

Mantel-Haenszel法

- 複数(k個)のスタディが行われ、複数(k個)の分割表(下記)が得られた場合

	曝露あり	曝露なし	合計
case	a_i	b_i	n_{1i}
control	c_i	d_i	n_{0i}
TOTAL	m_{1i}	m_{0i}	n_i

$i=\{1,2,\dots,k\}$

(参照: www.genome.med.kyoto-u.ac.jp/func-gen-redirect/wiki/)

19

Mantel-Haenszel法

- 統合オッズ比は下記のように求められる。

1. 各研究のオッズ比を計算する。
$$OR_i = \frac{a_i d_i}{b_i c_i}$$
2. オッズ比の標準誤差を計算する。
$$SE_i = \sqrt{\frac{n_i}{b_i c_i}}$$
3. 各研究の重み付け ω_i を計算する。
$$\omega_i = \frac{1}{SE_i^2} = \frac{b_i c_i}{n_i}$$
4. 統合オッズ比及びその95%信頼区間を計算する。

$$\hat{OR}_{MH} = \frac{\sum_{i=1}^k \frac{a_i d_i}{n_i}}{\sum_{i=1}^k \frac{b_i c_i}{n_i}} \quad 95\%CI: \exp\{\log \hat{OR}_{MH} \pm 1.96\sqrt{\text{Var}(\log \hat{OR}_{MH})}\}$$

(参照: www.genome.med.kyoto-u.ac.jp/func-gen-redirect/wiki/)

20

Mantel-Haenszel法

5. 有意性の検定を行う。

$$Q_{MH} = \frac{\{|\sum_{i=1}^k (\text{Obs}(a_i) - \text{Exp}(a_i))|\}^2}{\sum_{i=1}^k \text{Var}(a_i)} \sim \chi^2_1$$

$$\text{Obs}(a_i) = a_i, \text{Exp}(a_i) = \frac{n_{1i} m_{1i}}{n_i}, \text{Var}(a_i) = \frac{n_{1i} m_{0i} m_{1i} m_{0i}}{n_i^2 (n_i - 1)}$$

- 得られる統計量は両側検定となる。連続性の補正を行う場合は、分子を二乗する前に0.5を行う。

6. 均質性の検定を行う。

$$Q_{Hetero} = \sum_{i=1}^k \omega_i (\log OR_i - \log \hat{OR}_{MH})^2 \sim \chi^2_{K-1}$$

$$\omega_i = \left(\frac{1}{a_i} + \frac{1}{b_i} + \frac{1}{c_i} + \frac{1}{d_i} \right)^{-1}$$

(参照: www.genome.med.kyoto-u.ac.jp/func-gen-redirect/wiki/)

21

異質性 (Heterogeneity)

- 複数の臨床研究における結論の不一致 (inconsistency) は何に由来するか？
同じ対象に同じデザインで研究を行えば同じ結果が得られるはずなのに何故か一致しないことがある。
- 不一致の要因
 1. 一次研究におけるバイアス Bias
 2. 偶然 Chance
 3. 異質性 Heterogeneity
 - » 研究デザインの違い
 - » 患者集団の違い

(参考) 野口善令 福原俊一. はじめてのメタアナリシス(2009).

22

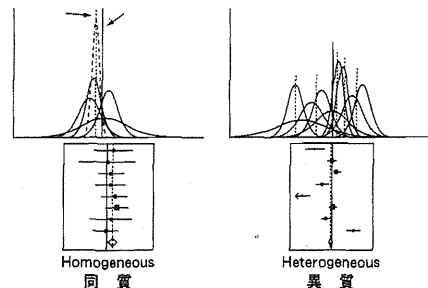
概念的異質性

- 対象となった患者集団の違い Clinical Heterogeneity
 - 年齢、性別、人種
 - 対象疾患の重症度
 - 合併症 comorbidity
- 研究デザインの違い Methodological Heterogeneity
 - 薬物の投与法 - 経口、静注
 - 薬物の投与量
 - 観察期間
 - エンドポイント

(参考) 野口善令 福原俊一. はじめてのメタアナリシス(2009).

23

統計学的異質性



(参考) 野口善令 福原俊一. はじめてのメタアナリシス(2009).

24

Heterogeneityの検定

- メタアナリシスにおいて、対象となる各研究の効果の大きさが均質 (homogeneity) であるか異質 (heterogeneity) であるかを検定する方法である。
- Heterogeneityを表す指標としてはCochranのQ統計量、 I^2 、Breslow-Day検定P値、等がある。
 - CochranのQ統計量、Breslow-Day検定P値は対象となるスタディ数が少ない時、Heterogeneityに対する検出力が落ちる点に注意が必要である。

(参照: www.genome.med.kyoto-u.ac.jp/func-gen-redirect/wiki/)

25

Heterogeneityの検定

- I^2 はCochranのQ統計量、対象スタディ数Kを使用して下記のように表される。

$$I^2 = \frac{Q - (K - 1)}{Q}$$

- I^2 は0%~100%の値をとり、 $\frac{Q}{Q}$ の間に認められた分散のうち、heterogeneityによって説明される割合を示している。
- I^2 の評価基準
 - 25%以下を「absence」
 - 50%以下を「moderate」
 - 75%以下を「large」
 - 100%以下を「extreme」として捉えることが推奨されている。
- 但し、Q統計量程ではないものの、やはり I^2 もスタディ数が少ない時には検出力が低下する傾向がある点に注意が必要である。
 - 実際には、定義式により導かれる I^2 は負の値となる場合がある。その際は $I^2=0$ として採用される例が多い。

(参照: www.genome.med.kyoto-u.ac.jp/func-gen-redirect/wiki/)

26

I^2 による異質性の比較

Table 1 Heterogeneity statistics for examples of meta-analyses from the literature. Meta-analyses were conducted using either meta or mean in STATA¹⁵

Topic	Outcome / analysis	Effect measure	No of studies	Heterogeneity test			I ² (95% uncertainty interval) ¹⁶
				Q	df	P	
① Tamoxifen for breast cancer ¹⁷	Mortality	Peto odds ratio	55	55.9	54	0.4	3 (0 to 26)
② Streptolisinase after myocardial infarction ¹⁸	Mortality	Odds ratio	33	39.5	32	0.17	19 (0 to 46)
③ Selective serotonin reuptake inhibitors for depression ¹⁹	Drop-out	Odds ratio	135	179.9	134	0.005	26 (7 to 40)
④ Magnesium for acute myocardial infarction ²⁰	Death	Odds ratio	16	40.2	15	0.0004	63 (39 to 78)
⑤ Magnetic fields and leukaemia ²¹	All studies	Odds ratio	6	15.9	5	0.007	89 (28 to 87)
⑥ Amantadine ²²	Prevention of influenza	Odds ratio	8	12.44	7	0.09	44 (0 to 75)

df=degrees of freedom

¹⁶Values of I^2 are percentages. 95% uncertainty intervals are calculated as proposed by Higgins and Thompson¹⁴

(参考)野口善令 福原俊一. はじめてのメタアナリシス(2009).

27

参考資料

- 野口善令 福原俊一. はじめてのメタアナリシス(2009). NPO法人健康医療評価研究機構 (iHope).
- シリウス先生の心理統計学
 - <http://homepage2.nifty.com/handemoarchive/index.htm>
- The SPELL
 - <http://spell.umin.jp/index.htm>

28

Nifedipine in the management of preterm labor: a systematic review and metaanalysis

Am J Obstet Gynecol
2011;204:134.e1-20.

29

OBSTETRICS

Nifedipine in the management of preterm labor: a systematic review and metaanalysis

Agustín Conde-Agudelo, MD, MPH; Roberto Romero, MD; Juan Pedro Kusanovic, MD

OBJECTIVE: To determine the efficacy and safety of nifedipine as a tocolytic agent in women with preterm labor.

STUDY DESIGN: A systematic review and metaanalysis of randomized controlled trials.

RESULTS: Twenty-six trials (2179 women) were included. Nifedipine was associated with a significant reduction in the risk of delivery within 7 days of initiation of treatment and before 34 weeks' gestation, respiratory distress syndrome, necrotizing enterocolitis, intraventricular hemorrhage, neonatal jaundice, and admission to the neonatal intensive care unit when compared with β_2 -adrenergic-receptor agonists.

There was no difference between nifedipine and magnesium sulfate in tocolytic efficacy. Nifedipine was associated with significantly fewer maternal adverse events than β_2 -adrenergic-receptor agonists and magnesium sulfate. Maintenance nifedipine tocolysis was ineffective in prolonging gestation or improving neonatal outcomes when compared with placebo or no treatment.

