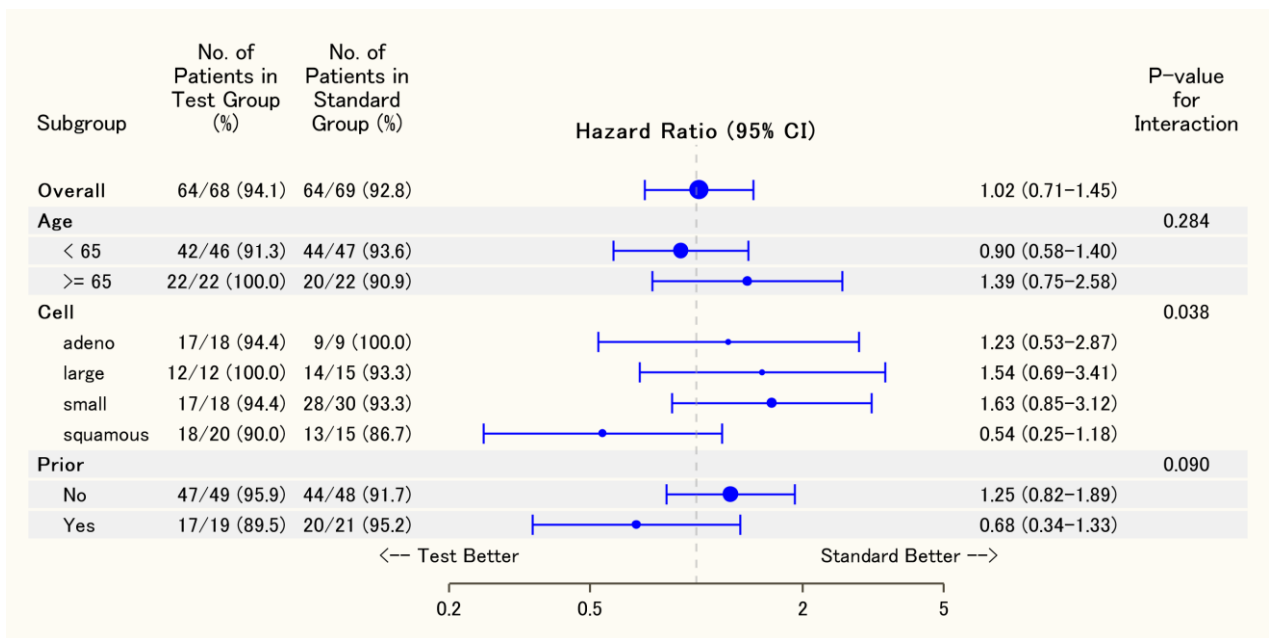


図 7 : SCATTER ステートメントを用いた Forest プロット

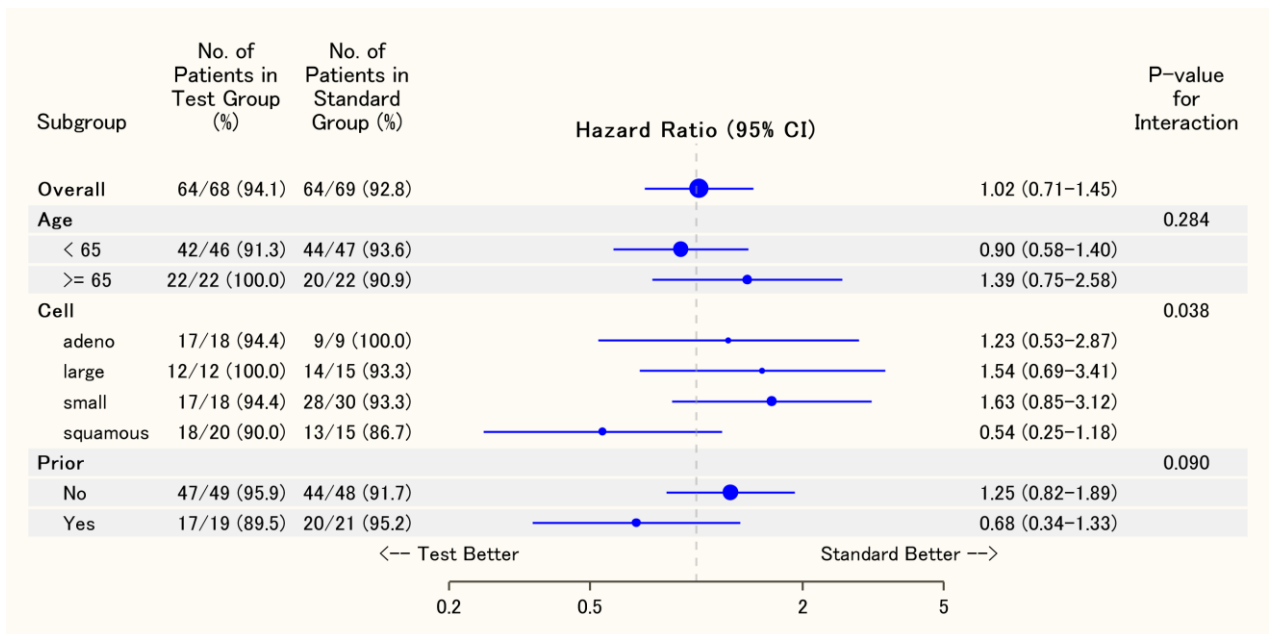


図 8 : HIGHLOW ステートメントを用いた Forest プロット

## 4 まとめ

本稿では、医薬品開発で用いられるよく用いられるグラフとして、Kaplan-Meier プロットと Forest プロットに焦点を当て、修飾を加えたプロットの作成方法を報告した。SAS による Kaplan-Meier プロットの作成に関してはいくつかの報告が行われているが、本稿では SAS/STAT12.1 以降、V9.4 以降で実施できる方法を示し、X2AXIS ステートメントの役割についても詳述した。なお、本稿で報告した Kaplan-Meier プロットの作成方法は、大橋ら (2016) の一部を引用している。また、Forest プロットの作成に関しては、Kaplan-Meier プロットと異なり SAS による実装報告が少ないトピックであったため、SGPLOT プロシジャを実行するにあたって、どのようなデータセットを用意しなければならないかについても詳述した。グラフの下段に数値出力を行う Kaplan-Meier プロットでは XAXISTABLE ステートメントによる実行手順を示したが、グラフの左右に数値出力を行う Forest プロットの作成では YAXISTABLE ステートメントを用いた。図 7 および図 8 のグラフは、見栄えとしてはグラフの出力範囲外への出力といえるが、YAXISTABLE ステートメントによるプログラム上は LOCATION = INSIDE を指定することで、グラフの出力範囲内への出力をしている。本稿で取りあげた Forest プロットの作成手順は、比例ハザードモデルによるハザード比の出力に限らず、ロジスティックモデルによるオッズ比の出力にも応用できる。なお、本稿で取りあげた生存時間データに対する Forest プロットの作成方法は、付録 B に SAS プログラムを詳細している。本稿の内容が国内外の多くの臨床試験の報告にお役に立てれば幸いである。

## 参考文献

- [1] Borghaei H, Paz-Ares L, Horn L, *et al.* Nivolumab versus Docetaxel in Advanced Nonsquamous Non-Small-Cell Lung Cancer. *New England Journal of Medicine*. **373**:1627–1639, 2015.
