

生存分析

基本概念：生存曲線

臨床や動物実験において帰結 (outcome) が生存時間ということは良くあります。その場合の調査目的は治療が生存時間に影響を及ぼしたかどうかを判定することにあります。Prism は Kaplan, Meier が開発した product limit method を用いて生存曲線を作成、それをログランク (logrank) 検定と Gehan-Wilcoxon 検定を用いて比較します。

打ち切りデータ

生存曲線を描くことはそれほど簡単なことではありません。難しい点は個々の対象の生存時間がきちっとわかっているケースが少ないという点にあります。

対象の中には調査終了時点で生存しているものがあります。どれだけ生きてきたかはわかって、まだどれだけ生存を続けるかは知る由もありません。

調査からのドロップアウト — 抜ける理由としてはいろいろとあるでしょう。別の町に移転したとか、プロトコル上許されない医療を受けなくなったとかが考えられます。プロトコルに従ってどれだけの期間生きてきたかはわかりますが、その後どれだけ生存を続けたかはわかりません(あるいはわかっているプロトコルから離脱しているためにその情報を利用することができないこともあります)。いずれにしてもこれらの患者に関する情報は censor (打ち切り) されたと言います。

これら censor された観測データを分析から消去したいとは思わないでしょう。必要なことはその旨適切に説明することです。“censored” という用語は対象が何か不適切なことをしたかのような印象を与えます。しかしそうではありません。“censored” という言葉は単に、ある時点以降の生存について情報がない、あるいは利用できないことを意味するものです。Prism は生存曲線を作成し比較する際に censor されたデータを適切に処置します。

「生存」という用語について

生存曲線というのは少々制限のきつい表現です。帰結 (outcome) は死に限ったことではなく、対象ごとに一度だけ起る任意の終端点であって良いからです。それは移植血管の閉塞、腫瘍の最初の転移、移植された腎臓の拒絶反応といったものでも良いのです。もちろんこのような暗い結末に限ったものではありません。腎臓機能の回復、病院からの退院、卒業といった事象であっても構いません。

別の手法

ある種の生存データは非線形回帰の手法を用いた方がうまく分析できます。例えば細胞の生存曲線の分析に際し、生存率 (Y) を放射線被爆量 (X) の関数でプロットした場合、本セクション

で記述する手法は使用すべきではありません。本セクションで説明する分析法は X が時間であり、かつ個々の対象の生存時間がわかっている場合にのみ有効となるものです。

比例ハザード回帰

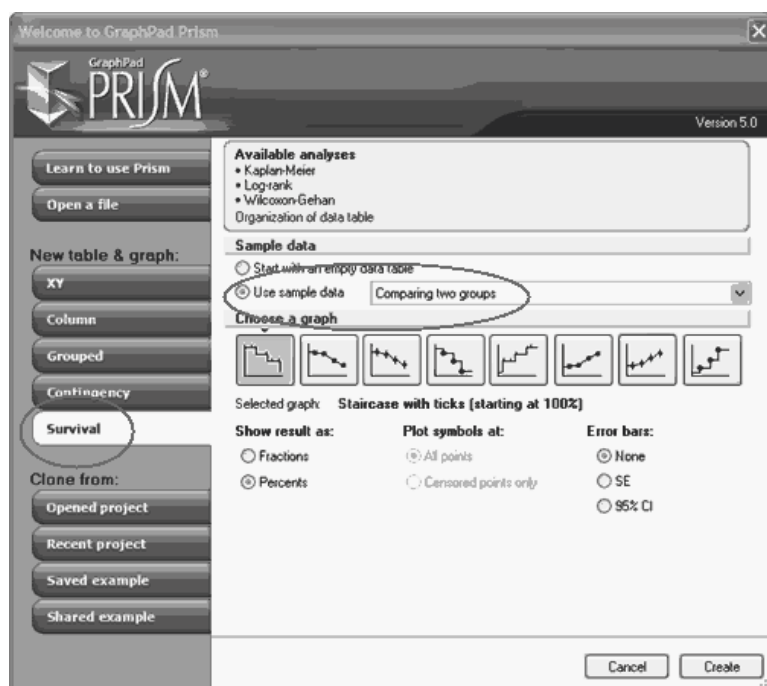
Prism に組み込まれている分析法では複数グループの生存曲線を比較することができます。しかしこれらの手法 (logrank 検定、Gehan-Breslow-Wilcoxon 検定) は、対象間に対応関係があったり、年齢、性別等によって調整を行いたいようなケースを扱うことができません。この種の問題に対しては比例ハザード回帰 (proportional hazards regression) が有効ですが、Prism ではサポートされていません。

操作手順：生存分析

1. 生存表の作成

下に示すダイアログ中で "Survival" タブを選択、目的に合致したグラフを選択します。

データ入力の準備が整っていないときにはサンプルデータを使用してください。データセットとしてはどれでも構いません。



2. 生存時間の入力

次のガイドラインに従って対象ごとのデータを表の各行に入力します。

死(またはトラッキングしている事象)、または censor までの時間を X カラムに入力します。時間の単位としては都合の良いもの(日、月、等)を適宜使用できます。時刻 0 は特定の日付である必要はありません。それは個々の対象が調査に加わった日時として定義されるため、対象によって日付が異なっても構いません。臨床的な調査の場合には患者を募集するのに時間を要するため、時刻 0 が数年にまたがることもあります。時間の値は数値で入力してください。日付を直接指定することはできません。

- ・ 個々の対象を特定するための行タイトルを入力することもできます。
- ・ X カラムに示される時点で対象が死亡(あるいは特定の事象が発生)した場合には、該当する Y カラムに”1”を入力します。対象データが打切られた(censor)場合には”0”を入力します。
- ・ 治療グループごとに異なる Y カラムを使用してください。最初のグループに属する対象の X データはテーブルの先頭から順に入力して行きます。その際、Y コードは最初の Y カラムに入力します。第2のグループに属する対象の X データは第 1 のグループに続く形で入力します(X データはソートしてある必要はありませんし、同じ値が何回も出現しても構いません)。対応する Y コードは2番目の Y カラムに入力します。以下の例の場合、グループ A のデータは最初の 10 行に入力され、グループ B のデータは行 11 から始まっています。

Table format:		X	A	B
Survival		Days after randomization	Control	Treated
	ID	X	Y	Y
1	AB	34	1	
2	GT	66	1	
3	RF	64	0	
4	ED	89	1	
5	CD	98	1	
6	TT	111	1	
7	RV	123	1	
8	TV	145	1	
9	VD	134	1	
10	BM	145	0	
11	UJ	88		1
12	UV	143		1
13	IT	78		1
14	TO	111		0
15	AT	95		0
16	TU	134		1
17	XX	167		1
18	XY	198		1
19	XO	211		1
20	HO	234		1

- ・ 治療グループのデータが本質的に順序付けられている場合(恐らく用量順)、データ入力に際してもその順序をキープするようにしてください。またカラム A から B, C への遷移は治療グループの自然なオーダに従うようにしてください。治療グループにそのような順序性がない場合にはどう並べ変えても構いません。

生存分析のデータ入力はトリッキーな場合があります。Q & A、臨床研究事例、動物実験事例を参照ください。

3. グラフと結果の表示

データ入力が終了したら「グラフ」フォルダを選択、作成された生存曲線が確認できます。また「結果」フォルダ中には曲線間の比較を行うログランク検定の結果も収納されています(データセットが複数ある場合)。