CONCLUSION: Nifedipine is superior to β_2 -adrenergic-receptor agonists and magnesium sulfate for tocolysis in women with preterm labor. **Key words:** calcium channel blocker, neonatal morbidity, pregnancy, premature birth, preterm birth, tocolysis, uterine contractility

30

研究の調査～解析用DB (MATERIALS AND METHODS)

- Search
- Study selection
- Outcome measures
- Study quality assessment
- Data extraction
- Statistical analysis (メタアナリシス)

31

Search(研究検索)

- 検索語とデータベース
 - “nifedipine,” “calcium channel blocker,” “calcium antagonist,” “tocolysis,” “preterm labor,” “premature,” and their associated medical subject headings (MeSH)
 - (without language restrictions)
 - 対象データベース
 - MEDLINE, EMBASE, CINAHL, and LILACS, the Cochrane Central Register of Controlled Trials, ISI Web of Science, Research Registers of Ongoing, and Google scholar.
- その他の調査
 - Proceedings of the Society for Maternal-Fetal Medicine and international meetings on preterm birth and tocolysis, reference lists of identified studies, textbooks, previously published systematic reviews, and review articles

32

Study selection

- 選択
 - We included randomized controlled trials in which nifedipine was used for tocolysis in patients with preterm labor compared with alternative tocolytic agents, placebo, or no treatment.
- 除外
 - Trials were excluded if they were quasi-randomized, if they compared only different doses of nifedipine or other calcium channel blockers, or if nifedipine was given in addition to or following failure of another tocolytic drug.
- 分類
 - We classified trials according to the aim of the treatment with nifedipine into 2 groups: acute tocolysis and maintenance.
- 選択者
 - Two reviewers independently evaluated studies for inclusion

33

Outcome measures

- Primary outcomes
 - delivery within 48 hours and 7 days of treatment for acute tocolysis;
 - delivery before 34 and 37 weeks' gestation for maintenance tocolysis;
 - and perinatal death, admission to neonatal intensive care unit (NICU), neurodevelopmental disability at 2 years of age, and severe maternal adverse drug reactions for both acute and maintenance tocolysis.
- Secondary outcomes
 - the interval between trial entry and delivery, gestational age at delivery, maternal adverse events, discontinuation of treatment because of adverse events, birthweight, Apgar score at 5 minutes, respiratory distress syndrome, intraventricular hemorrhage, necrotizing enterocolitis, retinopathy of prematurity, neonatal jaundice, neonatal sepsis, fetal death, neonatal death, length of stay in the NICU, longterm psychosocial and motor function, and pregnancy/neonatal outcomes among women enrolled at less than 32 weeks' gestation.

34

PICO (or PECO)

- 疑問を定式化したもの
 - P(Patient): 患者
 - E(Exposure) or I(Intervensio): 曝露又は介入
 - C(Comparison): 比較
 - O(Outcome): 結果

35

メタアナリシスにおけるバイアス

- 一次研究レベルのバイアス
 - 質の低いRCT
- メタアナリシスを行う過程のバイアス
 - 出版バイアス Publication bias
 - 選択バイアス Selection bias

(参考)野口晋令 福原俊一. はじめてのメタアナリシス(2009).

36

Study quality assessment

- We conducted quality assessment according to a modified scoring system proposed by Jadad et al,²⁴ which considers 4 items:
 - randomization,
 - Blinding,
 - follow-up,
 - and concealment of allocation.
- Two investigators (A.C.-A. and J.P.K.) independently assessed study quality, and discrepancies were resolved through discussion.

37

一次研究(RCT)の質の評価

質の低い一次研究はバイアスが大いではないか？

- 一次研究のQuality Assessmentの必要性
 1. 質の低い一次研究は除外する
(Pre-selection Inclusion / Exclusion criteriaによる排除)
 2. 質の低い一次研究を除外する／しない
で感受性分析 (sensitivity analysis) する
 3. 研究の質に応じてweightをかける

(参考)野口善令 福原俊一. はじめてのメタアナリシス(2009).

38

質の低い一次研究を除く (modified Jadad scoring system)

- Jadad (ハダッドと読みます)
- (1) quality of randomization
 - 2 points: computer-generated random numbers or similar
 - 1 point: not described
 - 0 points: quasi-randomized or not randomized
 - (2) double blinding
 - 2 points: neither the person doing the assessments nor the study participant could identify the intervention being assessed
 - 1 point: not described
 - 0 points: no blinding or inadequate method
 - (3) follow-up
 - 2 points: number or reasons for dropouts and withdrawals described and assessment of primary outcomes in 95% or more of randomized women
 - 1 point: number or reasons for dropouts and withdrawals described but assessment of primary outcomes in less than 95% of randomized women
 - 0 points: number or reasons for dropouts and withdrawals not described
 - (4) concealment of allocation
 - 2 points: adequate method [central randomization; or drug containers or opaque, sealed envelopes that were sequentially numbered and opened sequentially only after they have been irreversibly assigned to the participant]
 - 0 points: no concealment of allocation or inadequate method or not described

39

Table2: Methodological quality assessment of included studies

- modified Jadad scoring system
 - 合計:0-8点
 - 質が高い:6点以上
- 今回の論文では13/26(50%)の研究でmodified Jadad scoreが6点以上だった。
 - Randomization: 19/26が適切な割付
 - Blinding: 1/26が実施
 - Follow-up: 18/26が95%以上のランダム化された患者で実施。
 - Allocation concealment (記述なし)

40

出版バイアス

- 定義

positiveな結果(統計学的有意差をもって効果あり)は, negativeな結果(効果なし,あるいは,統計学的有意差なし)よりも発表されやすい。
- 発表された報告は, 現在までに行われたすべての研究を代表していない。

(参考)野口善令 福原俊一. はじめてのメタアナリシス(2009).

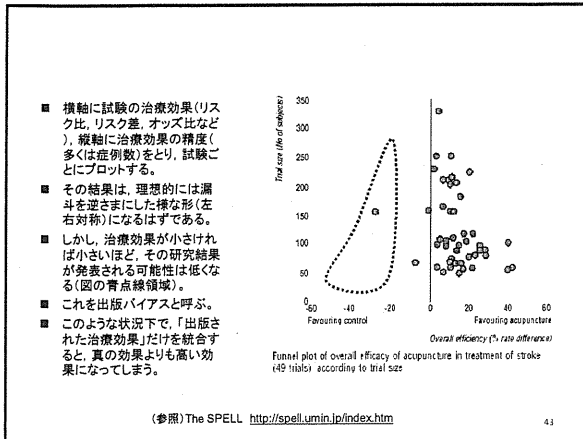
41

出版バイアスの検出

- 視覚的
 - Funnel plot
- 統計学的
 - Rank correlation test
 - (Begg, CB. Operating characteristics of a rank correlation test for publication bias. Biometrics. 1994;50:1088-101)
 - Linear regression test
 - (Egger. Bias in meta-analysis detected by a simple, graphical test. BMJ. 1997;315:629-34)

(参考)野口善令 福原俊一. はじめてのメタアナリシス(2009).

42



本研究における出版バイアスの検出

- visually by examining the symmetry of funnel plots
- statistically by using the Egger test
 - The larger the deviation of the intercept of the regression line from zero, the greater was the asymmetry and the more likely it was that the meta-analysis would yield a biased estimates of effect. We considered $P < 0.1$ to indicate significant asymmetry, as suggested by Egger.
- (結果: ファンネル・プロットや解析結果は非表示)
 - All funnel plots showed no asymmetry, either visually or in terms of statistical significance ($P > 0.10$ for all, by Egger test).

Data extraction

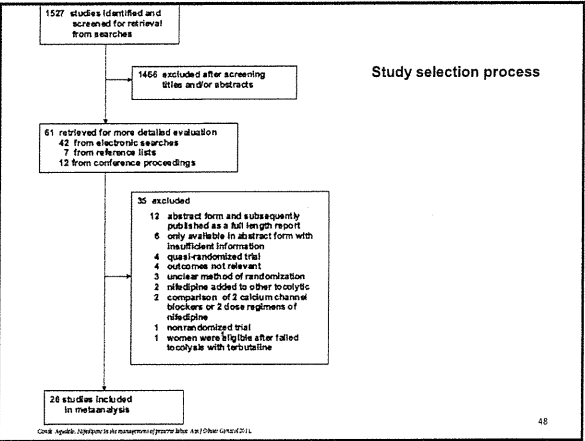
- Two reviewers (A.C.-A. and J.P.K.) independently extracted data from each eligible study using a standardized data abstraction form.

Statistical analysis

- intention-to-treat basis
- Four prespecified subgroup analyses were performed to compare nifedipine with other tocolytic agents for acute tocolysis.
 - 2-adrenergic- receptor agonists,
 - magnesium sulfate,
 - atosiban,
 - and nitric oxide donors
- One to compare nifedipine with placebo or no treatment for maintenance tocolysis.

Statistical analysis

- Heterogeneity of the results among studies was tested with the quantity I^2
 - 0% indicates no observed heterogeneity
 - 50% or more indicate a substantial level of heterogeneity
 - $I^2 < 50\%$: fixed-effects modelsを使用する
 - $I^2 \geq 50\%$: random-effects modelsを使用する
- Sensitivity analysis
 - performed by excluding trials with a modified Jadad score less than 6.



