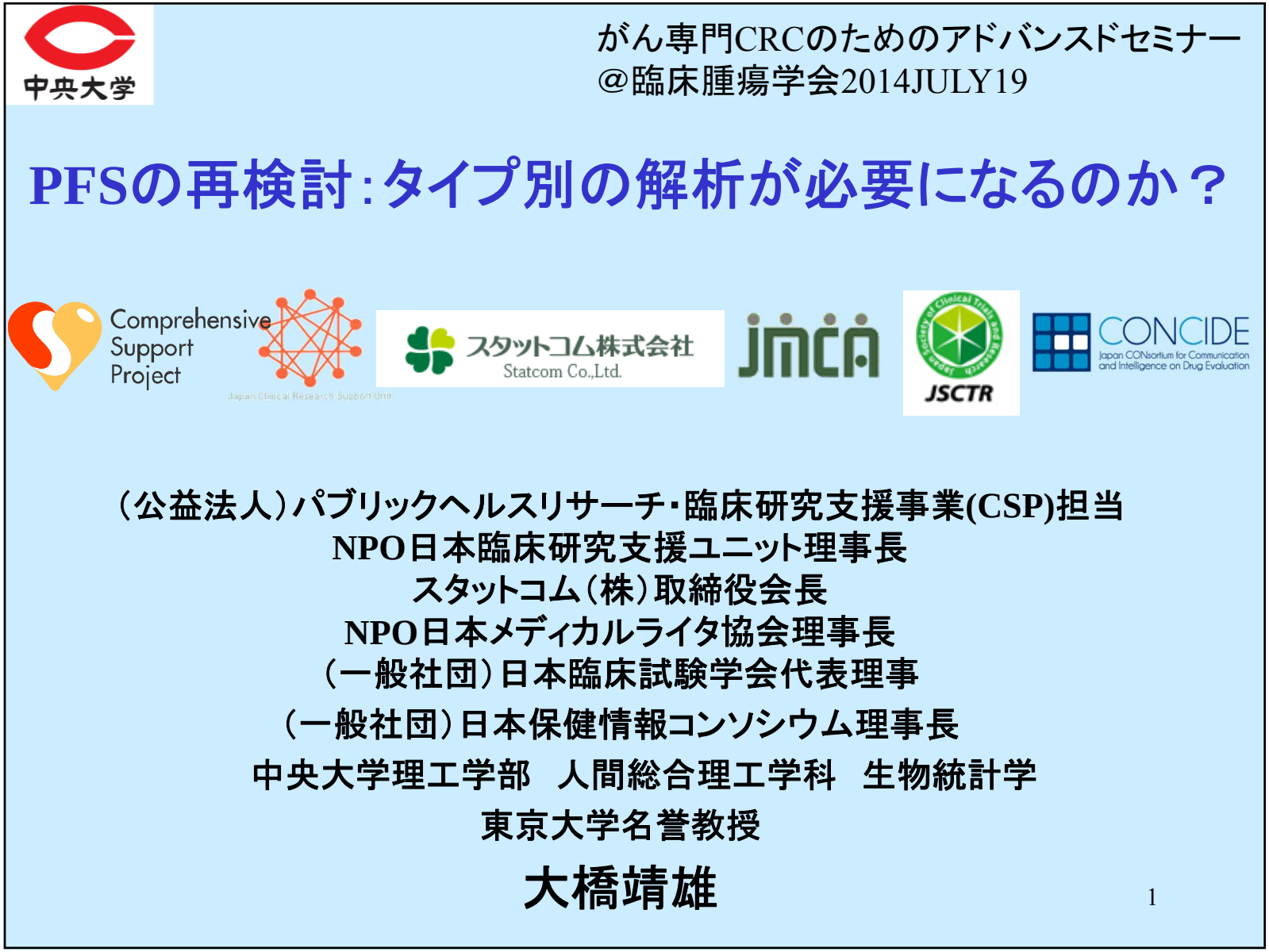


発表者の利益相反開示事項

研究費の財源	□科学研究費　□受託　□寄付 □その他（　　）☑該当なし	財源のスポンサー	
発表者氏名	大橋靖雄	所属／身分	中央大学教授

	該当なし	該当有りの場合：企業名等
企業の職員・法人の代表	□	スタットコム（株）会長、NPO　J-CRSU理事長、社団　CONCIDE理事長
企業等の顧問職	☑	
株式など	□	スタットコム（株）
講演料など	□	中外製薬
原稿料など	☑	
研究費（治験等）	☑	
寄付金	□	武田製薬、興和創薬、アステラス、協和発酵キリン
専門的助言・証言	☑	
臨床試験実施法人の代表	□	NPO J-CRSU
その他（贈答品等）	☑	

2

がん臨床試験のエンドポイントと奏効判定

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

- **エンドポイント**

  試験の目的にそって測定される評価項目だが、しばしば研究者から試験の目的そのものと混同され、またPMDAによって不適切なものに差し替えられる

- **奏効率**

  昔の二方向性の計測で45%とか、今のRECISTで25%とかいう微妙な縮小効果の時に、「もう一度きちんと測り直してみい」と上司に関西弁で言われて測定し直した結果、出る数値

3

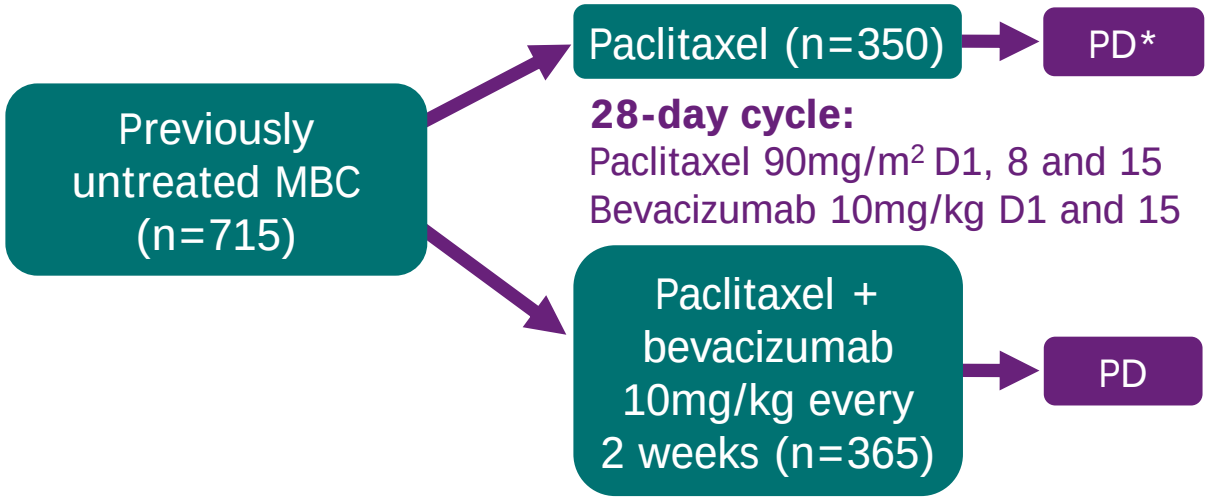
PFS とOS

- **Bevacizumab　は特殊な例？　ますます混迷？**
  MBCでも卵巣がんでも
- **治験でも研究者主導研究でもPFSが採用される例が増加**
- **FDAがPFSに基づいて近年承認した事例**
  gemcitabine in ovarian cancer
  sorafenib in advanced renal cancer
  bevacizumab in metastatic breast cancer　　？
  rituximab in Non-Hodgkin's lymphoma
  ・・・・

Guidance for Industry: Clinical Trial Endpoints for the Approval of Cancer Drugs and Biologics, May 2007

4

Phase III trial of 1st-line bevacizumab in MBC (E2100)

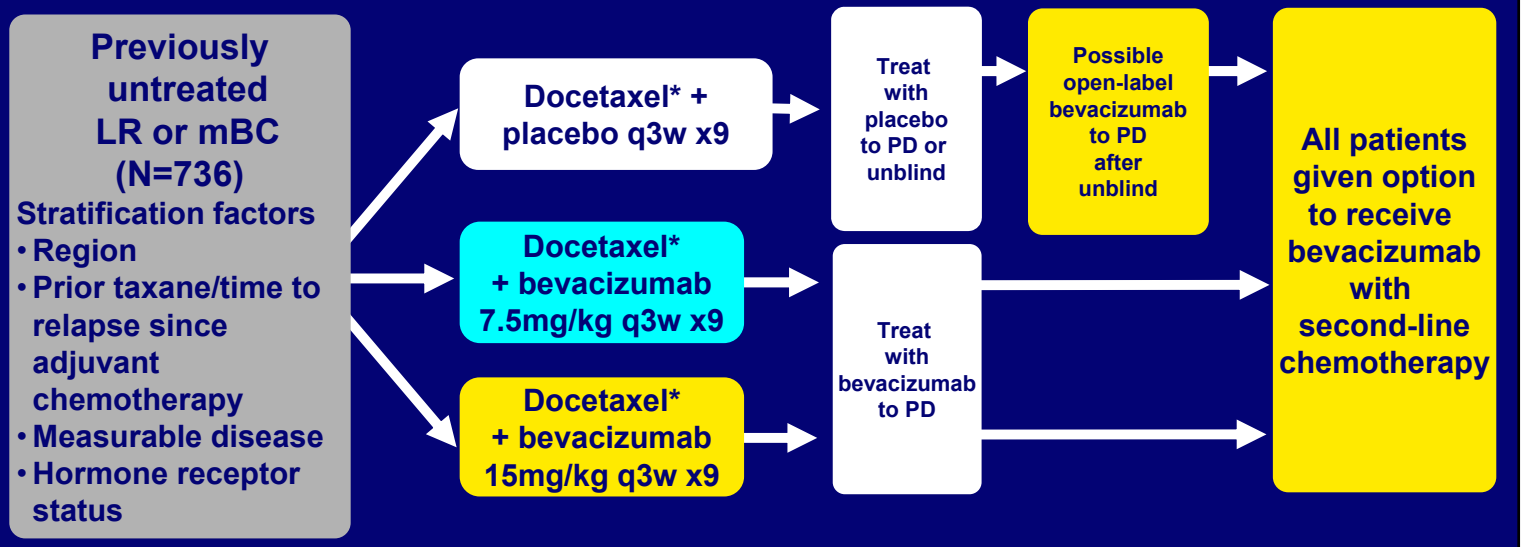


Stratify:

