

試験デザインとエビデンスの解釈



(財)パブリックヘルスリサーチセンター臨床研究支援事業(CSP)
 NPO日本臨床研究支援ユニット理事長
 スタットコム(株)取締役会長
 NPO日本メディカルライタ協会理事長
 (社)日本臨床試験研究会代表理事

東京大学医学系研究科 公共健康医学専攻 生物統計学

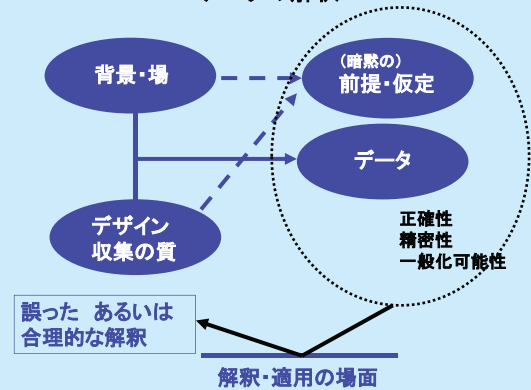
大橋靖雄

試験デザインとは混沌たる現実からrelevantな情報を切り出すための型あるいはフィルターであり、一方、複雑な現実を解釈するための枠組みと同時に限界を提供するものである

内容

- ◆ 説明的試験と実践的試験 MEGA
- ◆ 多因子要因実験とその解釈 NSAS BC02
 仮説検証とサンプルサイズ設計
- ◆ 非劣性試験とパラメータへの要約 Iressa製剤後試験
- ◆ 医療情報の適切な発信とは？ しばらく前の降圧薬試験

データの解釈

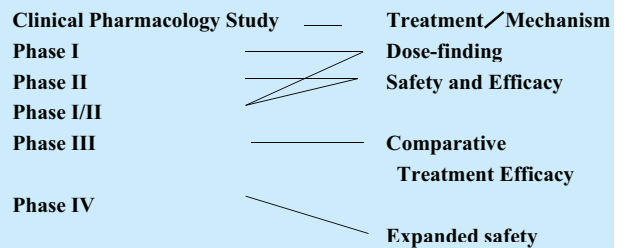


UMINでの解説

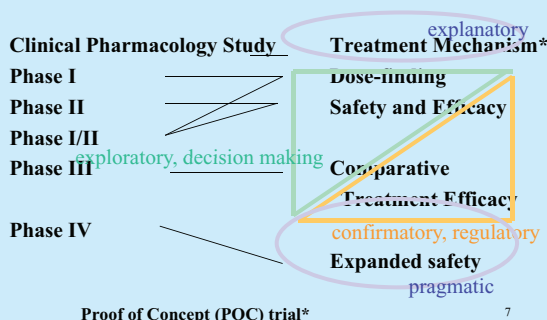
- ◆ 検証的 Confirmatory すでに探索的試験などにより仮説が形成されており、その仮説を検証するために実施される試験
- ◆ 探索的 Exploratory 検証的試験の実施の前に、仮説を形成するために実施する試験
- ◆ 説明的 Explanatory 介入法の作用機序などを解明する目的で、実施条件をある程度厳しく設定して実施する試験
- ◆ 実践的 Pragmatic 実施条件をゆるく設定し、日常診療に近い状況で介入法を評価するために実施する試験

Clinical Trial: Design Types

Piantadosi A(1997), *Clinical Trials*, Wiley



Statistical Approach in Drug Development



Explanatory vs Pragmatic

| Explanatory | Pragmatic | |
|----------------------------|---------------------------------|----------------|
| 目的は明確で広く理解されている | 科学者にはよく理解されない | Famant et al. |
| 実験室での実験の拡張 | 臨床実践への指針を与える | Famant et al. |
| 仮説検定 | 治療方針選択 | Famant et al. |
| method-effectiveness | use-effectiveness | Piantadosi |
| | 両者はcompetesする | Piantadosi |
| analysis of compliers only | analysis by intention-to-treat | Pocock |
| PPS | FAS | ICH E9 |
| 欠損での処理は what if ? で | total impactを評価 | Redmond&Colton |
| proof of concept 試験 | simple and large-scale evidence | |

例: 実践的試験と説明的試験

- ◆ 肺癌患者に対する術前化学療法の効果を比較したい。当然ランダム化試験を行うものとする。
- ◆ 純粋に化学療法の効果を評価したいとすれば、割り付けを行って一方の群には化学療法の後に手術を、一方の群には化学療法施行時間は待機してその後手術をすべきである。発癌処理を行ったネズミに対する動物実験ならこうすべきである。
- ◆ しかし臨床の場合なら手術群はすぐ手術を受けるべきであろう。倫理的な問題もあるし、化学療法単独の効果を抽出することよりも、治療方針を実践の場で評価することが目的であるから。

Simple and large scale evidence 厳しい選択条件がよい試験の条件とは限らない

ATLAS Adjuvant Tamoxifen Longer and Shorter

適格規準 explanatoryなら

乳癌であることが組織学的に確認され治癒切除された、ホルモン陽性、年齢何歳、リンパ節転移は問わない、現在までにtamoxifenを4.5年以上服用し、これこれの異常がなく、.....

pragmaticな本試験では

5年間tamoxifenを服用し、続けるべきか止めるべきか決定できない医師および乳癌術後患者

Nakamura et al.,
Lancet 2006; 368:
1155-63.

Primary prevention of cardiovascular disease with pravastatin in Japan (MEGA Study): a prospective randomised controlled trial

Hiroyuki Nakamura, Kiyohito Arakawa, Hirohiko Nakaya, Akira Kikuchi, Yoshio Goto, Takayoshi Toyama, Noriaki Nakaya, Shoji Nakahara, Masaharu Mizumaki, Akira Yamamoto, Kazuhiko Mizuno, Yasuo Ohshiki for the MEGA Study Group

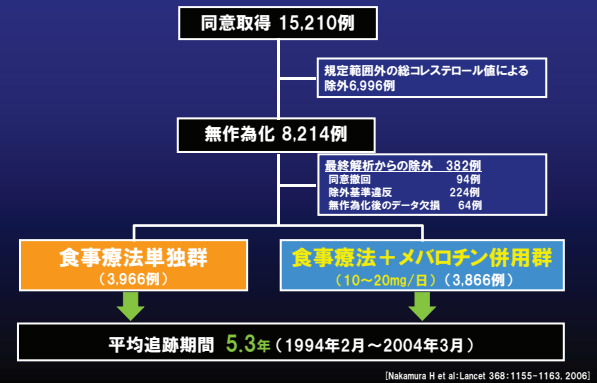
Summary
Background Evidence-based treatment for hypercholesterolemia in Japan has been hindered by the lack of direct evidence in this population. Our aim was to assess whether evidence for treatment with statins derived from western populations can be extrapolated to the Japanese population.

Methods In this prospective, randomised, open-labelled, blinded study, patients with hypercholesterolemia (total cholesterol 5.09-6.98 mmol/l) and no history of coronary heart disease or stroke were randomly assigned diet or diet plus 10-20 mg pravastatin daily. The primary endpoint was the first occurrence of coronary heart disease. Statistical analyses were done by intention to treat. This trial is registered at ClinicalTrials.gov, number NCT00211765.

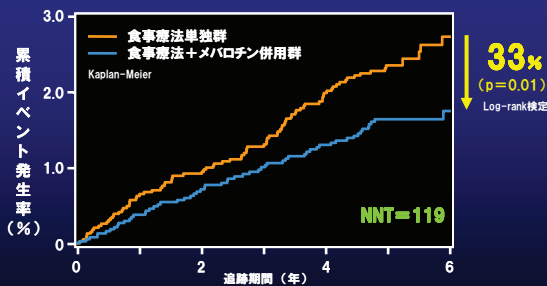
Findings 3966 patients were randomly assigned to the diet group and 3866 to the diet plus pravastatin group. Mean follow-up was 5.3 years. At the end of study, 471 and 522 patients had withdrawn, died, or been lost to follow-up in the diet and diet plus pravastatin groups, respectively. Mean total cholesterol was reduced by 2.1% (from 6.27 mmol/l to 6.13 mmol/l) and 11.5% (from 6.27 mmol/l to 5.55 mmol/l) and mean LDL cholesterol by 3.2% (from 4.05 mmol/l to 3.96 mmol/l) and 18.0% (from 4.05 mmol/l to 3.33 mmol/l) in the diet and the diet plus pravastatin groups, respectively. Coronary heart disease was significantly lower in the diet plus pravastatin group than in the diet alone group (66 events vs 101 events; HR 0.67, 95% CI 0.49-0.91; p=0.01). There was no difference in the incidence of malignant neoplasms or other serious adverse events between the two groups.

Interpretation Treatment with a low dose of pravastatin reduces the risk of coronary heart disease in Japan by much the same amount as higher doses have shown in Europe and the USA.

