2014JUNE10@がん研究センター

臨床試験の質とは何か？

（財）パブリックヘルスリサーチ・臨床研究支援事業(CSP)担当
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
日本保健情報コンソシウム理事長
中央大学理工学部　人間総合理工学科　生物統計学

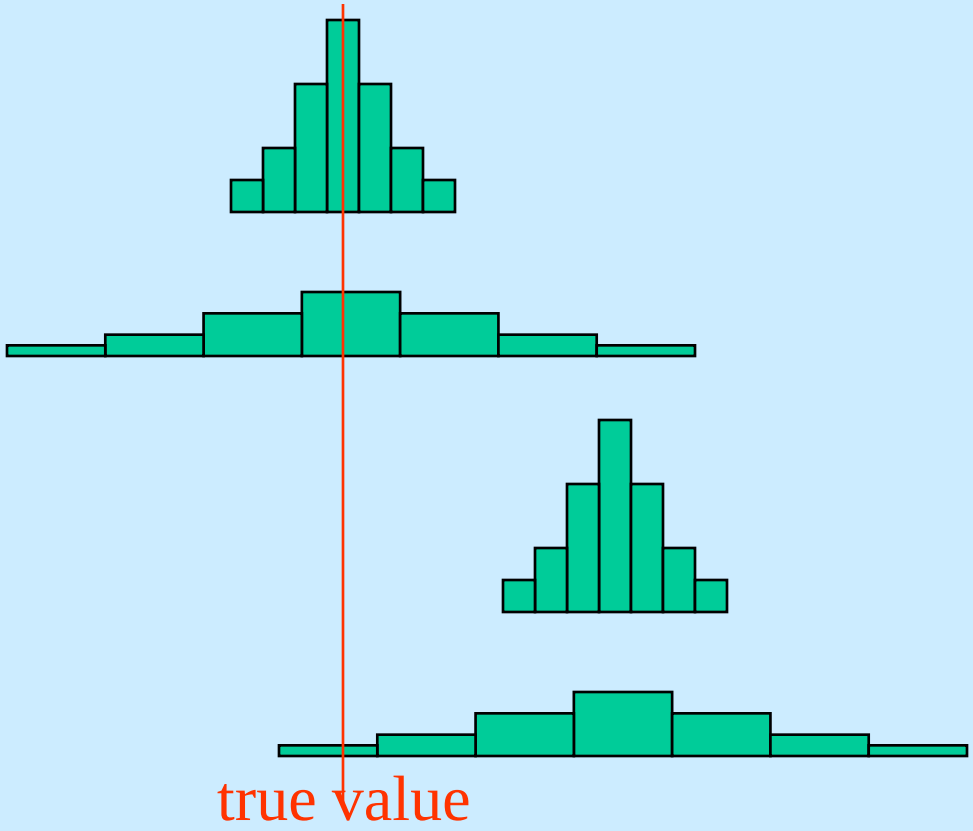
大橋靖雄

本発表は2014年3月14日の日本臨床試験研究会（現学会）で行なった講演に修正追加をおこなったものである

精密度と正確度 Precision and accuracy

観測値・研究結果　＝　真の値＋バイアス＋誤差的バラツキ

observed　　true　　bias　　error



	accuracy	precision
	○	○
	○	×
	×	○
	×	×

臨床（試験）研究の目標（G.Koch)

- 誤差的バラツキを小さくすること（精度を高くすること）

  Clarity Minimizing random error

- 偏り(バイアス）を小さくすること

  Comparability Minimizing bias, Internal validity

- 広い対象に適用できる結論を得ること、一般化可能性

  Generalizability External validity

臨床試験の質 Quality of Clinical Trials?

- 臨床試験の質とは？
  - 捏造・改竄は論外であるが・・・
  - バイアスがない（小さい）こと、当然の条件ではあるが？
  - プロトコルが精緻である？
  - 試験治療のコンプライアンスが高い？
  - ハードエンドポイントの試験である？
  - （研究者主導試験の場合）企業からの影響がない？
  - SDV（Source Document Verification) が徹底している？
  - 品質管理が徹底している 何のための品質管理？
- 質を高めるためにはどうすべきか？

品質管理と品質保証

ランダム化手法が不適切であった例

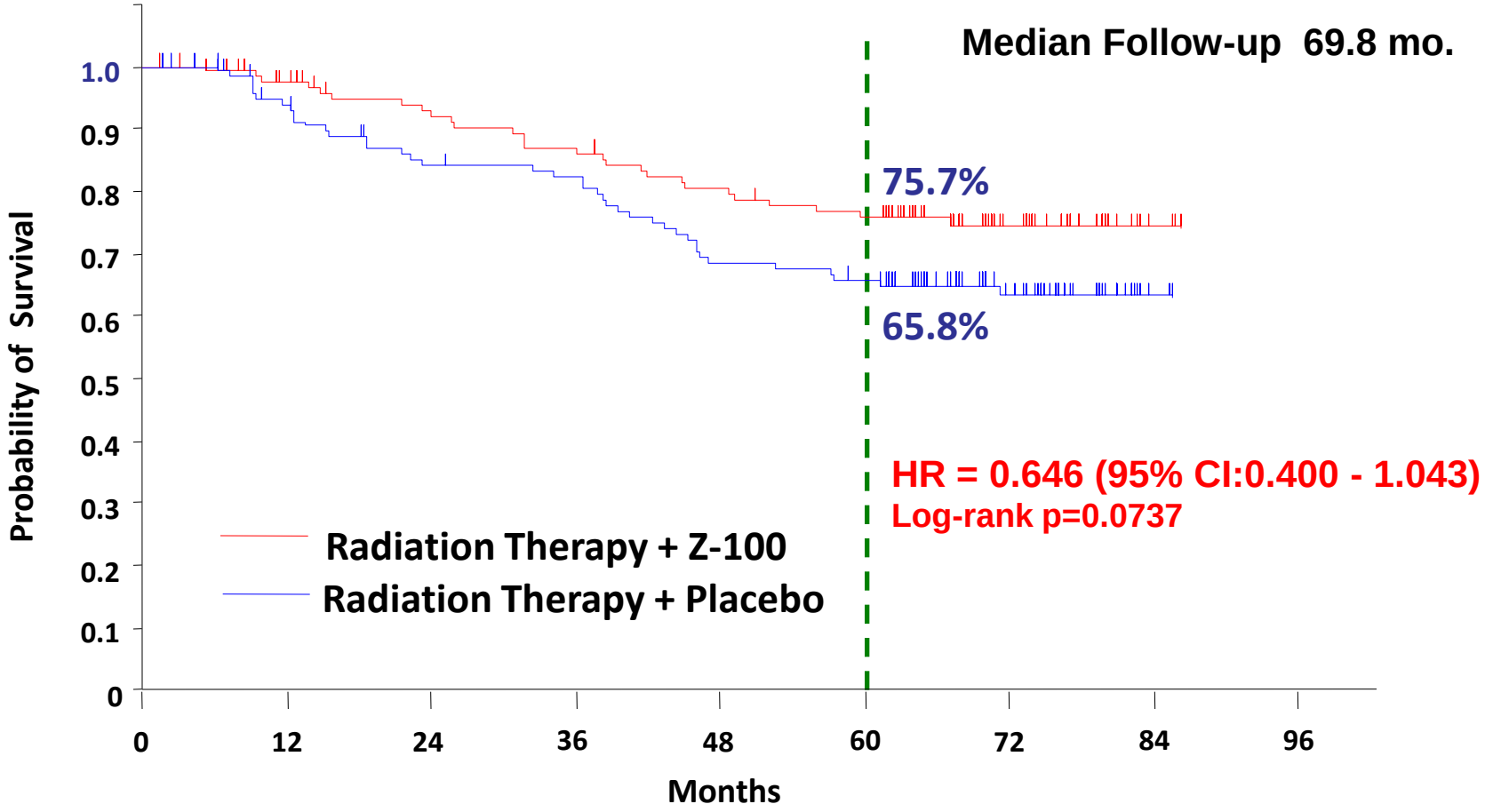
消化器癌を主体とする化療122例/12施設
非治癒切除適格例　A：化療対照　41例
　　B：化療＋SSM　40例
年齢，性，部位，ステージ，検査値などに群間差は認められない

封筒法割付けのA，Bの分布

	施設#	
	6	BBBBBBBB
	5	B
AAAA	10	BBBBBBBBBBBB
AAA	9	BBBB
AAAAA	3	BBBBB
AA	1	BB
A	12	B
	7 *	* 世話人施設
AAAAAAA	8	BBBBB
AAAAAAA	2 *	B
AAAAAAAAAA	11	B
AA	4 *	
41	計	40

Primary : Overall Survival

ASCO2014
Ann Oncology 2014FEB25
（Advanced Access)



Number at risk									
Z-100	121	109	99	93	85	79	34	3	0
Placebo	122	107	93	90	75	71	36	2	0

EUROPEAN SOCIETY OF CARDIOLOGY®

European Heart Journal (2013) **34**, 1018–1023
doi:10.1093/eurheartj/eht063

THE EDITOR'S PAGE

The codex of science: honesty, precision, and truth—and its violations

Thomas F. Lüscher*

Editorial Office, *European Heart Journal*, Zurich Heart House, Moussonstreet 4, 8091 Zürich, Switzerland

Figure 1 Giants of scientific discovery: (*A*) Gregor Mendel (1822–1884) who set the basis for modern genetics (reproduced by kind permission of Keystone), (*B*) Charles Darwin (1809– 1882), the father of evolution [reproduced with permission from John van Wyhe ed. The Complete Work of Charles Darwin Online. (http://darwin-online.org.uk/)], and (*C*) Isaac Newton (1642–1726), the discoverer of gravity and founder of modern physics (reproduced by kind permission of the Trustees of the Portsmouth Estates).

Mendelの逸話は浜田知久馬先生「数理統計学：医薬研究における統計学の役割」で詳しくわかる

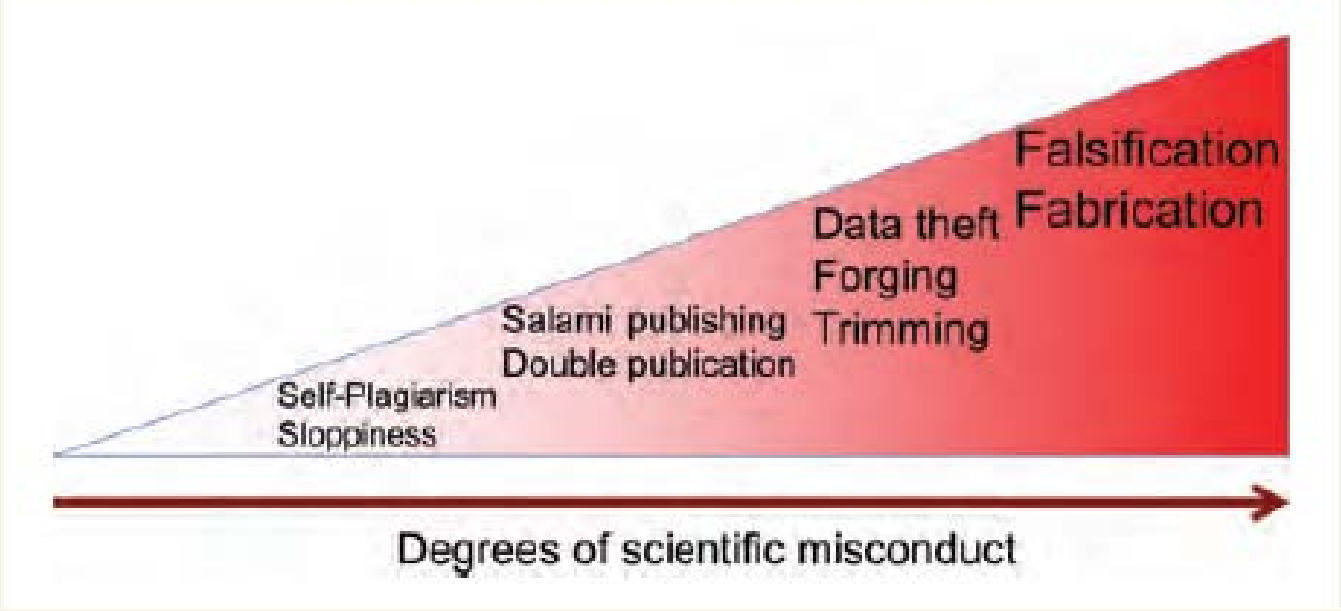


Figure 2 The spectrum of scientific misconduct.

Falsification : 改竄
Fabrication : 捏造
Theft : 盗み
Forging : でっち上げ
Trimming : トリミング
Plagiarism : 盗作・盗用
Sloppiness : ずさん

え！上には上がいる

最近の改竄事例：

転移性乳癌に対する大量化学療法　*Lancet*　2000;　*355*:999-1003

南アフリカBezwoda医師によるASCO　plenary session発表(1999)
無効とする他2演題に対して唯一の有効データ
2000年1月に監査実施　大量化学療法　17/75　確認できず
標準治療はまったく確認できず
プロトコルは監査直前に作成
同意書存在せず、倫理委員会承認書偽造、
不適格例登録
***J.Clin Onco* 1995; *13*:2483 (90例：354回引用)の比較試験もほぼ捏造**

European Heart Journal (2009) 30, 2461–2469
doi:10.1093/eurheartj/ehp363

FASTTRACK
ESC HOT LINE

Effects of valsartan on morbidity and mortality in uncontrolled hypertensive patients with high cardiovascular risks: KYOTO HEART Study

Takahisa Sawada[1]*, Hiroyuki Yamada[1], Björn Dahlöf[2], and Hiroaki Matsubara[1] for the KYOTO HEART Study Group

[1]Department of Cardiovascular Medicine, Kyoto Prefectural University School of Medicine, Kajicho 465, Kamigyoku, Kyoto 602-8566, Japan; and [2]Department of Medicine, Sahlgrenska University Hospital/ Östra, Göteborg, Sweden

Received 4 August 2009; accepted 13 August 2009; online publish-ahead-of-print 31 August 2009

See page 2427 for the commentary on this article (doi:10.1093/eurheartj/ehp364)

Aims The objective was to assess the add-on effect of valsartan on top of the conventional treatment for high-risk hypertension in terms of the morbidity and mortality.

Methods and results The KYOTO HEART Study was of a multicentre, Prospective Randomised Open Blinded Endpoint (PROBE) design, and the primary endpoint was a composite of fatal and non-fatal cardiovascular events (clintrials.gov NCT00149227). A total of 3031 Japanese patients (43% female, mean 66 years) with uncontrolled hypertension were randomized to either valsartan add-on or non-ARB treatment. Median follow-up period was 3.27 years. In both groups, blood pressure at baseline was 157/88 and 133/76 mmHg at the end of study. Compared with non-ARB arm, valsartan add-on arm had fewer primary endpoints (83 vs. 155; HR 0.52, 95% CI 0.42–0.72, $P = 0.00001$).

Conclusion Valsartan add-on treatment to improve blood pressure control prevented more cardiovascular events than conventional non-ARB treatment in high-risk hypertensive patients in Japan. These benefits cannot be entirely explained by a difference in blood pressure control.

Keywords High-risk hypertension • Angiotensin receptor blocker • Cardiovascular mortality–morbidity • Valsartan

Introduction

Cardiovascular disease is the leading cause of mortality worldwide.[1] Hypertension is the most common cause of coronary heart disease and heart failure in Japan; however, cerebrovascular disease is still more prevalent in Japan than in Western societies.[2] The percentage of cerebral bleeding is two or three times greater than in white people, and cerebral infarction is mostly caused by lacunar-type ischaemic stroke due to hypertensive small vessel disease.[3]

The renin–angiotensin system (RAS) plays a major role in the homeostasis of blood pressure, electrolytes, and fluid balance.[4] However, chronic activation of RAS contributes to the development of hypertension and cardiovascular organ damage.[5] Numerous trials have investigated the benefits of ACEI, e.g. The Heart Outcomes Prevention Evaluation (HOPE) Study reported that ACE inhibitors significantly reduced mortality, myocardial infarction, and stroke in high-risk patients.[6] Another important study, in this case with ARB, was the Losartan Intervention For Endpoint (LIFE) reduction in hypertension study, where losartan-based therapy prevented more cardiovascular morbidity and death, in particular stroke, than atenolol-based regimen despite similar blood pressure control.[7] There are now numerous studies showing beneficial effects of RAS blockers on cardiovascular outcomes, in particular with ARBs, in various stages of the CV continuum.[8] However, these studies have included as maximum a few percent of Asian patients in general and very few Japanese in particular.