- Disease-free interval ≤24 vs >24 months
- <3 vs ≥3 metastatic sites
- Adjuvant chemotherapy yes vs no
- ER+ vs ER- vs ER unknown

Miller et al ASCO 2005

AVADO: double-blind, placebo-controlled trial

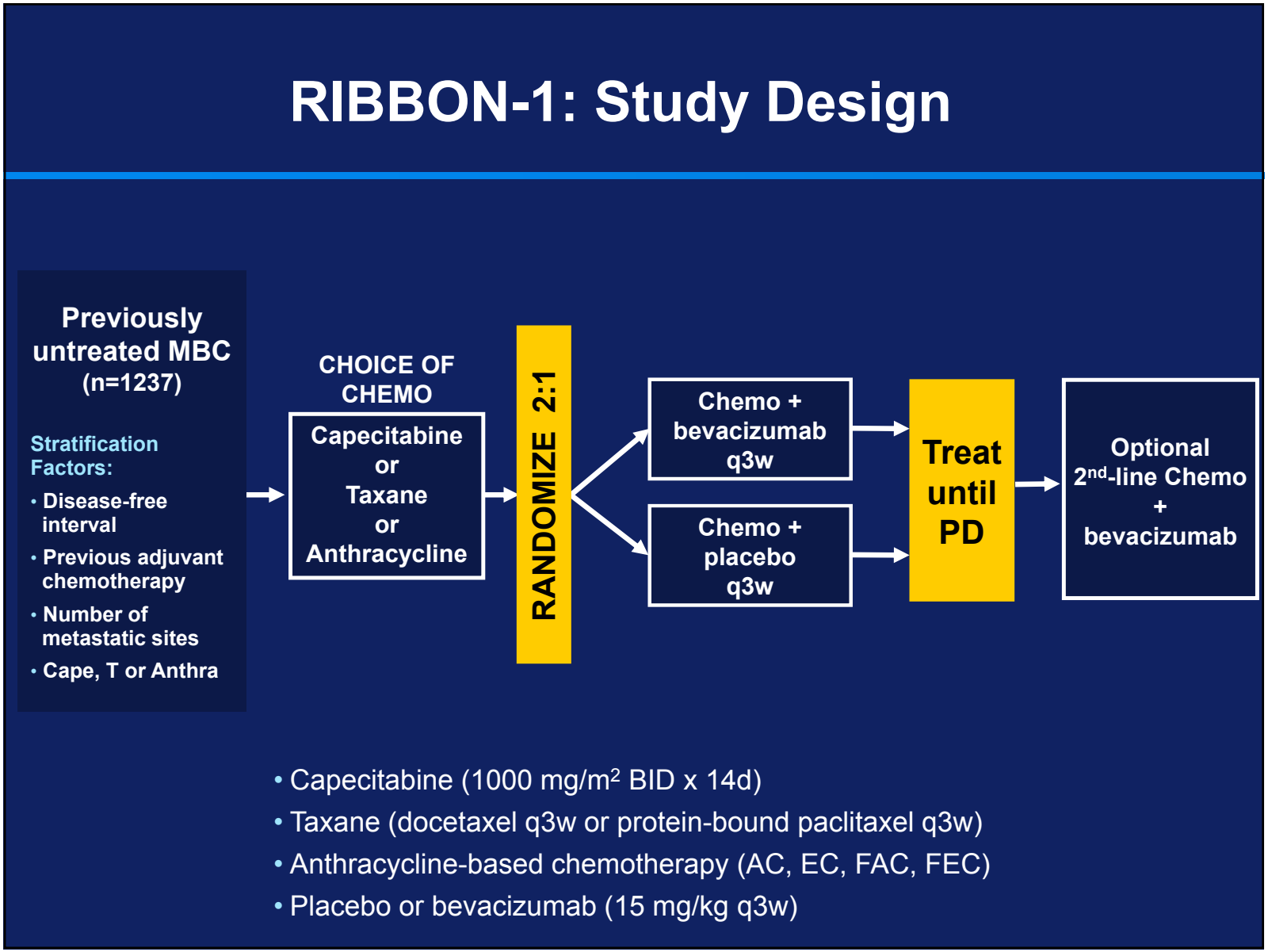


- Primary endpoint: PFS
- Secondary endpoints: ORR, 1-year survival, OS, TTF, duration of response, quality of life, safety

*Docetaxel: 100mg/m^2 q3w

Miles DW, et al. ASCO 2008

RIBBON-1: Study Design



Avastin in 1st line setting for MBC

Trial	E2100[1,2] (n=722)	AVADO[3] (n=736)	RIBBON-1[4] Cape cohort (n=615)	RIBBON-1[4] A/T cohort (n=622)
Placebo controlled	No	Yes	Yes	Yes
Chemo	weekly Paclitaxel	Q3w Docetaxel	Capecitabine	Anthracycline or Taxane
Dose of Avastin	10mg/kg q2w	7.5 or 15mg/kg q3w	15mg/kg q3w	15mg/kg q3w
Primary Endpoint	PFS	PFS	PFS	PFS
IRF* review	Retrospective	Yes	Yes	Yes

*IRF:Independent Review Facility

[1] Robert Gray et al. J Clin Oncol 2009; 27:4966-4972. [2]Kathy Miller et al. N Engl J Med 2007; 357:2666-76
[3] Miles D et al. SABCS2009 abstr#41.
[4]N. J. Robert et al. ASCO2009 abstr#1005.

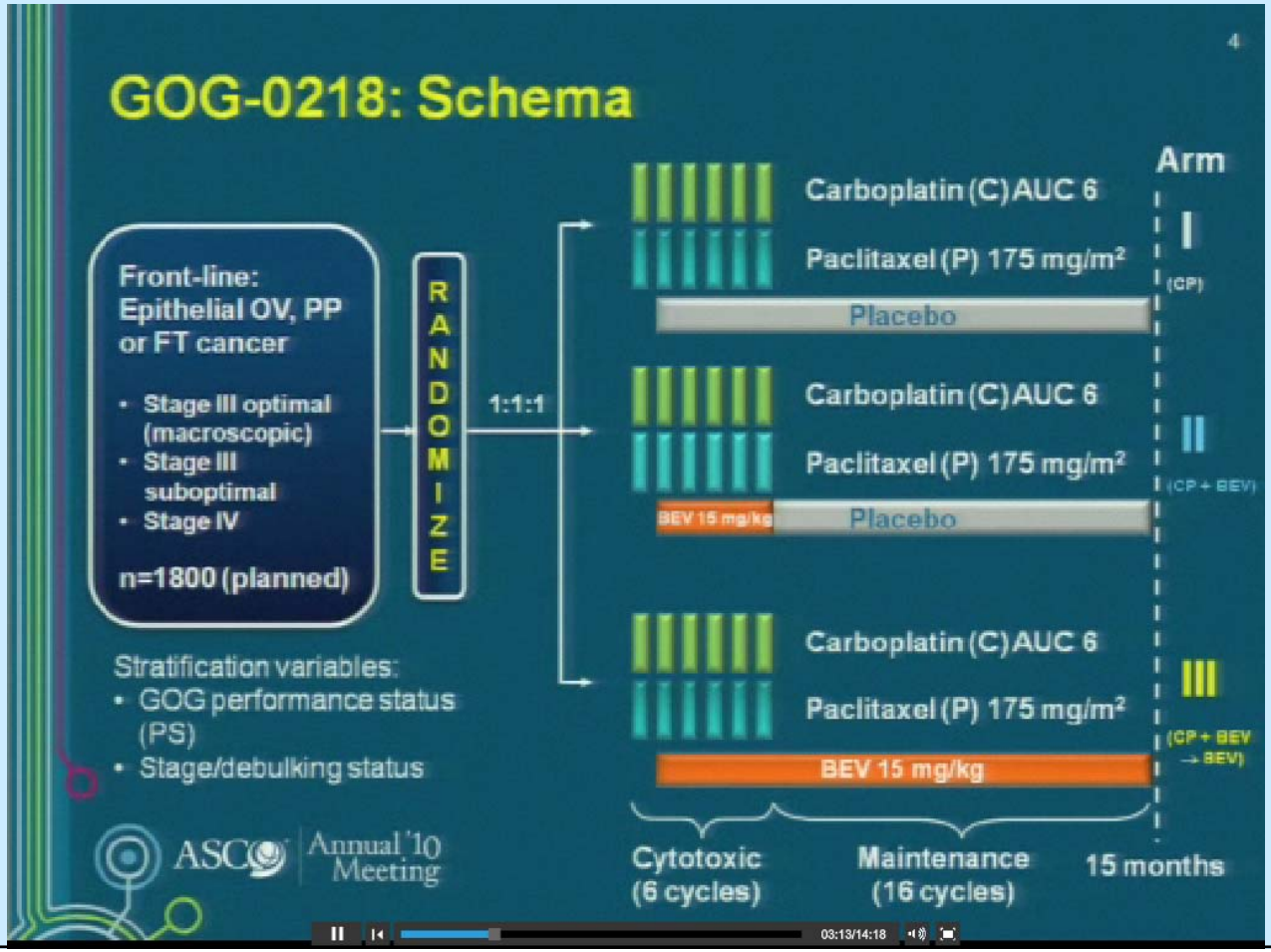
Avastin in 1st line setting for MBC <Efficacy>