- [2] Hebbbar P. Lost in the Forest Plot? Follow the GTL AXISTABLE Road! *Proceedings of the SAS Global Forum*. SAS Institute Inc., Cary, NC, 2015. Available at <http://support.sas.com/resources/papers/proceedings15/SAS1748-2015.pdf>.
- [3] Matange S. Graphs are Easy with SAS® 9.4. *Proceedings of the SAS Global Forum*. SAS Institute Inc., Cary, NC, 2015. Available at <http://support.sas.com/resources/papers/proceedings15/SAS1780-2015.pdf>.
- [4] Motzer RJ, Escudier B, McDermott DF, *et al.* Nivolumab versus Everolimus in Advanced Renal-Cell Carcinoma. *New England Journal of Medicine*. **373**:1803–1813, 2015.
- [5] SAS Institute Inc. *SAS/STAT(R) 12.1 User's Guide*. SAS Institute Inc., Cary, NC, 2012.
- [6] SAS Institute Inc. *SAS(R) 9.3 ODS Graphics: Procedures Guide* (3rd edn.). SAS Institute Inc., Cary, NC, 2012.
- [7] SAS Institute Inc. *SAS(R) 9.4 ODS Graphics: Procedures Guide* (4th edn.). SAS Institute Inc., Cary, NC, 2015.
- [8] Turner NC, Ro J, André F, *et al.* Palbociclib in Hormone-Receptor-Positive Advanced Breast Cancer. *New England Journal of Medicine*. **373**:209–219, 2015.
- [9] 魚住龍史, 浜田知久馬. SG (Statistical Graphics) Procedures による Kaplan-Meier プロットの作成. SAS ユーザー総会 論文集 2011, 185–199.
- [10] 魚住龍史, 浜田知久馬. がん臨床試験における腫瘍縮小効果の検討に有用なグラフの作成 -SGPLOT プロシジャの最新機能を活用-. SAS ユーザー総会 論文集 2012, 151–165.