試験フローチャート



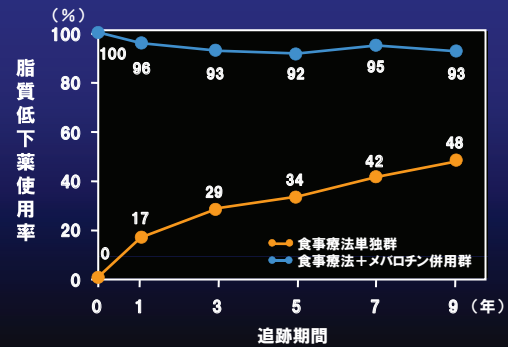
一次評価項目 - 冠動脈疾患 (CHD) -



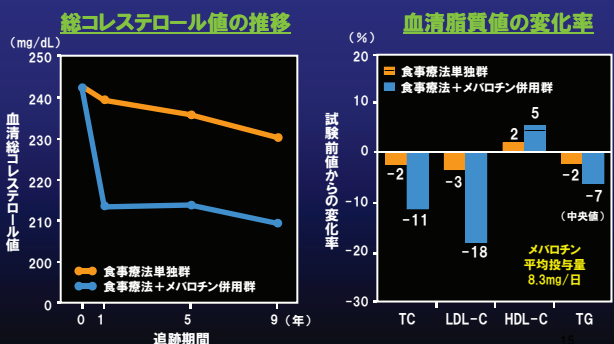
| リスク症例数 | 0 | 1 | 3 | 5 | 6 | | |
|---------------|-------|-------|-------|-------|-------|-------|-----|
| 食事療法単独群 | 3,966 | 3,758 | 3,648 | 3,529 | 3,430 | 2,476 | 830 |
| 食事療法+メバロチン併用群 | 3,866 | 3,642 | 3,490 | 3,385 | 3,307 | 2,434 | 859 |

[Nakamura H et al: Lancet 368:1155-1163, 2006]

患者の脂質低下薬の使用状況



血清脂質値の変化



モニターに寄せられた(回答困難なやや意地の悪い) MEGAに関する質問

- ◆ MEGAは2重盲検を採用していないので、バイアスが入っているのではないかと。2重盲検の海外のエビデンスに比べ質が落ちるのではないかと?
- ◆ MEGAではITTの原則に従うといっておきながら、割付後に除外を行っている。恣意的に結果を薬物群に有利なように解釈しているのではないかと?
- ◆ MEGAのエンドポイントのうち統計的に水準に達していないものがある(全期間での脳卒中、女性の結果など)。これらの点では試験は成功していないのではないかと。メバロチンの治療意義は確立していないのではないかと?
- ◆ 5年の時点で同意を取り直している。バイアスは生じていないのか?
- ◆ MEGAの結果のNNTは119である。治療効率は良くないのではないかと?

バイアスの要因とMEGA Studyでの対処

| 重要なバイアス | MEGAでは？ |
|--------------------|----------------|
| ・患者選択 解釈と一般化可能性 | 患者背景は一般診療を反映 |
| ・割り付け(ランダム化) | 中央登録と厳密な適用 |
| ・治療過程 | 試験に関する考え方の違い* |
| ・試験継続・中止の判断 | 基本的には同様 |
| ・評価(情報バイアス) 盲検化で対応 | 評価は盲検化(PROBE) |
| | 血管再建は議論の要ありか？ |
| ・交絡 ランダム化と一部解析で対応 | 厳密なランダム化、補助的解析 |
| ・データ管理 | 独立なデータセンター |
| ・統計解析 | 事前の解析計画書策定 |

*: explanatoryかpragmaticか

17

MEGAで証明された優越性とは何か？

- ◆ プラバスタチンの薬理的作用 × これは既に証明済み
- ◆ プラバスタチンを最初から用いるという治療方針
実は食事療法群で体重減少が達成された集団では
プラバスタチンより強いイベント抑制あり

18

Intensity Score法によるコンプライアンス考慮

Tanaka et al. Stat. in Med. 2008; 27: 1718-33.

Table IV. Estimates of treatment effect for CHD and stroke events.

| Method | CHD | | | | Stroke | | | |
|-------------------|---------|----------------|----------|----------------|---------|----------------|----------|----------------|
| | 5 years | | 10 years | | 5 years | | 10 years | |
| | HR | 95 per cent CI | HR | 95 per cent CI | HR | 95 per cent CI | HR | 95 per cent CI |
| ITT | 0.70 | 0.50, 0.97 | 0.67 | 0.49, 0.91 | 0.65 | 0.43, 0.97 | 0.83 | 0.57, 1.21 |
| Intensity score 1 | 0.68 | 0.44, 1.05 | 0.59 | 0.36, 0.99 | 0.44 | 0.25, 0.79 | 0.51 | 0.28, 0.95 |
| Intensity score 2 | 0.68 | 0.46, 1.02 | 0.66 | 0.27, 1.60 | 0.53 | 0.31, 0.90 | 0.45 | 0.17, 1.21 |
| g-estimation | 0.65 | 0.30, 0.91 | 0.64 | 0.39, 0.83 | 0.54 | 0.26, 0.87 | 0.63 | 0.33, 1.26 |

HR, hazard ratio; CI, confidence interval; intensity score 1, constant treatment effect from model (3); intensity score 2, cumulative treatment effect from model (7); g-estimation, semi-parametric randomization-based analysis using model (10).

今後はランダム化試験を補う解析がふつうに行われるであろう
コンプライアンスを上げられたら？
打ち切りが無かったら？ (IPCW)

19

MEGA Studyの意義

- ◆ (循環器領域でも) 日本からのエビデンス発信が可能であることを証明
- ◆ 日本人に対するプラバスタチンの効果を実験的試験の中で検証
- ◆ 臨床試験の実施体制・品質管理の重要性をアピール
- ◆ 日本人(アジア人)に対する用量設定の意義の強調
- ◆ 方法論開発に対する刺激
- ◆ 臨床試験結果の解釈に関する議論

20

内容

- ◆ 説明的試験と実践的試験 MEGA
- ◆ 多因子要因実験とその解釈 NSAS BC02
仮説検証とサンプルサイズ設計
- ◆ 非劣性試験とパラメータへの要約 Iressa製剤後試験
- ◆ 医療情報の適切な発信とは？ しばらく前の降圧薬試験

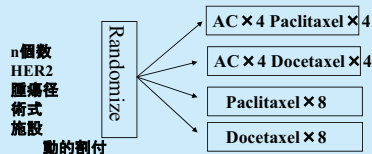
21

割付け(多因子要因実験)

N-SAS BC02

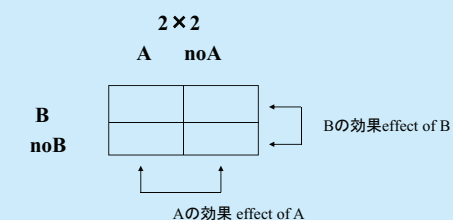
研究代表者 渡辺亨

- ◆ ホルモン療法低感受性、n+ 症例に対する術後補助療法
- ◆ エンドポイント
無病生存期間
神経毒性を中心としたHRQOL
- ◆ ACの有無、Taxaneの種類に関する2×2要因実験
- ◆ AC有無に関する非劣性non-inferiority試験、1200例



多因子要因実験

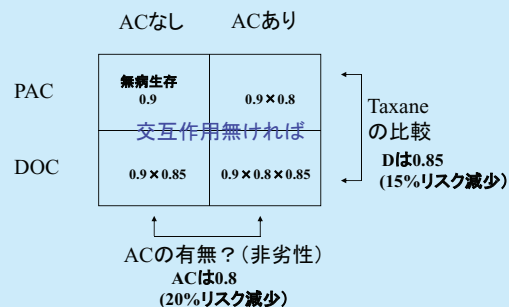
Factorial experiments



AとBの間に交互作用無ければ、2つの実験分の情報量

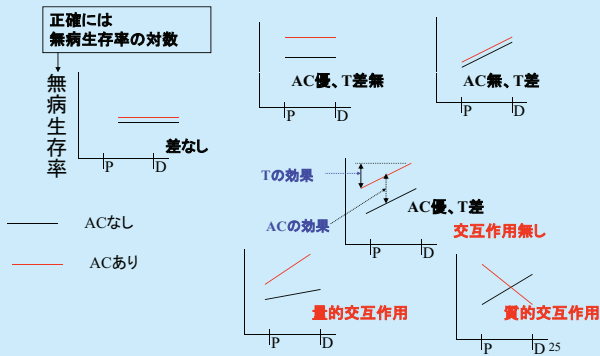
23

2×2要因実験: NSAS BC02



24

2×2要因実験: NSAS BC02



多因子要因実験の応用

Application of Factorial experiments

完全実施と一部実施 complete/fractional
工業・農業実験では常識化 線点図と直交表

臨床研究ではこれまでは稀
異なるメカニズムの薬剤、とくに予防研究

Blotによる中国臨牀のがん予防研究 cancer prevention trial
a,b,c,d 4薬剤群 0,ab,ac,ad,bc,bd,cd,abcdの1/2実施

異なる疾患(エンドポイント)

Physicians' Health Study
Aspirin(心筋梗塞MI予防) × β-carotene(癌予防)

推進派と慎重派

多因子要因実験の応用

Application of Factorial experiments

- 臨床試験での応用 Piantadosiの教科書から (Table 15.6)
- SWOG8300 Miller et al. 1995 PBI × Chemo for NSCLC
 - PHS Hennekens and Eberlein 1985
 - ATBC Heinonen et al. 1987
atocopherol × β-carotene for cancer prevention
 - ISIS4 ISIS4 Research Group 1996
captopril × mononitrate × magnesium for MI
 - Linxian Nutrition Trial Li et al. 1993 2⁴ for cancer prevention
 -
 - CALGBの最近の乳癌補助療法臨床試験
 - CALGB9344 Dose of AC × Taxol(P)
 - CALGB9741 Order of ACP × 2 or 3 weeks interval

1988Jan27 New York Times

Physicians' Health Study

心筋梗塞予防のアスピリン × 癌予防のβカロテン

22000名の医師が参加

中間解析によって
アスピリン部分の中止

Steering Committee of PHS,
NEJM 1987;321:129-35



2×2要因実験

- ◆ 交互作用ある場合(量的交互作用)には「平均」した効果を推定
- ◆ 交互作用の存在はエンドポイントをどう取るかに依存
生存時間をエンドポイントにした場合、通常は無病生存率の対数 ~ 対数ハザード
- ◆ 慎重派と積極派
「交互作用が存在すると厄介、その検出は困難」
「要因実験を行わなければそもそも交互作用判らない」
- ◆ 乳がん補助療法領域では要因実験好まれる
CALGB, NSABP

Green論文

Green et al., Factorial Design Considerations, JCO 2003; 20: 3424-30.