試験	E2100[1,2] (n=722)		AVADO[3] (n=736)		RIBBON-1[4] Cape cohort (n=615)		RIBBON-1[4] A/Tcohort (n=622)	
arm	Paclitaxel	Paclitaxel + BV	Docetaxel + PL	Docetaxel + BV*	Cape + PL	Cape + BV	A/T + PL	A/T + BV
PFS (m)	5.8	11.3	8.2	10.1	5.7	8.6	8.0	9.2
HR	0.48 P<0.0001		0.77 P=0.0061		0.69 P=0.0002		0.64 P<0.0001	
RR	22%	49%	46%	64%	24%	35%	38%	51%
	P<0.0001		P=0.0003		P=0.0097		P=0.0054	
OS (m)	25.2	26.7	31.9	30.2	21.2	29.0	23.8	25.2
HR	0.88		1.03		0.85		1.03	

*Bevacizumab15mg/kg q3w

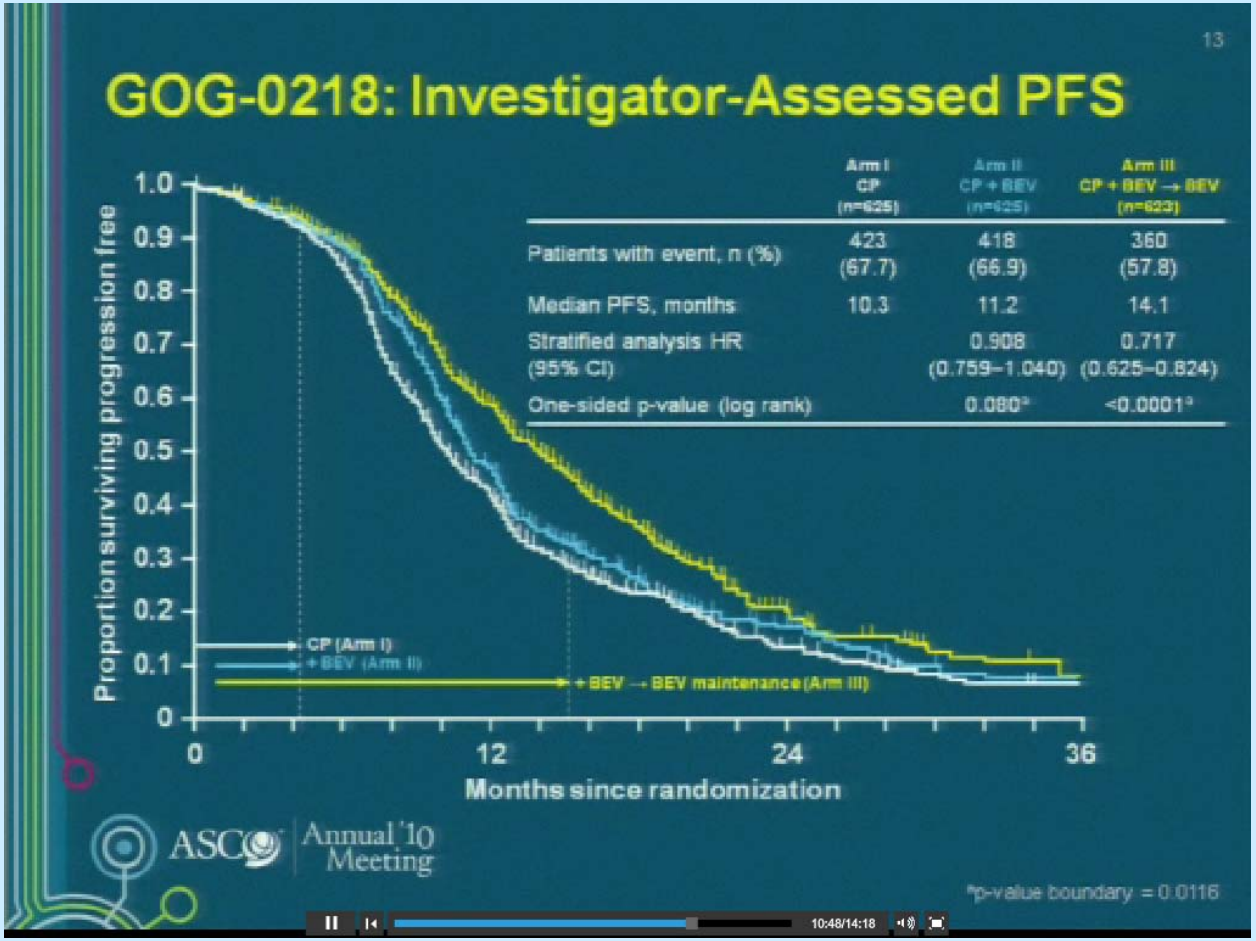
[1] Robert Gray et al. J Clin Oncol 2009; 27:4966-4972. [2]Kathy Miller et al. N Engl J Med 2007; 357:2666-76
[3] Miles D et al. SABCS2009 abstr#41.
[4]N. J. Robert et al. ASCO2009 abstr#1005.

ASCO2010　LBA1　Burger et al.
Phase III trial of bevacizumab (BEV) in the primary treatment of advanced epithelial ovarian cancer (EOC), primary peritoneal cancer (PPC), or fallopian tube cancer (FTC): A Gynecologic Oncology Group study.