- [11] 大橋靖雄・浜田知久馬. 生存時間解析 –SAS による生物統計. 東京大学出版会, 1995.
- [12] 大橋靖雄・浜田知久馬・魚住龍史. 生存時間解析 応用編 –SAS による生物統計. 東京大学出版会, 2016.
- [13] 高浪洋平. SG プロシジャと GTL によるグラフの作成と ODS PDF による統合解析帳票の作成 ～TQT 試験における活用事例～. SAS ユーザー総会 論文集 2011, 201-219.
- [14] 高浪洋平. SAS と HTML アプリケーションによる CDISC ADaM 形式の解析用データセットを用いた有害事象の解析帳票・グラフ簡易作成ツールの開発事例. SAS ユーザー総会 論文集 2012, 185-205.
- [15] 丹後俊郎. メタ・アナリシス入門 –エビデンスの統合をめざす統計手法. 朝倉書店, 2002.
- [16] 豊泉樹一郎・財前政美・北西由武・都地昭夫. ODS GRAPHICS を用いた臨床試験データの可視化への挑戦. SAS ユーザー総会 論文集 2014, 307-323.
- [17] 長島健悟, 佐藤泰憲. Kaplan-Meier プロットに付加情報を追加するマクロの作成. SAS ユーザー総会 論文集 2010, 285-294.
- [18] 西本尚樹. データの可視化を加速する SAS/JMP のグラフ新機能 –ODS Graphics と JMP グラフビルダー–. SAS ユーザー総会 2013.
- [19] 平井隆幸・吉田早織・叶健・魚住龍史. ベクター形式を用いたグラフの作成と有用性. SAS ユーザー総会 論文集 2015, 303-310.
- [20] 堀田真一. Let's make Forest Plot by SAS. SAS ユーザー総会 論文集 2014, 675-679.
- [21] 吉田早織・平井隆幸・叶健・魚住龍史. ODS POWERPOINT の活用: SAS から Microsoft PowerPoint へのエクスポート. SAS ユーザー総会 論文集 2015, 295-302.

## 連絡先

E-mail : [uozumi@kuhp.kyoto-u.ac.jp](mailto:uozumi@kuhp.kyoto-u.ac.jp)

付録 A: 図 2 で省略したデータハンドリングの詳細

```

*=====*;
* リスク集合出力時間間隔 ;

%let TimeInterval=200;
*=====*;
* リスク集合の出力最大時間;

%let TimeMax=1000;
*=====*;
* フォーマット ;

proc format;
  value Cellf 1='adeno' 2='large' 3='small' 4='squamous';
run;
*=====*;
* データセットSurvivalplotからSurvivalplot2の作成 ;

data form;
do StratumNum=1 to 4; do time=0 to &TimeMax by &TimeInterval; output; end; end;
data atrisk0; merge form Survivalplot; by StratumNum time;
data Survivalplot2; set atrisk0;
  if AtRisk=. then do; AtRisk=0; tAtRisk=time; end;
run;