Strengths of this study

- (1) the use of rigorous methodology for the performance of a systematic review of randomized controlled trials;
- (2) a comprehensive literature search to identify relevant studies;
- (3) the inclusion of a relatively large number of studies in metaanalyses;
- (4) the quality assessment of trials included in the review was based on a widely used and validated scale;
- (5) the relatively narrow confidence intervals obtained making our estimates of effect size;
- (6) the evidence of clinical and statistical homogeneity in the results of the trials for most of the outcomes evaluated;
- (7) the sensitivity analyses restricted to high-quality trials that were consistent with (and thus supportive of) the overall findings;
- (8) the subgroup and metaregression analyses that did not show any significant influence of study characteristics on effect size;
- (9) the symmetrical funnel plots suggesting absence of publication and related biases in our metaanalyses.

55

生存時間解析の論文を読み解くための予備知識

生物統計学セミナー
第2回
Feb. 14, 2012

本日のお話

- 生存時間の指標
- カプラン・マイヤー法
- 生存時間の2群比較
 - ログランク検定
 - 一般化ウイルコクソン検定
- 比例ハザードモデル
 - Cox回帰, ハザード比

生存時間の指標

和名	英名	起算日 (追跡試験)	イベント	打ち切り (打ち切り日)	備考
生存期間	overall survival	登録日	全死亡	生存 (最終生存確認日)	
無増悪生存期間	progression-free survival	登録日または治療開始日	全死亡 増悪	増悪なく生存 (最終無増悪生存確認日)	
無再発生存期間	relapse-free survival	登録日または手術日	全死亡 再発	再発なく生存 (最終無再発生存確認日 または最終生存確認日)	
無病生存期間	disease-free survival	登録日または手術日	全死亡 再発 二次がん	再発なく生存 (最終無再発生存確認日 または最終生存確認日) 再発なく二次がんもなく (最終無再発生存確認日 または最終生存確認日)	乳がんの術後補助療法の試験等で見られる
治療成功期間	time to treatment failure	登録日または治療開始日	全死亡 増悪 治療中止	治療中または治療終了後、増悪なく生存 (最終無増悪生存確認日)	白血球領域ではevent-free survival*
薬効期間 効果持続期間	response duration	CR または PR が得られた日	全死亡 増悪	増悪なく生存 (最終無増悪生存確認日)	対象は薬効例 (CR+PR)
有効期間 完全有効期間	complete response duration	CR が得られた日	全死亡 再発	増悪なく生存 (最終無増悪生存確認日)	対象はCR例 CR例の無再発生存期間と同じ

* event-free survival: 治療失敗 (treatment-failure), CR 後の再発, 全死亡のうち早いものをイベントとし、いずれのイベントも観察されていない場合、最終生存確認日で打ち切り。「治療失敗」の定義はプロトコル毎に明記する。

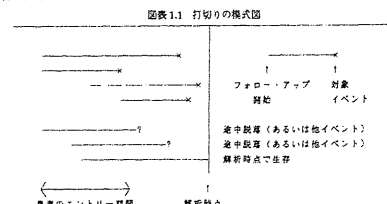
解析手法の分類

データの型	2値	連続値	time-to-event
解析の目的 分布の要約	頻度表	ヒストグラム, 要約統計量	Kaplan-Meier法
2群間比較	χ^2 検定 Fisher直接確率検定	t検定 Wilcoxon検定	ログランク検定
多群間比較	分割表の χ^2 検定	分散分析	ログランク検定
予後因子の調整 層別(化)調整	OMH	分散分析	ログランク検定の応用
統計モデル当てはめ	ロジスティック回帰	重回帰	Cox回帰

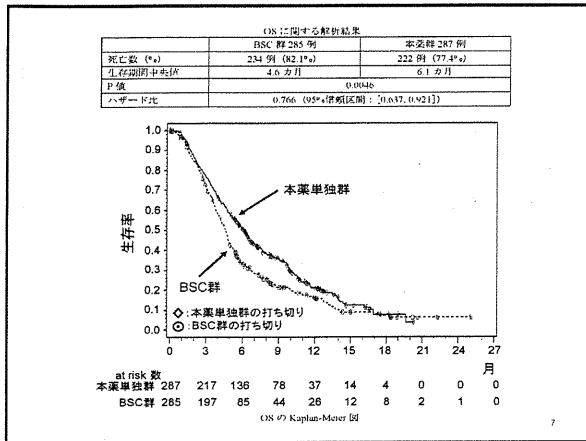
参考: 大橋謙雄/荒川雅弘 臨床試験の道の方

打ち切り(censoring)

- 打ち切りの原因としては、大きく分けて2種類が考えられる。
 - 途中脱落(drop out)の場合
 - 解析時点でイベント(死亡など)がおきていない場合



参考: 大橋謙雄・浜田 知久馬, 生存時間解析 SASによる生物統計



生存時間曲線 Kaplan・マイヤー法

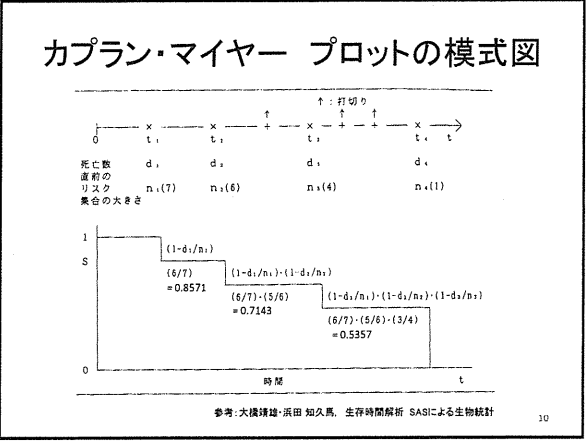
分布評価: Kaplan・マイヤー法

生存関数 $\hat{S}(t) = (1 - d_1/n_1) \times (1 - d_2/n_2) \times \dots$
 $= \prod (1 - d_i/n_i)$ (2.1.1)

標準誤差: グリーンウッドの公式より求める。

S(t): 生存関数の推定値
 t_i : 死亡が起きているある時点*i*における時間
 d_i : 同死亡者数
 n_i : 同リスク集合の大きさ(死亡が起こる直前に死亡のリスクにさらされていた人数)
 d_i (死亡数)/ n_i (リスク集合): ハザード成分

参考: 大橋謙雄・浜田 知久馬, 生存時間解析 SASによる生物統計



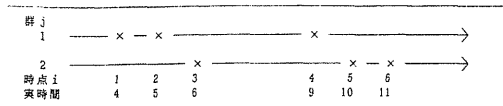
生存時間の2群比較

- ・ログランク検定
- ・一般化ウイルクソン検定

生存時間の2群比較

- ・生存時間分布のノンパラメトリック検定
 - ログランク検定 (log-rank test)
 - 一般化ウイルクソン検定 (Generalized Wilcoxon test)
- ・2つの検定法の違い
 - 計算過程で使用する重み (Weight) づけの方法が異なる

打ち切りが存在しない場合(例1)



群1では時点1,2,4で死亡例
群2では時点3,5,6で死亡例

参考: 大橋謙雄・浜田 知久典, 生存時間解析 SASによる生物統計

13

ノンパラメトリック検定の計算手順 (例1: 打ち切りがない場合)

時点	i=1	i=2	i=3	i=4	i=5	i=6
群1	1	1	1	1	1	1
群2	1	1	1	1	1	1
生存	3	2	2	2	0	0
死亡	1	1	0	1	0	1
計	3	2	2	1	2	0

時点*i*を数値添え字(*i*=1,2,...,6)と群を数値添え字(*j*=1,2)で生存と死亡のクロス表を作成すると左図のようになる。

$$\begin{aligned} \textcircled{1} \quad d_{ij} &= \begin{matrix} 1 & 1 & 0 & 1 & 0 & 0 \\ 2/6 & 1/2 & 1/4 & 1/3 & 0/2 & 0/1 \end{matrix} \\ d_{i.} &= \begin{matrix} 3 & 2 & 2 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{matrix} \\ e_{.j} &= \begin{matrix} 3/6 & 2/5 & 2/4 & 2/3 & 2/2 & 1/1 \end{matrix} \\ \textcircled{2} \quad u_{.i} &= \begin{matrix} 1-3/6 & 1-2/5 & 1-2/4 & 1-1/3 & 1-0/2 & 1-0/1 \\ 1/6 & 3/5 & 1/2 & 2/3 & 1 & 1 \end{matrix} \\ v_{.i} &= \begin{matrix} 3 \times 3 & 2 \times 3 & 2 \times 3 & 1 \times 3 & 0 \times 2 & 0 \times 1 \\ 6^2 & 5^2 & 4^2 & 3^2 & 2^2 & 1^2 \end{matrix} \\ w_{.i} &= \begin{matrix} 3 \times 1 & 2 \times 1 & 2 \times 1 & 1 \times 1 & 0 \times 2 & 0 \times 1 \\ 6^2 & 5^2 & 4^2 & 3^2 & 2^2 & 1^2 \end{matrix} \\ w_{.i} &= \begin{matrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 5 & 4 & 3 & 2 & 1 & 1 \end{matrix} \end{aligned}$$

