

胃癌手術で D1 と D2 はどちらがよいか？

1) 胃癌手術で D1 と D2 はどちらがよいか？

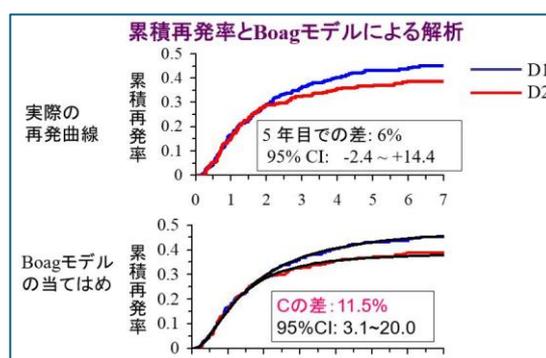
オランダの胃癌グループは、15 年経って始めて D2 手術が D1 手術より再発が少ないことを認めました (Lancet Oncol 2010; 11:439-49).

この結果は、Boag モデルを使えば 10 年前にわかっていました。

2) 癌手術の方法が大論争を巻き起こした

	パラメータ	D1	D2
全治例	全治率	45.4%	56.95
	95% 信頼区間	39.4~51.3%	50.9 ~ 62.9%
	全治率の差	11.5%	
	95%信頼区間	3.1 ~ 20.0%	
非治癒例	対数癌死時間の平均	3.145	2.945
	95% 信頼区間	2.978 ~ 3.313	2.791 ~ 3.098
	平均の差	0.201	
	95% 信頼区間	-0.027 ~ 0.428	
	対数癌死時間のSD	0.903	0.746
	95%信頼区間	0.788 ~ 1.018	0.632 ~ 0.860
	SDの差	0.157	
95%信頼区間	-0.05 ~ 0.318		

時間の単位は月1999年11月解析, 追跡期間136年に限定



オランダの D1 と D2 郭清手術における Boag パラメータの推定 (1999)

元来胃癌の手術は日本の外科医のお家芸であり、欧米の追随を許さなかった。そこで、オランダの胃癌グループは日本のがんセンターから胃癌専門の外科医を招き、日本で確立した広範リンパ節郭清 D2 と、欧米で常用されている D1 郭清の効果を RCT で比較した。

治験では、日本人の指導医の下ではあったが、D2 手術により手術死、合併症、在院日数共に有意に増加した。一方、log-rank 検定やハザード比では両群に全く差がみられなかった。

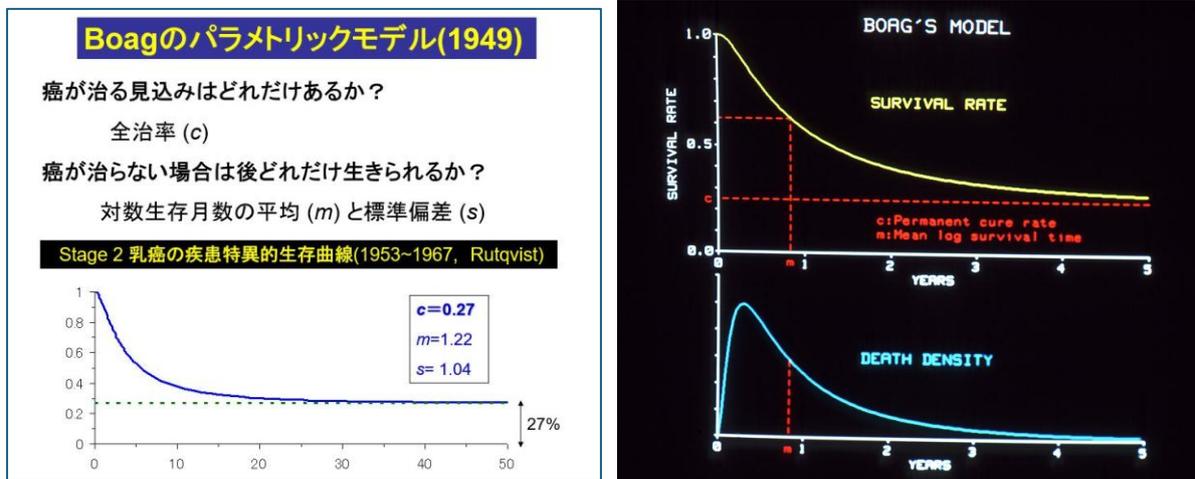
その結果、彼らは「D2 は常用すべきでない」と日本外科医にとってはショッキングな結論を 1999 年の N Engl J Med に発表した。