- ◆ Purpose
通常、多因子要因実験の解析は交互作用を無視、複数仮説検定の多重性も無視。多くの質問に答えられるというメリットと、どれくらいの頻度で誤った結論に達するかのトレードオフを検討する。
- ◆ Methods
シミュレーションによる検討。2×2の下、効果と交互作用の有り無しで幾つかのシナリオを用意。解析方法は1. 交互作用無視、2. まず交互作用を検定、3. 交互作用を検定さらに多重性を考慮。
- ◆ Results
どの方法も一概には優れることなし。ある設定では効率上昇、しかし別の状況ではいずれの方法も正しい結論に至る確率小。
- ◆ Conclusion
効率上昇の可能性はあるものの、良い状況かどうか予測することは困難。上記のトレードオフを考慮すべき。(あたしは嫌い。)₃₀

Phase III study of doxorubicin-cyclophosphamide followed by paclitaxel or docetaxel given every 3 weeks or weekly in patients with axillary node-positive or high-risk node-negative breast cancer: results of North American Breast Cancer Intergroup Trial E1199

Sparano JA¹, Wang M¹, Martino S², Jones V³, Perez EA⁴, Saphner T¹, Wolff AC¹, Sledge GW¹, Wood WC, & Davidson NE¹
From the Eastern Cooperative Oncology Group (ECOG)¹, Southwest Oncology Group (SWOG)², Cancer & Acute Leukemia Group B (CALGB)³, & the North Central Cancer Treatment Group (NCCTG)⁴.

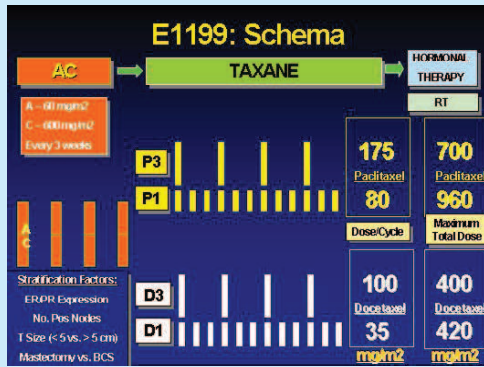


Eastern Cooperative Oncology Group



Background and Rationale

- Before E1199 study initiation
 - AC ⇒ Paclitaxel (q3 weeks) a reasonable standard
 - more effective than AC alone in May, 1998 (C9344)
 - Docetaxel more effective than paclitaxel?
 - Indirect comparisons of phase II & III trials in met. disease
 - Weekly taxane more effective than a every 3 week therapy?
 - Cytotoxic & antiangiogenic effects seen at low concentrations
 - Metronomic dosing schedules optimize antiangiogenic effects
 - Weekly taxane schedules feasible and effective
- After E1199 study completion
 - Docetaxel more effective than paclitaxel for metastatic disease (Ravdin, 2005)
 - Weekly paclitaxel more effective than every 3 week paclitaxel for metastatic disease (Seidman, 2004)



33

Patient Eligibility

- T 1-3, N1-2 or T2-3N0 invasive breast cancer
- Treated by lumpectomy or mastectomy + axillary dissection (at least 6 nodes removed) or negative sentinel node biopsy
- Her2/neu positive disease not excluded, but encouraged to enroll on N9831 (activated May, 2000)
- < 12 weeks since final surgical procedure
- Negative surgical margins (≥ 1 mm)
- Normal hematologic, hepatic, renal function
- No history of MI, CHF, or significant ischemic or valvular heart disease

34

Primary Endpoint and Primary and Secondary Comparisons

- **Primary endpoint:**
 - Disease free survival (DFS) – relapse, second primary breast cancer, or death without recurrence
- **Primary comparisons:**
 - Paclitaxel vs. docetaxel
 - Every 3 week vs. weekly schedule
- **Secondary exploratory comparisons:**
 - Standard P3 arm versus other experimental arms

35

Statistical Considerations

- **Primary DFS comparisons:**
 - 86% power to detect a 17.5% \downarrow in the failure HR for either primary comparison (2-sided p value of 0.05)
 - Corresponds to \uparrow 5-year DFS rate from 73.3% \Rightarrow 77.4%
 - Analysis projected after 1042 DFS events
- **Secondary DFS comparisons:**
 - 80% power to detect a 22% \downarrow in the failure HR for any of the experimental arms compared with the standard P3 arm (2-sided nominal 5% level tests, corrected for multiple comparisons)
 - Corresponds to \uparrow 5-year DFS rates from 73.3% \Rightarrow 78.5%
 - Analysis projected after 1400 DFS events
- **Secondary Overall Survival (OS) Endpoint:**
 - 81% power to detect an \uparrow 5-year OS from 81.0% to 84.0%

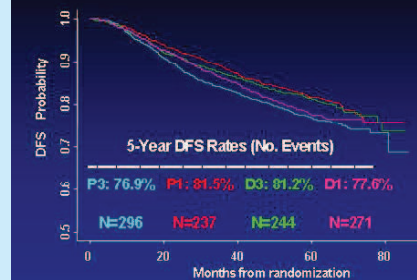
36

Primary Comparisons

| Comparison | HR | 95% C.I. | P Value |
|--------------------------|-------|------------|---------|
| Paclitaxel vs. Docetaxel | 1.032 | 0.91, 1.17 | 0.61 |
| Every 3 weeks vs. Weekly | 1.062 | 0.94, 1.20 | 0.33 |

37

Secondary Comparisons: Disease Free Survival



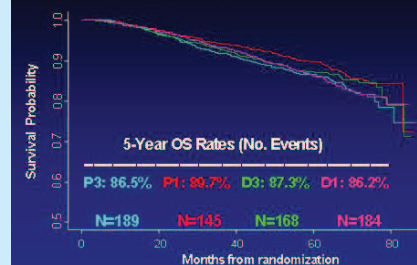
38

P3 Arm vs. Other Arms: Secondary DFS Comparisons

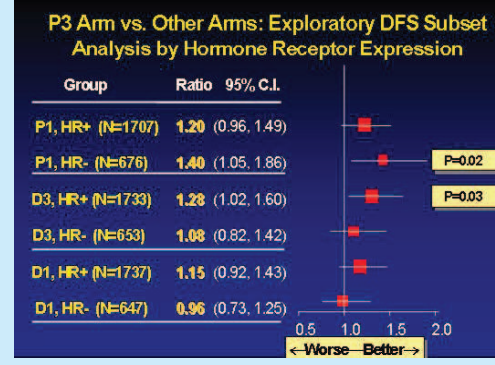
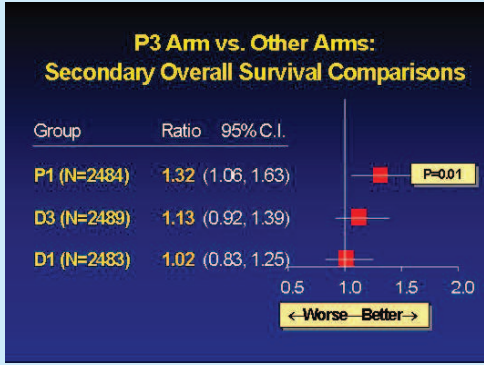


39

Secondary Comparisons: Overall Survival



40



Conclusions – Primary Comparisons

- Compared with paclitaxel
 - docetaxel did not improve DFS
- Compared with every 3 week therapy
 - weekly therapy did not improve DFS
- Comparisons confounded by unanticipated interaction between taxane & schedule



Conclusions – Secondary Comparisons Compared with Standard P3 Arm

- Efficacy
 - P1 and D3 arms associated with improved DFS
 - About 25% reduction in hazard rate
 - Similar trends observed in HR-positive disease
 - P1 arm associated with improved OS
 - About 30% reduction in hazard rate for death
- Toxicity
 - D3 arm associated with more severe neutropenia, febrile neutropenia, and infection
 - P1 arm associated with more neuropathy

Other Findings & Implications

- Results support the choice of AC ⇒ weekly paclitaxel in intergroup trial E5103 (chemotherapy +/- bevacizumab)
- Event rates lower than projected
 - 73% DFS projected in P3 arm compared with 77% observed
 - Difference corresponds to the treatment effect being sought
 - Consistent with recent reports showing declining event rates (Berry et al. NEJM 2005) and more indolent behavior for mammographically detected disease (Shen et al. JNCI, 2005)
 - Other contributing factors include concurrent N9831 trial (for Her2/pos disease), and use of A.I.s beginning in 2003
- Potential need for using molecular markers to enrich for high-risk subjects in subsequent trials
 - Goldstein et al (abstract 526) discussed at poster discussion session yesterday

臨床家は解釈にとまどい

- ◆ 有意差があったと判断していいの？ (secondary comparison)
- ◆ Regulatoryな試験なら検証とはいえないが...
- ◆ いずれの治療もほぼ日常臨床、処方には情報を活用すべき
- ◆ P1はP3より優れ、D3はD1より優れる示唆が既にあり(進行癌)
- ◆ 曖昧な解釈を避けるために、解析計画を手直すべきではなかったか？

Separano et al. NEJM 2008; 358: 1663-71

Weekly Paclitaxel in the Adjuvant Treatment of Breast Cancer

Joseph A. Sparano, M.D., Mollin Wang, Ph.D., Silvana Martino, D.O., Vicki Jones, M.D., Edith A. Perez, M.D., Tom Saphner, M.D., Antonio C. Wolff, M.D., George W. Sledge, Jr., M.D., William C. Wood, M.D., and Nancy E. Davidson, M.D.