① 死亡数 d_{ij} 、群別のリスク集合の大きさ $n_{i.}$ 、群を併合したリスク集合の大きさ $e_{.j}$ を用いて、時点 i 群 j ごとに期待死亡数 $e_{ij} = d_{ij} \times n_{i.} / n_{.i}$ を算出します。
② 期待死亡数 e_{ij} と死亡数 d_{ij} の差に時点の重み $w_{.i}$ (①)を掛けたものを時点 i における群のスコア $u_{.i}$ とします。

d_{ij} : 死亡数
 $n_{i.}$: 死亡時のリスク集合の大きさ
 $e_{.j}$: 死亡時のリスク集合の大きさ
 $v_{.i}$: 死亡時のリスク集合の大きさ

参考: 大橋謙雄・浜田 知久典, 生存時間解析 SASによる生物統計

14

スコアの計算手順

$$\begin{aligned} u_1 &= w_1(1-3/6) + w_2(1-2/5) + w_3(0-1/4) + w_4(1-1/3) \\ &\quad + w_5(0-0/2) + w_6(0-0/1) \\ u_2 &= w_1(0-3/6) + w_2(0-3/5) + w_3(1-3/4) + w_4(0-2/3) \\ &\quad + w_5(1-2/2) + w_6(1-1/1) \end{aligned}$$

Wi: 重み

- ログランク検定: 常に一様 (=1)
- 一般化Wilcoxon検定: 死亡時点でのリスク集団の患者数合計で重みをつける

参考: 大橋謙雄・浜田 知久典, 生存時間解析 SASによる生物統計

15

ログランク検定と 一般化ウィルコクソン検定の統計量 (例1: 打ち切りがない場合)

ログランク検定 ($w_{.i} = 1$)

$$\begin{aligned} u_{.i} &= 1 \cdot (1-3/6) + 1 \cdot (1-2/5) + 1 \cdot (0-1/4) + 1 \cdot (1-1/3) + 1 \cdot (0-0/2) + 1 \cdot (0-0/1) \\ &= 1.51667 \\ v_{.i} &= 1^2 \frac{3 \times 3}{6^2} + 1^2 \frac{2 \times 3}{5^2} + 1^2 \frac{1 \times 3}{4^2} + 1^2 \frac{1 \times 2}{3^2} + 1^2 \frac{0 \times 2}{2^2} + 1^2 \frac{0 \times 1}{1^2} \\ &= 0.89972 \\ \chi^2 &= 1.51667^2 / 0.89972 = 2.55665 \quad [= u_{.i}^2 / v_{.i}] \end{aligned}$$

一般化ウィルコクソン検定 ($w_{.i} = n_{i.}$)

$$\begin{aligned} u_{.i} &= 6 \cdot (1-3/6) + 5 \cdot (1-2/5) + 4 \cdot (0-1/4) + 3 \cdot (1-1/3) + 2 \cdot (0-0/2) + 1 \cdot (0-0/1) \\ &= 7 \\ v_{.i} &= 6^2 \frac{3 \times 3}{6^2} + 5^2 \frac{2 \times 3}{5^2} + 4^2 \frac{1 \times 3}{4^2} + 3^2 \frac{1 \times 2}{3^2} + 2^2 \frac{0 \times 2}{2^2} + 1^2 \frac{0 \times 1}{1^2} \\ &= 20 \\ \chi^2 &= 7^2 / 20 = 2.45000 \quad [= u_{.i}^2 / v_{.i}] \end{aligned}$$

③ スコア $u_{.i}$ を群ごとに時点について足し合わせ、スコア統計量 $u_{.i}$ を計算し、スコア統計量 $v_{.i}$ を以下の式で計算します。

$$v_{.i} = \sum_{j=1}^2 \frac{(n_{i.} - n_{ij})n_{ij}(n_{.j} - n_{ij})}{n_{.i}(n_{.j} - 1)}$$

④ 検定のためのカイ二乗統計量は、スコア統計量 $u_{.i}$ の二乗を分数で割ったもの(以下の式)になります。この結果からP値を求めます。

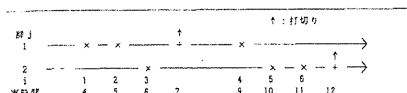
$$\chi^2 = u_{.i}^2 / v_{.i}$$

Test Chi-Square DF Pr > Chi-Square
Log-Rank 2.5567 1 0.1098
Wilcoxon 2.4500 1 0.1175

参考: 大橋謙雄・浜田 知久典, 生存時間解析 SASによる生物統計

16

打ち切りがある場合(例2)



時点3と4の間に1例の打ち切りが、また時点6の後に1例の打ち切りがそれぞれ発生しています。

打ち切りが起きた場合は、時点の時点でまだ打ち切りを受けていない個体(患者)は時点のリスク集団に含まれ、 e_{ij} の計算式の分母にのみ寄与します。

計算は「打ち切りが存在しない場合」の①~④の手順と同じです。

$$\textcircled{1} \quad d_{ij} = \begin{matrix} 1 & 1 & 0 & 1 & 0 & 0 \\ 4/8 & 3/7 & 2/6 & 1/4 & 0/3 & 0/2 \\ 0 & 0 & 1 & 0 & 1 & 1 \\ 4/8 & 4/7 & 4/6 & 3/4 & 3/3 & 2/2 \end{matrix}$$

17

ノンパラメトリック検定の計算手順 (例2: 打ち切りがある場合)

ログランク検定 ($w_{.i} = 1$)

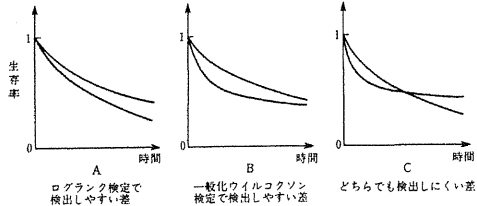
$$\begin{aligned} \textcircled{2} \quad u_{.i} &= 1 \cdot (1-4/8) + 1 \cdot (1-3/7) + 1 \cdot (0-2/6) + 1 \cdot (1-1/4) + 1 \cdot (0-0/3) + 1 \cdot (0-0/2) \\ &= 1.48810 \\ \textcircled{3} \quad v_{.i} &= 1^2 \frac{4 \times 4}{8^2} + 1^2 \frac{3 \times 4}{7^2} + 1^2 \frac{2 \times 4}{6^2} + 1^2 \frac{1 \times 3}{4^2} + 1^2 \frac{0 \times 3}{3^2} + 1^2 \frac{0 \times 2}{2^2} \\ &= 0.90462 \\ \textcircled{4} \quad \chi^2 &= 1.48810^2 / 0.90462 = 2.44791 \end{aligned}$$

一般化ウィルコクソン検定 ($w_{.i} = n_{i.}$)

$$\begin{aligned} \textcircled{2} \quad u_{.i} &= 8 \cdot (1-4/8) + 7 \cdot (1-3/7) + 6 \cdot (0-2/6) + 4 \cdot (1-1/4) + 3 \cdot (0-0/3) + 2 \cdot (0-0/2) \\ &= 9 \\ \textcircled{3} \quad v_{.i} &= 8^2 \frac{4 \times 4}{8^2} + 7^2 \frac{3 \times 4}{7^2} + 6^2 \frac{2 \times 4}{6^2} + 4^2 \frac{1 \times 3}{4^2} + 3^2 \frac{0 \times 3}{3^2} + 2^2 \frac{0 \times 2}{2^2} \\ &= 39 \\ \textcircled{4} \quad \chi^2 &= 9^2 / 39 = 2.07692 \end{aligned}$$

18

各種ノンパラメトリック検定の特徴



A ログランク検定で検出しやすい差
 B 一般化ウイロコクソン検定で検出しやすい差
 C どちらでも検出しにくい差

一般化ウイロコクソン検定 ($w_1=n_1$) は相対的に初期の死亡を重く評価しているのに対して、ログランク検定 ($w_1=1$) では、後期の死亡も重く評価します。

参考: 大橋 謙雄・浜田 知久馬, 生存時間解析 SASによる生物統計

19

比例ハザードモデル Cox回帰, ハザード比

血管新生 (Micro Vessel Count) と乳癌患者 (n_0) の予後 Heimann R. et al. JNCI 1996;88:1764-9

ログランク検定結果 Cox回帰結果

Table 2 Univariate and multivariate analysis of the prognostic value of tumor size, age, nuclear grade, estrogen receptor status, and microvessel count for disease-free survival*

Factor	Univariate		Multivariate	
	p	Hazard rate	95%CI	p
Tumor size, ≤ 2 cm; > 2 cm	.62	0.75	0.31-1.82	.53
Age, < 60 y; ≥ 60 y	.07	0.55	0.23-1.30	.17
Nuclear grade, 1; 2 and 3 combined	.21	4.10	0.54-31.03	.17
Estrogen receptor status, positive; negative	.37	1.44	0.69-3.55	.42
Microvessel count, low (< 15); high (≥ 15 per high-power field)	.018	4.51	1.04-19.55	.04

*Results of univariate analysis using the logrank test and multivariate analysis using the Cox proportional hazards model of prognostic factors for disease-free survival. A stepwise variable selection process was used, and the final model is shown. Only the microvessel count is a significant prognostic variable. CI=confidence interval.