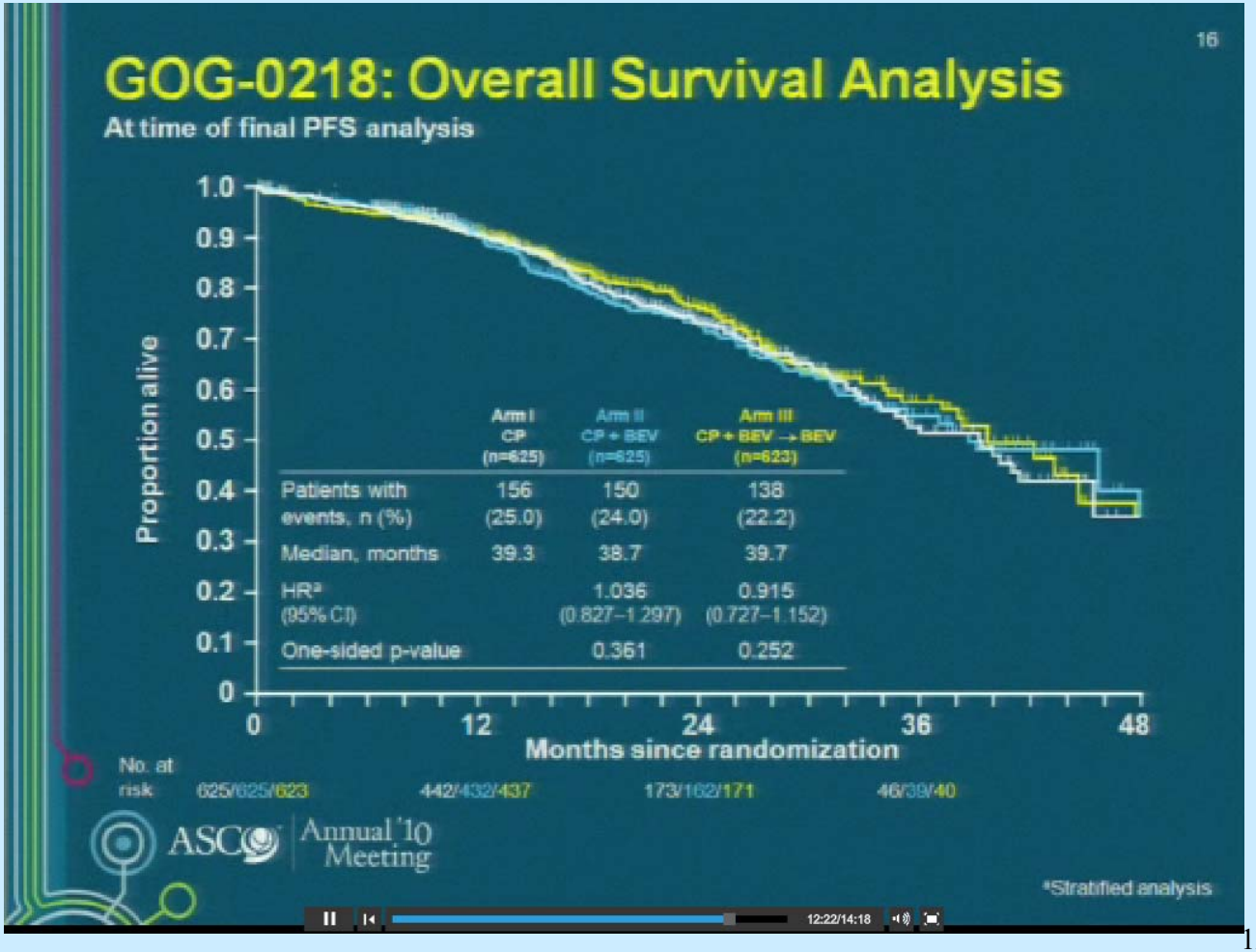
10

ASCO2010　LBA1　Burger et al.



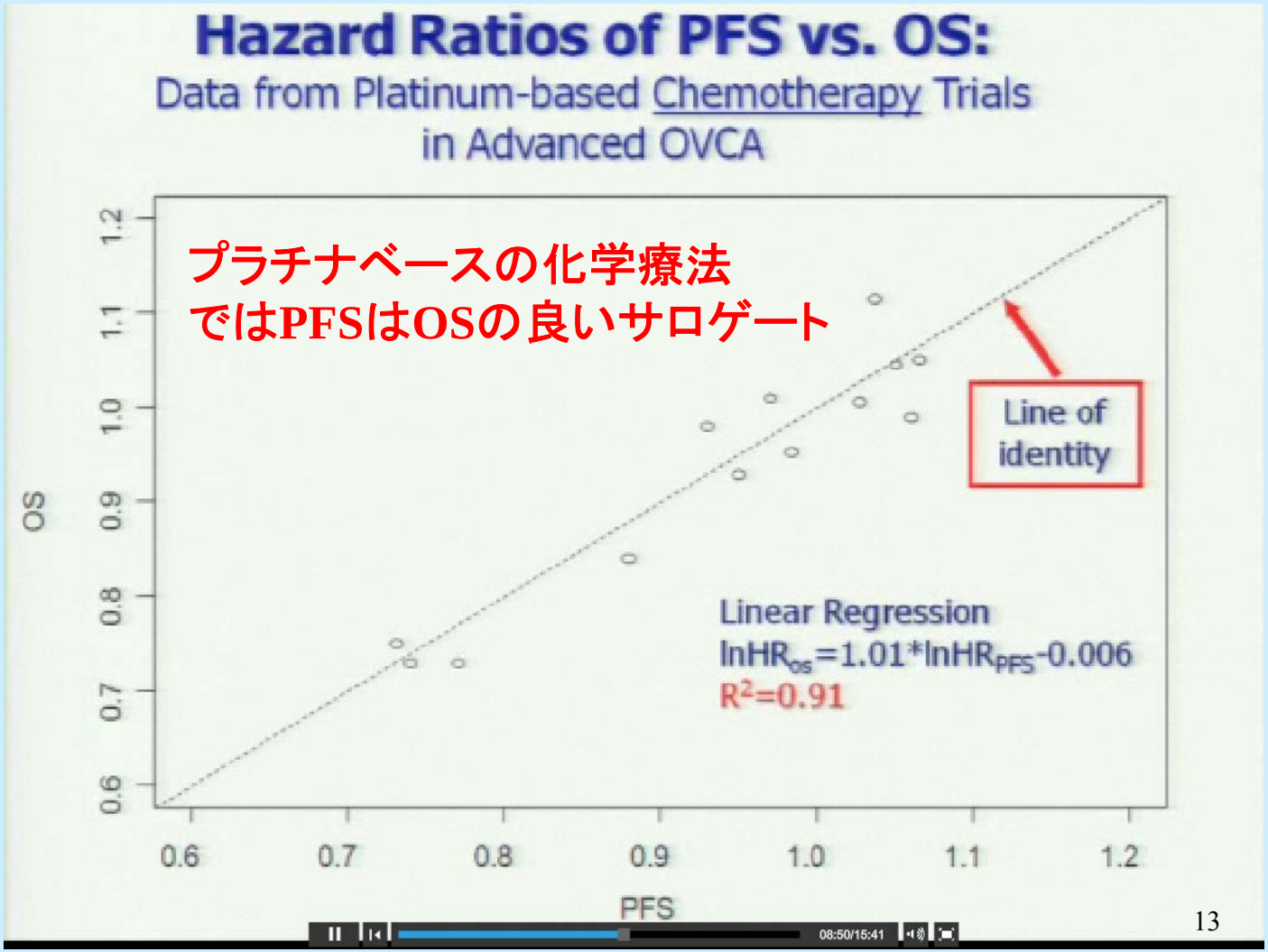
11

ASCO2010　LBA1　Burger et al.

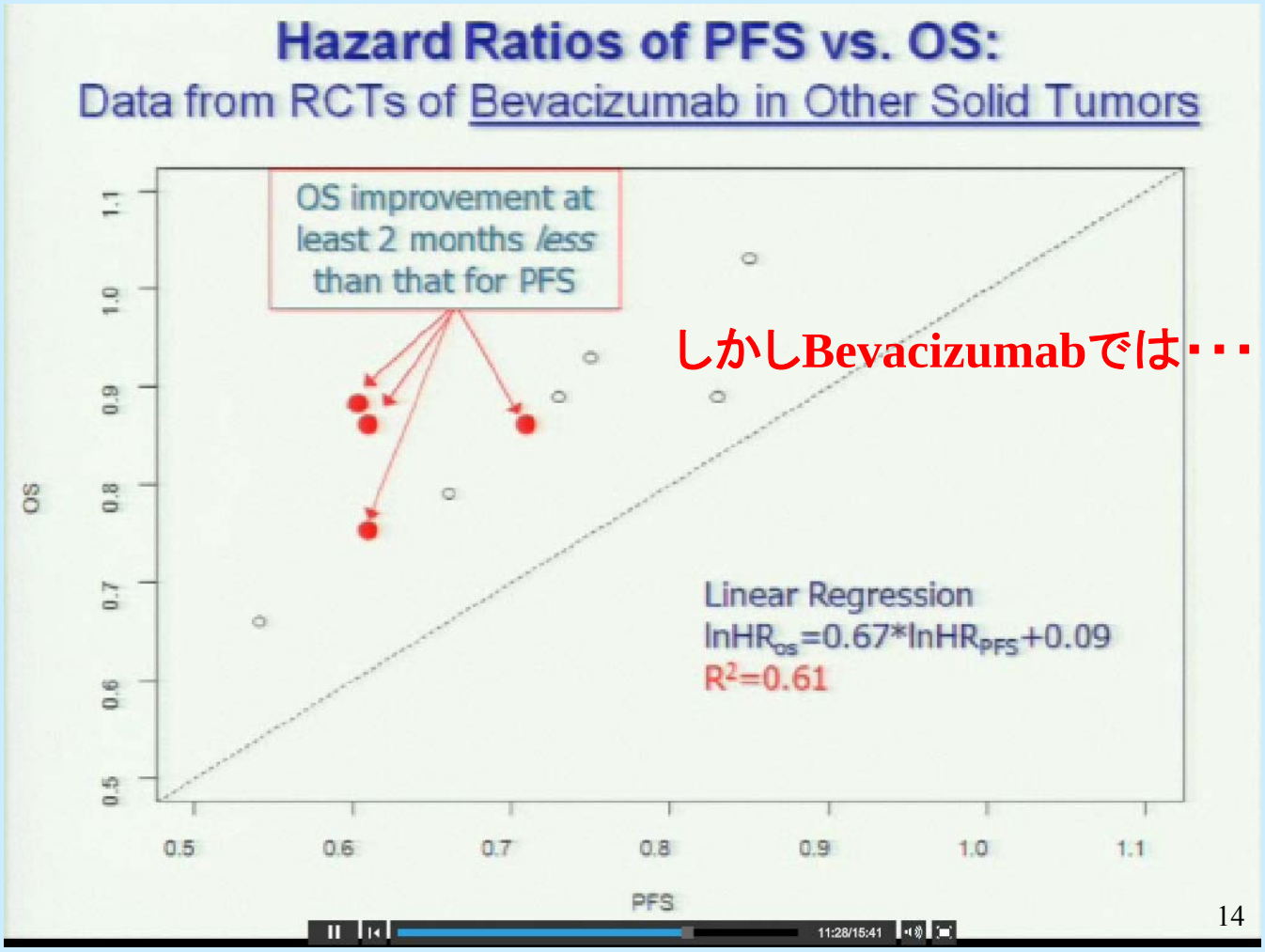


12

ASCO2010　LBA1　Discussion to Burger et al.