Cardiovascular disease incidence in Japan differs from those in Western countries. CAD mortality is one-third of that in the USA, and cerebrovascular disease mortality is ~1.5 times higher than in the USA.[9] The dietary habits in Japan differ from

業　　第 13646 号

発行所：株式会社じほう http://www.jiho.co.jp/
本社/〒101-8421 東京都千代田区猿楽町 1-5-15 猿楽町 SS ビル
○編集：TEL 03-3233-6351　FAX 03-3233-6359
○購読：TEL 03-3233-6336　FAX 0120-657-751
支局/〒541-0044 大阪市中央区伏見町 2-1-1 三井住友銀行高麗橋ビル
○購読：TEL 06-6231-7061　FAX 0120-189-015
◆FAX 版購読料　1年：83,160 円(税込)　6カ月：46,410 円(税込)
◆未送信・落丁などの場合は販売管理グループ TEL 03-3233-6336 まで
・スマートフォン対応　日刊薬業WEB　検索
・最新記事随時更新　http://nk.jiho.jp

2013（本号 11 頁）

有力エビデンスを使用中止

ノバルティス　ディオバンの論文撤回受け

ノバルティス ファーマは5日までに、主力品のARBディオバン」（一般名＝バルサルタン）に関する日本人を対象とした大規模臨床試験「KYOTO HEART Study」の結果が発表された2009年の論文について、欧州心臓病学会（ESC）学会誌から撤回が発表されたことを受け、プロモーション資材での使用を中止した。同論文は、ディオバンの有意性を証明する有力なエビデンスとして、同社が情報提供活動で使用していた。

撤回されたのは、ディオバンを非ARBの降圧薬と併用した高血圧症の患者群について、複数の心血管イベントの発症を非ARB降圧薬群と比較した試験に関する論文。09年9月にスペイン・バルセロナで開催されたESCでは、ディオバン群が主要評価項目の「複合心血管イベント」を

なぜ不祥事は起きたか

- EBMの名を借りて大規模臨床試験が（とくに競争の激しい抗圧薬で）次々に実施され、その「成果」が販売拡張に使われた。しかし
- 2007年から学会発表が宣伝に使えなくなり、販売拡張のために大規模試験の論文化が必要であった
- がん領域を除けば、研究者主導の臨床試験を実施する基盤はなかった
- がん領域など一部を除けば、臨床試験の品質管理・品質保証の概念はPIを務める医師にはなかった
- 「製薬会社と研究者あるいはそれを支える財団等が契約を交わし中立の立場で臨床試験を行う」という考え方は一般的でなかった。大学紛争以来、委任経理金という寄付で臨床試験は実施されることが多かった
- 寄付金が小額でとてもCROを本格的に使うような試験はできなかった
- 試験を請け負った小規模CROが破綻した？
- 中立で経験のある統計家の参加はなかった
- 日本ではプラセボ試験は市販後にはできない、という思い込みがあった。もちろん製薬会社の営業の立場からプラセボ試験は好ましくない

日本計量生物学会の「統計家行動基準」

- 情報を適切にあつかう
- 法やガイドラインを遵守する
- 不正行為を予防する
- 利益相反による弊害を防ぐ

日本計量生物学会（1980年設立）

会員数約500

企業、アカデミアの試験統計家が最も多く参加する学会

NHLBIモデルによる臨床試験

NHLBI : National Heart Lung Blood Institute

大橋靖雄 : 医学研究のための情報システム. 医学のあゆみ,1988;146:11-3.
もともとはGreenberg Report (1967)
http://sph.unc.edu/files/2013/07/greenberg_report.pdf

Bernard Greenberg

Coordinating Center = Data Center

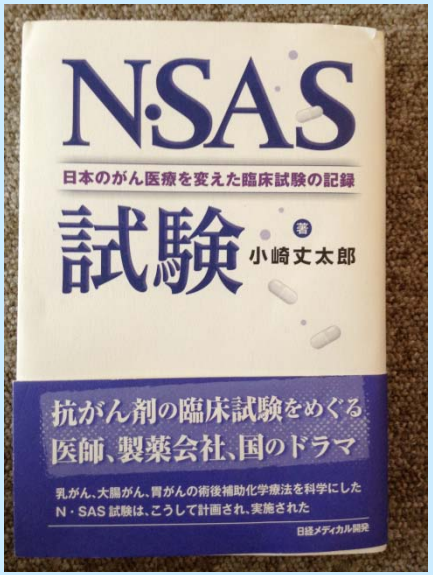
薬剤疫学的手法研究事業
（当時の厚生省1994-）

- 高脂血症のメバロチン（MEGAスタディ）
- がん術後補助療法のUFT（NSAS試験）
- 骨粗しょう症のグラケー（OFスタディ）

独立・中立なデータセンターと統計家

2013日経メディカル開発

「N・SAS試験」から



NSAS試験：厚生省薬剤疫学的研究事業、
メバロチンのMEGAの次の事業
乳癌、直腸癌・結腸癌、あとで胃癌
乳癌については1996年から2001年登録　結果はJCOに
NSASBC01：CMF（データなしで承認）とUFTの非劣性試験
患者団体の抗議活動など実施は困難を極めた

- N・SAS試験計画を後押ししたのが、国民と行政双方からの「大型医薬」への不信だった。「どうしようもない」と東大の大橋が斬って捨てる80年代の抗痴呆薬・・・
- 「UFTを潰してください」　：小林大鵬社長から白坂哲彦氏への激励
- 「（免疫賦活薬の）売り上げを半分にしてやる」：　がんセンター西條先生
- 「厚生省はUFTをつぶすつもりですよ」：大橋から大鵬幹部へ
- 「もし効かないということになったらどうする、えらいことだ」：厚生省黒川氏
- 「SDVをやりたかった」：EPS長屋氏

統計学者、生物統計学者？

今回の臨床試験PIの見解「統計家がいなかった」は本当か？

よい臨床試験を組むには必須というのは皆知っているが、さてどこにいけばいるのかよく分からず、統計家自身も「よい統計家に相談しなさい」とは言うけど「私がそうではない」と言うのみで、どこで誰に相談しろ、といってはくれないので殆ど幻ではないかと思われるような存在

里見・吉村：誰も教えてくれなかった癌臨床試験の正しい解釈、
中外医学社、2011

一連の試験開始は2001頃

BIOS開始	1989
東大疫学・生物統計学	1992
厚生省薬剤疫学事業（MEGA,NSAS)	1994
北里大学講座	1999
京都大学講座	2000
理科大修士コース	2002
久留米バイオ統計	2003

なぜ日本の医療アカデミアに統計家が存在しなかったか？

1990年ころまでは、医師研究者への文献検索・統計解析サービスは製薬会社営業部門の重要な仕事

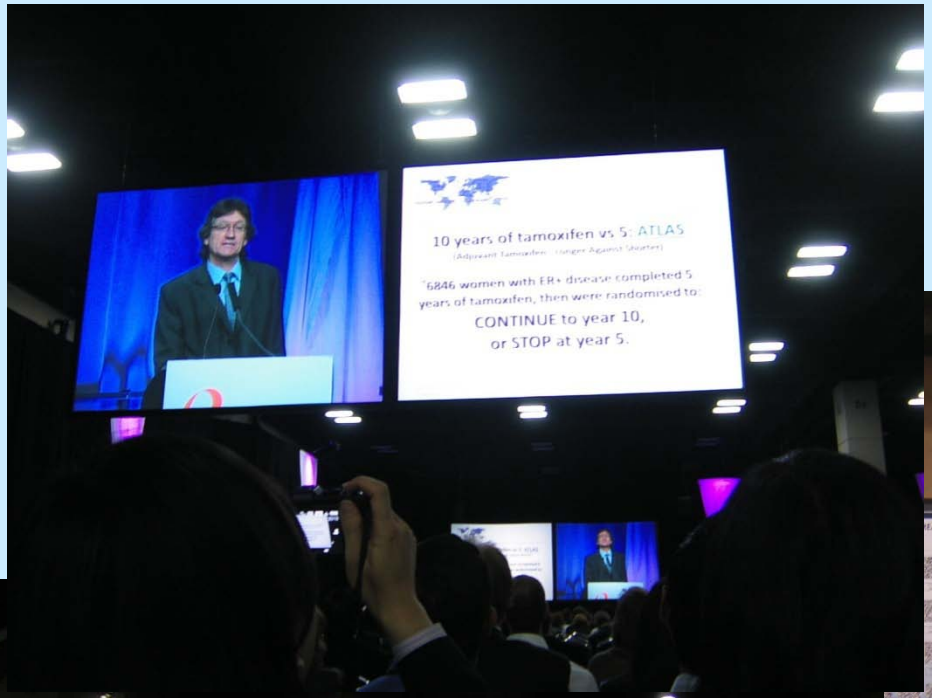
1991年の改正独占禁止法の施行、その具体的な業界版の製薬協公正競争規約が作成され、医療関係者に対して、無償でサービスを提供することを禁止した不当景品類防止法が発効し、上記のようなサービスは影をひそめたはずであったが・・・

臨床試験の中立性の確保

- 企業資金と学術的独立性
  - Montaner et al. “Industry-sponsored clinical research :a double-edged sword”, Lancet 2001; 358:1893-5.
  - 研究資金がアカデミアから民間・CROへシフト
  - 計画・解析・解釈・出版への資金提供者の影響
- ICMJE(International Committee of Medical Journal Editors)
  - 臨床試験登録の推進（2005July1からの試験は登録義務化）
  - 2010年にCOI宣言を含んだ統一投稿規程
- JAMAの投稿条件（2008）
  - ICMJEの認めるサイトへの事前登録（日本はUMIN）
  - ガイドラインCONSORTに沿った論文記述
  - アカデミア統計家による解析（現在は緩和されている）

San Antonio Breast Cancer Symposium 35 2012DEC04-08

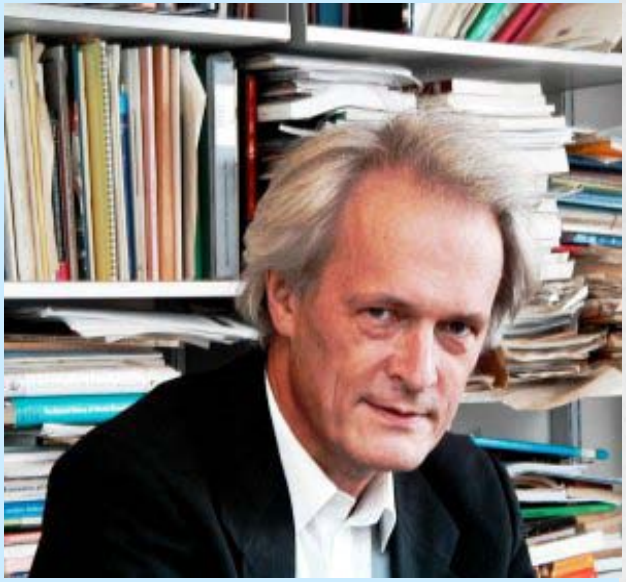
San Antonio Breast Cancer Symposium 35 2012DEC04-08



乳癌補助療法のメタアナリシス： メタアナリシスの嚆矢
Meta Analysis of Early Breast Cancer Trials

乳がん補助療法
Tamoxifen・卵巣摘出・
多剤化学療法の延命効果
Tamoxifen投与年数

Petoの執念
5年毎の再解析
研究者は競ってデータ提供
標準治療の発信



Systemic treatment of early breast cancer by hormonal, cytotoxic, or immune therapy

133 randomised trials involving 31 000 recurrences and 24 000 deaths among 75 000 women

EARLY BREAST CANCER TRIALISTS' COLLABORATIVE GROUP

In a worldwide collaboration, information was sought and centrally checked on mortality and recurrence for each woman in any randomised trial that began before 1985 of any aspect of systemic adjuvant therapy for early breast cancer. Checked data were available for 75 000 women (about 90% of those ever randomised), of whom 32% had died and another 10% had experienced recurrence. The parts now reviewed include 30 000 women in tamoxifen trials, 3000 in ovarian ablation trials, 11 000 in polychemotherapy trials, 15 000 in other chemotherapy comparisons, and 6000 in immunotherapy trials.

Highly significant reductions in the annual rates both of recurrence and of death are produced by tamoxifen (25% SD 2 recurrence and 17% SD 2 mortality: $2p < 0{\cdot}00001$), by ablation below age 50 (26% SD 6 recurrence and 25% SD 7 mortality: $2p = 0{\cdot}0004$), and by polychemotherapy (28% SD 3 recurrence and 16% SD 3 mortality: $2p < 0{\cdot}00001$), but not by ablation at older ages or by immunotherapy. (Tamoxifen also reduced the risk of development of contralateral breast cancer by 39% SD 9: $1p < 0{\cdot}00001$.) For tamoxifen and for polychemotherapy the avoidance of recurrence is chiefly during years 0–4 (this difference being maintained but not increased afterwards), but the avoidance of mortality is highly significant both during and after years 0–4, so the cumulative differences in survival produced by these relatively brief treatments (median: 2 years tamoxifen, 1 year polychemotherapy) are larger at 10 than at 5 years. There is little information beyond year 10 (except for ovarian ablation, which produces separately significant mortality reductions both during and after years 0–9). Both direct and indirect randomised comparisons show long-term polychemotherapy (eg, 12 months) to be no better than shorter (eg, 6 months) regimens, but do show polychemotherapy to be significantly better than single-agent chemotherapy. Indirect randomised comparisons do not reveal significant differences between different forms of polychemotherapy, or differences between different tamoxifen doses, but do show that long-term tamoxifen (eg, 2 years, or even 5 years) is significantly more effective than shorter tamoxifen regimens.

In old age (70+) tamoxifen is of demonstrated efficacy, but chemotherapy has not been evaluated. Between ages 50 and 69 direct comparisons show that chemotherapy plus tamoxifen is better ($1p < 0{\cdot}00001$) than chemotherapy alone both for recurrence and for mortality, and better ($1p < 0{\cdot}00001$) than tamoxifen alone for recurrence. In women aged under 50 chemotherapy and ovarian ablation appear, by an indirect comparison, to be of comparable efficacy, and the combination may be still better.

The 30–40% proportional risk reductions that can be produced by combined chemo-endocrine therapy in middle age are similar for node-positive and for node-negative patients, but the absolute improvement in 10-year survival is about twice as great for the former (at least 12 deaths avoided per 100 women treated) as for the latter.

Lancet 1992; **339**: 1–15, 71–85.

Collaborators are listed at the end of the report. Correspondence: EBCTCG Secretariat, ICRF/MRC Clinical Trial Service Unit, Nuffield Department of Clinical Medicine, Radcliffe Infirmary, Oxford OX2 6HE, UK.

タモキシフェンは
乳癌再発率を20-40%
減らす！

Tamoxifen reduces the recurrence of BC by 20-40%

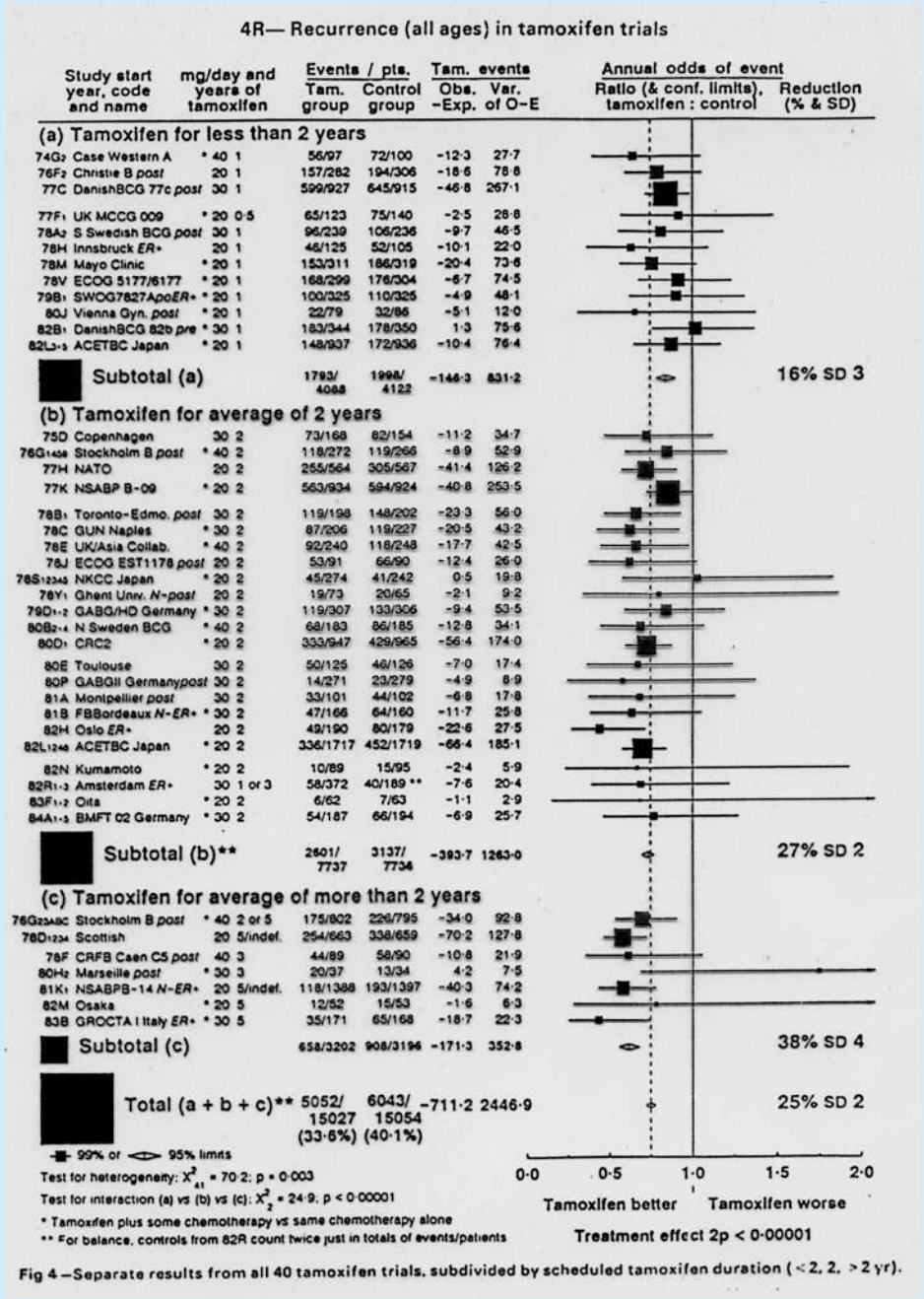


Fig 4—Separate results from all 40 tamoxifen trials, subdivided by scheduled tamoxifen duration (< 2, 2, > 2 yr).