21

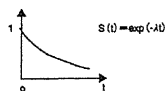
生存率 (累積生存率)

- 1年あたりの死亡率が5%の場合, 1年後の生存率は95% ($= (1-0.05) * 100$)となる。
- 2年後は, 1年後に残った95%の集団の5%が死亡するので, その生存率 (累積生存割合) は,
 - $0.95 \times 0.95 = (1-0.05)^2 = 0.9025$

22

生存関数

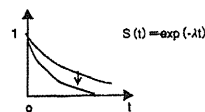
- t年後の累積生存率: $S(t)$
 - $S(t) = \exp(-\lambda t)$ [λ : ハザード]
 - 累積生存率のことを, 「生存関数: $S(t)$ 」と呼ぶ
 - 1年後生存率 $= \exp(-\lambda * 1) = 0.95$ を解くと, $\lambda = 0.05129$
 - λ が小さい場合は, $\exp(-\lambda t) \sim 1 - \lambda t$ と近似できる
 - ハザード (λ) は瞬間速度
 - 単位時間あたりの死亡率



23

ハザードと生存曲線の関係

- ハザード (λ) が k 倍になると
 - $\exp(-k\lambda t) = (\exp(-\lambda t))^k$
 - ハザードが2倍 ($\lambda_1 / \lambda_2 = 2$) = 生存関数の2乗
 - $0.95^2 = 0.9025$ (\approx 死亡率が0.05から0.10に)
 - $0.6^2 = 0.36$ (\approx 死亡率が0.4から0.64に)



24

比例ハザードモデル

- Cox回帰の基礎となる統計モデルは、比例ハザードモデルで、ハザード関数($hazard(t, x)$)として以下の式を想定します。

$$hazard(t, x) = hazard(t) \times \psi(x)$$

$$\psi(x) = \exp(\beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k)$$

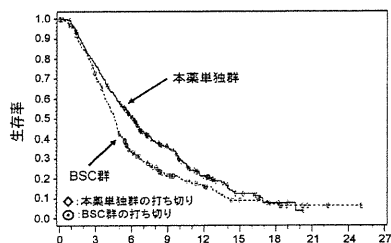
$$= \exp(\beta_1 x_1) \exp(\beta_2 x_2) \dots \exp(\beta_k x_k)$$

- ここで、 $hazard(t)$ は基準ハザード関数、 $\psi(x)$ は共変量を意味します。つまり、ハザードを時間に関する部分と共変量の部分に分解して構成しています。
- β_i が推定すべき未知パラメータです。例えば、 $x_i=1$ (試験治療)又は0(対照群)とし、解析の結果 $\beta < 0$ であれば試験治療の効果あり、 $\beta=0$ ならば試験治療の効果なしとなります。また、 $(\beta/\text{標準誤差})^2 > 3.84$ ならば 5%有意と判断されます。

25

生存時間解析結果の要約

OSに関する詳細結果		
	BSC群 285例	本薬群 287例
症例数 (n)	284例 (98.1%)	282例 (97.9%)
生存期間中央値	4.6カ月	6.1カ月
P値	0.0046	
ハザード比	0.766 (95%信頼区間: [0.637, 0.921])	



at risk 数	0	3	6	9	12	15	18	21	24	27
本薬単独群	287	217	136	78	37	14	4	0	0	0
BSC群	285	197	85	44	26	12	8	2	1	0

OSのKaplan-Meier図

27

MSTとハザード比(信頼区間)

- 生存時間解析の結果は通常以下に示す2つの指標で要約され提示される
 - 生存時間中央値 (Median survival time; MST)
 - ハザード比 (及びその信頼区間)

28

生存時間中央値 (Median survival time; MST)

- 生存時間中央値は、 Kaplan・マイヤー曲線のような生存時間分布を要約するための代表的な指標のひとつ。
- 50%点とも呼ばれ、全対象者の50%が生きている時点(時間)を意味する。
- セツキシマブの例
 - 「生存期間中央値」と表現された結果 (BSC群: 4.6カ月, 本薬群: 6.1カ月) がこれに相当する。

29

ハザード比(及びその信頼区間)

- ハザード比は、群ごとに推定されるハザード(瞬間死亡率)の比で、通常は試験薬のハザードを対照薬のハザードで割った値になる。
- ハザード比は、Cox回帰で計算される。
- ハザード比は、1を超えると対照薬よりも死亡のリスクが相対的に大きくなり、1未満では死亡のリスクが相対的に小さくなる。
- 通常、ハザード比は信頼区間とともに表示する。
- セツキシマブの例
 - ハザード比: 0.766 (95%信頼区間: [0.637, 0.921])
 - ハザード比の信頼区間の上限値(0.921)が1よりも小さいので、統計的に有意に死亡のリスクが小さいと判断される。
 - この結果は、同じ表にあるP値(層別ログランク検定の結果)で有意であることとほぼ符合します。

30

Long-Term Results of a Randomized Controlled Trial in Childhood IgA Nephropathy

- Clin J Am Soc Nephrol. 2011 Jun;6(6):1301-7

31

本報告の元となった研究 A Controlled Trial of Combined Therapy for Newly Diagnosed Severe Childhood IgA Nephropathy

- J Am Soc Nephrol 10: 101-109, 1999

32

1. 研究背景と目的

- The most appropriate treatment for patients with IgA nephropathy is controversial.
- Treatment with prednisolone, azathioprine, heparin-warfarin, and dipyridamole early in the course of disease may prevent immunologic renal injury in children with severe IgA nephropathy.
- To determine whether similar results can be obtained with a combination of just heparin-warfarin and dipyridamole, the effects of such treatment were compared to those of treatment with prednisolone, azathioprine, heparin-warfarin, and dipyridamole in 78 children with newly diagnosed IgA nephropathy showing diffuse mesangial proliferation.

33

2. 方法

- The patients were randomly assigned to receive either prednisolone, azathioprine, heparin-warfarin, and dipyridamole for 2 yr (group 1) or heparin-warfarin and dipyridamole for 2 yr (group 2).
- All of the 40 patients in group 1 and 34 of the 38 patients in group 2 completed the trial.

34

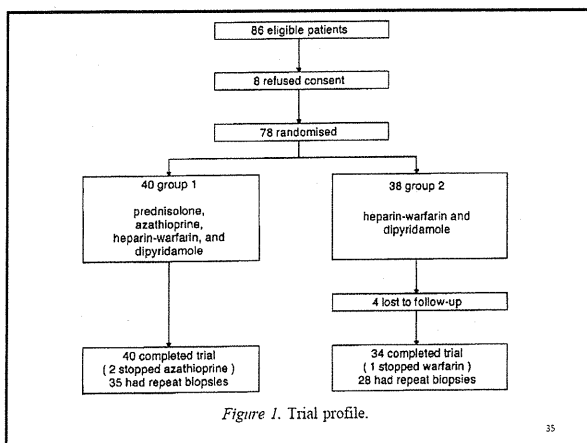


Figure 1. Trial profile.

35

Table 1. Clinical and pathologic data of the study patients at the start of treatment with prednisolone, azathioprine, heparin-warfarin, and dipyridamole (group 1) and heparin-warfarin and dipyridamole only (group 2)^a

Clinical and Pathologic Data	Group 1 (n = 40)	Group 2 (n = 38)	P Value
Age (yr)	12.2 ± 3.0	11.6 ± 2.3	0.16
MF	22/18	29/9	0.06
Months of disease	11 ± 11	11 ± 1	0.71
Months from biopsy	1.3 ± 1.4	1.8 ± 1.5	0.10
Initial presentation (no. of patients)			
asymptomatic proteinuria and hematuria	34	30	0.56
macroscopic hematuria	6	8	
Blood pressure (mmHg)			
systolic	113 ± 20	115 ± 12	0.99
diastolic	63 ± 12	66 ± 12	0.32
Urinary protein excretion (g/d)	1.35 ± 1.01	1.02 ± 1.00	0.03
Hematuria in morning urine ^b	2.9 ± 0.8	2.6 ± 0.9	0.16
Blood urea nitrogen (mg/dl)	15.5 ± 5.6	13.9 ± 2.8	0.07
Serum creatinine (mg/dl)	0.64 ± 0.27	0.56 ± 0.18	0.41
Creatinine clearance (ml/min per 1.73 m ² of body surface area)	144 ± 52	152 ± 47	0.35
Serum IgA (mg/dl)	290 ± 115	276 ± 98	0.75
Glomeruli showing sclerosis (%)	5.7 ± 8.6	4.1 ± 5.8	0.65
Glomeruli showing crescents (%)	23.6 ± 20.7	21.2 ± 17.9	0.68
Glomeruli showing capsular adhesions (%)	9.0 ± 10.0	7.0 ± 7.1	0.40
No. of patients in whom ≥50% of glomeruli showed sclerosis, crescents, or adhesions	23	21	1.00

^a Plus-minus values are means ± SD.

^b Hematuria was quantified using dipsticks and macroscopic hematuria was quantified as 4.

36

3. 結果

- The mean urinary protein excretion fell in group 1 patients ($P < 0.0001$), but remained unchanged in group 2 patients.
- The mean serum IgA concentration was reduced in group 1 patients ($P < 0.0002$), but was unchanged in group 2 patients.
- BP and creatinine clearance were normal at the end of the trial in all but one group 2 patient, who developed chronic renal insufficiency.
- The percentage of glomeruli showing sclerosis was unchanged in group 1 patients, but increased in group 2 patients ($P < 0.006$).
- The intensity of mesangial IgA deposits decreased in group 1 patients ($P < 0.02$), but remained unchanged in group 2 patients.