ASCO2010　LBA1　Discussion to Burger et al.



Burger et al.
NEJM 2011; 365: 2473-83.

ORIGINAL ARTICLE

Incorporation of Bevacizumab in the Primary Treatment of Ovarian Cancer

Robert A. Burger, M.D., Mark F. Brady, Ph.D., Michael A. Bookman, M.D., Gini F. Fleming, M.D., Bradley J. Monk, M.D., Helen Huang, M.S., Robert S. Mannel, M.D., Howard D. Homesley, M.D., Jeffrey Fowler, M.D., Benjamin E. Greer, M.D., Matthew Boente, M.D., Michael J. Birrer, M.D., Ph.D., and Sharon X. Liang, M.D., for the Gynecologic Oncology Group*

Perren et al.
NEJM 2011; 365: 2484-96.

ORIGINAL ARTICLE

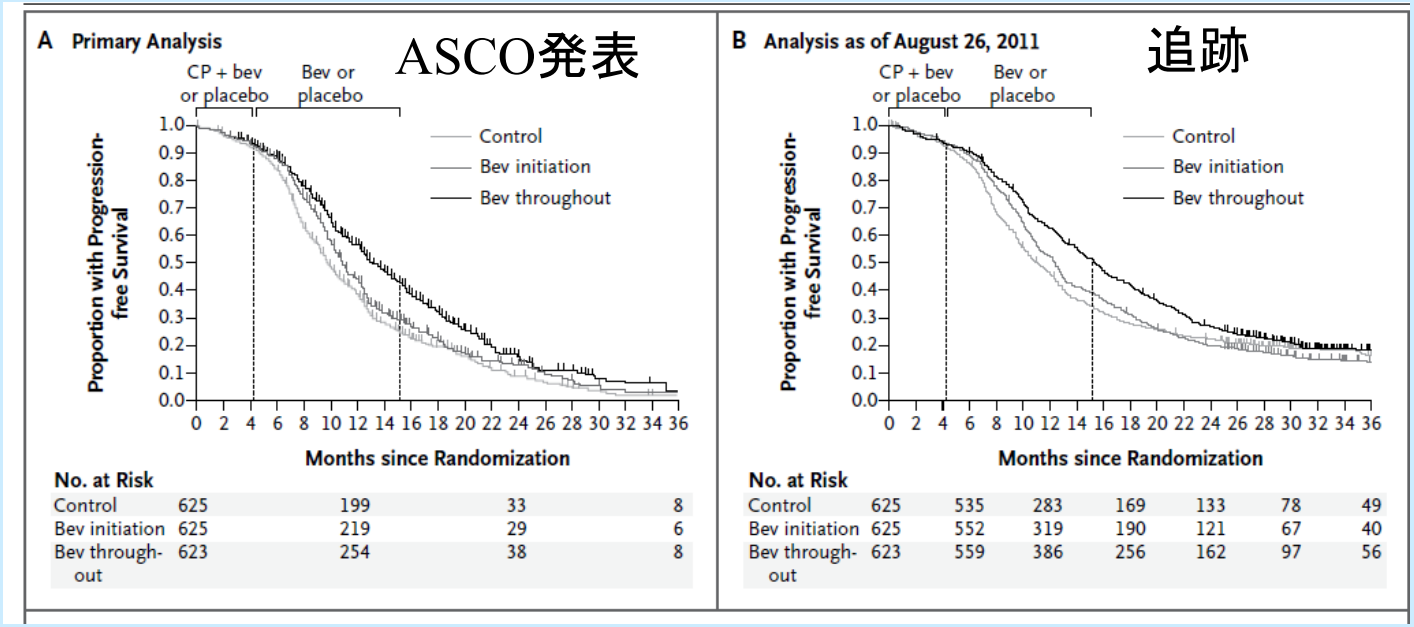
A Phase 3 Trial of Bevacizumab in Ovarian Cancer

Timothy J. Perren, M.D., Ann Marie Swart, M.D., Jacobus Pfisterer, M.D., Jonathan A. Ledermann, M.D., Eric Pujade-Lauraine, M.D., Gunnar Kristensen, M.D., Mark S. Carey, M.D., Philip Beale, M.D., Andrés Cervantes, M.D., Christian Kurzeder, M.D., Andreas du Bois, M.D., Jalid Sehouli, M.D., Rainer Kimmig, M.D., Anne Stähle, M.D., Fiona Collinson, M.D., Sharadah Essapen, M.D., Charlie Gourley, M.D., Alain Lortholary, M.D., Frédéric Selle, M.D., Mansoor R. Mirza, M.D., Arto Leminen, M.D., Marie Plante, M.D., Dan Stark, M.D., Wendi Qian, Ph.D., Mahesh K.B. Parmar, Ph.D., and Amit M. Oza, M.D., for the ICON7 Investigators*

15

Burger et al.
NEJM 2011; 365: 2473-83.

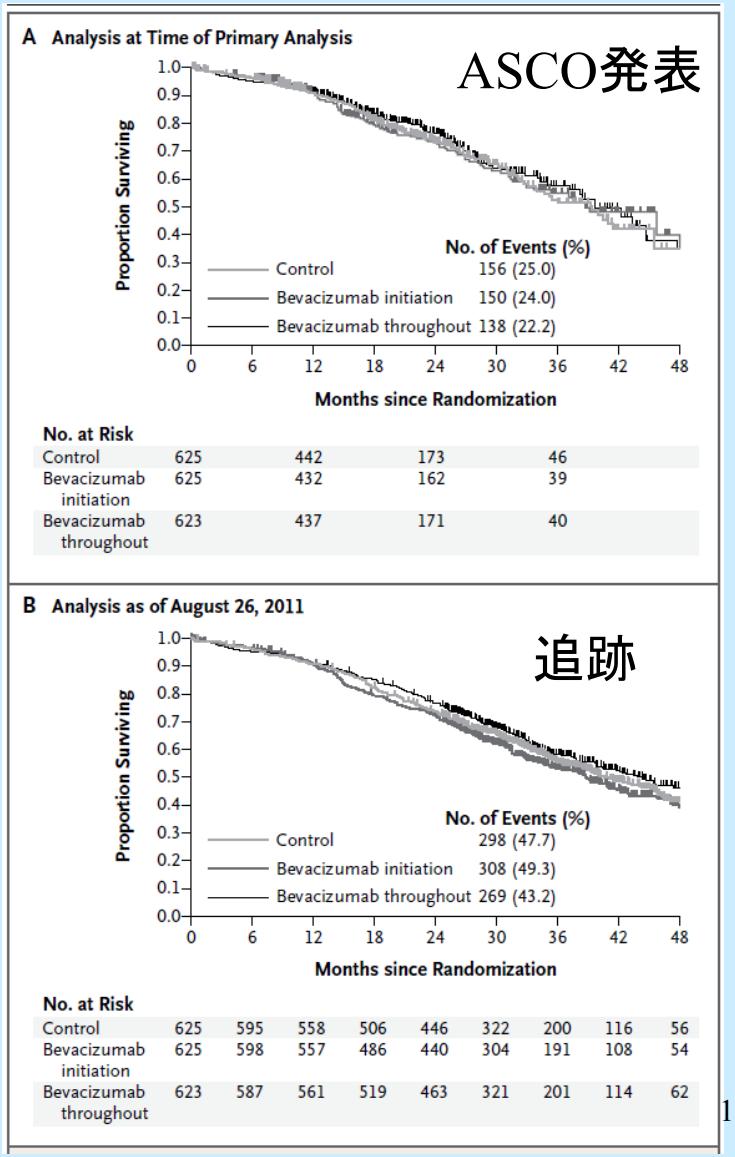
PFSの有意性は持続



16

Burger et al.
NEJM 2011; 365: 2473-83.

