

## 【統計一口メモ 第38話】

### ＜ログランク検定とは？＞

名古屋市立大学大学院医学研究科 非常勤講師 薬学博士 松本一彦

ログランク検定 log rank test は、生存時間解析で一般化 Wilcoxon 検定、ハザード比と一緒に生存時間解析で使われています。したがって、がん関連の研究をしている基礎研究者と臨床家には馴染み深い検定ですが、いったい何を検定しているのでしょうか。

例題1. ステージ3と4のリンパ腫患者の生存時間に差はあるだろうか？

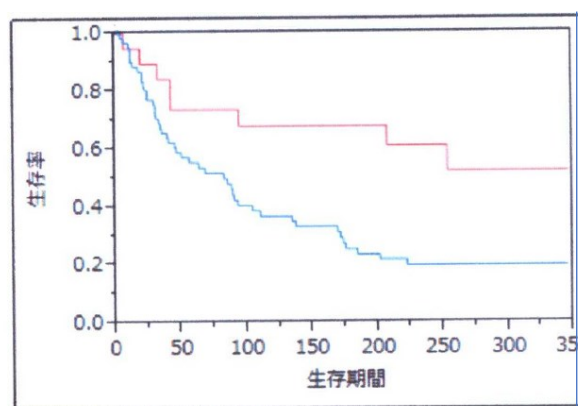
【医学研究のための統計的方法、P.Armitage, G.Berry. サイエントリスト社】<sup>1)</sup>

表1. 腫瘍のステージに応じたリンパ腫患者の死亡/打ち切り日

	生存日	n数
ステージ 3	6 19 32 42 42 43* 94 126* 169* 207 211* 227* 253 255* 270* 310* 316* 335* 346*	19
ステージ 4	4 6 10 11 11 11 13 17 20 20 21 22 24 24 29 30 30 31 33 34 35 39 40 41* 43* 45 46 50 56 61* 61* 63 68 82 85 88 89 90 93 104 110 134 137 160* 169 171 173 175 184 201 222 235* 247* 260* 284* 290* 291* 302* 304* 341* 345*	61

\* 途中打ち切り例

下図は生存時間解析で見る Kaplan-Meier 曲線です。



#### §1. ログランク検定とは

表1のデータおよび図は、リンパ腫患者の生存時間を示しています。赤線(上)がステージ3群で青線(下)がステージ4群です。この生存曲線で2群間に差があるか否かを比較検討するにはどうしたらよいのでしょうか？

その1:最終日の時点で生存率を比較する→それまでの経過を無視することになり、データを有効に使っていない。

その2:半数が死亡するまでの時間すなわち「50%死亡率」を比較する→半数を超えてからの死亡率が無視される。

その3:全期間を通して得られた情報を使って解析する→それが、**ログランク検定**です。

概念的には、カプランマイヤー曲線を描く各時点で求めたログランク統計量を全時点にまとめて2群間で $\chi^2$ 検定をする手法と言えます。帰無仮説は「2群間の生存曲線は等しい」となります。単純な例で示すと次のようになります<sup>2)</sup>。

各群3例 時点 群	ログランク・重み付きWilcoxonの計算手順											
	1日		2日		3日		4日		5日		6日	
	1群	2群	1群	2群	1群	2群	1群	2群	1群	2群	1群	2群
生存	2	3	1	3	1	2	0	2	0	1	0	0
死亡	1	0	1	0	0	1	1	0	0	1	0	1
計	3	3	2	3	1	3	1	2	0	2	0	1

浜田知久馬「学会・論文発表のための統計学」真興交易(株)図書出版部

カプラン・マイヤー曲線の全体の分散を使った $\chi^2$ 検定で、どの時点においても死亡率が同じ(重みづけなし)として検定する手法です。

それでは、実際のデータから「ログランク統計量」の計算を見ていきましょう。

## § 2. ログランク統計量を求める手順(エクセル解:Armitage 法<sup>1)</sup>)

手順1. 患者毎に打切/死亡と生存期間を求める

表2	打切=0 死亡=1		
患者番号	打切/死亡	生存期間	情報1
1	1	4	ステージ4
2	1	6	ステージ3
3	1	6	ステージ4
4	1	10	ステージ4
60	1	207	ステージ3
61	0	211	ステージ3
62	1	222	ステージ4
63	0	227	ステージ3
64	0	235	ステージ4
65	0	247	ステージ4
66	1	253	ステージ3