英国における乳癌死亡率の減少 Reduction of BC Mortality
早期発見、早期治療の影響とくにタモキシフェン

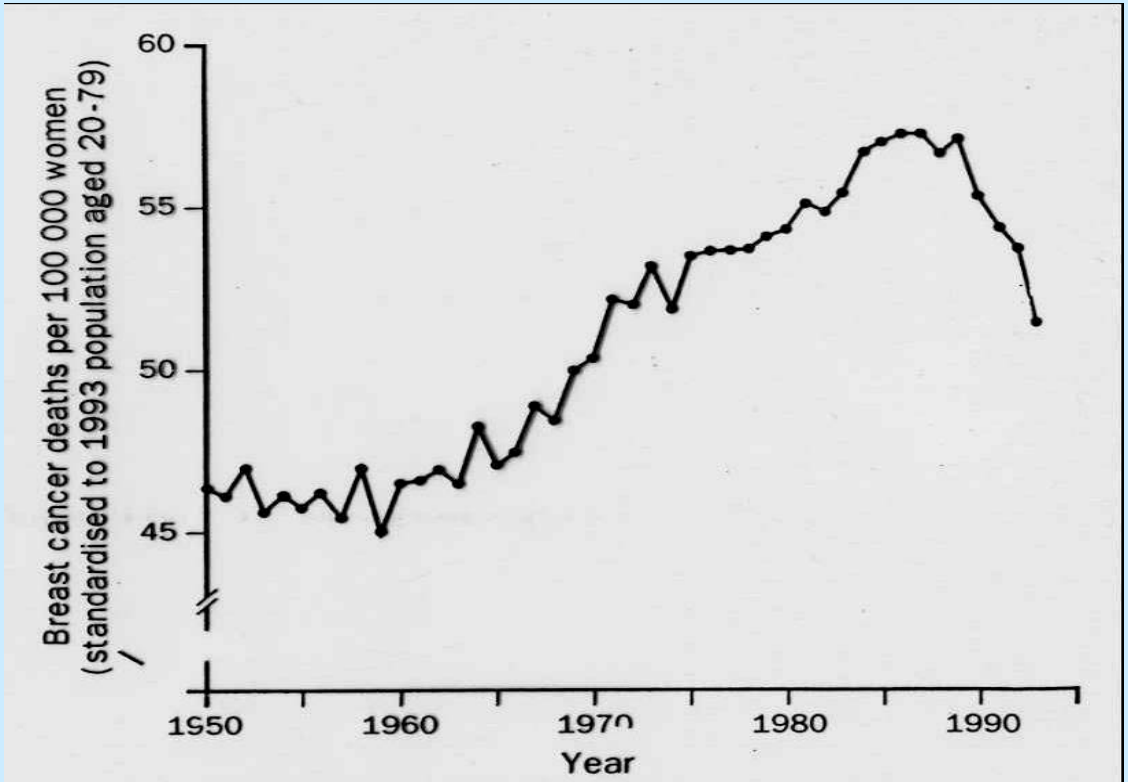
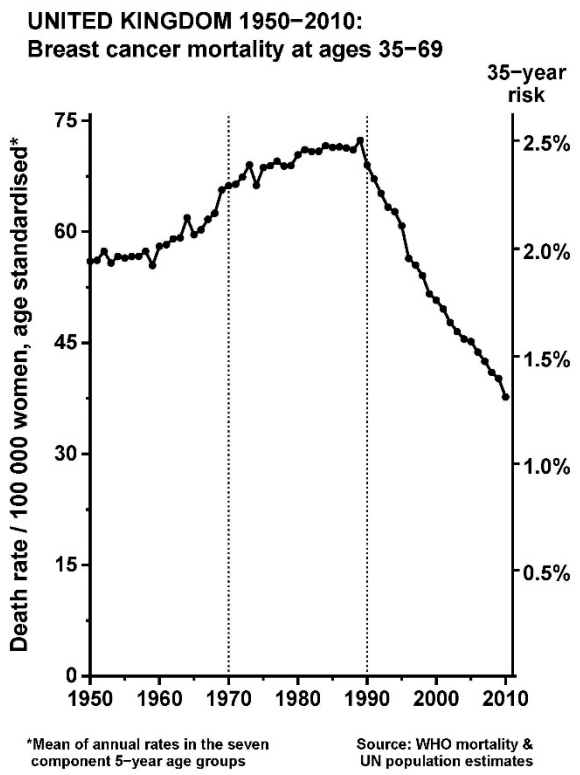
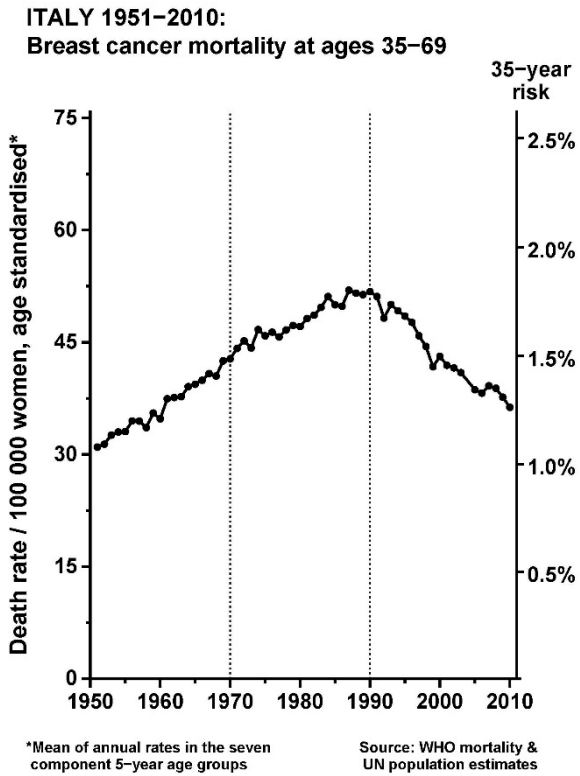
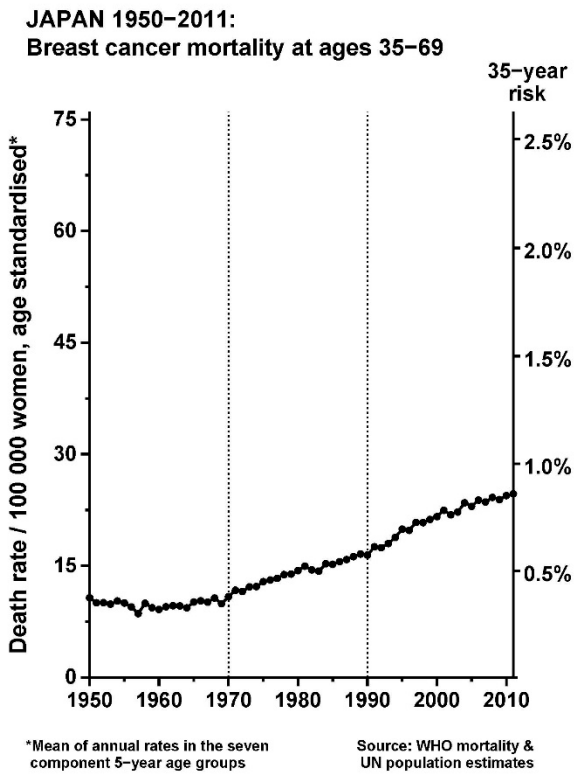


Figure: **Breast cancer mortality in England and Wales, 1950–93**

英国　　イタリア　　日本







乳癌死亡率の年次変化

Long-term effects of continuing adjuvant tamoxifen to 10 years versus stopping at 5 years after diagnosis of oestrogen receptor-positive breast cancer: ATLAS, a randomised trial

*Christina Davies, Hongchao Pan, Jon Godwin, Richard Gray, Rodrigo Arriagada, Vinod Raina, Mirta Abraham, Victor Hugo Medeiros Alencar, Atef Badran, Xavier Bonfill, Joan Bradbury, Michael Clarke, Rory Collins, Susan R Davis, Antonella Delmestri, John F Forbes, Peiman Haddad, Ming-Feng Hou, Moshe Inbar, Hussein Khaled, Joanna Kielanowska, Wing-Hong Kwan, Beela S Mathew, Bettina Müller, Antonio Nicolucci, Octavio Peralta, Fany Pernas, Lubos Petruzelka, Tadeusz Pienkowski, Balakrishnan Rajan, Maryna T Rubach, Sera Tort, Gerard Urrútia, Miriam Valentini, Yaochen Wang, Richard Peto, for the Adjuvant Tamoxifen: Longer Against Shorter (ATLAS) Collaborative Group**

Summary

Background For women with oestrogen receptor (ER)-positive early breast cancer, treatment with tamoxifen for 5 years substantially reduces the breast cancer mortality rate throughout the first 15 years after diagnosis. We aimed to assess the further effects of continuing tamoxifen to 10 years instead of stopping at 5 years.

Methods In the worldwide Adjuvant Tamoxifen: Longer Against Shorter (ATLAS) trial, 12 894 women with early breast cancer who had completed 5 years of treatment with tamoxifen were randomly allocated to continue tamoxifen to 10 years or stop at 5 years (open control). Allocation (1:1) was by central computer, using minimisation. After entry (between 1996 and 2005), yearly follow-up forms recorded any recurrence, second cancer, hospital admission, or death. We report effects on breast cancer outcomes among the 6846 women with ER-positive disease, and side-effects among all women (with positive, negative, or unknown ER status). Long-term follow-up still continues. This study is registered, number ISRCTN19652633.

Findings Among women with ER-positive disease, allocation to continue tamoxifen reduced the risk of breast cancer recurrence (617 recurrences in 3428 women allocated to continue *vs* 711 in 3418 controls, p=0·002), reduced breast cancer mortality (331 deaths *vs* 397 deaths, p=0·01), and reduced overall mortality (639 deaths *vs* 722 deaths, p=0·01). The reductions in adverse breast cancer outcomes appeared to be less extreme before than after year 10 (recurrence rate ratio [RR] 0·90 [95% CI 0·79–1·02] during years 5–9 and 0·75 [0·62–0·90] in later years; breast cancer mortality RR 0·97 [0·79–1·18] during years 5–9 and 0·71 [0·58–0·88] in later years). The cumulative risk of recurrence during years 5–14 was 21·4% for women allocated to continue versus 25·1% for controls; breast cancer mortality during years 5–14 was 12·2% for women allocated to continue versus 15·0% for controls (absolute mortality reduction 2·8%). Treatment allocation seemed to have no effect on breast cancer outcome among 1248 women with ER-negative disease, and an intermediate effect among 4800 women with unknown ER status. Among all 12 894 women, mortality without recurrence from causes other than breast cancer was little affected (691 deaths without recurrence in 6454 women allocated to continue versus 679 deaths in 6440 controls; RR 0·99 [0·89–1·10]; p=0·84). For the incidence (hospitalisation or death) rates of specific diseases, RRs were as follows: pulmonary embolus 1·87 (95% CI 1·13–3·07, p=0·01 [including 0·2% mortality in both treatment groups]), stroke 1·06 (0·83–1·36), ischaemic heart disease 0·76 (0·60–0·95, p=0·02), and endometrial cancer 1·74 (1·30–2·34, p=0·0002). The cumulative risk of endometrial cancer during years 5–14 was 3·1% (mortality 0·4%) for women allocated to continue versus 1·6% (mortality 0·2%) for controls (absolute mortality increase 0·2%).

Interpretation For women with ER-positive disease, continuing tamoxifen to 10 years rather than stopping at 5 years produces a further reduction in recurrence and mortality, particularly after year 10. These results, taken together with results from previous trials of 5 years of tamoxifen treatment versus none, suggest that 10 years of tamoxifen treatment can approximately halve breast cancer mortality during the second decade after diagnosis.

Funding Cancer Research UK, UK Medical Research Council, AstraZeneca UK, US Army, EU-Biomed.

Published Online
December 5, 2012
http://dx.doi.org/10.1016/S0140-6736(12)61963-1

See Online/Comment
http://dx.doi.org/10.1016/S0140-6736(12)62038-8

*Members listed at end of paper

Clinical Trial Service Unit and Epidemiological Studies Unit (CTSU), University of Oxford, UK (C Davies MBChB, H Pan PhD, Prof R Gray MSc, Prof R Collins FMedSci, A Delmestri PhD, Y Wang MD, Prof R Peto FRS); **Glasgow Caledonian University, Glasgow, UK** (J Godwin DPhil); **Institut Gustave-Roussy, Villejuif, France** (Prof R Arriagada MD); **Institute Rotary Cancer Hospital, All-India Institute of Medical Sciences, New Delhi, India** (Prof V Raina MD); **Instituto Cardiovascular Rosario (ICR), Rosario, Argentina** (M Abraham MD); **Instituto do Cancer do Ceará (ICC), Fortaleza, Brazil** (V H Medeiros Alencar MD); **National Cancer Institute, Cairo University, Cairo, Egypt** (Prof H Khaled MD, A Badran PhD); **Sant Pau Biomedical Research Institute (IIB Sant Pau-CIBERESP), Barcelona, Spain** (X Bonfill MD, S Tort MD, G Urrútia MD); **School of Public Health and Preventive Medicine, Monash University, Melbourne, Victoria, Australia** (Prof S R Davis MBBS, J Bradbury); **Queens University, Belfast, UK** (Prof M Clarke DPhil); **Australia and New Zealand Breast Cancer Trials Group, University of**

ATLAS
onlie published
Lancet, Dec5, 2012

ER+早期乳癌
に対するタモキシフェン
5年投与の後

5年継続かストップか

1996-2005年に36ヶ国
から15000人以上が登録

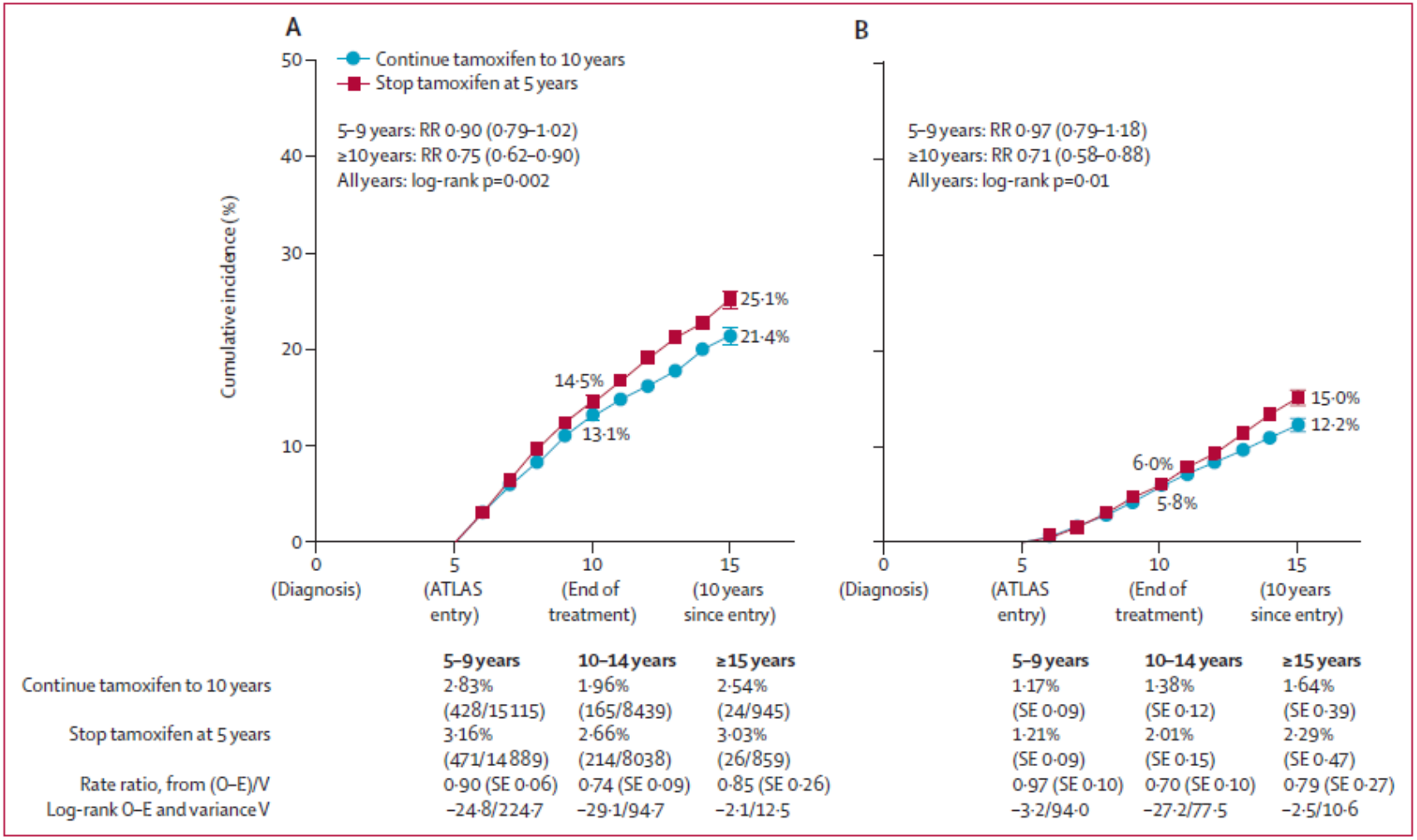
ER+6846人が解析対象

ATLAS Steering Committee 2005

ATLAS coordinating centre and CTSU supporting staff (1995–2012)

Principal investigator: Christina Davies. Administrative office: Jenny Sayer, Valerie Collett. Central randomisation: Jill Crowther, Angela Radley. Analysts/programmers: Antonella Delmestri, Jon Godwin, Yaochen Wang. Statisticians: Richard Gray, Hongchao Pan, Richard Peto. Former staff: A Beighton, M Forster, A Headon, C Hope, S Knight, P McGale, S Mozley, H Monaghan, A Muldal, A Naughten, S Turner. CTSU also prepared the software that randomised in four regional or national coordinating centres (Australia/New Zealand: National Health & Medical Research Council Clinical Trials Unit; Italy: Consorzio Mario Negri Sud; Japan: Tokyo University Department of Epidemiology and Biostatistics, then from 2001 Japan Clinical Research Support Unit; Spain: Institut d'Investigació Biomèdica Sant Pau, Barcelona [FIS PI020391]). CONAC (Corporación Nacional del Cáncer: Director C Agosin) distributed Nolvadex in Chile.