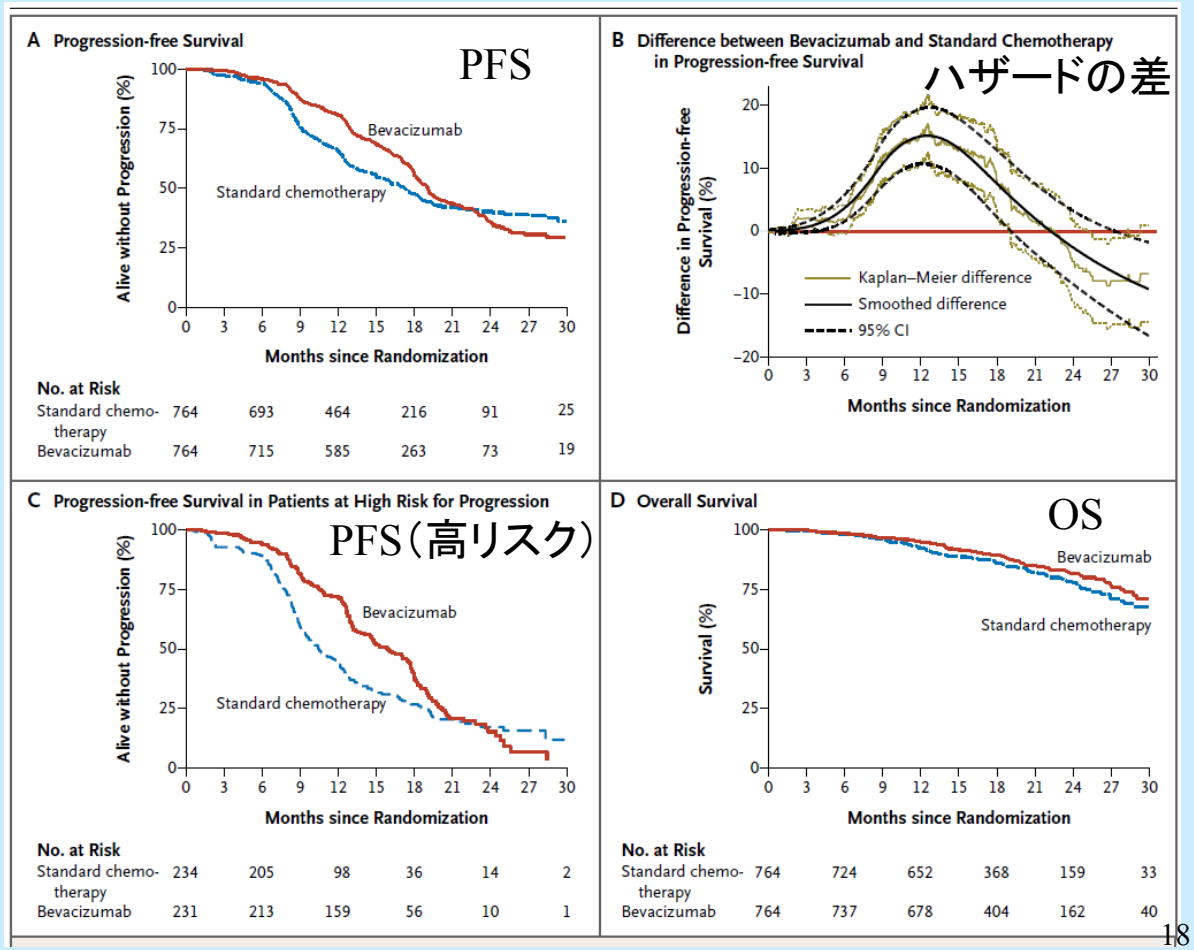
OSの「差なし」も持続



17

Perren et al.
NEJM 2011; 365: 2484-96.

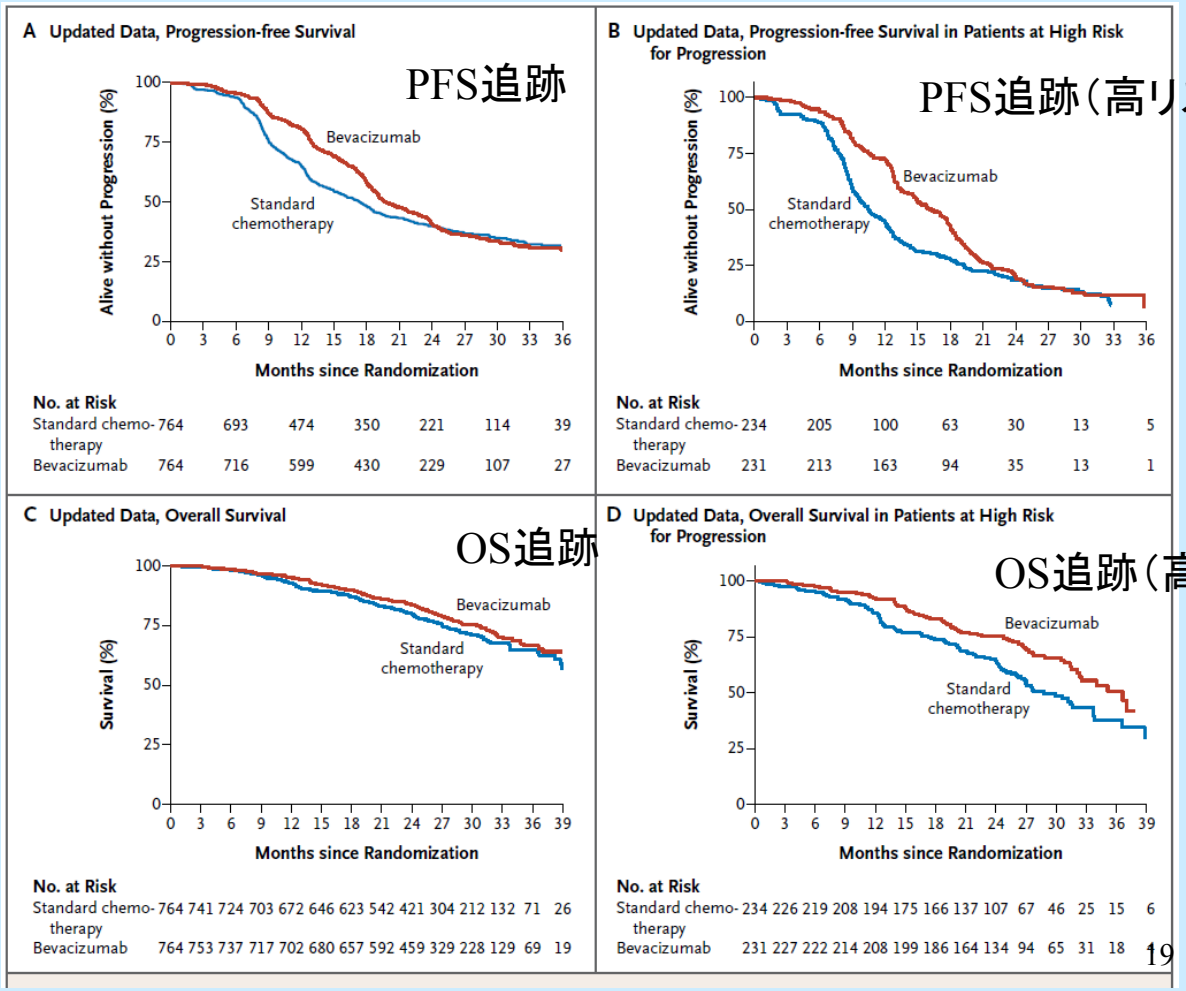
比例ハザード性がPFSでも成立しない



18

Perren et al.
NEJM 2011; 365: 2484-96.

ハイリスクではOSにも差



19

PFS とOS

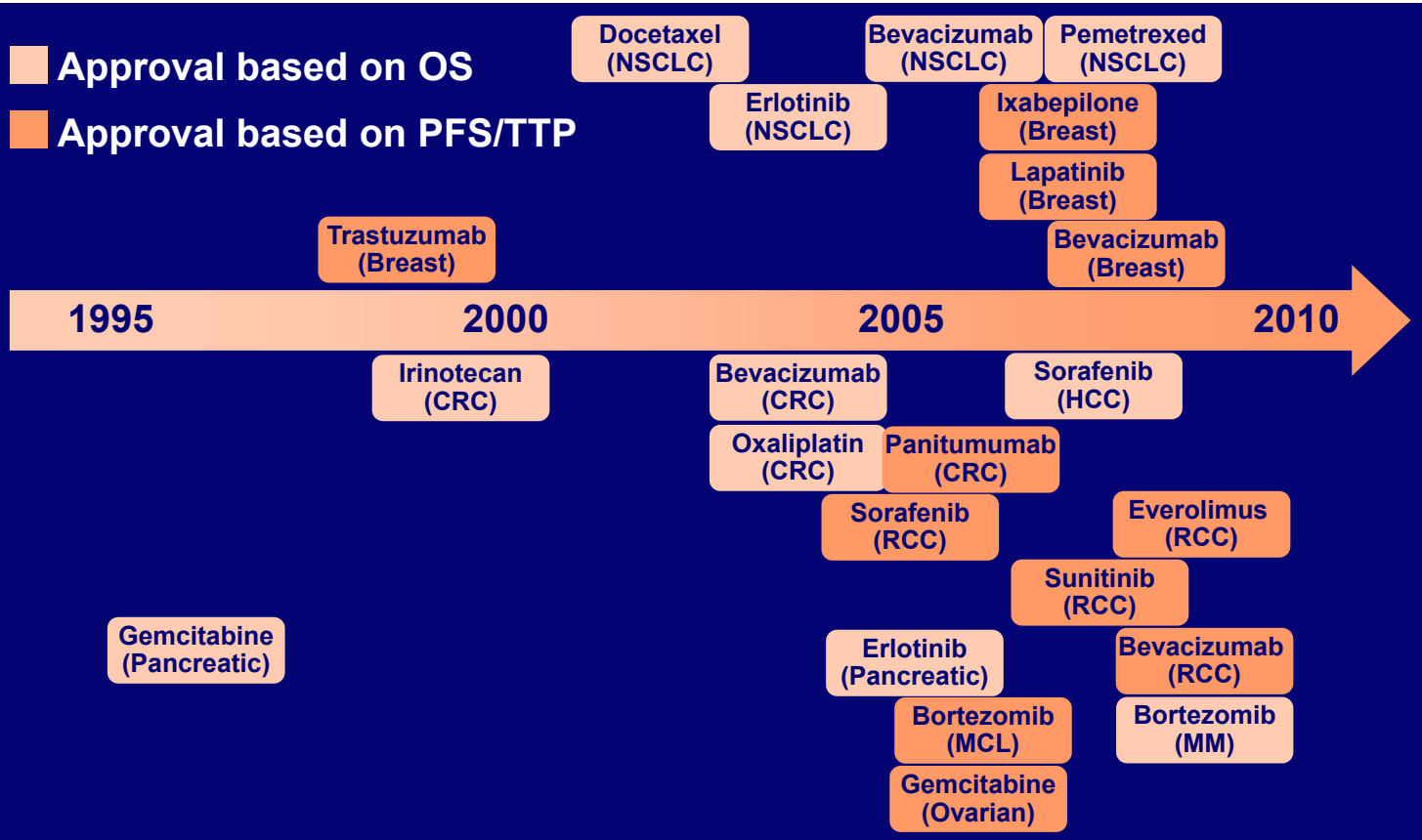
- Bevacizumab　は特殊な例？　ますます混迷？
- 治験でも研究者主導研究でもPFSが採用される例が増加
- FDAがPFSに基づいて近年承認した事例
  - gemcitabine in ovarian cancer
  - sorafenib in advanced renal cancer
  - bevacizumab in metastatic breast cancer
  - rituximab in Non-Hodgkin's lymphoma
  - ....