255 日以降は両群ともに打切りデータなのでカットする

手順2. 死亡発生日毎の個体数を求める

表3	個体数		合計(n <sub>i</sub> )
	ステージ3	ステージ4	
死亡発生日	n <sub>3</sub>	n <sub>4</sub>	合計(n <sub>i</sub> )
4	19	61	80
6	19	60	79
10	18	59	77
11	18	58	76
207	10	11	21
222	8	11	19
253	7	8	15

ステージ4で 4 日目に死亡が見られた。したがって、それまで、生存していた個体数は両群合わせて総数80となる。6 日目は両群で死亡が見られたが、ステージ3では初めてなので、ここではカウントされず、ステージ4のn数が 61 から 60 になる。

手順3. 死亡発生日毎に死亡数を求める

表4	死亡数		合計(d <sub>i</sub> )
	ステージ3	ステージ4	
死亡発生日	d <sub>3</sub>	d <sub>4</sub>	合計(d <sub>i</sub> )
4	0	1	1
6	1	1	2
10	0	1	1
11	0	3	3
207	1	0	1
222	0	1	1
253	1	0	1
合計	8	46	54
	03	04	

死亡発生日毎に各ステージにおける死亡数をカウントし合計する。  
 ステージ毎の死亡数の合計を 03、04 とする。

手順4. 死亡発生日毎に期待値と分散を求める

表5 死亡発生日	期待値		分散
	E(d3)	E(d4)	Var
4	0.2375	0.7625	0.1811
6	0.4810	1.5190	0.3606
10	0.2338	0.7662	0.1791
11	0.7105	2.2895	0.5278
207	0.4762	0.5238	0.2494
222	0.4211	0.5789	0.2438
253	0.4667	0.5333	0.2489
合計	16.6609	37.3391	11.2380
	E3	E4	V3

① 期待値を求める: 死亡発生日の 4 日目のステージ3の場合、(個体数 x 合計死亡数)/合計個体数=(19 x 1)/80=0.2375となる。

② 分散を求める: 死亡発生日の 4 日目の場合、{合計死亡数 x (合計個体数-1) x ステージ3個体数 x ステージ4個体数} / {合計個体数の二乗 x (合計個体数-1)} =

$$\frac{1 \times (80-1) \times 19 \times 61}{80^2 \times (80-1)}$$

$$=0.1811$$

ステージ毎の期待値の合計を E3、E4 分散の合計を V3 とする。

手順5. ログランク検定計算法

① ログランク統計量を求める:

$$\frac{(O3 - E3)^2}{V3} = \frac{(8 - 16.66)^2}{11.24} = 6.672$$

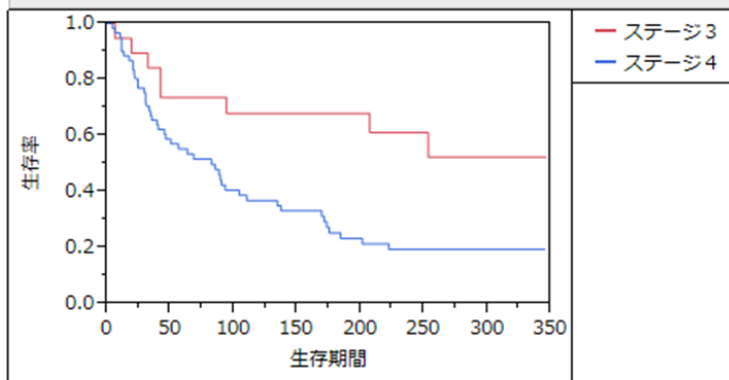
② p値を求める:

$$p = \text{CHIDIST}(6.672, 1) = 0.00979$$

### § 3. JMP での解析

#### Kaplan-Meier法によるあてはめ

##### 生存分析プロット



イベントまでの時間: 生存期間

打ち切り変数: 打切

打ち切りの値 1

グループ変数: 情報 1

##### 要約

グループ	故障数	打ち切り数	平均	標準誤差
ステージ3	8	11	181.701 バイアスあり	24.4974
ステージ4	46	15	101.644 バイアスあり	10.7041
組み合わせ	54	26	125.159 バイアスあり	11.2654