37

4. 結論

- In conclusion, the present study shows that treatment of children with severe IgA nephropathy with prednisolone, azathioprine, heparin-warfarin, and dipyridamole for 2 yr early in the course of disease reduces immunologic renal injury and prevents increase of sclerosed glomeruli.

38

この研究へのコメント

- ランダム化して比較可能な状況にしているにもかかわらず、背景情報以外の群間比較は行なっていません。
- プライマリーエンドポイントと思われる解析は、尿への蛋白排泄量の変化に関するベースラインからの治療前後変化です (Wilcoxon signed rank test)。
- しかし、ベースラインのUrinary protein excretionに群間で差があります ($p < 0.03$)。
- このため、望ましくは、(データの分布に問題なければ)ベースラインを共変量とした共分散分析を行ったほうが良かったと思います。

39

Long-Term Results of a Randomized Controlled Trial in Childhood IgA Nephropathy

- Clin J Am Soc Nephrol. 2011 Jun;6(6):1301-7

40

Background and objectives

- Children with IgA nephropathy showing diffuse (>80%) mesangial proliferation are at high risk for end-stage renal failure (ESRF).
- A previous controlled trial showed that combination therapy consisting of prednisolone, azathioprine, heparin-warfarin, and dipyridamole early in the course of disease reduces immunologic renal injury and prevents the progression of sclerosed glomeruli.
- The objective of this study was to evaluate the long-term effectiveness of combination therapy in children with IgA nephropathy showing diffuse mesangial proliferation.

41

Design, setting, participants, & measurements

- A secondary analysis of a multicenter, randomized, controlled trial involving 78 children with IgA nephropathy who received either 2-year combination therapy or heparin-warfarin and dipyridamole (control) therapy was conducted.

42

Clinical and pathological findings at the start of initial treatments

Table 1. Clinical and pathological findings at the start of initial treatments			
	Combination Therapy (n = 40)	Heparin-Warfarin and Dipyridamole Therapy (n = 38)	P
Gender (Men/Women)	22/18	29/9	0.06
Age at diagnosis (years)	12.2 ± 3.0	11.6 ± 2.3	0.10
Initial presentation (Chance/Gross hematuria)	34/6	30/8	0.56
Creatinine clearance (ml/min per 1.73 m ²)	144 ± 52	152 ± 47	0.35
Serum albumin (g/dl)	3.73 ± 0.65	3.94 ± 0.59	0.36
Serum cholesterol (mg/dl)	210 ± 116	194 ± 47	0.13
Serum IgA (mg/dl)	292 ± 115	276 ± 98	0.75
Urinary protein excretion (g/1.73 m ² /day)	2.09 ± 1.78	1.35 ± 1.26	0.02
Hematuria ^a	2.9 ± 0.8	2.6 ± 0.9	0.16
Glomeruli			
showing sclerosis (%)	5.7 ± 8.6	4.1 ± 5.8	0.65
showing crescents (%)	23.6 ± 20.7	21.2 ± 17.9	0.68
showing capsular adhesions (%)	9.5 ± 10.0	7.0 ± 7.1	0.40
showing mesangial proliferations (%)	91.7 ± 7.4	90.2 ± 7.0	0.11
Intensity of mesangial IgA deposits ^b	2.2 ± 0.6	2.3 ± 0.5	0.72

^aThe plus-minus values are the means ± SD.
^bHematuria was quantified using dipsticks, and macrohematuria was quantified as 4.
^cThe intensity of deposits on immunofluorescence microscopy was graded semiquantitatively on a scale from 0 to 3+: 0, no; 1+, slight; 2+, moderate; 3+, intense.

43

Clinical and pathological findings at the end of initial 2-year treatments

Table 2. Clinical and pathological findings of at the end of initial 2-year treatments			
	Combination Therapy (n = 40)	Heparin-Warfarin and Dipyridamole Therapy (n = 34)	P
Creatinine clearance (ml/min per 1.73 m ²)	147 ± 33	145 ± 44	0.89
Serum IgA (mg/dl)	229 ± 87	281 ± 92	0.33
Urinary protein excretion (g/1.73 m ² per day)	0.28 ± 0.36	1.07 ± 1.57	0.01
Patients with heavy proteinuria (≥1.0 g/3 m ² per day)	2 (5.0%)	10 (23.5%)	0.209
Patients without proteinuria (<0.2 g/1.73 m ² per day)	20 (50.0%)	6 (17.6%)	0.007
Hematuria ^a	0.5 ± 1.0	1.5 ± 1.1	0.0002
Glomeruli			
showing sclerosis (%)	5.0 ± 6.9	16.4 ± 23.0	0.07
showing crescents (%)	0.4 ± 1.1	4.4 ± 10.3	0.16
showing capsular adhesions (%)	8.1 ± 11.2	6.4 ± 10.4	0.86
showing mesangial proliferations (%)	30.7 ± 26.7	57.5 ± 36.4	0.07
Intensity of mesangial IgA deposits ^b	1.3 ± 1.1	2.2 ± 0.6	0.02

^aThe plus-minus values are the means ± SD.
^bHematuria was quantified using dipsticks, and macrohematuria was quantified as 4.
^cThe intensity of deposits on immunofluorescence microscopy was graded semiquantitatively on a scale from 0 to 3+: 0, no; 1+, slight; 2+, moderate; 3+, intense.

44

Outcome definitions

- 主要評価項目
 - 腎置換療法を要するend-stage renal failure (ESRF)
- 副次評価項目
 - 腎不全: e-GFR<60ml/min per 1.73m²
 - 尿蛋白
 - 最初の2年: g/1.73m²
 - その後: early morning protein/creatinin ratio(uP/Cr [g/g])
 - 判定
 - Heavy: ≥=1.0 g/1.73m² or ≥=1.0 g/g
 - Mild: 0.2 to 1.0 g/1.73m² or 0.2 to 1.0 g/g
 - No protein urea: <0.2 g/1.73m² or <0.2 g/g

45

統計解析 Statistical analysis

- 連続変数 (continuous variables)
 - Mann-Whitney U-test (=Wilcoxon rank-sum test)
- カテゴリカル変数 (categorical variables)
 - Fisher's exact test
- 生存時間解析
 - 生存時間曲線: Kaplan-Meier法
 - 生存時間群間比較: log-rank test
 - 多変量解析: Cox比例ハザードモデル
 - 変数選択: Stepwise法(p<0.20)
- 有意水準 p<0.05 (両側?)
- 解析ソフト
 - JMP8.0 (SAS institute Japan Ltd., Tokyo, Japan)

46

Treatments used after the initial 2-year study period

Table 3. Treatments used after the initial 2-year study period		
	Combination Therapy (n = 40)	Heparin-Warfarin and Dipyridamole Therapy (n = 34)
Combination		
with ACEI or ARB	0 (0.0%)	2 (5.9%)
without ACEI or ARB	0 (0.0%)	5 (14.7%)
Steroid and immunosuppressant		
with ACEI or ARB	3 (7.5%)	0 (0.0%)
without ACEI or ARB	3 (0.0%)	1 (2.9%)
Steroid		
with ACEI or ARB	6 (15.0%)	1 (2.9%)
without ACEI or ARB	1 (2.5%)	0 (0.0%)
ACEI or ARB (partly with anti-platelets and/or warfarin and/or Sairei-to)	11 (27.5%)	3 (8.8%)
Anti-platelets and/or warfarin and/or Sairei-to	6 (15.0%)	4 (11.8%)
Warfarin alone	0 (0.0%)	1 (2.9%)
Sairei-to alone	6 (15.0%)	4 (11.8%)
No treatment	7 (17.5%)	9 (26.5%)
Unknown	0 (0.0%)	4 (11.8%)

Combination treatment included prednisolone, azathioprine, heparin-warfarin, and dipyridamole. The immunosuppressant used was azathioprine or mizoribine. Sairei-to is a Chinese herbal medicine.

47

Renal outcome at the last observation

Table 4. Renal outcome at the last observation		
	Combination Therapy (n = 40)	Heparin-Warfarin and Dipyridamole Therapy (n = 34)
No proteinuria (uP/Cr <0.2 g/g)	24 (60.0%)	18 (52.9%)
Mild proteinuria (≤0.2 uP/Cr <1.0 g/g)	12 (30.0%)	6 (17.6%)
Heavy proteinuria (uP/Cr ≥1.0 g/g)	1 (2.5%)	4 (11.8%)
Renal insufficiency (eGFR <60 ml/min per 1.73 m ²)	1 (2.5%)	1 (2.9%)
ESRF (requiring renal replacement therapy)	2 (5.0%)	5 (14.7%)

48

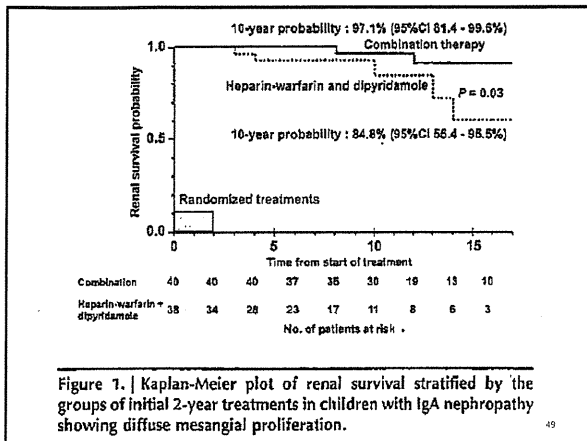


Figure 1. | Kaplan-Meier plot of renal survival stratified by the groups of initial 2-year treatments in children with IgA nephropathy showing diffuse mesangial proliferation.