ABSTRACT

BACKGROUND

We compared the efficacy of two different taxanes, docetaxel and paclitaxel, given either weekly or every 3 weeks, in the adjuvant treatment of breast cancer.

METHODS

We enrolled 4950 women with axillary lymph node-positive or high-risk, lymph node-negative breast cancer. After randomization, all patients first received 4 cycles of intravenous doxorubicin and cyclophosphamide at 3-week intervals and were then assigned to intravenous paclitaxel or docetaxel given at 1-week intervals for 4 cycles or at 1-week intervals for 12 cycles. The primary end point was disease-free survival.

RESULTS

As compared with patients receiving standard therapy (paclitaxel every 3 weeks), the hazard ratio for disease-free survival was 1.27 among those receiving weekly paclitaxel (P=0.006), 1.23 among those receiving docetaxel every 3 weeks (P=0.045), and 1.09 among those receiving weekly docetaxel (P=0.29) (with a hazard ratio 91 favoring the groups receiving experimental therapy). As compared with standard therapy, weekly paclitaxel was also associated with improved survival (hazard ratio, 1.32; P=0.01). An exploratory analysis of a subgroup of patients whose tumors expressed no human epidermal growth factor receptor type 2 protein found similar improvements in disease-free and overall survival with weekly paclitaxel treatment, regardless of hormone-receptor expression. Grade 2, 3, or 4 neuropathy was more frequent with weekly paclitaxel than with paclitaxel every 3 weeks (27% vs. 20%).

CONCLUSIONS

Weekly paclitaxel after standard adjuvant chemotherapy with doxorubicin and cyclophosphamide improves disease-free and overall survival in women with breast cancer. (ClinicalTrials.gov number, NCT00004125.)

From the Eastern Cooperative Oncology Group, Philadelphia (J.A.S., M.W., T.S., A.C.W., G.W.S., W.C.W., N.E.D.); the Southwest Oncology Group, Ann Arbor, MI (J.M.); the Cancer and Leukemia Group B, Chicago (V.J.); and the North Central Cancer Treatment Group, Rochester, MN (E.A.P.). Address reprint requests to Dr. Sparano at Memorial Sloan-Kettering Cancer Center, Memorial Division, 1275 York Ave., Box 208, 132 Eastchester Rd., 2 South, Box 40, Bronx, NY 10461, or at sparano@mskcc.org.

N Engl J Med 2008;358:1663-71.
Copyright © 2008 Massachusetts Medical Society.

DISCUSSION

This trial was designed to compare the efficacy of paclitaxel with that of docetaxel and to compare the standard taxane schedule (every 3 weeks) with a weekly schedule in nearly 5000 women with axillary lymph node-positive or high-risk, lymph node-negative breast cancer. We found no significant differences in survival between the groups treated with paclitaxel and those treated with docetaxel or between the groups treated weekly and those treated every 3 weeks. However, the interpretation of our results is complicated by an unanticipated interaction between the type of taxane administered and the treatment schedule for the group receiving weekly docetaxel. In a com-

CONCLUSIONS

Weekly paclitaxel after standard adjuvant chemotherapy with doxorubicin and cyclophosphamide improves disease-free and overall survival in women with breast cancer. (ClinicalTrials.gov number, NCT00004125.)

NSAS-BC02 症例数の設定

- ◆ 本試験において検討の対象となる研究仮説は二つ
 仮説1: Taxane単独はAC(EC)+Taxaneに劣らないか
 仮説2: 二つのTaxaneのいずれが優れているか
- ◆ 主たるエンドポイントは殆どのadjuvant試験において採用されているDFS(無病生存)であり、2次的エンドポイントとしてOS(全生存)、RFS(無再発生存)、HRQOL、医療経済効率、有害事象を選択する。QOL、有害事象に対しては仮説1は通常の優越性の仮説(単純な比較)に置き換えられる。
- ◆ 主たるエンドポイントに対する仮説1は、本試験単独で検証すべき(非劣性)仮説と位置づける。
- ◆ 仮説2は、主たるエンドポイントにおいては本試験のみでは検証が困難な探索的な仮説、あるいは他試験とのメタナリシスによって検証されるべき仮説と位置づける。

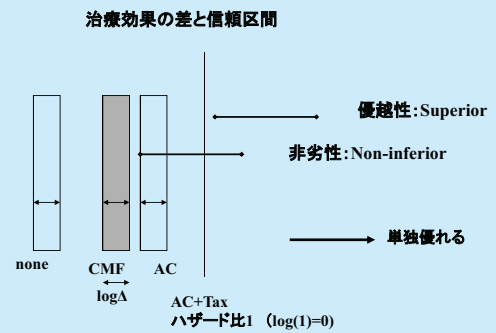
統計的検討: 症例数の設定(続き)

- ◆ したがって症例数の設定は、主たるエンドポイントに対する仮説1に対して行う。
- ◆ 主な有害事象(血液毒性、神経毒性、浮腫、脱毛など)、およびHRQOLに対する検出力(仮説1、仮説2)は、400-800例で十分。

統計的検討: 症例数の設定(続き)

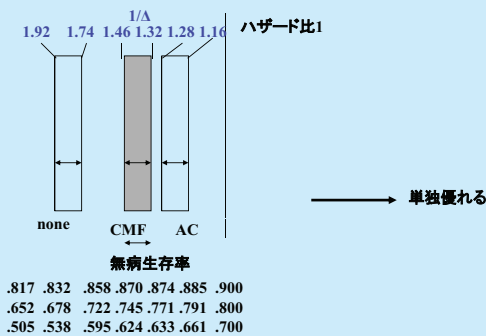
- ◆ 検証的な解析は、NSAS-BC01と同様に対数ハザード比の推定値とその信頼区間に基づいて行う。なお、仮説1、仮説2のいずれに対しても主たる解析は2x2要因実験の主効果に対する比較をもって行う。2次的に交互作用項の検討も行うが検出力は十分ではない。(大きな交互作用あるいは質的交互作用が存在することは考えにくい。)
- ◆ EBCTCGおよびCALGB研究から、無治療に対するCMF、CMFに対するEC(AC)、EC(AC)に対するAC(EC)+Taxaneのハザード比を、それぞれ0.76、0.88、0.78-0.86と想定する。
- ◆ 非劣性仮説検証のための閾値(Δ)は、Taxane単独が少なくともCMFを凌駕することを、最低の有効性保証の条件として設定する。すなわち、AC(EC)+TaxaneのTaxane単独に対するハザード比の下限を $0.88*0.78=0.686(1/1.457)$ から $0.88*0.86=0.756(1/1.321)$ 以上とする。⁵¹

非劣性試験の説明



非劣性試験の説明

治療効果の差と信頼区間

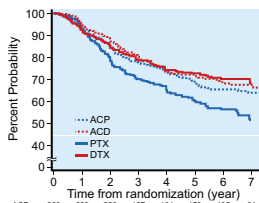


統計的検討: 症例数の設定(続き)

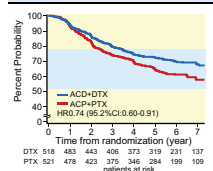
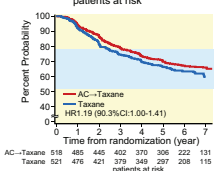
- ◆ Δ を0.756、真のハザード比を1、0.95(Taxane単独がやや劣る)、 $\alpha=0.05$ (片側、90%信頼区間に対応)、 $\beta=0.20$ とした場合、検証に必要なイベント数は、対数ハザード比の推定値の正規近似により両群で320、480となる。なお、 Δ を0.686とした場合には、それぞれ180、240が対応する数となる。
- ◆ n+例の年間再発率を8%と見積もると、320、480のイベント数を観察するために必要な年数は4000-6000であり、全例最低5年の追跡を行うと必要症例数は800-1200である。
- ◆ n+例の割合を45%、うちER、PgR共に陽性例を60%とすると、対象例の全切除例に対する割合は $0.45*0.40=18\%$ となる。同意に応ずる割合を30-40%とすると、3年で登録を完了するために必要な年間手術件数は、 $(800-1200)/3/(0.18*(0.30-0.40))=3700-7400$ となる。

Disease-free Survival (DFS)

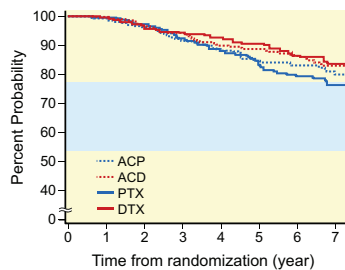
SABCS2011



| Summary of events (DFS) | | | | |
|---|-----------|-------------|-----------|-----------|
| | ACP | ACD | PTX | DTX |
| No. of pts | 260 | 258 | 261 | 260 |
| Hypothesis 1: A taxane alone is not inferior to AC + a taxane | | | | |
| Hazard ratio (AC + a taxane as standard) | | 1.19 | | |
| 90.3% CI | | 1.00 - 1.41 | | |
| p value | | 0.31 | | |
| Hypothesis 2: Whether PTX or DTX is more effective | | | | |
| Hazard ratio (PTX as standard) | | 0.74 | | |
| 95.2% CI | | 0.60 - 0.91 | | |
| p value | | 0.0037 | | |
| Comparison of four arms (DFS) | | | | |
| | ACP | ACD | PTX | DTX |
| Hazard ratio | reference | 0.89 | 1.40 | 0.86 |
| 95% CI | | 0.66-1.20 | 1.06-1.84 | 0.64-1.17 |