##### 分位点

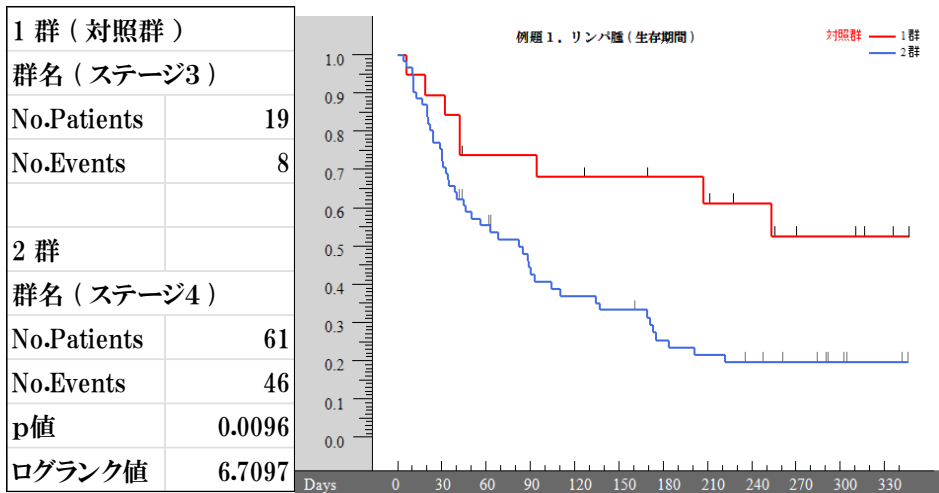
グループ	中央値時間	下側95%	上側95%	25%寿命	75%寿命
ステージ3	.	42	.	42	.
ステージ4	82	40	110	30	184
組み合わせ	90	50	171	31.5	.

##### グループ間での検定

検定	カイ2乗	自由度	p値(Prob>ChiSq)
ログランク	6.7097	1	0.0096*
Wilcoxon	5.0998	1	0.0239*

ログランク値は 6.7097 でエクセル解は 6.672、また p 値は 0.0096 でエクセル解の 0.0098 といずれも若干異なった値を示しています。

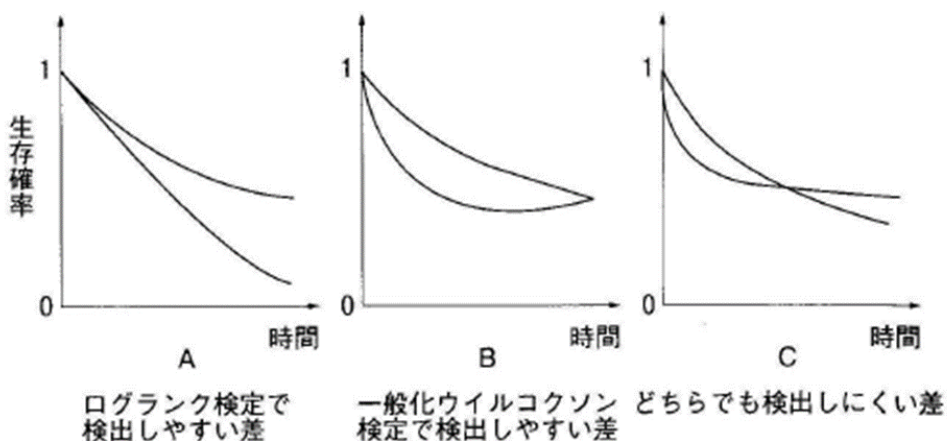
#### § 4. Pharmaco での解析<sup>3)</sup>



ログランク値およびp値はJMPと同値を示しています。

#### § 5. ログランク検定と一般化 Wilcoxon 検定

JMP および Pharmaco 解析では、生存時間解析の結果にログランク検定と一般化 Wilcoxon 検定が併記されています。初めて見ると、どちらを使えばいいのか迷います。実は計算過程で重み付けが両方で異なります。浜田先生は著書の中で次のように述べています<sup>2)</sup>。「一般化 Wilcoxon 検定は、「観察度数－期待度数」に時間 $t_f$ での全群のリスク数  $n_f$  で重み付けをする。つまり、リスクのある人がより多い生存曲線のはじめの方の情報をより重視し、遅い死亡よりも早い死亡の重みを大きくする。この重み付けは、生存に対する治療の効果が治療の初期にもっとも大きく、時間が経過するにつれて効果が弱まっていく傾向にあるかどうかを検討するのに向いている」



一般化 Wilcoxon とよばれる理由は、打切りがない場合、この検定が Wilcoxon 順位和検定に一致するからと言われています。なお、ログランク検定は Mantel-Haenzel 検定、Cox-Mantel-Haenzel 検定と呼ばれることがあります。

- 1) P.Armitage, G.Berry. 医学研究のための統計的方法、サイエンティスト社 2001 年
- 2) 浜田知久馬、新版「学会・論文発表のための統計学」真興交易(株)図書出版部 2016 年
- 3) Pharmaco 工房 : <https://pharmaco.club/>