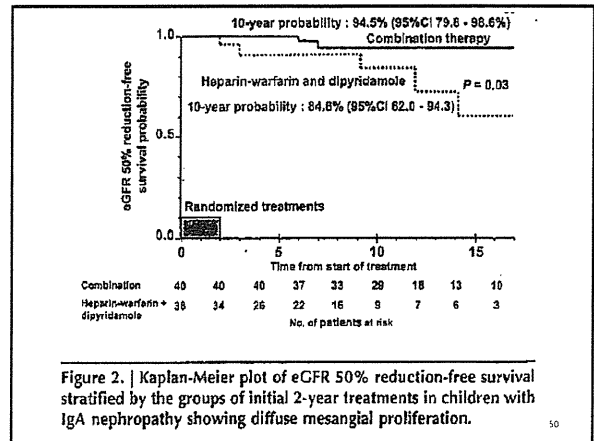


Figure 2. | Kaplan-Meier plot of eGFR 50% reduction-free survival stratified by the groups of initial 2-year treatments in children with IgA nephropathy showing diffuse mesangial proliferation.

Predicted risk factors for ESRF

Table S. Predicted risk factors for ESRF

	Univariate			Multivariate		
	HR	95% CI	P	HR	95% CI	P
Initial treatment (combination or control)	0.2	0.0 to 0.9	0.03	0.03	0.0 to 0.3	0.002
Urinary protein excretion at diagnosis (g/1.73 m ² per day)	2.3	1.0 to 5.0	0.06	2.7	1.1 to 6.5	0.03
Glomeruli showing global sclerosis at diagnosis (%)	1.1	1.0 to 1.1	0.04	1.1	1.0 to 1.2	0.02

HR, hazard ratio.

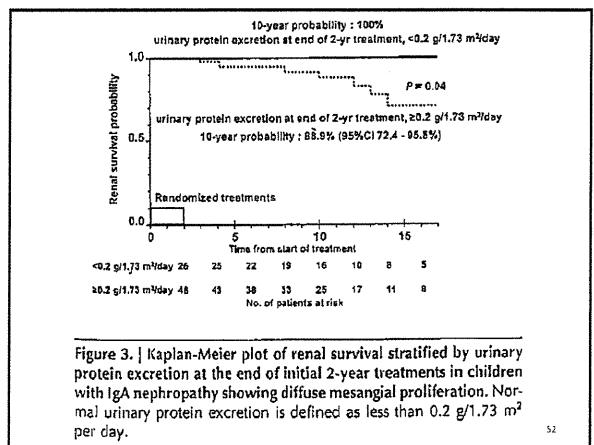


Figure 3. | Kaplan-Meier plot of renal survival stratified by urinary protein excretion at the end of initial 2-year treatments in children with IgA nephropathy showing diffuse mesangial proliferation. Normal urinary protein excretion is defined as less than 0.2 g/1.73 m² per day.

Results

- The median duration of observation was 10 years (range, 0.5 to 18). Two of 40 patients (5%) who received combination therapy and five of 34 patients (14.7%) who received control therapy developed ESRF.
- A Kaplan-Meier plot of renal survival showed that the outcomes of patients in the combined therapy group were better than those in the control therapy group (log-rank $P = 0.03$).
- The 10-year renal survival probability of each group was 97.1% (95% confidence interval, 81.4 to 99.6%) and 84.8% (95% confidence interval, 55.4 to 95.5%), respectively.
- The Cox proportional hazards model showed that the 2-year combination therapy was significantly associated with renal survival in both univariate and multivariate analyses.

Conclusions

- Two-year combination therapy not only ameliorated the activity of the acute phase of nephritis but also improved the long-term outcome of severe childhood IgA nephropathy.

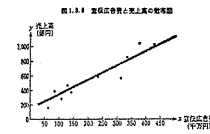
予後因子解析の論文を読み解くための予備知識

生物統計学セミナー
第3回
Feb. 28, 2012

宣伝広告費から売上高を予測する

- 一流企業10社における宣伝広告費と売上高

社名	宣伝広告費 (百万円)	売上高 (億円)
1	107	200
2	259	853
3	223	259
4	92	509
5	81	156
6	278	1927
7	129	400
8	313	653
9	142	272
10	428	1020



- いったいいくらぐらい宣伝や広告のために費用を使えば良いか?
- 宣伝広告費と売上高との関係調べて、宣伝広告費から売上高を予測したい。

参考: 石村真矢, 統計解析のはなし

回帰直線の計算

【基本手順】

手順1) 回帰直線を求めるため、あらかじめ次の表を作成する。

社名	x	y	x ²	xy	y ²
1	107	200	11449	21400	40000
2	259	853	67081	220929	728609
3	223	259	49729	57757	67041
4	92	509	8464	46828	259081
5	81	156	6561	12636	24336
6	278	1927	77324	536406	3713829
7	129	400	16641	51600	160000
8	313	653	97969	203489	426409
9	142	272	20164	38624	74084
10	428	1020	183184	436560	1040400
合計	2107	7720	649921	1822200	3222200



手順2) 回帰係数(傾き)と切片を次の式で計算する。

$$b = \frac{N \sum xy - (\sum x)(\sum y)}{N \sum x^2 - (\sum x)^2}$$

$$a = \frac{(\sum y) - b(\sum x)}{N}$$

【数値例】

社名	x	y	x ²	xy	y ²
1	107	200	11449	21400	40000
2	259	853	67081	220929	728609
3	223	259	49729	57757	67041
4	92	509	8464	46828	259081
5	81	156	6561	12636	24336
6	278	1927	77324	536406	3713829
7	129	400	16641	51600	160000
8	313	653	97969	203489	426409
9	142	272	20164	38624	74084
10	428	1020	183184	436560	1040400
合計	2107	7720	649921	1822200	3222200

$$b = \frac{10 \times 1822200 - 2107 \times 7720}{10 \times 649921 - 2107^2} = 2.146$$

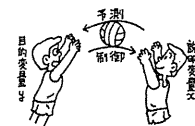
$$a = \frac{7720 - 2.146 \times 2107}{10} = 99.027$$

よって、回帰方程式は $y = 2.146x + 99.027$ となった。

参考: 石村真矢, 統計解析のはなし

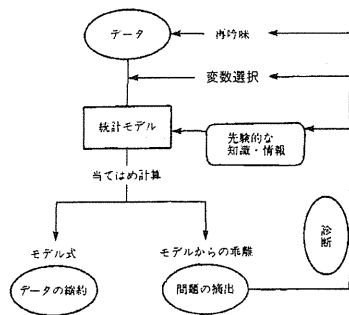
回帰分析の応用

- 予測と制御
 - 宣伝広告費を200(千万円)にすると売上高は、次の通り528,227(億円)と予測される。
 $2.146 \times 200 + 99.027 = 528.227$ (億円)
 - 売上高を700(億円)に伸ばそうと思えば、以下の方程式をとり、宣伝広告費を280(千万円)にすればよい(制御)。
 $700 = 2.146x + 99.027$



参考: 石村真矢, 統計解析のはなし

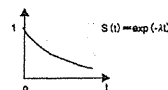
モデルのあてはめと診断



(参考: 大槻, 浜田, 生存時間解析 - SASによる生物統計 -)

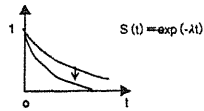
生存関数

- t年後の累積生存率: $S(t)$
 - $S(t) = \exp(-\lambda t)$ [λ : ハザード]
 - 累積生存率のことを、「生存関数: $S(t)$ 」と呼ぶ
 - 1年後生存率 $= \exp(-\lambda \cdot 1) = 0.95$ を解くと, $\lambda = 0.05129$
 - λ が小さい場合は, $\exp(-\lambda t) \sim 1 - \lambda t$ と近似できる
 - ハザード(λ)は瞬間速度
 - 単位時間あたりの死亡率



ハザードと生存曲線の関係

- ハザード(λ)がk倍になると
 - $\exp(-k\lambda t) = (\exp(-\lambda t))^k$
 - ハザードが2倍 ($\lambda_1/\lambda_2=2$)=生存関数の2乗
 - $0.95^2 = 0.9025$ (≒死亡率が0.05から0.10に)
 - $0.6^2 = 0.36$ (≒死亡率が0.4から0.64に)