ATLAS: タモキシフェンには残存効果、継続の効果は10年（ランダム化から5年）後から



	5–9 years	10–14 years	≥15 years	5–9 years	10–14 years	≥15 years
Continue tamoxifen to 10 years	2·83% (428/15 115)	1·96% (165/8439)	2·54% (24/945)	1·17% (SE 0·09)	1·38% (SE 0·12)	1·64% (SE 0·39)
Stop tamoxifen at 5 years	3·16% (471/14 889)	2·66% (214/8038)	3·03% (26/859)	1·21% (SE 0·09)	2·01% (SE 0·15)	2·29% (SE 0·47)
Rate ratio, from (O–E)/V	0·90 (SE 0·06)	0·74 (SE 0·09)	0·85 (SE 0·26)	0·97 (SE 0·10)	0·70 (SE 0·10)	0·79 (SE 0·27)
Log-rank O–E and variance V	−24·8/224·7	−29·1/94·7	−2·1/12·5	−3·2/94·0	−27·2/77·5	−2·5/10·6

***Figure 3*: Recurrence (A) and breast cancer mortality (B) by treatment allocation for 6846 women with ER-positive disease**
Bars show SE. Recurrence rates are percentage per year (events/patient-years of follow-up). Death rates (overall rate – rate in women without recurrence) are percentage per year (SE). ATLAS=Adjuvant Tamoxifen: Longer Against Shorter.

推定される継続投与の効果（対無投与）は　10年後には0.52、つまりがん死亡半減！
ガイドライン、教科書への影響必至、治療標準が変わる

	A: effects in meta-analyses of the trials of 5 years of tamoxifen *vs* none[1] (n=10 645)	B: effects in the ATLAS trial of continuing tamoxifen to 10 years *vs* stopping at 5 years (n=6846)	C: estimated effects in a trial of 10 years of tamoxifen *vs* none (product of A and B)
Recurrence			
0–4 years	0·53 (0·48–0·57)*	1	0·53 (0·48–0·57)*
5–9 years	0·68 (0·60–0·78)*	0·90 (0·79–1·02)	0·61 (0·51–0·73)*
≥10 years	0·94 (0·79–1·12)	0·75 (0·62–0·90)†	0·70 (0·54–0·91)†
Breast cancer mortality			
0–4 years	0·71 (0·62–0·80)*	1	0·71 (0·62–0·80)*
5–9 years	0·66 (0·58–0·75)*	0·97 (0·79–1·18)	0·64 (0·50–0·82)‡
≥10 years	0·73 (0·62–0·86)‡	0·71 (0·58–0·88)§	0·52 (0·40–0·68)*

(A) Trials of 5 years of tamoxifen (n=10 645; ~80% complied). (B) ATLAS trial of 10 years *vs* 5 years of tamoxifen (n=6846; ~80% difference in tamoxifen use [figure 2]). (C) Hypothetical trial of 10 years of tamoxifen *vs* none (with ~80% compliance). Two-sided p values in this table relate to particular time periods; values elsewhere combine all time periods. ER=oestrogen receptor. *p<0·00001. †p<0·01. ‡p=0·0001. §p=0·0016.

Table 3: **Event rate ratios (95% CIs) in ER-positive disease, by time period from diagnosis**

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.

わが国では最初と言っても良い循環器臨床試験のエビデンス

日本発薬剤プラバスタチン有無のランダム化試験

エンドポイントはCHD発症

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

Haruo Nakamura, Kikuo Arakawa, Hiroshige Itakura, Akira Kitabatake, Yoshio Goto, Takayoshi Toyota, Noriaki Nakaya, Shoji Nishimoto, Masaharu Muranaka, Akira Yamamoto, Kyoichi Mizuno, Yasuo Ohashi, for the MEGA Study Group

Summary

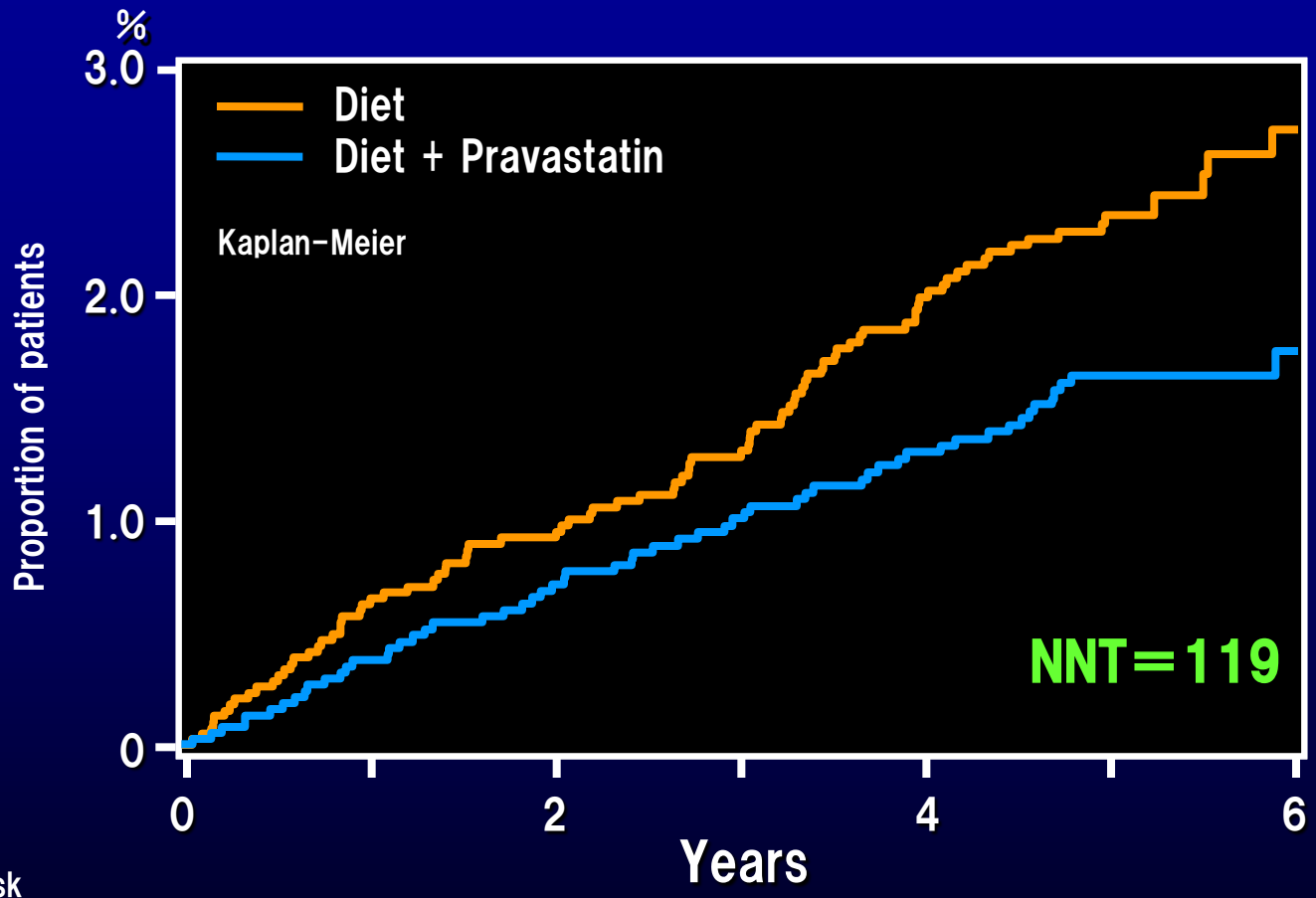
Background Evidence-based treatment for hypercholesterolaemia in Japan has been hindered by the lack of direct evidence in this population. Our aim was to assess whether evidence for treatment with statins derived from western populations can be extrapolated to the Japanese population.

Methods In this prospective, randomised, open-labelled, blinded study, patients with hypercholesterolaemia (total cholesterol 5·69–6·98 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 NCT00211705.

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 3·90 mmol/L) and 18·0% (from 4·05 mmol/L to 3·31 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.

Primary Endpoint - CHD -



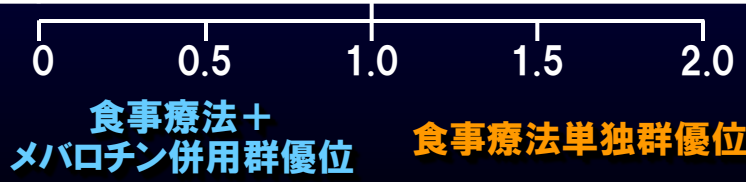
33%
(p=0.01)
Log-rank test

Number at risk							
Diet	3,966	3,758	3,648	3,529	3,430	2,476	830
Diet + Pravastatin	3,866	3,642	3,490	3,385	3,307	2,434	859

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

一次評価項目

評価項目	例数（/1,000人・年） 食事療法単独群（n＝3,966）	例数（/1,000人・年） 食事療法＋メバロチン併用群（n＝3,866）	ハザード比（95%信頼区間）	p値*
冠動脈疾患	101（5.0）	66（3.3）	0.67（0.49-0.91）	0.01
心筋梗塞	33（1.6）	17（0.9）	0.52（0.29-0.94）	0.03
致死性	3（0.1）	2（0.1）		
非致死性	30（1.5）	16（0.8）		
心臓死/突然死※	10（0.5）	5（0.2）	0.51（0.18-1.50）	0.21
狭心症	57（2.8）	46（2.3）	0.83（0.56-1.23）	0.35
冠動脈血行再建術	66（3.2）	39（2.0）	0.60（0.41-0.89）	0.01

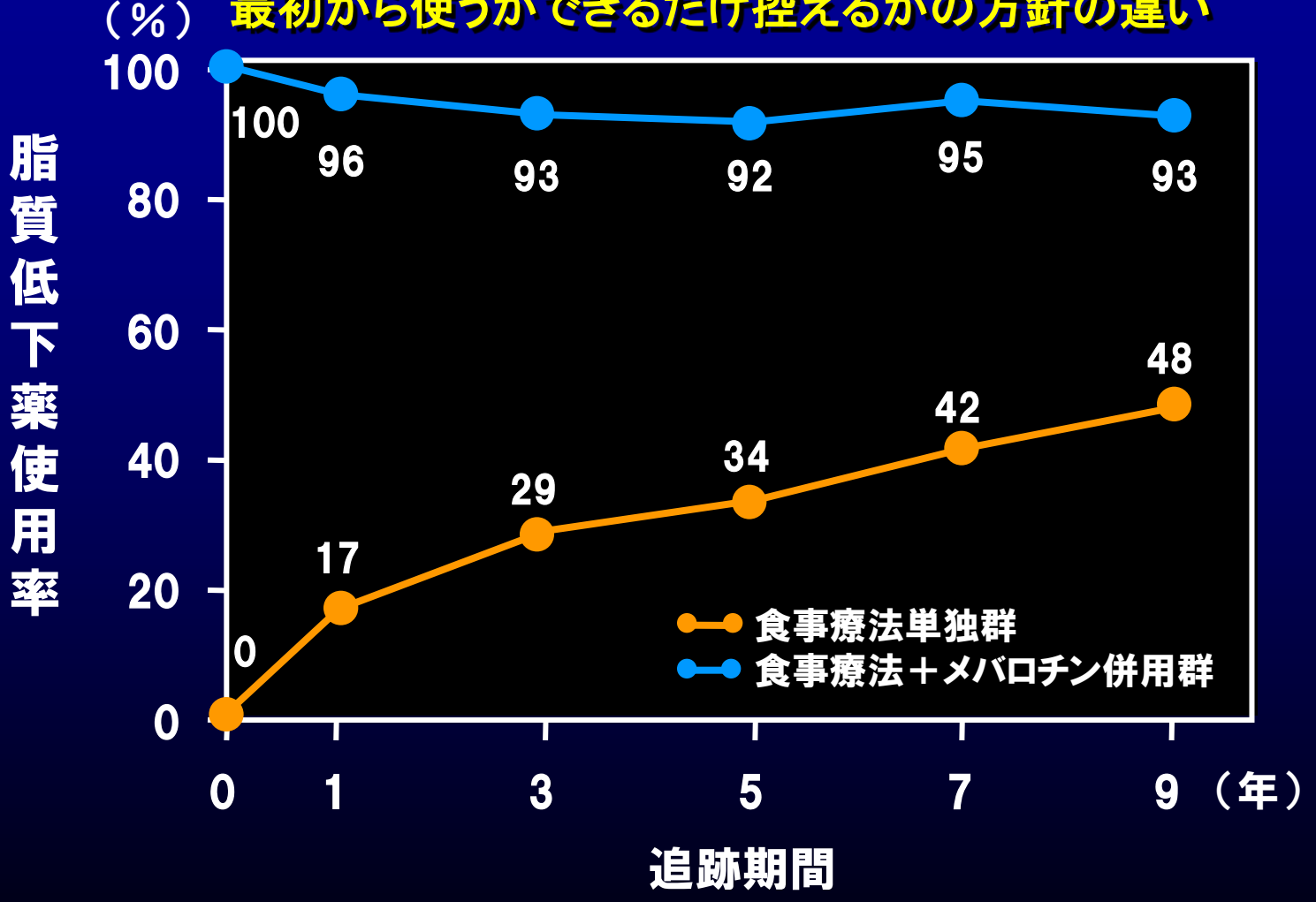


*：Log-rank検定
※：24時間以内に生じた原因不明の死亡
ハザード比の正方形の大きさは、イベント発生数に比例

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

患者の脂質低下薬の使用状況

コンプライアンスが低いといって質が低いわけではない
比較しているのはﾌﾟﾗﾊﾞｽﾀﾁﾝを
最初から使うかできるだけ控えるかの方針の違い



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

UMIN Registgration
Type of Clinical Trials

- 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

Explanatory vs Pragmatic

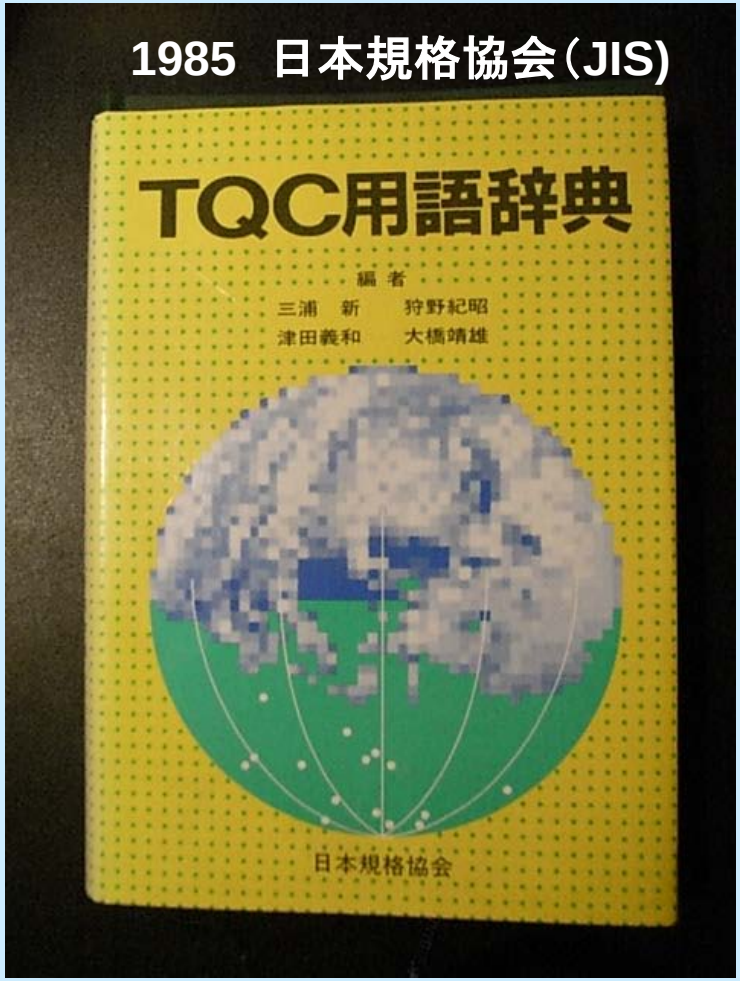
Explanatory	Pragmatic	
目的は明確で広く理解されている	科学者にはよく理解されない	Flamant et al.
実験室での実験の拡張	臨床実践への指針を与える	Flamant et al.
仮説検定	治療方針選択	Flamant et al.
method-effectiveness	use-effectiveness	Piantadosi
両者はcompeteする		Piantadosi
analysis of compliers only	analysis by inten(tion)-to-treat	Pocock
PPS	FAS	ICH E9
欠損での処理は what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