Guidance for Industry: Clinical Trial Endpoints for the Approval of Cancer Drugs and Biologics, May 2007

20

OS or PFS

Increasing Acceptance of PFS as a Basis for FDA Approval



www.cancer.gov/cancertopics/druginfo

21

最近までの、とりあえずの結論は

- 2次治療以降に有効な治療が登場すれば
- SPP(survival post progression)が延長すれば
  PFSのOSに対する代替性は薄まる（ハザードは希釈される）

- PFSの曖昧さとそれに対する対処の必要性
- 大きなPFSの改善と優れたrisk/benefit profileなら（さらに
  経済的に許容できるなら）PFSによる承認は当然ありうる

FDA: Guidance for Industry Clinical Trial Endpoints for the Approval of Non-Small Cell Lung Cancer Drugs and Biologics, 2011 June

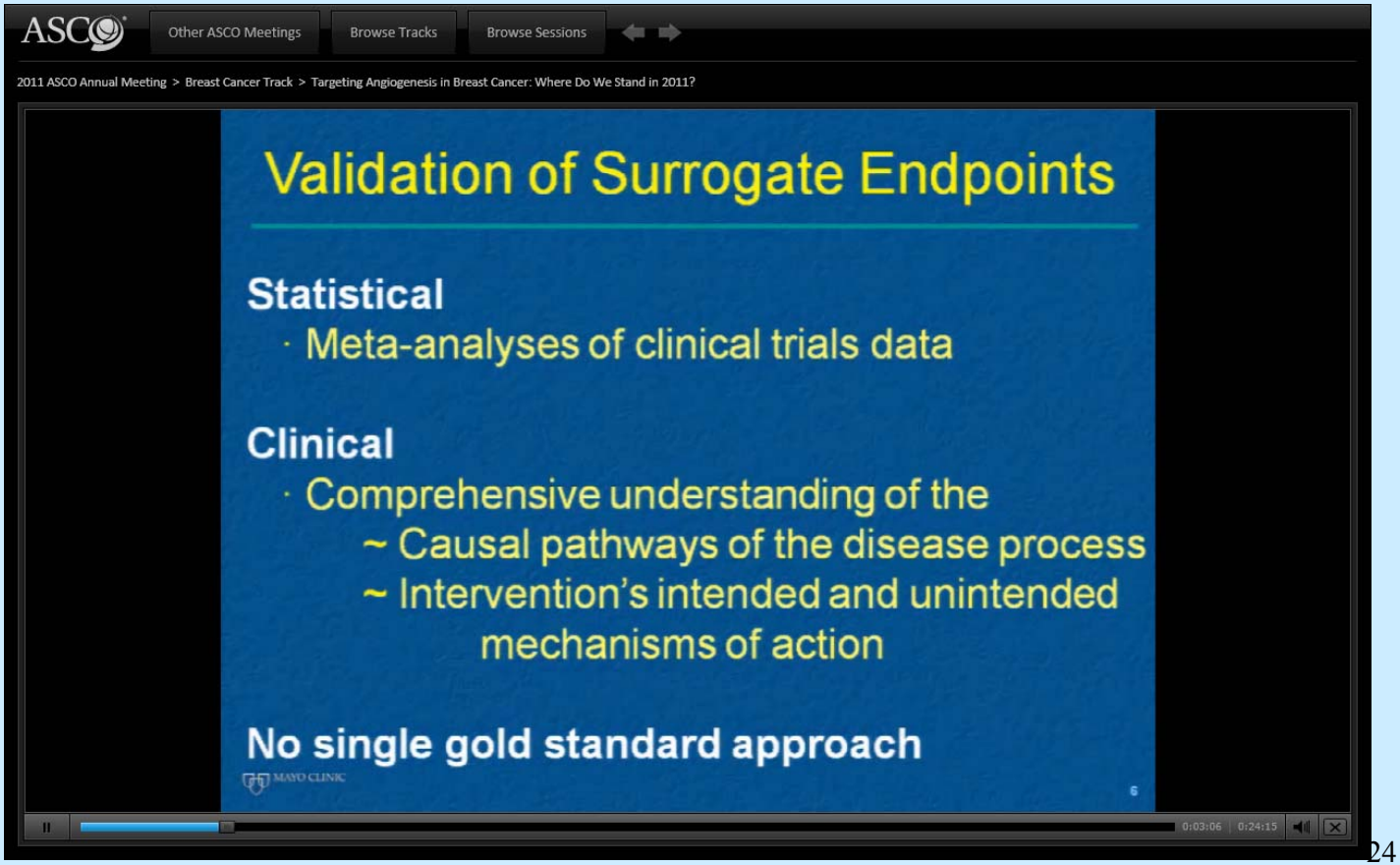
EMA: Appendix 1 to the guideline on the evaluation of anticancer medicinal products in man. EMA/CHMP/27994/2008/Rev.1, 2012, December

22

ASCO2011　Targeting Angiogenesis in Breast Cancer: Where Do We Stand in 2011? Appropriate Endpoints for Clinical Trials and Source of U.S. Food and Drug Administration/Oncologic Drugs Advisory Committee Controversy