そこで私は発表の基となったデータをオランダグループから取り寄せ、Boag モデルで再解析した。その結果、両群の累積再発率曲線は時間と共に開いて、末広りのパターンである。D2 は D1 より全治率 c が約 11% 高くその差は有意であった。この結果を担当教授に報告し、このままでは患者が不利な治療を受ける恐れがあるので、論文として発表したいと許可を求めたが、拒絶された。

ところが 2010 年彼らは 15 年追跡結果を Lancet Oncol に発表し、D2 が D1 に比べて癌再発を 11% 有意に減らすことを認めた。こうして「D2 を支持しない」から一転して「D2 を推奨する」に転換した。それにしても 15 年も経たねば本当のことがわからないのは、一体どのような統計的検定を用いたのが懸念された。それ以前の論文をもう一度読むと、彼らは比例ハザード性を検証しないまま、logrank 検定を用いたと述べている。

3) Boag モデルで何ができるか？

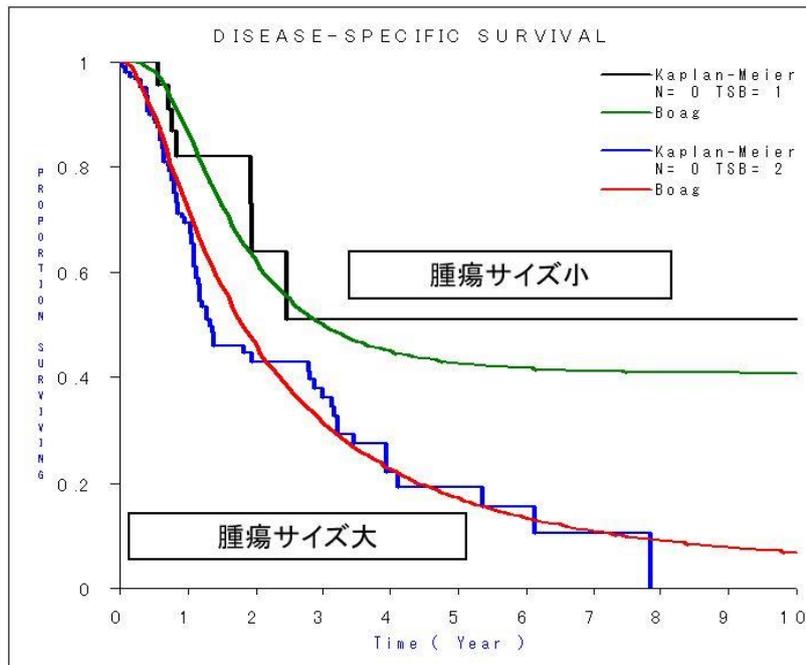
「その治療で癌が治るチャンスはどれだけあるのか？ もし治らないとすればどれだけ長く生きられるのか？」という情報を患者さんに伝えられます。「更に平均余命がどれだけか」を推測できます。