Quality 品質？（Z8101）

品物又はサービスが、使用目的を満たしているかどうかを決定するための評価の対象となる固有の性質・性能の全体

備考　1. 品物又はサービスが使用目的を満たしているかどうか判定する際に、その品物又はサービスが**社会に与える影響**を考慮する必要がある。

2. 品質は品質特性によって構成される。例えば・・・



Quality 質、品質？

- 複数の評価機軸によって評価される「構成概念Construct」
  他の例：知能、QOL

- 誰が評価するか　顧客
  製薬会社にとっては　　審査当局、そして国民
  医師研究者にとっては　医学コミュニティ、そして国民

  2重の顧客の存在がこの分野の難しさ

- 当たり前品質と魅力的品質（狩野紀昭）

Domain of Clinical Trial Quality Characteristics
臨床試験の品質特性

- 標準治療に対するインパクト
- 正確さと精密さ、一般化可能性
- プラン
- デザイン
- 実施
- 出版
- 品質保証のシステム

Domain of Clinical Trial Quality Characteristics
臨床試験の品質特性

- 標準治療に対するインパクト

- 正確さと精密さ、一般化可能性

- プラン　　研究組織
- デザイン　プロトコルの質(2013年1月にSPIRIT発表）
  - relevantな仮説　pragmaticかexplanatoryか
  - エンドポイントの設定
  - 症例数と適切な中間解析、適切な解析方法
- 実施　　治療コンプライアンス
  - 追跡と脱落、欠損データ
  - データ管理とそのシステム
- 出版　　（CONSORTにより品質特性は測定可能）
- 品質保証のシステム

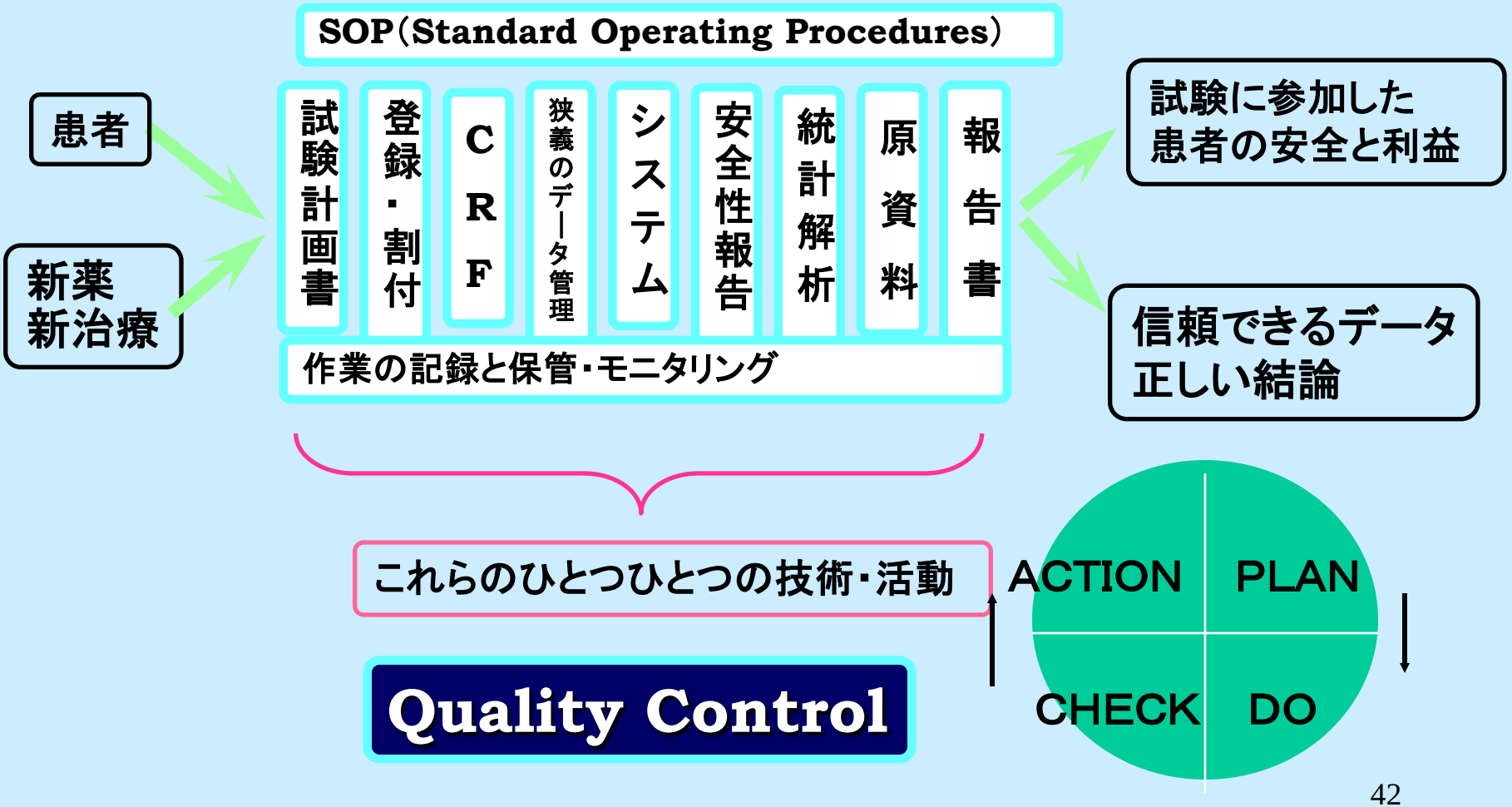
研究データ・論文の正しさを
いかに保証するか？

- プロセスとしての保証（ISO9000、14000の考え方）
  - 組織・プロセスの確立
  - SOP
  - 外部からの監査
- 中立性・透明性
- 参加者のモラル・教育

福田治彦先生による

Quality Control（品質管理）

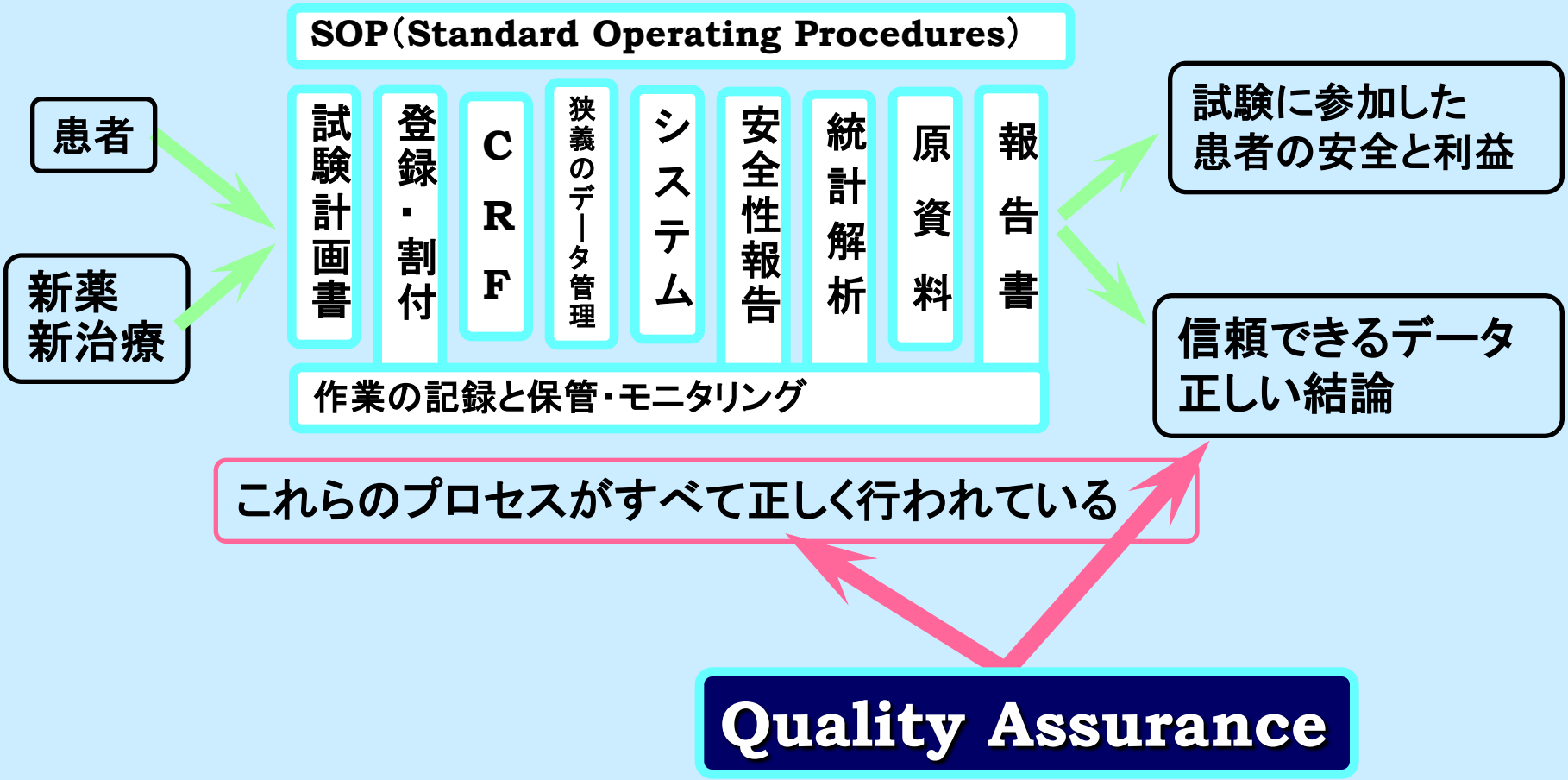
input **system** **output**



福田治彦先生による

Quality Assurance（品質保証）

input　system　output



http://www.equator-network.org/

CONSORT

CONSORT statement

- 臨床疫学者、統計学者、主要医学雑誌編集者グループICMJE（Int Comm Med J Editor)によるランダム化臨床試験（2群）の標準報告様式
- 患者の流れ図と25項目のチェックリスト

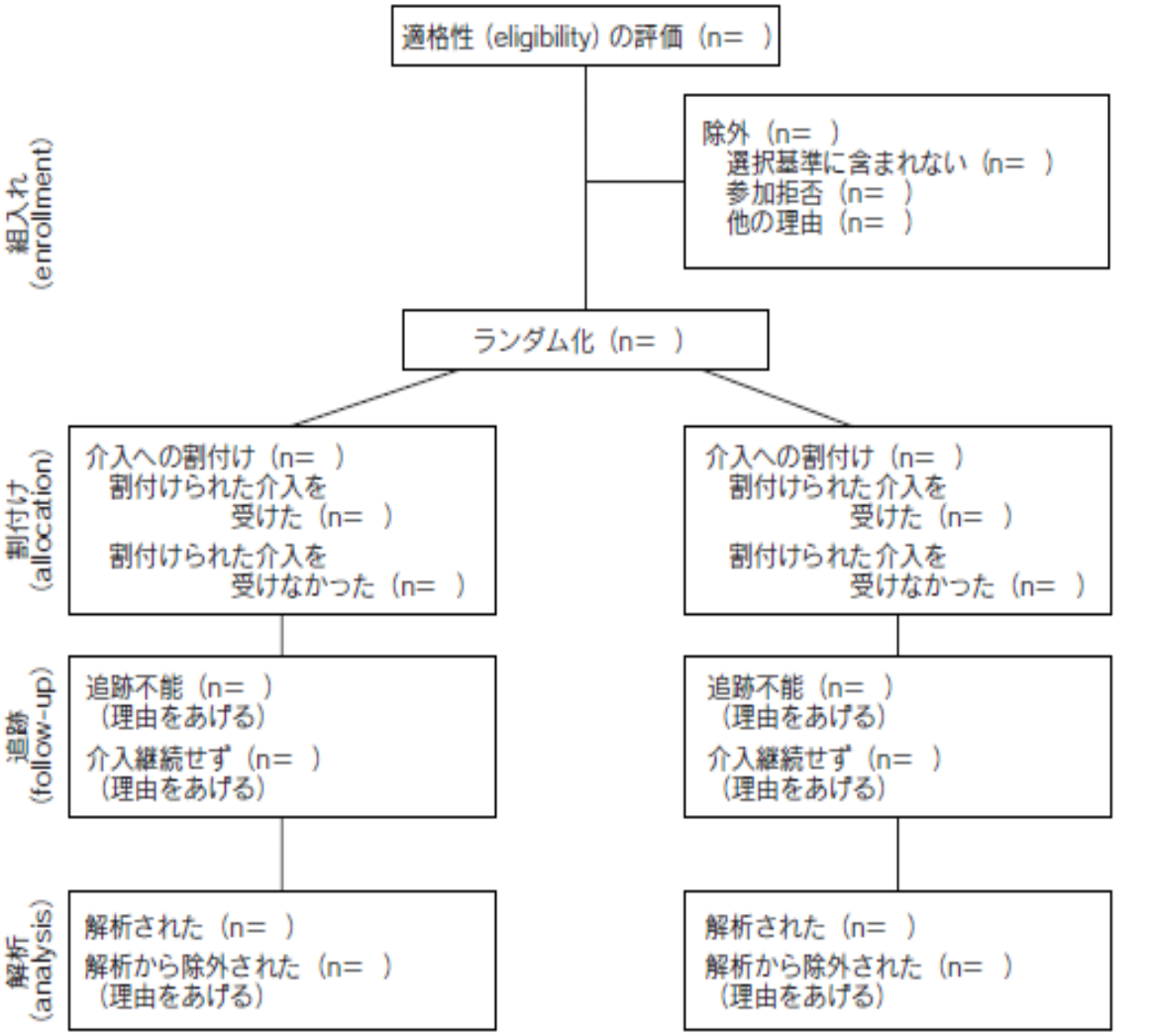
- NEJM、Lancet、BMJ、Ann Int Medなど一流誌が採用
- 近年の臨床試験の質向上に貢献したとの評価

＊医学研究の質を高めるための支援サイトhttp://www.equator-network.org/

*中山・津谷「臨床研究と疫学研究のための国際ルール集」、ライフサイエンス出版

CONSORT Statement フローチャート

図　ランダム化比較試験の各段階の被験者(subject)の数を示すフローチャート



適格性のラフな評価

↓ 一般化可能性？

ランダム化

ランダム化

↓ 同意？登録の質

治療開始

治療開始

↓

追跡の質

治療不遵守

評価・解析対象

CONSORT Statement チェックリスト（2010）

CONSORT

章／トピック (Section/Topic)	項目番号 (Item No)	チェックリスト項目 (Checklist Item)	報告頁 (Reported on page No)
タイトル・抄録 (Title and Abstract)			
	1a	タイトルにランダム化比較試験であることを記載。	
	1b	試験デザイン（trial design），方法（method），結果（result），結論（conclusion）の構造化抄録（詳細は「雑誌および会議録でのランダム化試験の抄録に対する CONSORT 声明」[21,31]を参照）。	
はじめに (Introduction)			
背景・目的 (Background and Objective)	2a	科学的背景と論拠（rationale）の説明。	
	2b	特定の目的または仮説（hypothesis）。	
方法 (Method)			
試験デザイン (Trial Design)	3a	試験デザインの記述（並行群間，要因分析など），割付け比を含む。	
	3b	試験開始後の方法上の重要な変更（適格基準 eligibility criteria など）とその理由。	
参加者 (Participant)	4a	参加者の適格基準（eligibility criteria）。	
	4b	データが収集されたセッティング（setting）と場所。	
介入 (Intervention)	5	再現可能となるような詳細な各群の介入。実際にいつどのように実施されたかを含む。	
アウトカム (Outcome)	6a	事前に特定され明確に定義された主要・副次的アウトカム評価項目。いつどのように評価されたかを含む。	
	6b	試験開始後のアウトカムの変更とその理由。	
症例数 (Sample size)	7a	どのように目標症例数が決められたか。	
	7b	あてはまる場合には，中間解析と中止基準の説明。	
ランダム化 (Randomization)			
順番の作成 (Sequence generation)	8a	割振り（allocation）順番を作成（generate）した方法。	
	8b	割振りのタイプ：制限の詳細（ブロック化，ブロックサイズなど）。	
割振りの隠蔽機構 (Allocation concealment mechanism)	9	ランダム割振り順番の実施に用いられた機構（番号付き容器など），各群の割付けが終了するまで割振り順番が隠蔽されていたかどうかの記述。	
実施 (Implementation)	10	誰が割振り順番を作成したか，誰が参加者を組入れ（enrollment）たか，誰が参加者を各群に割付けた（assign）か。	
ブラインディング (Blinding)	11a	ブラインド化されていた場合，介入に割付け後，誰がどのようにブラインドかされていたか（参加者，介入実施者，アウトカムの評価者など）。	
	11b	関連する場合，介入の類似性の記述。	
統計学的手法 (Statistical method)	12a	主要・副次的アウトカムの群間比較に用いられた統計学的手法。	
	12b	サブグループ解析や調整解析のような追加的解析の手法。	