Overall Survival (OS)



| Summary of events (OS) | | | | |
|---|-----|-------------|-----|-----|
| | ACP | ACD | PTX | DTX |
| No. of pts | 262 | 263 | 263 | 261 |
| Hypothesis 1: A taxane alone is not inferior to AC + a taxane | | | | |
| Hazard ratio (AC + a taxane as standard) | | 1.06 | | |
| 90.3% CI | | 0.83 - 1.36 | | |
| p value | | 0.14 | | |
| Hypothesis 2: Whether PTX or DTX is more effective | | | | |
| Hazard ratio (PTX as standard) | | 0.73 | | |
| 95.2% CI | | 0.54 - 0.98 | | |
| p value | | 0.035 | | |

Summary

- 8 cycles of a taxane is not 'non-inferior' to 4 cycles of AC followed by 4 cycles of a taxane in all analyzed patients in terms of disease-free survival.
- Docetaxel (75 mg/m²) is superior to paclitaxel (175 mg/m²) when given every three weeks in terms of disease-free survival and overall survival.
- Regarding adverse events:
 - Incidences of nausea and vomiting were higher with 4 cycles of AC followed by 4 cycles of a taxane than with 8 cycles of a taxane.
 - Incidences of edema and febrile neutropenia were higher with docetaxel (75 mg/m²) than with paclitaxel (175 mg/m²).
 - Incidence of sensory neuropathy was higher with paclitaxel (175 mg/m²) than with docetaxel (75 mg/m²).

57

AKA-0984-57

Conclusion

- As compared with paclitaxel, docetaxel is associated with improved efficacy given in tri-weekly schedule. The results are consistent with a study of metastatic breast cancer¹. The efficacy of taxanes seems to be schedule-dependent. Docetaxel is suitable for tri-weekly schedule but paclitaxel is more efficacious when administered in weekly dose-dense schedule. Our observations in this study are in a very good agreement with those in ECOG 1199 Trial².
- Eight cycles of Docetaxel (75 mg/m²) is an alternative regimen in patients who are intolerance for, or reluctance to treatment with anthracyclines because of concern of cardiac malfunction or fear of nausea and vomiting.

1. Jones SE et al. J Clin Oncol 2005;23:5542

2. Sparano JA et al. N Engl J Med 2008;358:1683

AKA-0984-58

内容

- ◆ 説明的試験と実践的試験 MEGA
- ◆ 多因子要因実験とその解釈 NSAS BC02
仮説検証とサンプルサイズ設計
- ◆ 非劣性試験とパラメータへの要約 Iressa製販後試験
- ◆ 医療情報の適切な発信とは？ しばらく前の降圧薬試験

59

イレッサ(ゲフィチニブ)臨床試験とその背景

- ◆ 2002年7月5日、優れた第II相試験結果を下に世界に先駆けてわが国で承認
- ◆ 発売当時から副作用である間質性肺炎については添付文書で注意。しかし・・・
- ◆ 「夢の抗癌剤」としての過剰な期待から非専門家による処方・治療も含め一気に使用が広がり、結果として間質性肺炎による死亡例が続発
- ◆ 厚生労働省は13例の死亡報告を受け、10月15日にアストラゼネカ社に緊急安全性情報発出を指示。12月25日には第1回のゲフィチニブ安全性問題検討
- ◆ 2003年5月5日には米国FDAがゲフィチニブを認可したものの、2005年1月4日、全体での延命効果は存在しないというプラセボ対照のISEL試験結果(AACR2005)を受け、アストラゼネカ社は欧州医薬品審査庁への申請を取り下げ
- ◆ しかし、ISEL試験では東洋人に対しては延命効果があることが示唆され、他の試験においても女性・腺癌・非喫煙という特定の部分集団に奏効するという示唆
- ◆ この示唆はゲフィチニブがEGFレセプター突然変異例に奏効するという独立な研究結果(Lynch et al. NEJM 2004;350:2129-39, Paez et al.Science 2004; 304:1497-500)と符合(突然変異例が女性・腺癌・非喫煙に多かった)
- ◆ ドセタキセルを対照とした市販後試験結果の概要は、2007年2月1日に開催された厚生労働省の安全対策調査会で検討
(http://www.mhlw.go.jp/houdou/2007/02/h0201-4.html)

60

イレッサ公式サイト

http://www.iressa.jp/ekigaku/eki_index.asp

IRESSA
イレッサ錠250 総合情報サイト

○イレッサの国内第III相比較試験の結果について

イレッサの国内第III相比較試験の結果について

アストラゼネカ株式会社は、平成18年度第2回薬事・食品衛生審議会医薬品等安全対策部会安全調査会(平成19年2月1日開催)に、イレッサの国内第III相比較試験の結果を報告しました(調査審については厚生労働省のウェブサイトに掲載されています)。

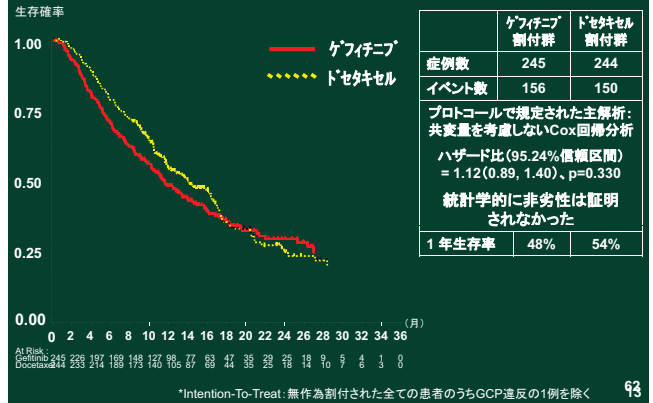
平成18年度第2回
薬事・食品衛生審議会医薬品等安全対策部会安全対策調査会における検討の結果について
<http://www.mhlw.go.jp/houdou/2007/02/h0201-4.html>

試験の概要

本試験は、1又は2レジメンの化学療法治療歴(少なくとも1レジメンは白金製剤を含む)を有する/転移性III期/IV期又は術後再発の非小細胞肺癌患者において、イレッサ(1日1回600mg/日)とドセタキセル(60mg/m²)を比較したものです。

61

主要評価項目 — 全生存期間 (ITT*)



63

イレッサ公式サイト: 結果の概要

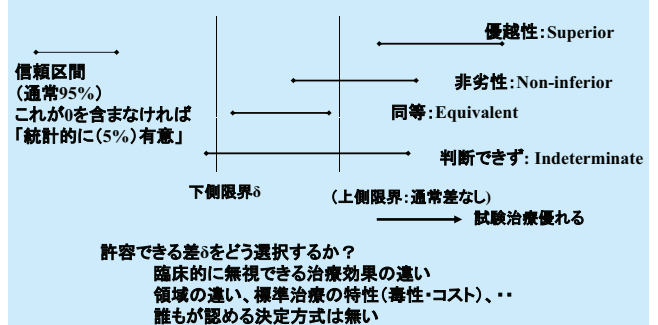
本試験の主要目的は、試験実施計画書で規定した統計学的な判定基準(Cox回帰分析を行い、ハザード比の信頼区間の上限が1.25以下であること)に従い、全生存期間におけるイレッサのドセタキセルに対する非劣性を証明することです。この規定された解析を行った結果、主要目的であるイレッサのドセタキセルに対する非劣性の証明には至りませんでした(ハザード比 1.12、95.24%信頼区間 0.89-1.40、p= 0.330)。

なお、この解析は、全生存期間におけるドセタキセル群に対するイレッサ群の効果が時間の経過にかかわらず一定であることを前提としていますが、第三者統計専門家が行った追加の解析では、当該効果が時間依存的に変化するという結果がでており、その前提が成り立っているとは言い難い結果でした。

患者は、処方する医師は、どう解釈？

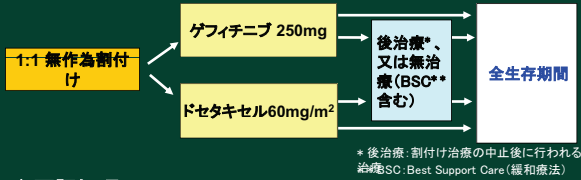
63

優越性試験と非劣性試験



64

試験デザイン



主要評価項目

- 全生存期間
 - 共変量を考慮しない比例ハザードモデルに基づいたハザード比の信頼区間の上限が1.25以下であれば非劣性が結論づけられる
 - 最終解析の目標死亡例: 296例

主な選択基準

- 進行/転移性(ⅢB期/Ⅳ期)又は術後再発の非小細胞肺癌患者
- 1又は2レジメンの化学療法治療歴(少なくとも1レジメンは白金製剤を含む)
- 年齢20歳以上
- 全身状態: WHO Performance Status (PS) が0~2

65

副次的評価項目

- 有効性
 - 無増悪生存期間 (PFS)
 - 治療成功期間 (TTF)
 - 奏効率 (ORR)
 - 病勢コントロール率 (DCR)
- 随伴症状及びQOL
 - 肺癌サブスケール (LCS) による随伴症状の変化
 - QOL: Functional Assessment of Cancer Therapy-Lung (FACT-L)*
- 安全性
 - 安全性及び忍容性