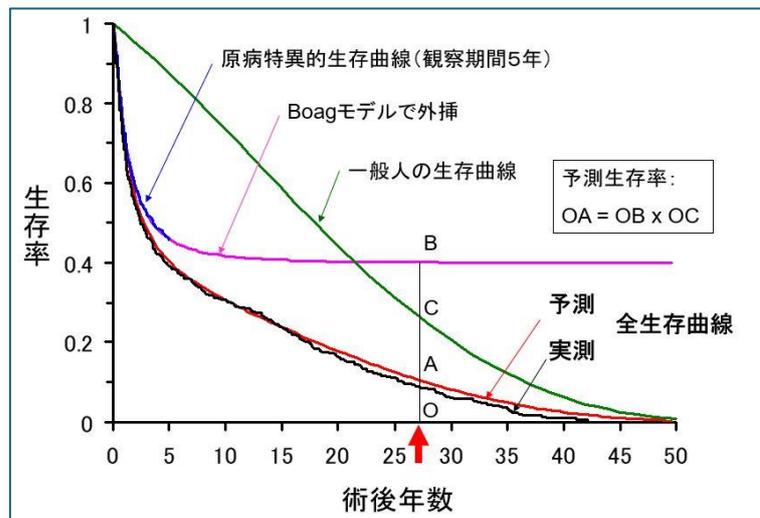
1949年 Boag はこのモデルを発表しました。彼は癌患者にとっての最大の関心事が何かを考えました。その一つは「自分の癌が治る見込みはどれだけあるのか」で、これが全治率 c というパラメータで表されます。もう一つの関心事は「もし癌が治らないとすれば、後どれだけ生きられるのか」です。これは2つのパラメータ、即ち余命の平均 m と標準偏差 s で、その対数値が正規分布をすると仮定しました。こうすると生存曲線の特徴が3つの数値で表されます。その解釈はきわめて簡単明瞭です。 c は全治患者の割合、 m は非治癒患者の平均余命、 s は余命のばらつきです。これが Boag のパラメトリックモデルです。さて乳癌では他の癌に比べて再発死亡までの期間が長いですが、全治する例はほとんどないという研究者もいました。つまり c は 0 です。これに対して手術で全治する例は確かにいるという研究者もあり、この議論は今も続いています。

1983年ノルウェーの Rutqvist が約 5000 人のステージ 2 乳癌登録患者のデータにこの Boag モデルを当てはめて c, m, s を求めました。それがこの値です。逆にこの値から生存曲線全体が再現できます。但し疾患特異的生存曲線と違って、乳癌で死亡したときだけ、生存曲線が下降します。時間がたつにつれて乳癌で死亡する患者は少なくなり、曲線は次第に水平になります。その漸近線の高さが全治率 c で、27%となります。つまり Boag モデルでは stage2 乳癌の 27% が手術で全治するとみなされます。

4) 新しい癌の生存分析法 (Boag モデル)



5) Boag モデルと競合リスクモデルに基づく全生存曲線と平均余命の予測

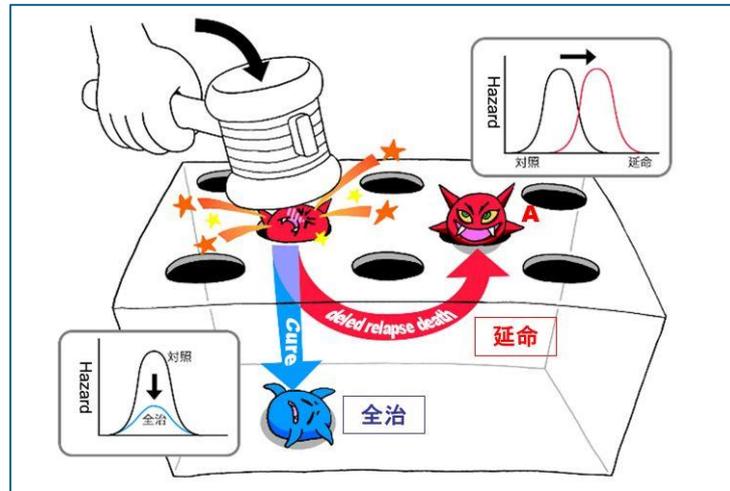


これは胃癌 T2 患者を死亡まで追跡した癌研病院のデータを借りて、追跡 5 年間の時点で全生存曲線を 50 年先まで予測したものです。(その原理は、任意の時点における overall survival rate とは、その時点までに原病死もせず、かつまた他病死もしないということが同時に起きる確率なので、他病死が独立と仮定します。) 先ず Boag モデルで 5 年以降の疾患特異的生存曲線を予測し、(第一の確率 OB は Boag のパラメータを外挿し、第二の確率 OC は患者群と同性・同時代の一般人の生命表で近似)、これと患者と同性、同時代の一般人の生存曲線とを掛け合わせると、赤色の予測曲線が得られます。黒の (Kaplan-Meier) 実測曲線とよく一致しています。この下の面積にはわずかの誤差しかみられません。その面積を一般人の AUC (曲線下の面積) と比較すると、T2 患者では半分近くの面積を失っていることがわかります。