CONSORT Statement チェックリスト（2010） 続き

CONSORT

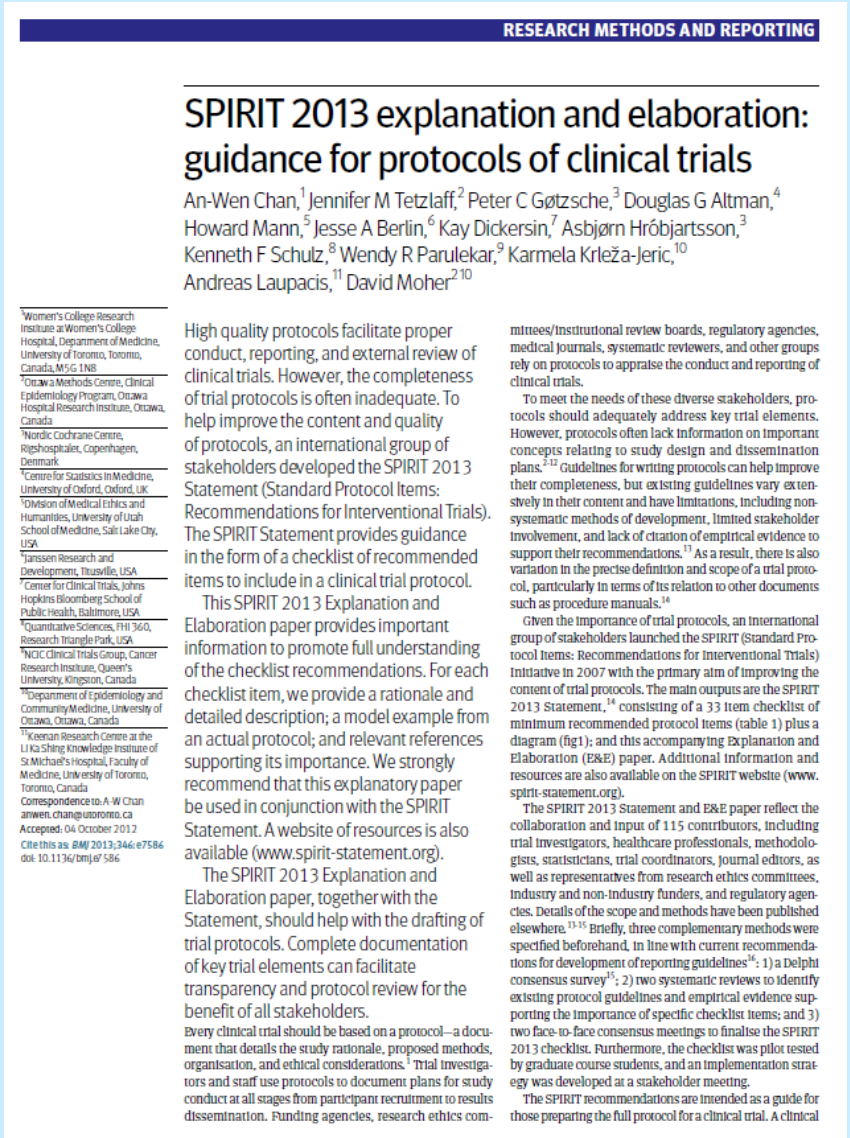
結果（Results）			
参加者の流れ（Participant flow）（フローチャートを強く推奨）	13a	各群について，ランダム割付けされた人数，意図された治療を受けた人数，主要アウトカムの解析に用いられた人数の記述。	
	13b	各群について，追跡不能例とランダム化後の除外例を理由とともに記述。	
募集（Recruitment）	14a	参加者の募集期間と追跡期間を特定する日付。	
	14b	試験が終了または中止した理由。	
ベースライン・データ（Baseline data）	15	各群のベースラインにおける人口統計学的（demographic），臨床的な特性を示す表。	
解析された人数（Number analyzed）	16	各群について，各解析における参加者数（分母），解析が元の割付け群によるものであるか。	
アウトカムと推定（Outcome and estimation）	17a	主要・副次的アウトカムのそれぞれについて，各群の結果，介入のエフェクト・サイズの推定とその精度（95％信頼区間など）。	
	17b	2項アウトカムについては，絶対エフェクト・サイズと相対エフェクト・サイズの両方を記載することが推奨される。	
補助的解析（Ancillary analysis）	18	サブグループ解析や調整解析を含む，実施した他の解析の結果。事前に特定された解析と探索的解析を区別する。	
害（Harm）	19	各群のすべての重要な害（harm）または意図しない効果（詳細は「ランダム化試験における害のよりよい報告：CONSORT 声明の拡張」[28]を参照）。	
考察（Discussion）			
限界（Limitation）	20	試験の限界，可能性のあるバイアスや精度低下の原因，関連する場合は解析の多重性の原因を記載。	
一般化可能（Generalisability）	21	試験結果の一般化可能性（外的妥当性，適用性）。	
解釈（Interpretation）	22	結果の解釈，有益性と有害性のバランス，他の関連するエビデンス。	
その他の情報（Other information）			
登録（Registration）	23	登録番号と試験登録名。	
プロトコール（Protocol）	24	可能であれば，完全なプロトコールの入手方法。	
資金提供者（Funding）	25	資金提供者と他の支援者（薬剤の供給者など），資金提供者の役割。	

＊本声明は，各項目についての重要な解説を記載したCONSORT 2010解説と詳細[13]とともに用いることを強く推奨する。クラスターランダム化比較試験[11]，非劣性・同等性試験[12]，非薬理学的治療[32]，ハーブ療法[33]，実用的試験[34]については，CONSORT声明拡張版を推奨する。そのほかの拡張版も近日発表予定（それらと本チェックリスト関連の最新情報はwww.consort-statement.orgを参照）。

SPIRIT

Standard Protocol Items Recommendations for Interventional Trials

http://www.spirit-statement.org/



RESEARCH METHODS AND REPORTING

SPIRIT 2013 explanation and elaboration: guidance for protocols of clinical trials

An-Wen Chan,[1] Jennifer M Tetzlaff,[2] Peter C Gøtzsche,[3] Douglas G Altman,[4] Howard Mann,[5] Jesse A Berlin,[6] Kay Dickersin,[7] Asbjørn Hróbjartsson,[3] Kenneth F Schulz,[8] Wendy R Parulekar,[9] Karmela Krleža-Jeric,[10] Andreas Laupacis,[11] David Moher[2 10]

[1]Women's College Research Institute at Women's College Hospital, Department of Medicine, University of Toronto, Toronto, Canada, M5G 1N8
[2]Ottawa Methods Centre, Clinical Epidemiology Program, Ottawa Hospital Research Institute, Ottawa, Canada
[3]Nordic Cochrane Centre, Rigshospitalet, Copenhagen, Denmark
[4]Centre for Statistics in Medicine, University of Oxford, Oxford, UK
[5]Division of Medical Ethics and Humanities, University of Utah School of Medicine, Salt Lake City, USA
[6]Janssen Research and Development, Titusville, USA
[7]Center for Clinical Trials, Johns Hopkins Bloomberg School of Public Health, Baltimore, USA
[8]Quantitative Sciences, FHI 360, Research Triangle Park, USA
[9]NCIC Clinical Trials Group, Cancer Research Institute, Queen's University, Kingston, Canada
[10]Department of Epidemiology and Community Medicine, University of Ottawa, Ottawa, Canada
[11]Keenan Research Centre at the Li Ka Shing Knowledge Institute of St Michael's Hospital, Faculty of Medicine, University of Toronto, Toronto, Canada
Correspondence to: A-W Chan anwen.chan@utoronto.ca
Accepted: 04 October 2012
Cite this as: BMJ 2013;346:e7586
doi: 10.1136/bmj.e7586

High quality protocols facilitate proper conduct, reporting, and external review of clinical trials. However, the completeness of trial protocols is often inadequate. To help improve the content and quality of protocols, an international group of stakeholders developed the SPIRIT 2013 Statement (Standard Protocol Items: Recommendations for Interventional Trials). The SPIRIT Statement provides guidance in the form of a checklist of recommended items to include in a clinical trial protocol.

This SPIRIT 2013 Explanation and Elaboration paper provides important information to promote full understanding of the checklist recommendations. For each checklist item, we provide a rationale and detailed description; a model example from an actual protocol; and relevant references supporting its importance. We strongly recommend that this explanatory paper be used in conjunction with the SPIRIT Statement. A website of resources is also available (www.spirit-statement.org).

The SPIRIT 2013 Explanation and Elaboration paper, together with the Statement, should help with the drafting of trial protocols. Complete documentation of key trial elements can facilitate transparency and protocol review for the benefit of all stakeholders.

Every clinical trial should be based on a protocol—a document that details the study rationale, proposed methods, organisation, and ethical considerations.[1] Trial investigators and staff use protocols to document plans for study conduct at all stages from participant recruitment to results dissemination. Funding agencies, research ethics committees/institutional review boards, regulatory agencies, medical journals, systematic reviewers, and other groups rely on protocols to appraise the conduct and reporting of clinical trials.

To meet the needs of these diverse stakeholders, protocols should adequately address key trial elements. However, protocols often lack information on important concepts relating to study design and dissemination plans.[2-12] Guidelines for writing protocols can help improve their completeness, but existing guidelines vary extensively in their content and have limitations, including non-systematic methods of development, limited stakeholder involvement, and lack of citation of empirical evidence to support their recommendations.[13] As a result, there is also variation in the precise definition and scope of a trial protocol, particularly in terms of its relation to other documents such as procedure manuals.[14]

Given the importance of trial protocols, an international group of stakeholders launched the SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) initiative in 2007 with the primary aim of improving the content of trial protocols. The main outputs are the SPIRIT 2013 Statement,[14] consisting of a 33 item checklist of minimum recommended protocol items (table 1) plus a diagram (fig1); and this accompanying Explanation and Elaboration (E&E) paper. Additional information and resources are also available on the SPIRIT website (www.spirit-statement.org).

The SPIRIT 2013 Statement and E&E paper reflect the collaboration and input of 115 contributors, including trial investigators, healthcare professionals, methodologists, statisticians, trial coordinators, journal editors, as well as representatives from research ethics committees, industry and non-industry funders, and regulatory agencies. Details of the scope and methods have been published elsewhere.[13-15] Briefly, three complementary methods were specified beforehand, in line with current recommendations for development of reporting guidelines[16]: 1) a Delphi consensus survey[15]; 2) two systematic reviews to identify existing protocol guidelines and empirical evidence supporting the importance of specific checklist items; and 3) two face-to-face consensus meetings to finalise the SPIRIT 2013 checklist. Furthermore, the checklist was pilot tested by graduate course students, and an implementation strategy was developed at a stakeholder meeting.

The SPIRIT recommendations are intended as a guide for those preparing the full protocol for a clinical trial. A clinical

33項目のプロトコルチェックリスト

管理情報
- タイトル、試験登録、
- 版、資金提供、役割と責任

イントロダクション
- 背景とrationale、目的、デザイン

方法
- 設定、適格条件、介入、アウトカム、
- タイムライン、サンプルサイズ、
- リクルート戦略、割付け、マスキング、
- データ収集、データ管理、統計解析、
- モニタリング、有害事象対応、監査

倫理と公表
- 倫理審査、プロトコル変更、同意、
- 秘密保持、COI、データへのアクセス、
- 試験後の患者対応、公表ポリシー

付録
- 説明・同意に用いる資料、
- 検体の収集・測定・保存についての資料

ICMJE INTERNATIONAL COMMITTEE of MEDICAL JOURNAL EDITORS

http://www.equator-network.org/

論文作成に関するガイドライン
- CONSORT：RCT
- STROBE：　疫学観察研究
- TREND：　非ランダム化研究
- PRISMA：　RCTメタアナリシス
- MOOSE：　疫学メタアナリシス
- STRAD：　診断
- ORION：　院内感染
- STREGA：　遺伝的相関研究

（SPIRIT:RCTプロトコル）

ICMJE統一規定
「論文の書き方」

ICMJE Recommendation
「出版とはどうあるべきか」

Recommendations for the Conduct, Reporting, Editing, and Publication of Scholarly Work in Medical Journals*

Updated August 2013

I. About the Recommendations

A. Purpose of the Recommendations

ICMJE developed these recommendations to review best practice and ethical standards in the conduct and reporting of research and other material published in medical journals, and to help authors, editors, and others involved in peer review and biomedical publishing create and distribute accurate, clear, unbiased medical journal articles. The recommendations may also provide useful insights into the medical editing and publishing process for the media, patients and their families, and general readers.

B. Who Should Use the Recommendations?

These recommendations are intended primarily for use by authors who might submit their work for publication to ICMJE member journals. Many non-ICMJE journals voluntarily use these recommendations (see www.icmje.org/journals.html). The ICMJE encourages that use but has no authority to monitor or enforce it. In all cases, authors should use these recommendations along with individual journals' instructions to authors. Authors should also consult guidelines for the reporting of specific study types (e.g., the CONSORT guidelines for the reporting of randomized trials); see http://equator-network.org.

* Developed by members of the ICMJE over the period 2011 to 2013.

ICMJE Recommendataion (2013Aug13): Authorship?

•Substantial contributions to: the conception or design of the work; or the acquisition, analysis, or interpretation of data for the work; AND
•Drafting the work or revising it critically for important intellectual content; AND
•Final approval of the version to be published; AND

•Agreement to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.
研究結果の正確性とintegrityに関する疑問に対する説明責任

質の高い臨床試験とは（検証的試験の場合）

- 目的・仮説が明確で
- その検証のためにデザインが適切で
- 適切に実施され
- その結果、得られたデータの品質保証が適正なレベルでなされ
- 目的が達成できたかどうかが明確で
- 以上のことが明確に論文化され
- 実地医療に（大きな）影響を与える試験