*肺癌治療の機能的評価

66

臨床試験の実施状況

- 実施施設: 全国計50施設
- 無作為割付け症例数: 490例
- 登録期間: 2003年9月~2006年1月
- 全生存期間のデータカットオフ日: 2006年10月31日
- データカットオフ日までの死亡: 306例(全症例の63%)
- 追跡期間(割付け日~死亡又は最終生存確認日)の中央値: 21ヵ月

67

患者背景 (ITT*)

| | ゲフィチニブ割付け群 (n=245) | ドセタキセル割付け群 (n=244) |
|--------------------|--------------------|--------------------|
| 年齢 64歳以下 | 56% | 55% |
| 65歳以上 | 44% | 45% |
| 性別 女性 | 38% | 38% |
| 男性 | 62% | 62% |
| WHO PS 0 | 35% | 38% |
| 1 | 61% | 58% |
| 2 | 4% | 4% |
| 喫煙歴なし | 29% | 36% |
| 喫煙歴あり | | |
| 試験開始前のみ | 65% | 55% |
| 試験開始時に喫煙あり(毎日ではない) | <1% | 0% |
| 試験開始時に喫煙あり(毎日) | 6% | 9% |

*Intention-To-Treat: 無作為割付けされた全ての患者のうちGCP違反の1例を除く

68

腫瘍の背景 (ITT*)

| | ゲフィチニブ割付け群 (n=245) | ドセタキセル割付け群 (n=244) |
|------------|--------------------|--------------------|
| 扁平上皮癌 | 15% | 17% |
| 腺癌 | 78% | 77% |
| その他 | 6% | 7% |
| 罹病期間 6ヵ月未満 | 29% | 25% |
| 6~12ヵ月 | 40% | 39% |
| 12ヵ月超 | 31% | 36% |
| 臨床病期 ⅢB | 19% | 20% |
| 臨床病期 Ⅳ | 65% | 61% |
| 再発 | 16% | 18% |
| 開始時の標的病変あり | 82% | 77% |

*Intention-To-Treat: 無作為割付けされた全ての患者のうちGCP違反の1例を除く

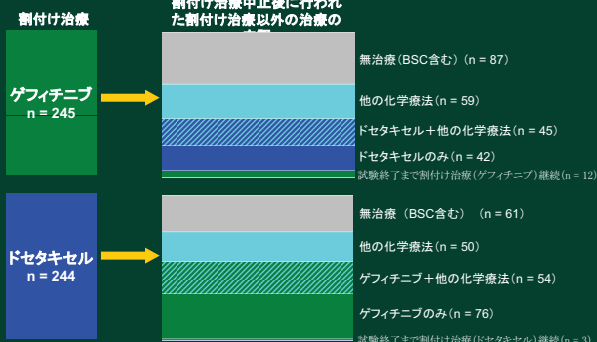
69

後治療

- ゲフィチニブ群に割付けられた症例:
 - 36%の症例に後治療としてドセタキセルが投与された
 - 40%の症例は無治療(BSC含む)又はゲフィチニブの投与が継続されていた
- ドセタキセル群に割付けられた症例:
 - 53%の症例に後治療としてゲフィチニブが投与された
 - 26%の症例は無治療(BSC含む)又はドセタキセルの投与が継続されていた

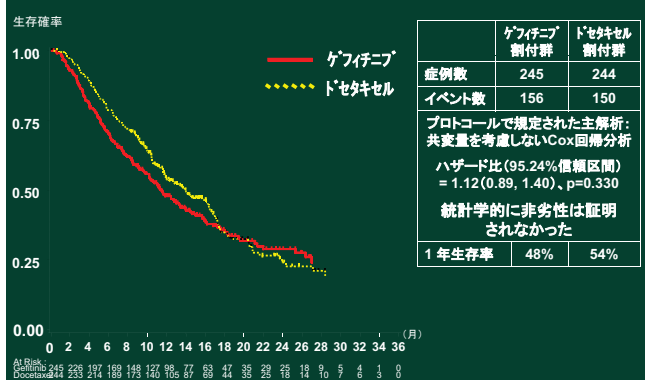
70

後治療の内訳



71

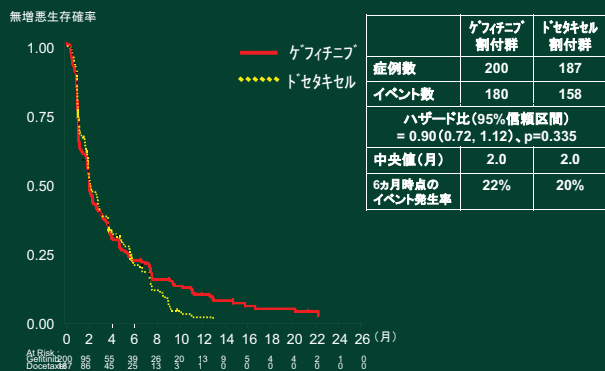
主要評価項目 - 全生存期間 (ITT*)



*Intention-To-Treat: 無作為割付けされた全ての患者のうちGCP違反の1例を除く

73

副次的評価項目(有効性) — 無増悪生存期間 (抗腫瘍効果評価対象症例)



73

副次的評価項目(有効性) — 治療成功期間 (抗腫瘍効果評価対象症例)



74

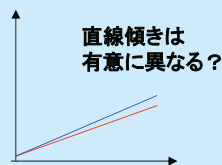
非劣性が検証できなかった とは

1. 劣っていた
2. 判断できるだけの情報量が存在しない
3. 非劣性を判断するために、2群の違いをハザード比に要約した。それが適切ではなかった!

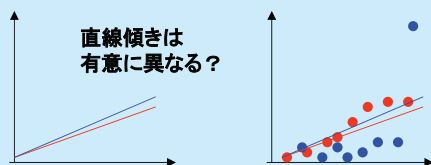
比例ハザード性が成立しなかった!

なぜ?

75



76



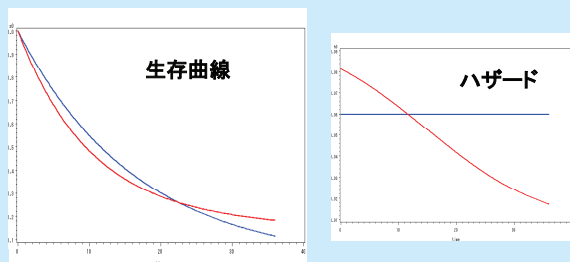
77

なぜイレッサ臨床試験で比例ハザード性が成立しなかったのか?

- ◆ 再発では比較的プロポーションハザード性がよく成り立つが...
- ◆ イレッサが奏効する部分集団の存在
ハザード比の交差
- ◆ 進行癌の生存がエンドポイントであった
そもそも比例ハザード的でないのかも
さらに後治療が比例ハザード性をくずした
- ◆ さらに
後治療の選択に群間差(ドセタキセル群で早く移行の傾向)
登録された患者は必ずしもイレッサに有利ではなかった
- ◆ 振り返ってみればデザインに問題、しかし...
当時、NSCLCの非劣性試験は生存ハザードがエンドポイント
後治療が生存に寄与することは殆ど予想されていなかった
予測因子に関する情報は試験途中で判明

78

奏効する部分集団が存在すると 生存曲線(ハザード)は交差



赤: 奏効集団20%(MST 10倍) 残り80%(MST 0.6倍)

79

イレッサを適切に評価するには

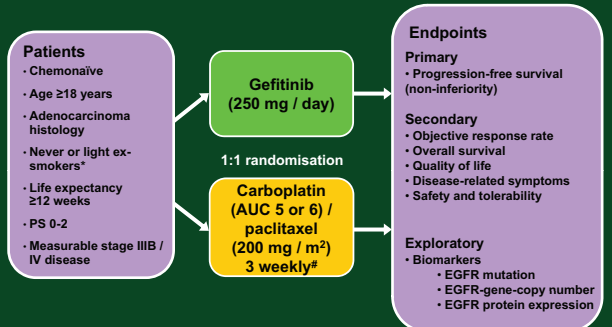
- ◆ 1次治療としての評価
- ◆ 適格条件の選定
- ◆ エンドポイントの選択
生存のハザード比は適切か?
PFS(進行までの時間)あるいは
QALY(QOLを考慮した質調整生存年)?
- ◆ 後治療の影響評価による解釈の補い
- ◆ IPASS(アジア)試験が実施
腺癌、非あるいは軽い喫煙者対象、カルボプラチン・パクリタキセルとの比較、PFSが主たるエンドポイント
しかし、ハザード比の限界 $\delta=1.2$ の非劣性試験

80

IPASS試験結果の公表
そして変異例を対象とした前向き試験結果の公表

やっぱりね・・・

IPASS: Study design



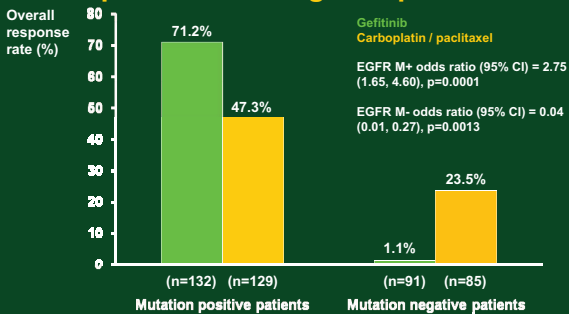
*Never smokers, <100 cigarettes in lifetime; light ex-smokers, stopped ≥15 years ago and smoked ≤10 pack years; #limited to a maximum of 6 cycles
Carboplatin / paclitaxel was offered to gefitinib patients upon progression
PS, performance status; EGFR, epidermal growth factor receptor

IPASS: Study conduct

- 87 centres in 9 countries in Asia
 - China, Hong Kong, Indonesia, Japan, Malaysia, Philippines, Singapore, Taiwan, Thailand
- 1217 patients randomised
- Randomisation period: March 2006 to October 2007
- Data cut-off: 14 April 2008
 - 950 PFS events observed in ITT population (78% maturity)
- Mean time on treatment
 - gefitinib, 6.4 months
 - carboplatin / paclitaxel, 3.4 months (median number of cycles*: 6)
- Final survival data (944 events) expected mid-2010



Objective response rate in EGFR mutation positive and negative patients



Odds ratio >1 implies greater chance of response on gefitinib

IPASS: Results

Tony S. Mok et al.
NEJM 2009;
361:947-57.