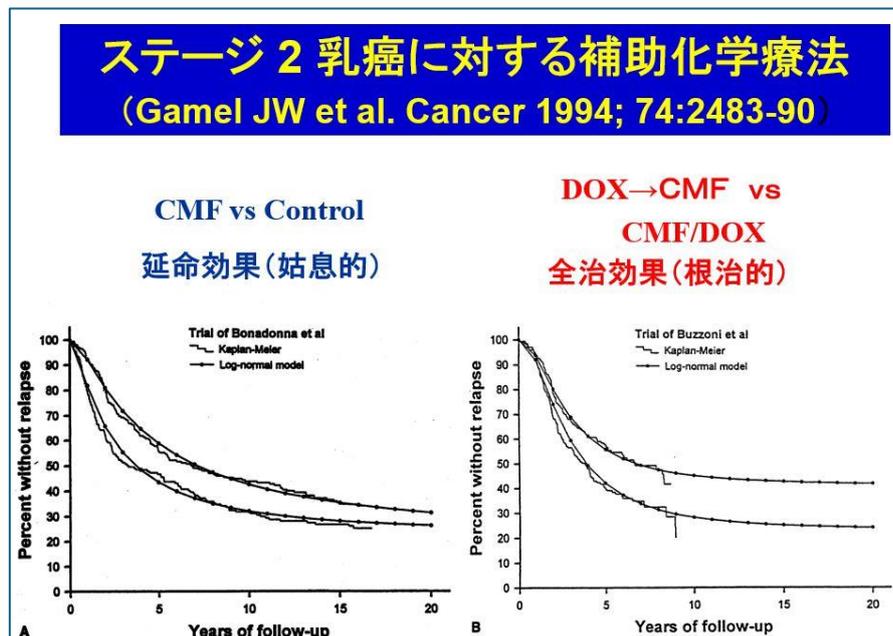
6) Cox モデルやログランク検定の問題点

延命効果しかもたない治療法をログランク検定値やハザード比で評価すると、追跡早期ほど延命効果を過大に評価する危険があり、ときには根治療法を上回る結果となります。

7) 全治と延命：Cox モデルやログランク検定は両者を区別できない



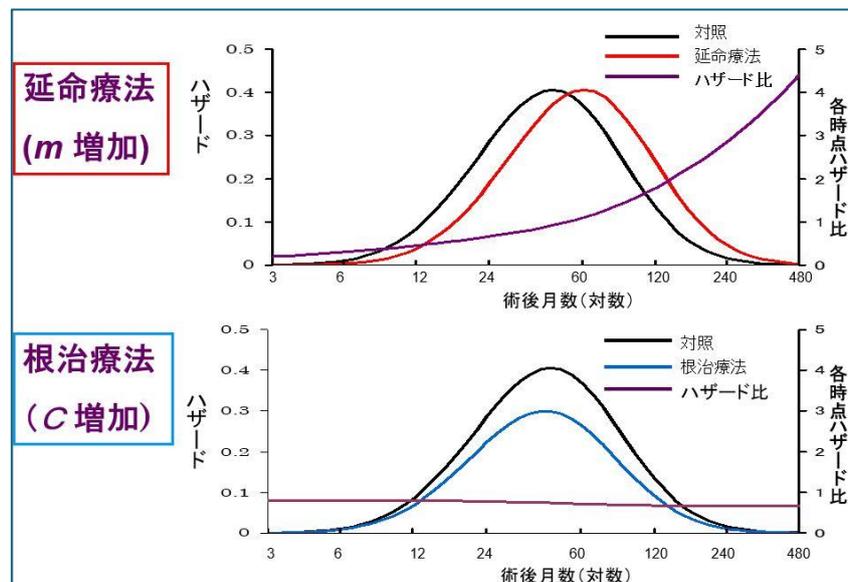
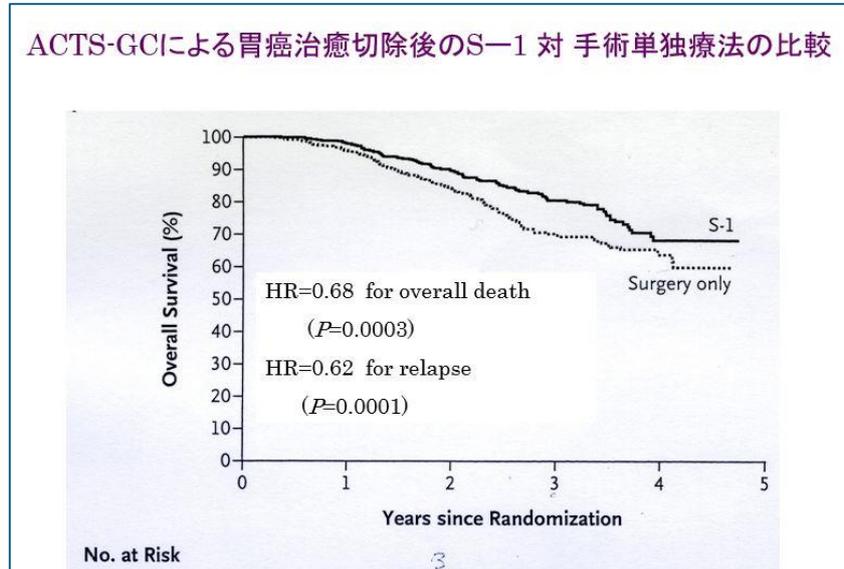
8) 全治効果のあるものと、延命効果しかないもの