7

比例ハザードモデル

- Cox回帰の基礎となる統計モデルは、比例ハザードモデルで、ハザード関数($hazard(t, x)$)として以下の式を想定する。

$$hazard(t, x) = hazard(t) \times \psi(x)$$

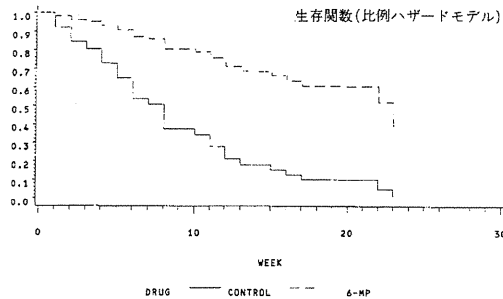
$$\psi(x) = \exp(\beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k)$$

$$= \exp(\beta_1 x_1) \exp(\beta_2 x_2) \dots \exp(\beta_k x_k)$$
- ここで、 $hazard(t)$ は基準ハザード関数、 $\psi(x)$ は共変量を意味する。つまり、ハザードを時間に関する部分と共変量の部分に分解して構成している。
- θ_i が推定すべき未知パラメータ。例えば、 $x_i=1$ (試験治療)又は0(対照群)とし、解析の結果 $\theta < 0$ であれば試験治療の効果あり、 $\theta=0$ ならば試験治療の効果なしとなる。また、 $(\theta/\text{標準誤差})^2 > 3.84$ ならば5%有意と判断される。
- ハザード比= $\exp(\beta)$ で計算される。

8

比例ハザード性

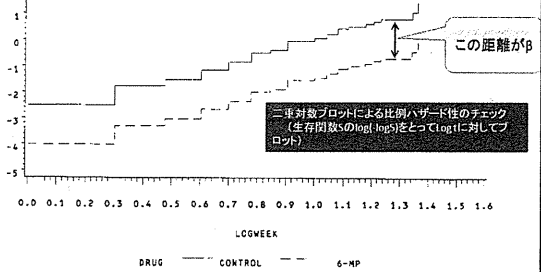
図表 3.2.5 生存曲線とその2重対数のプロット(GEHA)



9

比例ハザード性(つづき)

2重対数プロット(比例ハザードモデル)



10

変数選択

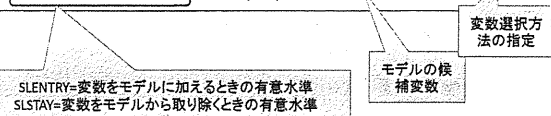
- Stepwise法(変数増減法)
 - ある段階のモデルに対して、次の新しい変数を取り入れるか、あるいはモデルにすでに入っている変数を落とすかを逐次的に行う方法。
- Score法(総当り法)
 - 候補としていた変数の組み合わせでできる可能なすべてのモデルから、モデルの良さを表す統計量(Phregではラオのスコア χ^2 乗統計量)の大きさに基づいてモデルを選択する方法。

11

Stepwise法

プログラム 3.4.1 PHREG プロシジャによる変数増減法の変数選択

```
PROC PHREG DATA=PCANCER;
  MODEL TIME=CENSOR(1)=AGE SEX TREAT BUI CH P STAGE PS / SELECTION=STEPWISE
  SLENTRY=0.20 SLSTAY=0.20 DETAILS; RUN;
```



12

Stepwise法(続き)

一番最初は何もモデルに取り込まれていない

出力 3.4.1 変数増減法による変数選択 1)

Analysis of Variables Not in the Model

Variable	Score	Chi-Square	Pr >
AGE	0.2742	0.2742	0.6005
SEX	0.0021	0.0021	0.9631
TREAT	5.6713	5.6713	0.0172
BUI	2.0415	2.0415	0.1531
CH	0.0198	0.0198	0.8882
P	2.8756	2.8756	0.0899
STAGE	3.7167	3.7167	0.0539
PS	0.5560	0.5560	0.4559

Residual Chi-square = 16.3798 with 8 DF (p=0.0373)

SLENTRY=0.20 よりも小さいので TREATが選択される

Stepwise法(続き)

出力 3.4.2 変数増減法による変数選択 2)

Step 1: Variable TREAT is entered. The model contains the following explanatory variables.

TREAT

Testing Global Null Hypothesis: BETA=0

Criterion	Without Covariates	With Covariates	Model Chi-Square
-2 LOG L	570.430	565.361	5.069 with 1 DF (p=0.0244)
Score			5.671 with 1 DF (p=0.0172)
Wald			5.507 with 1 DF (p=0.0189)

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr >	Risk Ratio
TREAT	1	-0.610339	0.26009	5.50658	0.0189	0.543

SLSTAY=0.20 よりも小さいのでTREATはモデルに残される

モデルに取り込まれる変数とモデルから取り除かれる変数の両方がなくなった段階で変数選択終了。

Score法

プログラム 3.4.2 PHREG プロシジャによる総当たり法の変数選択

```
PROC PHREG DATA=PCANCER;
MODEL TIME=CENSOR(1) AGE SEX TREAT BUI CH P STAGE PS
/SELECTION=SCORE BEST=2;RUN;
```

図表 3.4.1 候補変数が8の場合の可能なモデルの数

変数選択方法の指定	モデルに含まれる変数の数	可能なモデルの数
	0	1
	1	8
	2	28
	3	56
	4	70
	5	56
	6	28
	7	8
	8	1
合計		256

すべての変数の組み合わせのモデルについて当てはまりの良さを示すスコア、二乗統計量を計算する

Score法(続き)

出力 3.4.9 総当たり法による変数選択の結果

Regression Models Selected by Score Criterion

Number of Variables	Score Value	Variables Included in Model
1	5.6713	TREAT
1	3.7167	STAGE
2	9.4148	TREAT BUI
2	9.3957	TREAT STAGE
3	11.5226	TREAT BUI STAGE
3	10.8750	TREAT BUI P
4	12.8568	AGE TREAT BUI STAGE
4	12.6253	AGE TREAT BUI P
5	14.0891	AGE TREAT BUI P STAGE
5	13.5306	AGE TREAT BUI P PS
6	15.2356	AGE TREAT BUI P STAGE PS
6	14.3531	AGE SEX TREAT BUI P STAGE
7	15.8855	AGE TREAT BUI CH P STAGE PS
7	16.3910	AGE SEX TREAT BUI P STAGE PS
8	16.3798	AGE SEX TREAT BUI CH P STAGE PS

Score二乗値は8変数で最大。しかし、どのモデルを選んだらよいかの指標にはならない。

異なる変数の数同士を比較する際には、AICを計算して利用する。(AIC=Akaike Information Criterion)

Score法(続き)

図表 3.4.2 AICによるモデル選択

Number of Variables	Score Value	Variables Included in Model	-2 LOG L	AIC
1	5.6713	TREAT	565.361	567.361
2	9.4148	TREAT BUI	561.574	565.574
3	11.5226	TREAT BUI STAGE	559.434	563.434
4	12.8568	AGE TREAT BUI STAGE	558.310	566.310
5	14.0891	AGE TREAT BUI P STAGE	557.075	567.075
6	15.2356	AGE TREAT BUI P STAGE PS	556.200	568.200
7	15.8855	AGE TREAT BUI CH P STAGE PS	555.755	569.755
8	16.3798	AGE SEX TREAT BUI CH P STAGE PS	555.377	571.377

AIC=2×(最尤推定値の元での対数尤度)+2×(変数の数)
AICが小さいほうが良いモデル
3変数(TREAT BUI STAGE)が最もAICが小さいモデル
(ただし、最終的には医学的な解釈も含めてモデルを決定する)

Long-Term Results of a Randomized Controlled Trial in Childhood IgA Nephropathy (Cox重回帰)

Clin J Am Soc Nephrol. 2011 Jun;6(6):1301-7

Statistical analysis

- (前文略)
- For multivariate analysis, we utilized the Cox proportional hazard model in the stepwise fashion, entering and eliminating variables with a P value of 0.20.
- (後文略)

19

Predicted Risk for ESRF

- We evaluated predicted risk factors using univariate and multivariate analyses using the Cox proportional hazard model for ESRF (end-stage renal failure).
- The initial 2-year treatments (combination or control) and ratios of glomeruli showing global sclerosis at diagnosis were included as significant risk factors in both the univariate and multivariate analysis for ESRF (Table 5).
- The urinary protein excretion at diagnosis was significantly included as a risk factor in the multivariate analysis for ESRF (Table 5).
- The ratios of glomeruli showing crescents and mesangial proliferations at diagnosis were not selected as significant risk factors for ESRF with a multivariate Cox model in a stepwise fashion.

20

Predicted risk factors for ESRF

Table 5. Predicted risk factors for ESRF

	Univariate			Multivariate		
	HR	95% CI	P	HR	95% CI	P
Initial treatment (combination or control)	0.2	0.0 to 0.9	0.03	0.03	0.0 to 0.3	0.002
Urinary protein excretion at diagnosis (g/1.73 m ² per day)	2.3	1.0 to 5.0	0.06	2.7	1.1 to 6.5	0.03
Glomeruli showing global sclerosis at diagnosis (%)	1.1	1.0 to 1.1	0.04	1.1	1.0 to 1.2	0.02

HR, hazard ratio.

21