さて、ASCO2014

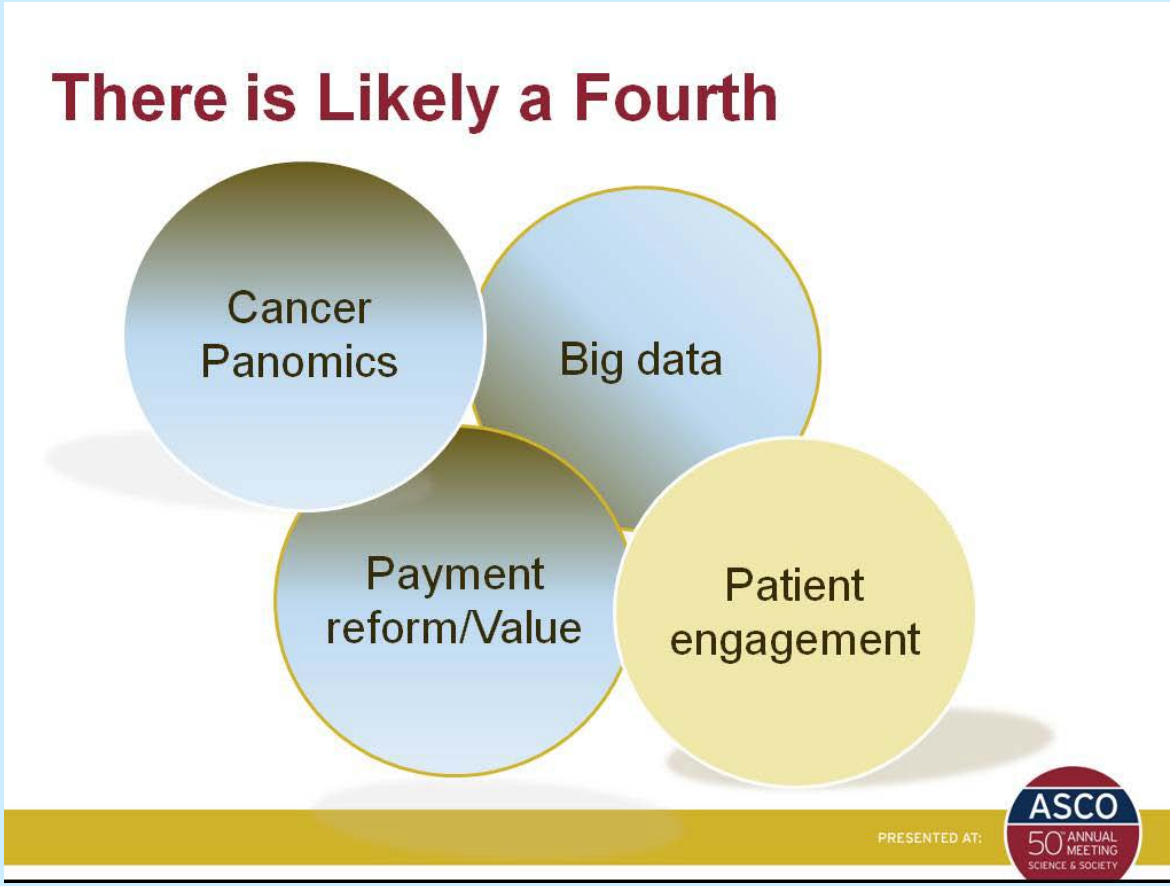
50周年

SELECT：進行再発乳癌1次治療
TS1 vs Taxane　非劣性試験

展示

ASCO2014 Keywords

VALUE
PRECISION MEDICINE

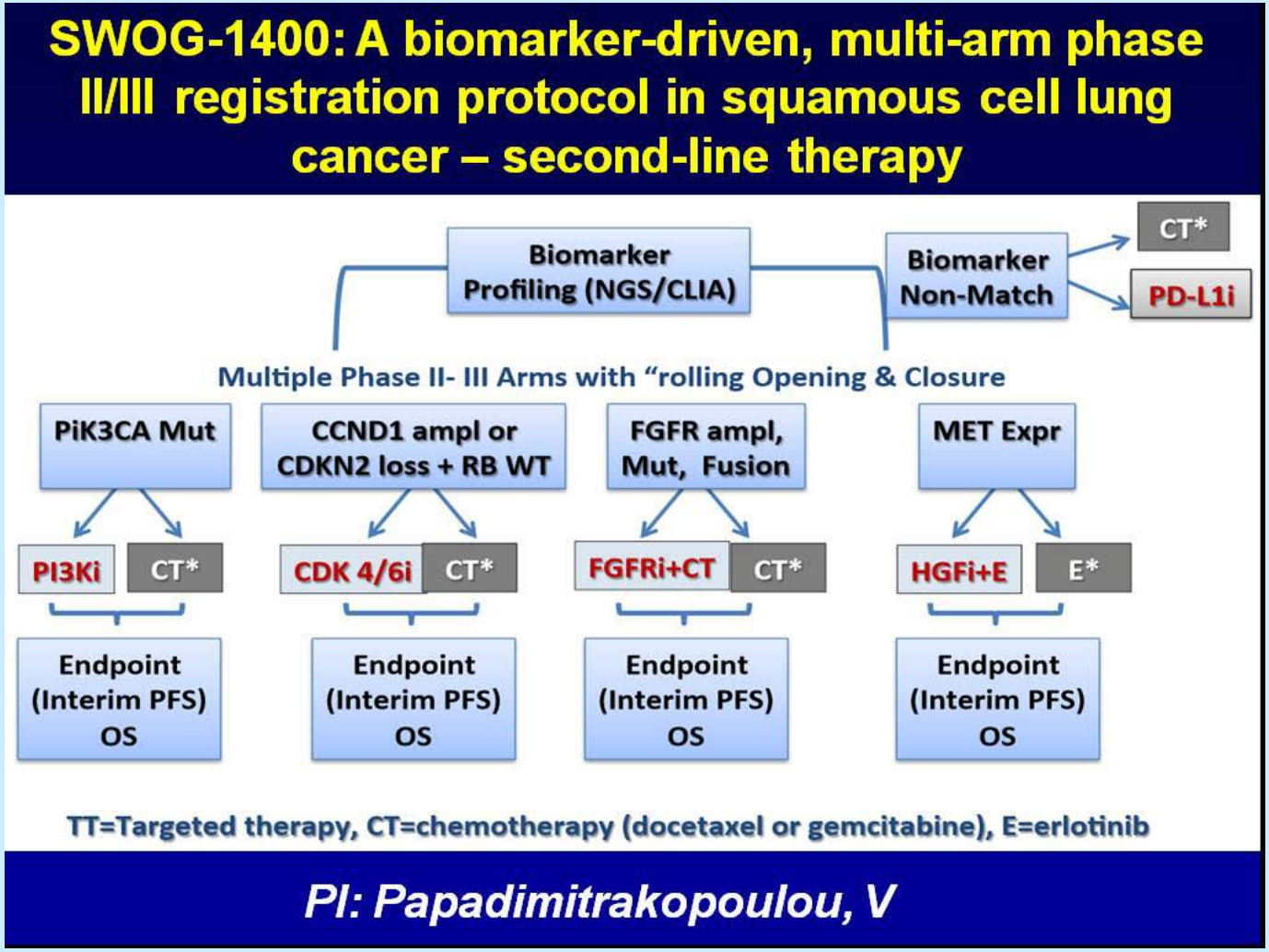


Presented By Richard Schilsky at 2014 ASCO Annual Meeting

Shilsky R: Precision cancer methods in the years ahead, ASCO2014 Educational Session: 50 years of precision medicine, advances in the field of prospective and innovative randomized clinical trials, 2014 June 01.

Preciseに！
紹介された試験

BATTLE-2
SWOG-1400
I-SPY2
SAFIR-02
WINTHER
SHIVA
M-PACT
IMPACT2



Presented By Apostolia Tsimberidou at 2014 ASCO Annual Meeting

Tsimberidou AM: Advances in precision of medicine: Prospective clinical trials, ASCO2014 Educational Session: 50 years of precion medicine, advances in the field of prospective and innovative randomized clinical trials, 2014 June 01.

仮説検証からより
良い医療の提供へ！

ベイズ統計学は有効か？

Future Major Shift in Clinical Trial Design

- From hypothesis testing, type I error and statistical power
- To explicit goal of delivering good medicine

39

Presented By Donald Berry at 2014 ASCO Annual Meeting

Berry D: Efficient clinical trials in the era of biomarkers, ASCO2014 Educational Session: Companion diagnostics and biomarker development in oncology drug development, 2014 May 31.

MSA Partners LLC

The Coming Bayesian Tsunami of Clinical Development

2006年6月30日発行 第138号

MSA Partners LLC
516 Fifth Avenue, 9th Fl.
10036
760
r.com
http://www.msapr.com

目次

ベイズ統計学パート2

臨床開発にベイズ統計学の「津波」が到来

テキサス大学アンダーソン癌センター生物統計学科主任、ドナルド・A・ベリー博士

30年以上もの間、ドナルド・ベリー博士は臨床試験のデザインと分析におけるベイズ統計学の利用を提唱してきた。この発想は当初、行政機関や製薬業界、および臨床研究における教育を受けておらず認識の乏しい統計学者から無視された。しかし、同氏が統計学の薬剤開発への適用に情熱を傾け続けた結果、数十年後、FDAと製薬会社が同氏に耳を傾け、ベイズ的アプローチの利点を理解するようになった。ベリー博士にこの変化について聞いた。

ドナルド・A・ベリー博士

――薬剤開発のデザインと分析において、統計は乱用されているとお考えですか。

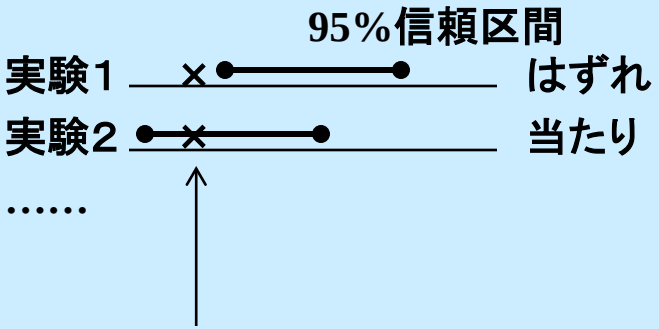
ベリー　統計学者を含め、臨床試験のデザ

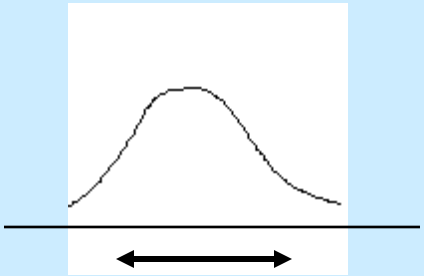
Frequentist 頻度論

ベイズ流（Bayesian)

未知の定数　　関心のあるパラメータ　　あいまいさ
（確率分布）

仮想的繰り返しに基づく相対頻度　　確率　　あいまいさ





パラメータが
存在する確率95%

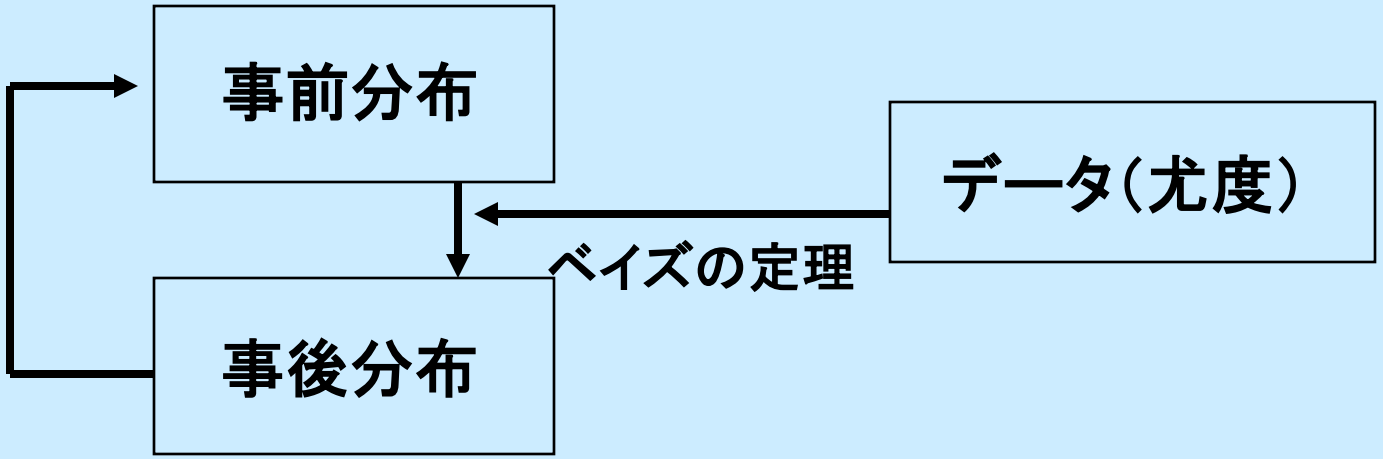
Sample spaceでの推測

Parameter spaceでの推測

ベイズ流の統計学（確率の考え方）

ベイズ流アプローチ～医師の診断過程

興味のあるパラメータの曖昧さを事前分布として定式化
観察データの出現確率（尤度Likelihood)を定式化



有効率（あるいは毒性）の逐次推定　事前分布は一様分布

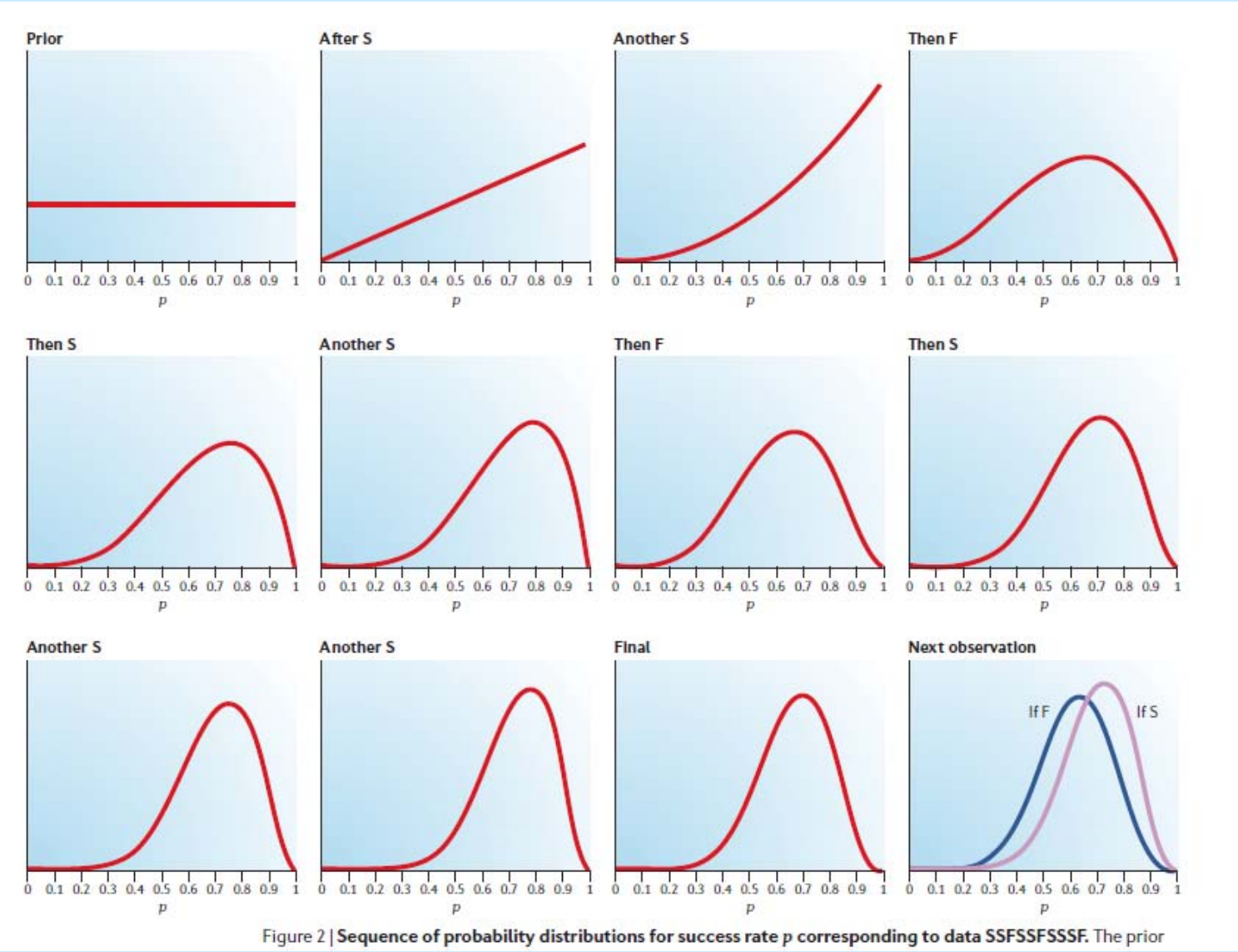


Figure 2 | **Sequence of probability distributions for success rate *p* corresponding to data SSFSSFSSSF.** The prior

生物統計を中心としたベイズ手法の発展

- **Bayes(1763)**
- **古典的教科書　Lindlay(1975)、Box and Chao(1973)**
- **J. Cornfieldの貢献(1960-)**
- **Weinstein and Fineberg(1980) *Clinical Decision Analysis***
- **Beal and Sheiner (1979-)NONMEM project**
- **Chiba group (Racine, Grieve, Smith) (1980's) の活躍**

  LD50, cross-0ver, bioequivalence, random effect
- **Gelfand and Smith(1990; JASA) Gibbs sampling**
- **BUGS 誕生(1992) MRC (Spiegelhalter and Thomas)**
- **Spiegelhalter and Freedman (1994; *JRSS A*) 臨床試験の解釈、中間解析**
- **Berry and Stangle (1996) *Bayesian Biostatistics***
- **Spiegelhalter et al. (2004) *Bayesian Approaches to Clinical Trials and Health-Care Evaluation***
- **FDA Medical Device Guidance(2006)**
- **Berry(2006; *Nature Review*)**
- **SAS: PROC MCMC発表(2008)**

A GUIDE TO DRUG DISCOVERY

Bayesian clinical trials

Donald A. Berry

Abstract | Bayesian statistical methods are being used increasingly in clinical research because the Bayesian approach is ideally suited to adapting to information that accrues during a trial, potentially allowing for smaller more informative trials and for patients to receive better treatment. Accumulating results can be assessed at any time, including continually, with the possibility of modifying the design of the trial, for example, by slowing (or stopping) or expanding accrual, imbalancing randomization to favour better-performing therapies, dropping or adding treatment arms, and changing the trial population to focus on patient subsets that are responding better to the experimental therapies. Bayesian analyses use available patient-outcome information, including biomarkers that accumulating data indicate might be related to clinical outcome. They also allow for the use of historical information and for synthesizing results of relevant trials. Here, I explain the rationale underlying Bayesian clinical trials, and discuss the potential of such trials to improve the effectiveness of drug development.