1994年 Gamel はイタリアで Bonadonna らが行ったステージ2乳癌の補助化学療法に関する5つのRCTのデータを彼のモデルで解析しました。その結果、対照に対するCMFの生存曲線は尻すぼみで、CMFは延命効果しか認められませんでした。これに対して、CMFの前にアドリマイシンを使うとこれがCを有意に増加させ、曲線は末広がりになりました。つまり補助化学療法にも全治効果のあるものと、延命効果しかないものがあることを認めたのです。(ところが、この画期的な結果は余り知られていません。)

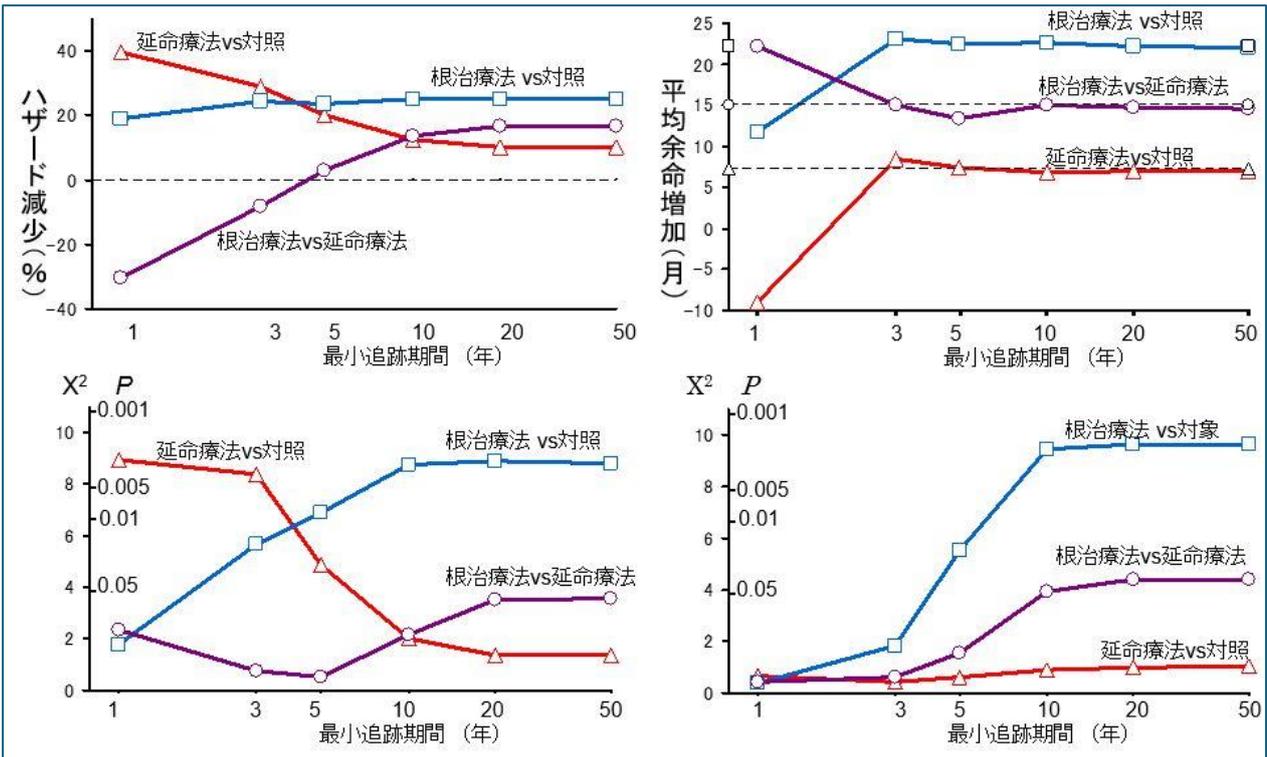
9) 延命療法と全治療法のハザードの比較

例題として ACTS-GC グループによるランダム化比較試験を取り上げます。これは胃癌ステージ II, III で治癒切除を行った 1000 例余の症例を手術単独と S-1 を 1 年間で内服の 2 群に分けその成績を比較した RCT です。その結果は 50% 追跡期間が 2 年の時点で両群の間に高度に有意な差を認め、試験は早期に中止となりました。



下の図は S1 が延命療法であった場合と根治療法であった場合の、両群のハザード曲線を比較したものです。2つの曲線から、各時点のハザード比をプロットしてみると、上段の紫の曲線で示したように「延命療法」では、ハザード比は早期には小さく、時間がたつほど上昇し、比例ハザードが成り立ちません。このため、通常使用される「平均ハザード比」は、術後早期には、小さい値をとることがわかります。根治療法では各時点のハザード比はほぼ一定になり、 c のみ増加するときに比例ハザード性を満足します。

10) 根治療法と延命療法の追跡期間がハザードと平均余命に及ぼす影響



これは治療効果の物差しとして、「ハザード比」と「平均余命」とを比較したものです（各点は100回の試行を繰り返しその中央値をとる）。下段は対応する χ^2 検定値とP値で、追跡期間を1年から50年まで6段階に増やして、それぞれの値がどう変わるかをみています。