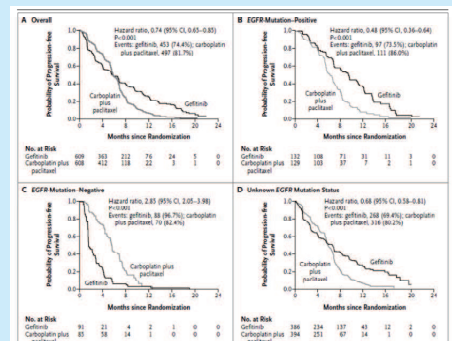
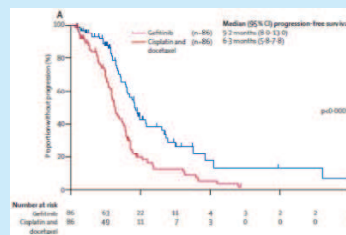


Figure 2. Kaplan-Meier Curves for Progression-Free Survival. Kaplan-Meier curves for progression-free survival are shown for the overall population (Panel A), patients who were positive for the EGFR mutation (Panel B), patients who were negative for the EGFR mutation (Panel C), and patients with unknown EGFR mutation status (Panel D). Analyses were performed on the basis of the intention-to-treat population. With respect to the overall population, results of the exploratory secondary analyses (including a log-rank test, which is valid under the null hypothesis even when hazards are not proportional, and analysis in the per-protocol population) were consistent with the result of the primary analysis. Hazard ratios were calculated with the use of a Cox proportional-hazards model, with the WHO performance status (0 or 1, or 2), smoking history (never or former light smoker), and sex as covariates. EGFR denotes epidermal growth factor receptor.

Gefitinib versus cisplatin plus docetaxel in patients with non-small-cell lung cancer harbouring mutations of the epidermal growth factor receptor (WJTOG3405): an open label, randomised phase 3 trial

Tarutomo Mitsuhashi, Susumu Morita, Yasuhiro Takeda, Shiroshi Nagano, Naomasa Ohama, Junji Yamazaki, Takashi Saito, Miyuki Sawashi, Haruhiko Tada, Takanori Hironaka, Kazuhiko Jojima, Naohiko Kobayashi, Shiroshi Takano, Hiroyuki Yoshida, Kazumasa Saito, Shiroshi Kozaki, Eiji Shimada, Haruhiko Saka, Shiroshi Toyooka, Kazuhiko Nakagawa, Masahiro Fukuhara, for the West Japan Oncology Group

Lancet Oncol 2010; 11: 121-28

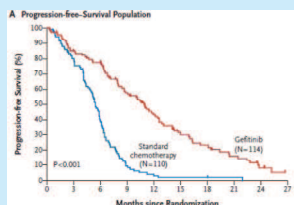


First-line trial with cisplatin+docetaxel EGFR-mutated 86+86 pts. PFS 9.2 vs 6.3 months HR 0.49 (95% CI 0.34-0.71)

ORIGINAL ARTICLE
Gefitinib or Chemotherapy for Non-Small-Cell Lung Cancer with Mutated EGFR

Makoto Maemondo, M.D., Ph.D., Akira Inoue, M.D., Ph.D., Kunihiko Kobayashi, M.D., Ph.D., Shunichi Sugawara, M.D., Ph.D., Satoshi Oizumi, M.D., Ph.D., Hiroshi Isobe, M.D., Ph.D., Akihiko Gemma, M.D., Ph.D., Masao Harada, M.D., Ph.D., Hirohisa Yoshizawa, M.D., Ph.D., Ichiro Kinoshita, M.D., Ph.D., Yuka Fujita, M.D., Ph.D., Shoji Okinaga, M.D., Ph.D., Haruto Hirano, M.D., Ph.D., Kozi Yoshimori, M.D., Ph.D., Toshiyuki Harada, M.D., Ph.D., Takashi Ogura, M.D., Masahiro Ando, M.D., Ph.D., Hitoshi Miyazawa, M.S., Tomoaki Tanaka, Ph.D., Yasuo Saijo, M.D., Ph.D., Koichi Hagihara, M.D., Ph.D., Satoshi Morita, Ph.D., and Toshihiro Nukiwa, M.D., Ph.D., for the North-East Japan Study Group*

NEJM 2010; 362: 2380-8.



trial with cisplatin+docetaxel EGFR-mutated 114+110 pts. PFS 10.8 vs 5.4 months HR 0.30 (95% CI 0.22-0.41)

内容

- ◆ 説明的試験と実践的試験 MEGA
- ◆ 多因子要因実験とその解釈 NSAS BC02
- ◆ 仮説検証とサンプルサイズ設計
- ◆ 非劣性試験とパラメータへの要約 Iressa製販後試験
- ◆ 医療情報の適切な発信とは？ しばらく前の降圧薬試験

2006年年末に、まだ学会発表レベルの大規模循環器系疾患臨床試験の結果が過剰なまでに報道された。

プロモーションコード改訂直前の駆け込み措置であったと聞く。今後このような事態は発生しないことを期待するが、報道内容についてその質が問われること、発信するライターに試験結果を「ただしく読む」能力が要求されることは変わらないであろう。

MEGA Studyを例にとり、解釈上の問題点をとりあげ議論する。なお、本内容については3月13日の日本循環器病学会ランチョン、JMCAサロで講演した内容とかなり重複する。

JMCA年会2007May22

降圧薬大規模試験(国内試験)の過剰報道

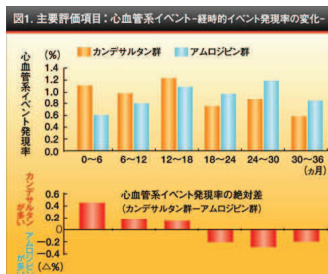
◆ 2006年年末に学会レベルの大規模試験の過剰報道
規制強化に対する駆け込み

- ◆ 研究計画と事後的解析の乖離
- ◆ 結果と結論(解釈)の乖離
- ◆ ソフトなエンドポイント・オープン試験(PROBE)に対する信頼性問題の提起

90

--- CASE-Jが立証したカンデサルタンのレニン・アンジオテンシン系抑制効果 ---

主要評価項目である心血管系イベントは両群ともに134例(5.7%)で、心血管系イベントを構成する突然死・脳血管イベント・心イベント・腎イベント・血管イベントいずれも両群間で有意差はなかった。なお、試験前半(18ヶ月まで)はアムロジピン群で、試験後半(18ヶ月以降)はカンデサルタン群で心血管系イベントが減少する傾向が認められた(図1)。カンデサルタンの持つレニン・アンジオテンシン(RA)系に対する強い抑制効果が、時間経過とともに発揮されるようになったことが示唆される。



92



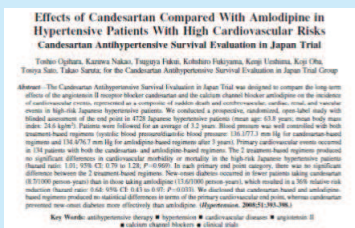
主要評価項目(心血管イベント)でカンデサルタン群とアムロジピン群は同等。

同試験において対象は、カンデサルタン4~8mg/日(最大12mg/日) アムロジピン2.5~5mg/日(最大10mg/日)の2群に無作為に割り付けられ、増量でも効果不十分の時は他の降圧薬が追加された。平均追跡期間は3.2年。試験デザインはPROBE法(非盲検無作為群間比較法)が採用され、登録・追跡から解析までを京都大学EBM研究センターが担当した。実施に際しては、Webを用いた新しい臨床試験システムが使われた。CASE-Jでは高齢者が約半数を占め、糖尿病合併例が約43%と他の大規模臨床試験よりも多く、また、平均BMIが24.5kg/m²前後ということから、日本人の高血圧の現状を強く反映していると考えられる。試験開始時の血圧はカンデサルタン群162/92mmHg、アムロジピン群163/92mmHgだった。達成血圧値は各々136/77mmHg、134/77mmHgとアムロジピン群で低かった。

研究計画と事後的解析の乖離

CASE-J

Ogihara et al. Hypertension, 2008; 51: 393-8
4703例/ハイリスク高血圧患者
PROBE試験
Candesartan vs Amlodipine



1次エンドポイントは突然死+脳イベント(脳卒中、TIAの新規発症・再発)+心イベント(心不全、狭心症、心筋梗塞の新規発症・再発・増悪)+腎機能不全+血管イベント
1次エンドポイントでは同等 (元来はCandesartanの優越性試験として計画された)