```

付録 B : Forest プロット作成時のデータセット FORESTDATA の作成プログラム

```

*=====*;
* フォーマット ;

proc format;
  value Cellf 1='adeno' 2='large' 3='small' 4='squamous';
  value Priorf 10='Yes' 0='No'; value AgeCf 1='< 65' 2='>= 65';
  value Allf 1='Overall';
run;
*=====*;
* MACRO: ForestTTE ;
*=====*;
* サブグループ解析およびForestプロット用のデータセット作成
*=====*;
* 引数の説明 ;
* data : 解析対象データセット ;
* sub : サブグループを表すカテゴリ変数 ;
* varname : Forestプロットに示したいサブグループ名 ;
* format : サブグループのカテゴリの (事前に定義した) フォーマット名 ;
* grp : 群を表す変数 (grp = 2 を対照群とする) ;
* tte : 生存時間の評価項目を表す変数 ;
* censor : 打ち切りを表す変数 ;
* censorval : 打ち切り値 ;
* num : サブグループに付与した識別値 ;
* out : 出力データセット名 ;
*=====*;

%macro ForestTTE (data, sub, varname, format, grp, tte, censor, censorval, num, out) ;

proc sort data=&data out=data00; by &sub;
run;
*-----*;
* 各群のイベント割合の計算 ;

ods listing close; ods output CensoredSummary=event00;
proc lifetest data=data00;
  time &tte*&censor(&censorval); strata &grp; by &sub;
run;
ods listing;

data event00;set event00; where &grp in(1,2);
  N=Total; Event=Failed; prop=Failed/Total;
  pevent=trim(left(put(cat(trim(left(Event)),"/",trim(left(N)), " ("),
    trim(left(put(round(prop*100,1e-1),8.1))),")"),20.)));
  keep &grp &sub pevent;
run;

proc transpose data=event00 out=event00; id &grp; by &sub; var pevent;
run;

data event00;set event00; rename _1=grp1 _2=grp2; drop _NAME_;
run;
*-----*;
* 各サブグループにおける例数の計算 ;

ods listing close; ods output CensoredSummary=total00;
proc lifetest data=data00; time &tte*&censor(&censorval); by &sub;
run;
ods listing;

data total00;set total00; N=Total; Event=Failed; keep &sub N Event;
run;
*-----*;
* ハザード比の計算 ;

ods listing close;
ods output ParameterEstimates=HR00;
proc phreg data=data00;
  class &grp(ref="2");

```

```

model &tte*&cursor(&cursorval) = &grp / rl ties=breslow; by &sub;
run;
ods listing;

data HR00;set HR00;
  where Parameter="&grp"; keep &sub HazardRatio HRLowerCL HRUpperCL HR;
  HR=trim(left(put(cat(trim(left(put(round(HazardRatio,1e-2),8.2))),
    "(",trim(left(put(round(HRLowerCL,1e-2),8.2))),"-",
    trim(left(put(round(HRUpperCL,1e-2),8.2))),")"),20.))););
run;
*-----;
* 交互作用のp値の計算 ;

ods listing close;
/* SAS/STAT14.1に対応, STAT/STAT13.2以前ではModelANOVAの代わりにType3を指定 */
ods output ModelANOVA=int00;
proc phreg data=data00;
  class &grp(ref="2") &sub;
  model &tte*&cursor(&cursorval) = &grp &sub &grp*&sub / rl ties=breslow;
run;
ods listing;

data int00; set int00; where Effect="&grp*&sub";
  intpval=trim(left(put(round(ProbChiSq,1e-3),pvalue8.3)));
  subgroup=&varname; subgroupc=&num; keep subgroup subgroupc intpval;
run;
*-----;
* すべての計算結果を集約 ;

data surv00; merge total00 event00 HR00; by &sub; &grp=1;
data surv00;set surv00; where &sub ^=.;
run;

data surv00;set surv00;
  subgroupc=&num; subgroup=put(&sub,&format); catc=&sub;
run;

data item00;set int00; valtype=1;
data cat00;set surv00; valtype=2; if &varname="Overall" then valtype=1;
data &out; length subgroup $ 50; set item00 cat00; drop &sub &grp;
run;

%mend ForestTTE;
*-----*;
* サブグループごとに解析実行 ;

%ForestTTE(VAlung,all,"Overall",Allf.,Therapy2,Time,Censor,1,0,all);
%ForestTTE(VAlung,AgeC,"Age",AgeCf.,Therapy2,Time,Censor,1,1,age);
%ForestTTE(VAlung,CellC,"Cell",Cellf.,Therapy2,Time,Censor,1,2,cell);
%ForestTTE(VAlung,Prior,"Prior",Priorf.,Therapy2,Time,Censor,1,3,prior);
*-----*;
* サブグループ解析結果の集約 ;

data forestdata00; length subgroup $ 50; set all age cell prior;
run;
*-----*;
* 全体の解析結果をマクロで実施した結果の調整 ;

data forestdata00;set forestdata00; if subgroupc=0 and N=. then delete;
run;
*-----*;
* インデントの設定 ;

data forestdata00;set forestdata00;
  forestval=_n_;
  if mod(subgroupc,2) ^=0 then ref=forestval;
  indent=1; if subgroup="Overall" or intpval ^= "" then indent=0;