左上のハザードの減少は「0」より大きいほど治療効果が大きいとみなされます。グラフで明らかのように、追跡早期には対照に対する「延命療法」のハザード減少（赤線）が最も大きく、あたかも延命療法が最も効果的な治療のような印象を与えます。しかし追跡を続けてゆくとその値は最下位になり、有意でもなくなっています。これに対して、（青線）の「根治療法」は、対照または延命療法と比較すると、最初はハザードの減少は相対的に小さくても、追跡期間と共に大きくなり、10年以後では延命療法より大きくなっています。

右側の平均余命の増加のグラフをみますと、追跡3年から、ほぼ安定した値となり、本来の値（点線）とほぼ同じになっており、根治療法が延命療法に勝っています。

1 1) Gamel-Boag の回帰分析

Gamel によるBoag モデルの多変量解析への拡張 (Gamel-Boagの回帰分析, 1994)	
$c = 1 / \{1 + \exp(-\mathbf{B}_c \mathbf{X}_c)\}$	C: 全治率
$m = \mathbf{B}_m \mathbf{X}_m$	m, s : 対数生存時間の平均, SD
$s = \exp(\mathbf{B}_s \mathbf{X}_s)$	$\mathbf{B}_c, \mathbf{B}_m, \mathbf{B}_s$: 回帰係数ベクトル
	$\mathbf{X}_c, \mathbf{X}_m, \mathbf{X}_s$: 予後因子ベクトル
<hr/>	
群間(治療)の比較	多変量(予後因子)の評価
ログランク検定	Cox の回帰分析
Boagのパラメトリック分析	Gamel-Boagの回帰分析

これがその式です。これを見ても何のことが分からないかもしれませんが、しかしログランク検定と Cox の回帰分析から類推すると理解しやすいと思います。ログランク検定は群間で治療を比較します。一方 Cox モデルは多変量の予後因子を評価するときに使います。それと同じことで、Gamel-Boag の回帰分析は予後因子の評価に利用できます。しかも全治をきたす予後因子を特定できるところが Cox との違いです。

1 2) 詳細は、次の PDF 「何故 Boag モデルか」をお読み下さい。