93

研究計画と事後的解析の乖離

口頭発表後1年半をかけてようやく論文化された。論文を読むと、学会発表では明らかにしていなかった大きな事柄も論文化によって明らかにになっており、研究は論文化されてこそ評価すべきものというのを改めて感じさせる。
まず驚きは、candesartan群とamlodipine群の降圧薬併用薬剤の違いである。candesartan群の方が明らかに多くの降圧薬を併用していることが明らかになったことの意味は大きい。すなわちcandesartanは他の薬剤をたくさん併用しなければamlodipine群と同等の降圧が得られないことを証明しているのである。
またエンドポイントの内訳をあらためて見直してみると、candesartan群が優位なのは狭心症、TIAといった客観性に乏しいエンドポイントばかりであり、脳卒中などの客観性のあるエンドポイントはむしろamlodipine優位に傾いているのである。狭心症やTIAが試験薬群に優位という傾向は、JIKEI-HEART試験でも同じように認められ、ここに企業支援によるPROBE法の問題点が明確に浮かんできているのである。
またトライアル初期から中期にかけてcandesartan群の方の脱落例が明らかに増えており、candesartan群はハイリスク症例で早期にイベントを起こしていることを示している。このような試験ではtime to eventまでの期間を比較するべきであり、その意味ではamlodipine群の方がcandesartan群よりもハイリスク症例のイベント発症をより先送りさせることができたことを示している。amlodipine群の方が治療中の血圧が低かったことが早期のイベント発症抑制に効果があったと考えられる。

(循環器トライアルデータベース 桑島) 94

JATOS結果速報

わが国初の高血圧治療に関するエビデンス-JATOSの最終解析結果が、第21回国際高血圧学会(ISH)にて発表されました。

福岡で開催された第21回国際高血圧学会(ISH)学術集会において、高齢者高血圧治療に関する日本初の大規模臨床試験JATOS(The Japanese Trial to Assess Optimal Systolic Blood Pressure in Elderly Hypertensive Patients)の結果が発表されました。

JATOSは、一般臨床医を中心に全国約1,000施設が参加し、高齢者高血圧に対する降圧治療をどのように行うべきかについて検討することを目的に実施されたPROBE法を用いた多施設無作為割り付け比較試験です。

ここでは、その最終解析結果をいち早くお届けします。

JATOSの結果は、高齢者高血圧の国際的ガイドラインを支持する

横浜市立大学 名誉教授 石井 富男 先生

JATOSでは収縮期血圧(SBP)160mmHg以上の高血圧患者(65~85歳)4,418例を、SBP140mmHg未満を目標とする厳密な治療群(A群)とSBP140mmHg以上160mmHg未満を目標とする緩和な治療群(B群)に無作為に、2~4週間の観察期間を経て、長時間作用型Ca拮抗薬・塩酸エモズピンを基礎薬とした降圧治療を2年間行った。その結果、2年後の平均血圧はA群135.9/74.8mmHg、B群145.6/78.1mmHgと、両群間の血圧差は9.7/3.3mmHgであった。

1次評価項目(脳心血管疾患および腎機能障害)の発症率、死亡率および2次評価項目(総死亡率、有害事象・副作用の発現率、脳拍動への影響、降圧効果など)については、いずれも両群間で有意差は認められなかった。

この結果は、これまで実臨床において一部懸念があった、高齢高血圧患者に対するSBP140mmHg未満への降圧に関する安全性が確認されたことを示しているといえる。さらに石井氏は、「JATOSは、高齢者の降圧目標値を140/90mmHg未満とした国際的ガイドラインを支持するわが国で初めてのエビデンスを提供した」と述べた。

95

2006 JIKEI HEART

日本初のARBエビデンス Online Conference Report
ISH 2006 - International Society of Hypertension 2006-

発表!

パルサルタンを使用した大規模臨床試験 The JIKEI HEART Studyが 日本人のエビデンスを世界に発信

日本初の高血圧介入大規模臨床試験The JIKEI HEART Study(3081例)の結果が、9月5日のHOT LINEセッションで報告された。報告者は同試験のチェアマンで東京慈恵会医科大学循環器内科教授の望月正武氏とコチェアマンでスウェーデンSahlgrenska大学のBjorn Dahlöf氏。

The JIKEI HEART Studyの対象は、カルシウム拮抗薬やアンジオテンシン転換酵素(ACE)阻害薬を含む従来治療を受けている虚血性心疾患または心不全を伴う日本人高血圧患者(22-79歳:平均65歳)。約33%に虚血性心疾患、11%に心不全の合併が認められた。試験開始時には67%にCa拮抗薬、35%にACE阻害薬が投与されていた。非アンジオテンシンII受容体拮抗薬(ARB)群ではこの従来治療が継続され、パルサルタン群ではパルサルタン(平均75mg/日)が投与されたとの比較となった。

96

ソフトなエンドポイントをオープン試験で評価する

Jikei Heart Study
Mochizuki et al.
Lancet 2007;
369: 1431-9

Valsartan in a Japanese population with hypertension and other cardiovascular disease (Jikei Heart Study): a randomised, open-label, blinded endpoint morbidity-mortality study

Saku Mochizuki, Eiji Usuki, Hiroyuki Shimizu, Etsuro Horiuchi, Makoto Yoshikawa, Kenjiro Takayama, Makoto Ohno, Taku Yasuda, Toshiaki Shimizu, Kazuhiko Nakano, Makoto Fusa, Tetsuo Sakai, Toshiaki Watanabe, Wataru Takemoto, Takao Yoshida, Naoki Tamura, Kenji Jikei (Heart Study group)

Summary
Background Drugs that inhibit the renin-angiotensin-aldosterone system benefit patients at risk for or with atherosclerotic cardiovascular disease. However, evidence for this effect in Asian populations is scarce. We aimed to investigate whether addition of an angiotensin receptor blocker, valsartan, to conventional cardiovascular treatment was effective in Japanese patients with cardiovascular disease.

Methods We initiated a multicentre, prospective, randomised controlled trial of 3001 Japanese patients, aged 50-79 years, from 41 (33 non-urban) sites who were undergoing conventional treatment for hypertension, coronary artery disease, heart failure, or a combination of these disorders. In addition to conventional treatment, patients were assigned either to valsartan (40-160 mg per day) or to other treatment without angiotensin receptor blockers. Our primary endpoint was a composite of cardiovascular morbidity and mortality. Analysis was by intention to treat. The study was registered at ClinicalTrials.gov with the identifier NCT00131328.

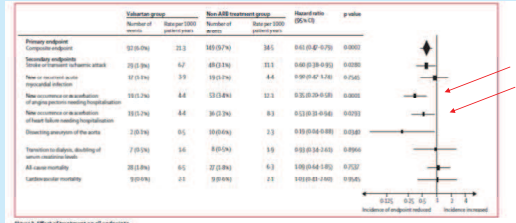
Findings After a median follow-up of 3.1 years (range 1-11) the primary endpoint was recorded in fewer individuals given valsartan than in controls (21.6% absolute risk, 2.2 (1.9-2.5) per 100 patient years; hazard ratio 0.61, 95% CI 0.47-0.79, $p < 0.002$). This difference was mainly attributable to lower incidences of stroke and transient ischaemic attack (9.6 vs 16.6, 6.9-10.4-9.5, $p < 0.01$), angina pectoris (9.5 vs 13.6, 6.2-20.6-9.6, $p < 0.001$), and heart failure (9.9 vs 12.1, 6.3-13.8-9.6, $p < 0.01$) in those given valsartan than in the control group. Mortality or cardiovascular mortality did not differ between groups.

Interpretation The addition of valsartan to conventional treatment prevented more cardiovascular events than supplementary conventional treatment. These benefits cannot be entirely explained by a difference in blood pressure control.

3081例高血圧患者
PROBE試験
現行治療+Valsartan vs
現行治療+ARB以外の追加治療

1次エンドポイントは心血管死亡+心血管合併症(脳卒中、TIAによる入院; 心筋梗塞; 心不全による入院; 狭心症による入院; ...)
1次エンドポイントで有意差 ハザード比92:149=0.61(0.47-0.79, $p=0.0002$) 97

ソフトなエンドポイントをオープン試験で評価する



ある意味で、本研究の結果は今日のわが国で公正な大規模臨床試験を行うことが非常に困難であることを示唆する内容である。まず本研究はdouble blind法ではなく、PROBE法で行われたこと、狭心症、一過性脳虚血性発作(TIA)という客観性に乏しいイベントがエンドポイントに含まれており、この2つのイベント発症の差異が一次エンドポイントの差異に大きく貢献していることである。本研究は慈恵会医科大学関連病院という単一グループでおこなわれ、しかもPROBE法という担当医が患者がどちらの群に割り当てられたかを知っている方法によって実施されていることから、狭心症やTIAという客観性に乏しいエンドポイントを設定することに問題がある。率直に言えば、経済的支援を行っているARBサイドに好意的なイベント発症を中央委員会へ報告するというバイアスが生じる可能性は否定できない。(循環器トライアルデータベース 桑島)

最近の過剰？報道

- CASE-J **そもそもの仮説は何だったの？**
BMI高値例のサブ集団解析結果: 計画してたの？
- JATOS **結果と解釈が一致しているの？**

デザインと解釈

- ◆ 検証的な試験なのか探索的な試験なのか
- ◆ 設定は？
実践への指針pragmatic それともメカニズム検討explanatory
- ◆ 検証なら
検出力は充分か
事前に計画されているのか
前提は満たされているか
多因子要因実験における交互作用の存在
time-to-eventなら比例ハザード性