```



```

run;

ods listing close; ods output summary=forestvalmax;
proc means data=forestdata00 max; var forestval;
run;
ods listing;
*=====*;
* 横軸の説明テキストの設定 ;

data axis; length text $ 50; set forestvalmax;
  forestval=forestval_Max+1; valtype=2;
  xl=0.2; yl=forestval; text='<-- Test Better'; output;
  xl=4; yl=forestval; text='Standard Better -->'; output;
run;

data forestdata;set forestdata00 axis;
run;
*=====*;
* 変数ラベルの指定 ;

data forestdata;set forestdata;
  label grp1="No. of Patients in Test Group (%)"
        grp2="No. of Patients in Standard Group (%)"
        intpval="P-value for Interaction" subgroup="Subgroup" HR=" ";
run;
*=====*;
* 付録DのデータセットATTRMAP作成プログラム ;

data attrmap; length textweight $10;
id='text'; value='1'; textcolor='Black'; textsize=7; textweight='bold'; output;
id='text'; value='2'; textcolor='Black'; textsize=5; textweight='normal'; output;
run;

```

付録 C : Forest プロット作成時のデータセット FORESTDATA

Subgroup	Int pval	Type	N	Grp1	Grp2	Hazard Ratio	HR Lower CL	HR Upper CL	HR	Indent	Forest val	Ref
Overall		1	137	64/68 (94.1)	64/69 (92.8)	1.016	0.713	1.448	1.02 (0.71-1.45)	0	1	.
Age < 65	0.284	1	.							0	2	2
>= 65		2	93	42/46 (91.3)	44/47 (93.6)	0.903	0.583	1.4	0.90 (0.58-1.40)	1	3	3
Cell adeno	0.038	2	44	22/22 (100.0)	20/22 (90.9)	1.393	0.751	2.582	1.39 (0.75-2.58)	1	4	4
large		1	.							0	5	.
small		2	27	17/18 (94.4)	9/9 (100.0)	1.231	0.528	2.872	1.23 (0.53-2.87)	1	6	.
squamous		2	27	12/12 (100.0)	14/15 (93.3)	1.536	0.692	3.409	1.54 (0.69-3.41)	1	7	.
Prior No	0.09	2	48	17/18 (94.4)	28/30 (93.3)	1.633	0.854	3.123	1.63 (0.85-3.12)	1	8	.
Yes		2	35	18/20 (90.0)	13/15 (86.7)	0.543	0.25	1.179	0.54 (0.25-1.18)	1	9	.
Prior No		1	.							0	10	10
Yes		2	97	47/49 (95.9)	44/48 (91.7)	1.248	0.822	1.894	1.25 (0.82-1.89)	1	11	11
		2	40	17/19 (89.5)	20/21 (95.2)	0.676	0.344	1.33	0.68 (0.34-1.33)	1	12	12

付録 D : SGPLOT プロシジャ (図 6) で用いるデータセット ATTRMAP

text	forestval_Max	forestval	valtype	xl	yl
<-- Test Better	12	13	2	0.2	13
Standard Better -->	12	13	2	4	13