Frequentist
An approach to statistical inference that is an inverse of the Bayesian approach. The focus is on the probability of results of a trial — usually including the observed data — assuming that a particular hypothesis is true. For example, a frequentist P-value is the probability of observing results as extreme as or more extreme than the observed results assuming that the null hypothesis is true.

Department of Biostatistics and Applied Mathematics, The University of Texas M. D. Anderson Cancer Center, 1515 Holcombe Boulevard, BOX 447, Houston, Texas 77030-4009, USA.
e-mail: dberry@mdanderson.org
doi:10.1038/nrd1927

Statistical thinking has had a central role in raising the scientific standards of clinical research over the last two centuries, especially during the past 50 years. A major reason has been the appreciation of statistical inference by drug- and medical-device-regulatory agencies. Traditional frequentist statistics has had the dominant, and often exclusive, role in this scientific renaissance. The greatest virtue of the traditional approach may be its extreme rigour and narrowness of focus to the experiment at hand, but a side effect of this virtue is inflexibility, which in turn limits innovation in the design and analysis of clinical trials. Because of this, clinical trials tend to be overly large, which increases the cost of developing new therapeutic approaches, and some patients are unnecessarily exposed to inferior experimental therapies.

Owing to such issues, there is increasing interest in Bayesian methods in medical research. Advances in computational techniques and power are also facilitating the application of these methods (BOX 1). More than 100 ongoing clinical trials at the University of Texas M. D. Anderson alone have been designed or are being monitored from the Bayesian perspective. And of recent medical device approvals by the Center for Devices and Radiological Health of the US FDA, ~10% are based on Bayesian designs and analyses, as compared with none 10 years ago. Furthermore, at least one drug (Pravigard Pac; Bristol-Myers Squibb) was approved by the FDA on the basis of Bayesian analyses of efficacy (BOX 2). And in May 2004, the FDA co-sponsored a workshop to address the role of Bayesian approaches in drug and medical device development, 'Can Bayesian Approaches to Studying New Treatments Improve Regulatory Decision-Making?' (The video/audio presentations are available as webcasts; see Further information).

After setting the context of the Bayesian approach by describing the frequentist perspective and relating the two approaches, this article discusses the Bayesian approach to the design and analysis of clinical trials, and to drug and medical device development more generally. The goal is to improve drug and medical device development, in terms of costs and the effective treatment of patients, both those in and those outside of clinical trials, and the Bayesian approach provides a better perspective, and a more efficient methodology, for accomplishing this goal. It should be emphasized though that I want to preserve the high scientific standards wrought by the hard and effective work of statisticians and other scientifically oriented clinical researchers during the past 50 years (indeed, the Bayesian approach is more closely in line with the scientific method[1]).

Statistical inference

Statistical inferences are based on mathematical models of experiments, including clinical trials. Each model corresponds to a 'state of nature', the underlying process that produces the experimental results. Candidate

Draft Guidance for Industry and FDA Staff

Guidance for the Use of Bayesian Statistics in Medical Device Clinical Trials

DRAFT GUIDANCE
This guidance document is being distributed for comment purposes only.
Draft released for comment on May 23, 2006

Comments and suggestions regarding this draft document should be submitted within 90 days of publication in the *Federal Register* of the notice announcing the availability of the draft guidance. Submit written comments to the Division of Dockets Management (HFA-305), Food and Drug Administration, 5630 Fishers Lane, rm. 1061, Rockville, MD 20852. Alternatively, electronic comments may be submitted to http://www.fda.gov/dockets/ecomments. All comments should be identified with the docket number listed in the notice of availability that publishes in the *Federal Register*.

For questions regarding this document, contact Dr. Greg Campbell at 240-276-3133 or greg.campbell@fda.hhs.gov.

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Devices and Radiological Health

Division of Biostatistics
Office of Surveillance and Biometrics

FDA: Medical Device Guidance (2006)

Guidance for the Use of Bayesian Statistics in Medical Device Clinical Trials

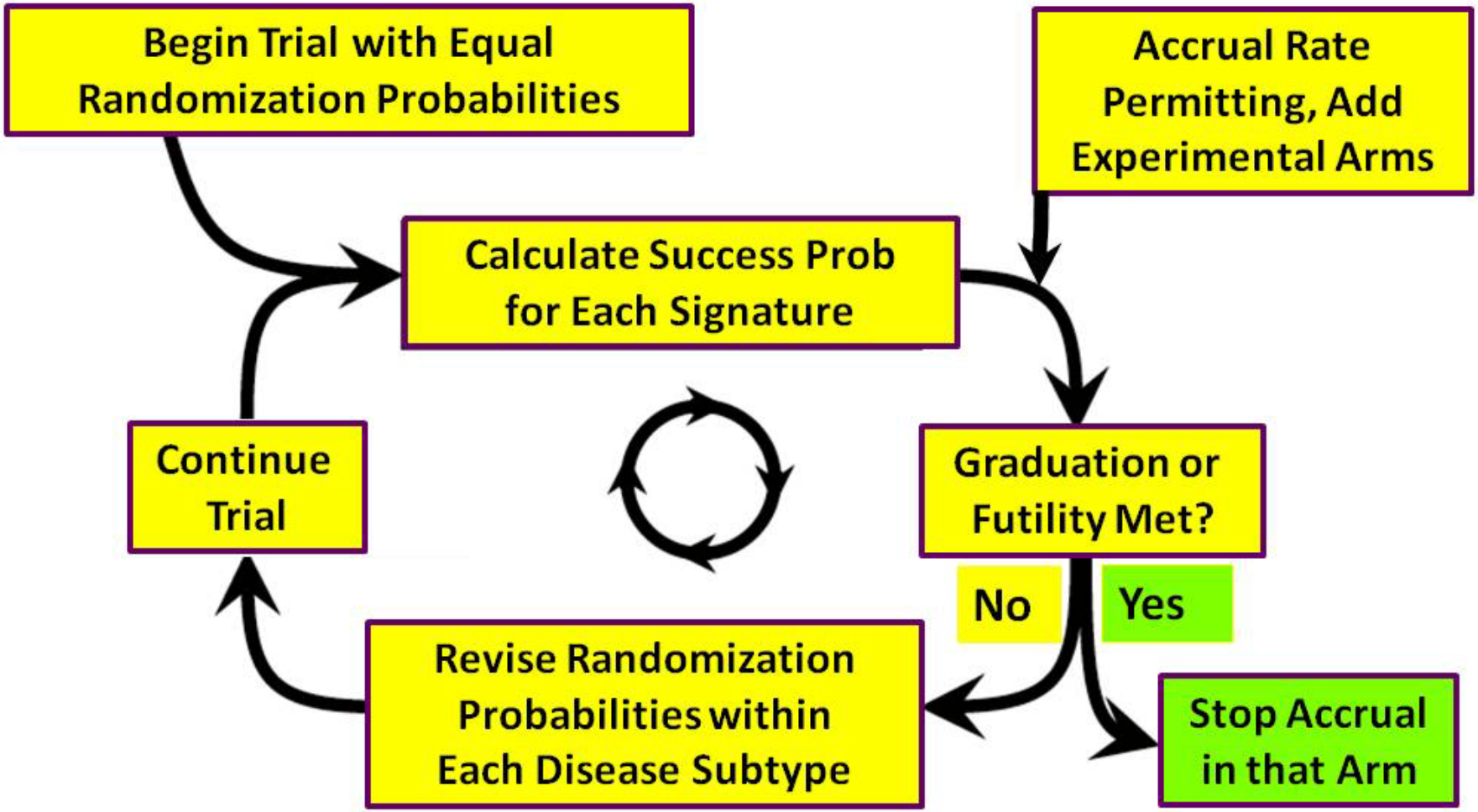
- なぜdevice？
  - 事前情報活用により小規模・短期試験で同様の結論に到達可能
  - 事前情報を活用しない場合でも柔軟に対処可能（中間解析、試験計画の変更）
  - 事前情報：前世代機器の情報、海外データ
  - メカニズムが物理的で局所的
- FDAとの事前協議・合意の必要性
  - 事前分布とモデルに対する合意
  - 計算結果の確認のため、データとプログラムをFDAは要求

I-SPY2デザイン

- ホルモン感受性（HR）とHER2の陽性・陰性、そして70遺伝子を組み合わせたMammPrint再発リスク得点（MM）の高・低によって層別された8つの乳癌患者サブ集団（ただしHR+ HER2- MM低は低リスク群として除外）
- ClinicalTrials.govに登録されている当初のプロトコルでは3つの現在未承認薬を含む7試験治療、SABCS（2013年12月）で発表された資料ではHER2+、HER2-それぞれ3つの試験治療が評価
- 術前のネオアジュバント投与であり、プライマリエンドポイントは6ヶ月時点でのpathological CR(pCR)
- 複数の試験治療は、逐次評価される成績により、期待できるものは最速60例で「卒業」して第III相比較試験に進む。一方期待ができないものは最速20例で試験治療から脱落する。試験治療毎の登録継続は最大120例か最長登録期間18ヶ月
- HER2+患者にはパクリタキセル12週＋トラスツズマブ、HER2-患者はパクリタキセル単剤12週投与後それぞれAC4サイクルが標準治療に設定されており、標準治療か試験治療にランダム割付け
- ランダム割付けにはadaptive randomization。患者のサブタイプ毎にpCRが得られる「確率」を逐次的に推定し、その「確率」が大きな試験治療に割り付けられる確率が高くなるように実施、標準治療には常に20%
- 「卒業」判定は、pCRをプライマリエンドポイントとして標準治療を対照とした第III試験を合計300例で新たに実施した場合、有意差が得られる「確率」が85%を超えた場合。見込みの低い試験治療は脱落

- 本スライドは非公開です。

I-SPY 2 Adaptive Process

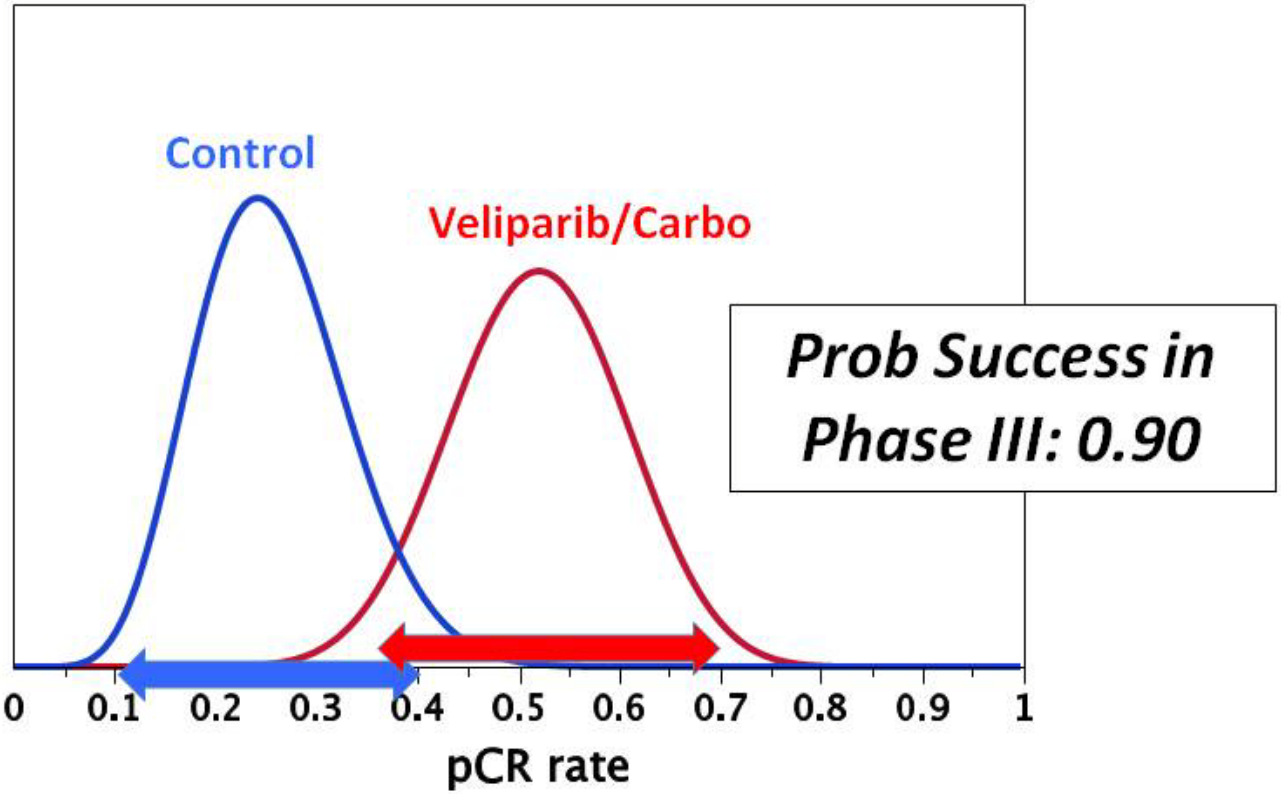


I-SPY2中間結果：卒業生誕生

- SABCS 2013 Decで第一号発表
- HR-HER2-（トリプルネガティブ）乳癌に対するヴァリパリブ（variparib）+カルボプラチン
- 卒業時点での試験治療（この2剤＋パクリタキセル）、標準治療（パクリタキセル）へのHER2-登録例数はそれぞれ72例と44例
- HR-HER2-に対する試験治療が標準治療を上回るベイズ流「確率」は92%、新たに第III相試験を実施した際の成功予想「確率」は90%であり「卒業」資格を満たしていた
- HR+HER2-例での試験治療が標準治療を上回る「確率」は28%でだめ

- AACR2014Aprに第二号発表
- HR-HER2+に対する2番目の「卒業」生としてネラチニブ（neratinib）

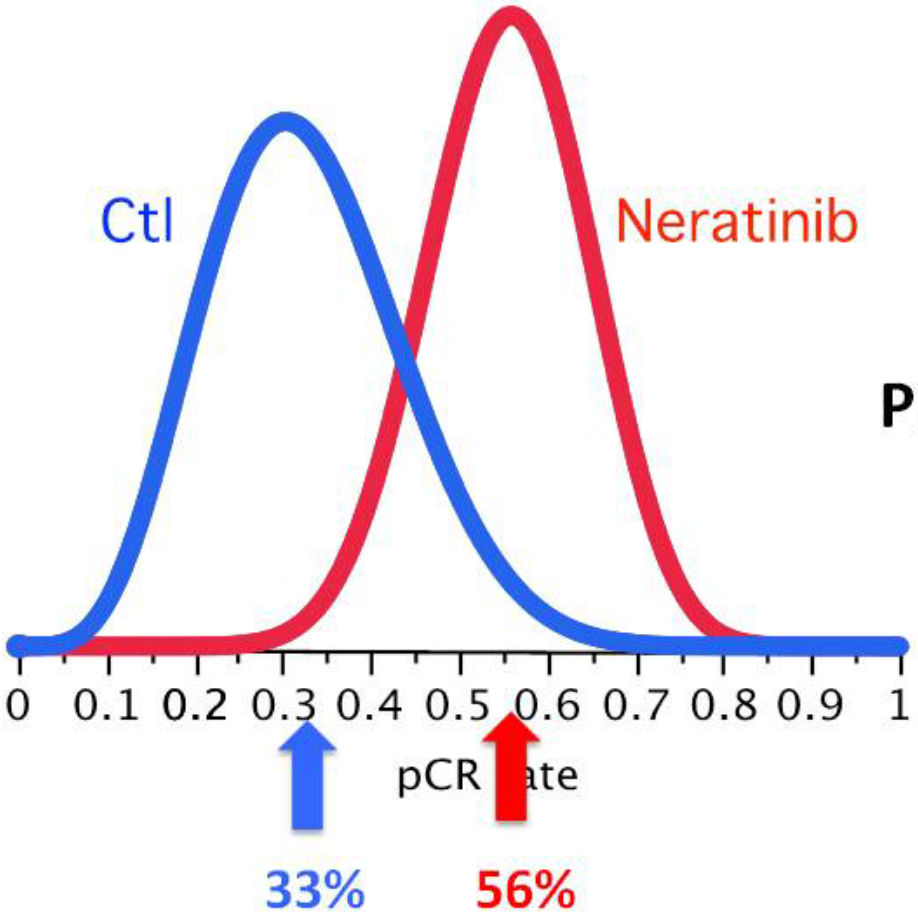
Veliparib/carboplatin graduated with Triple-negative Signature



Estimated pCR rate: 26%
95% interval: 11% to 40%

Estimated pCR rate: 52%
95% interval: 35% to 69%

Neratinib graduated in HR-/HER2+



Pred. prob.
79%

顧客あっての質

- 現在の（臨床試験に参加しようとする）患者と将来の患者
- 試験が進まなければ将来の患者への恩恵も遅れる
- 仮説検定によるアプローチ（生産者危険と消費者危険とのバランスに基づく）は、将来患者数＞＞現在患者数の場合には適切な意思決定規則
- 細分化された患者サブ集団、しかも治療とサブ集団最適組み合わせ（presicion medicine)には不確実さ
- （分子生物的プロフィルに基づく試験（basket trial）の概念）
- 第I相の拡大コホート、第I/II相での加速承認もありうる
- デバイスの事例（ヒストリカルコントロールの積極活用）

- ベイズ統計手法の応用は強力なオプション