Daniel J. Sargent, P　　**もっともスマートな癌の生物統計家**

PFSはOSのsurrogate
代替性の評価：唯一の標準はない

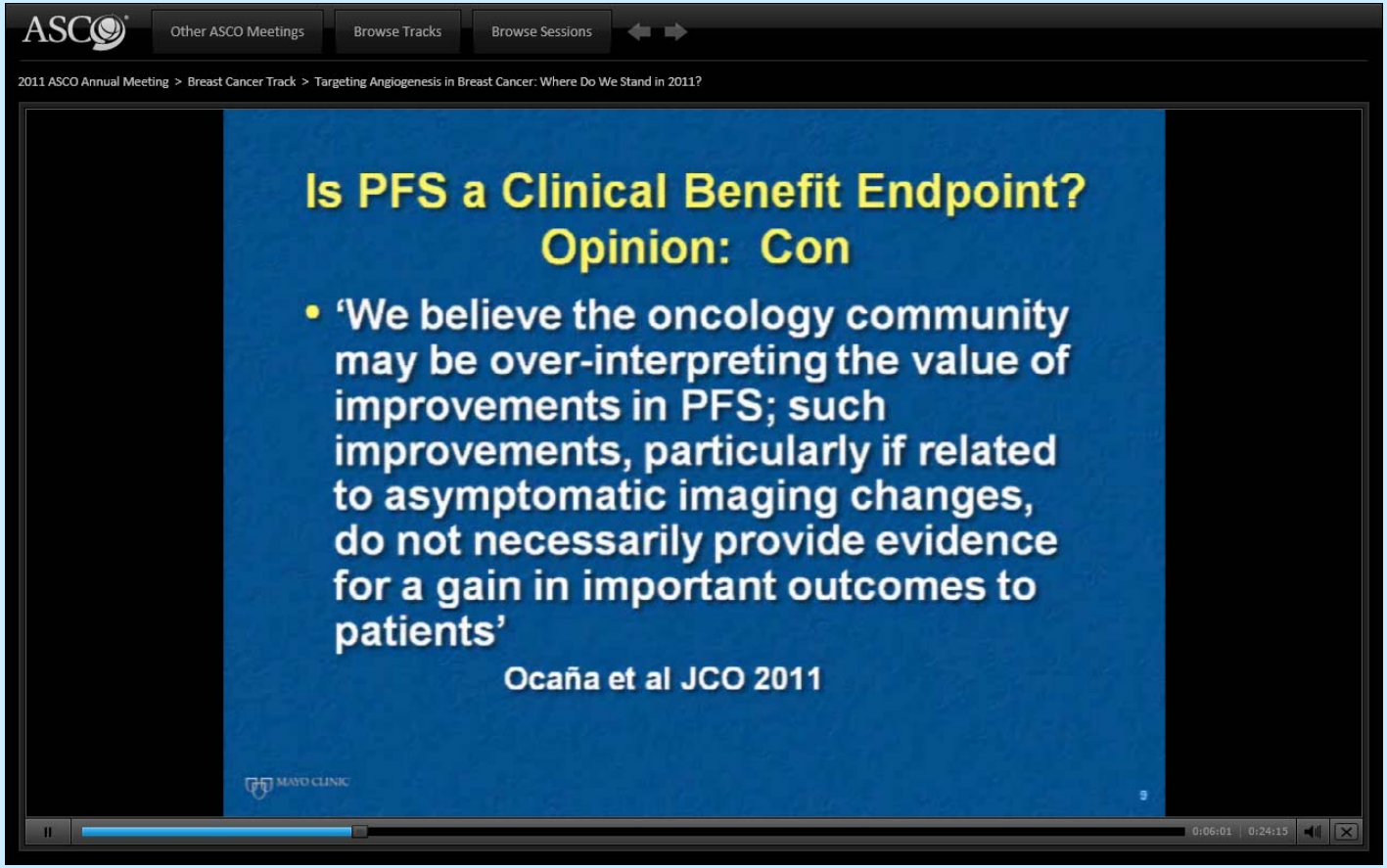


24

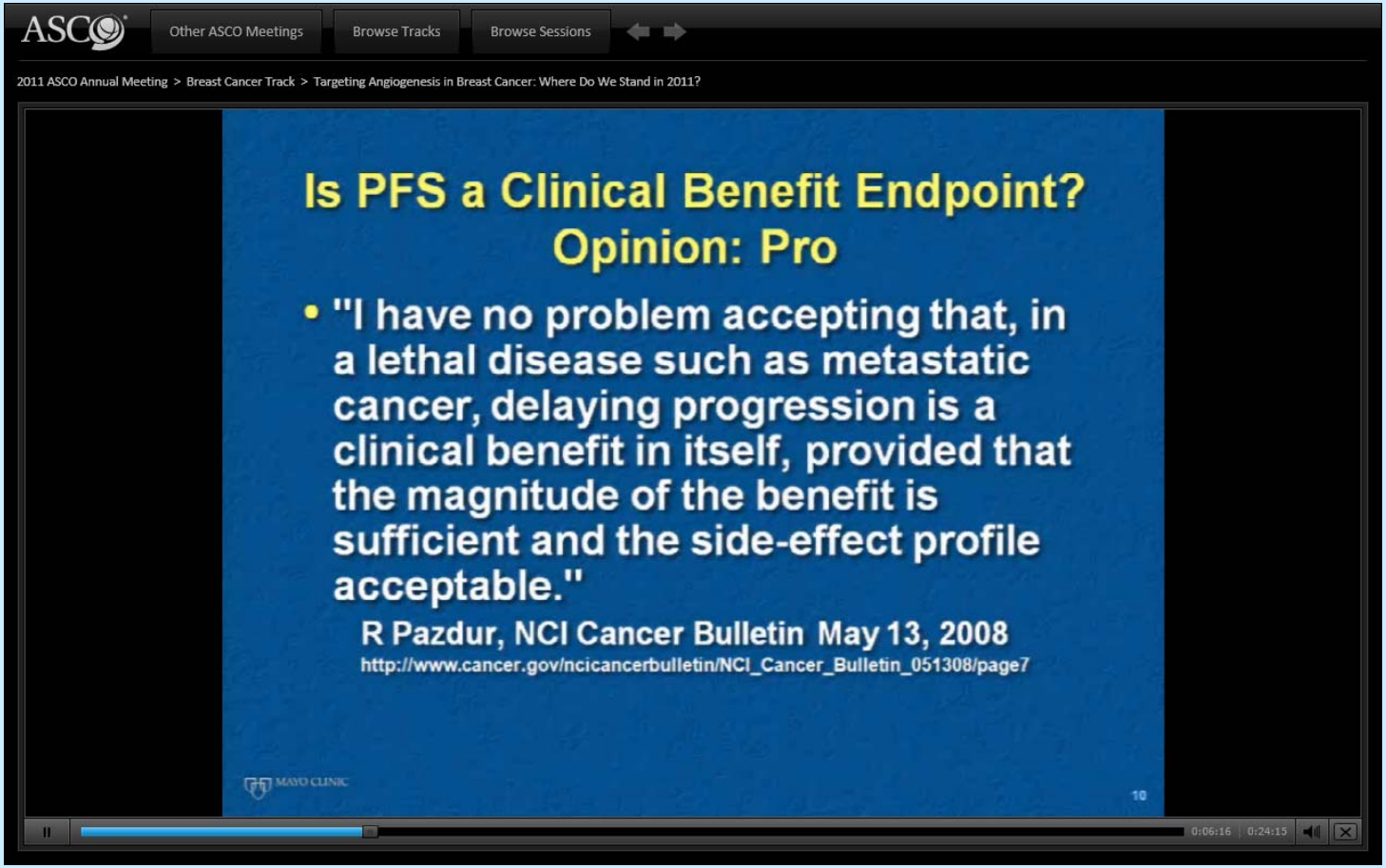
PFSの問題点

25

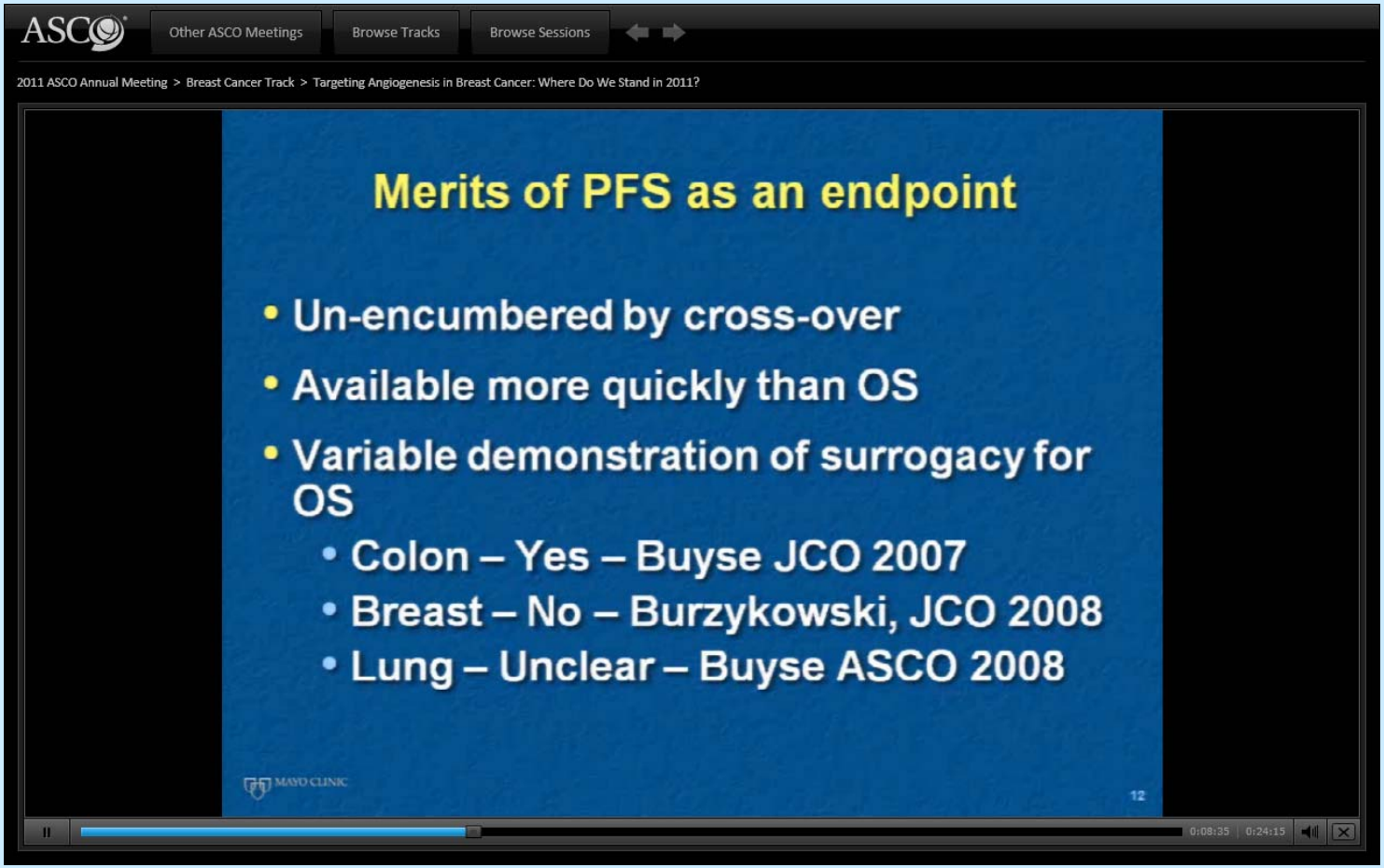
オンコロジストはPFS向上の意義を過大評価しがち



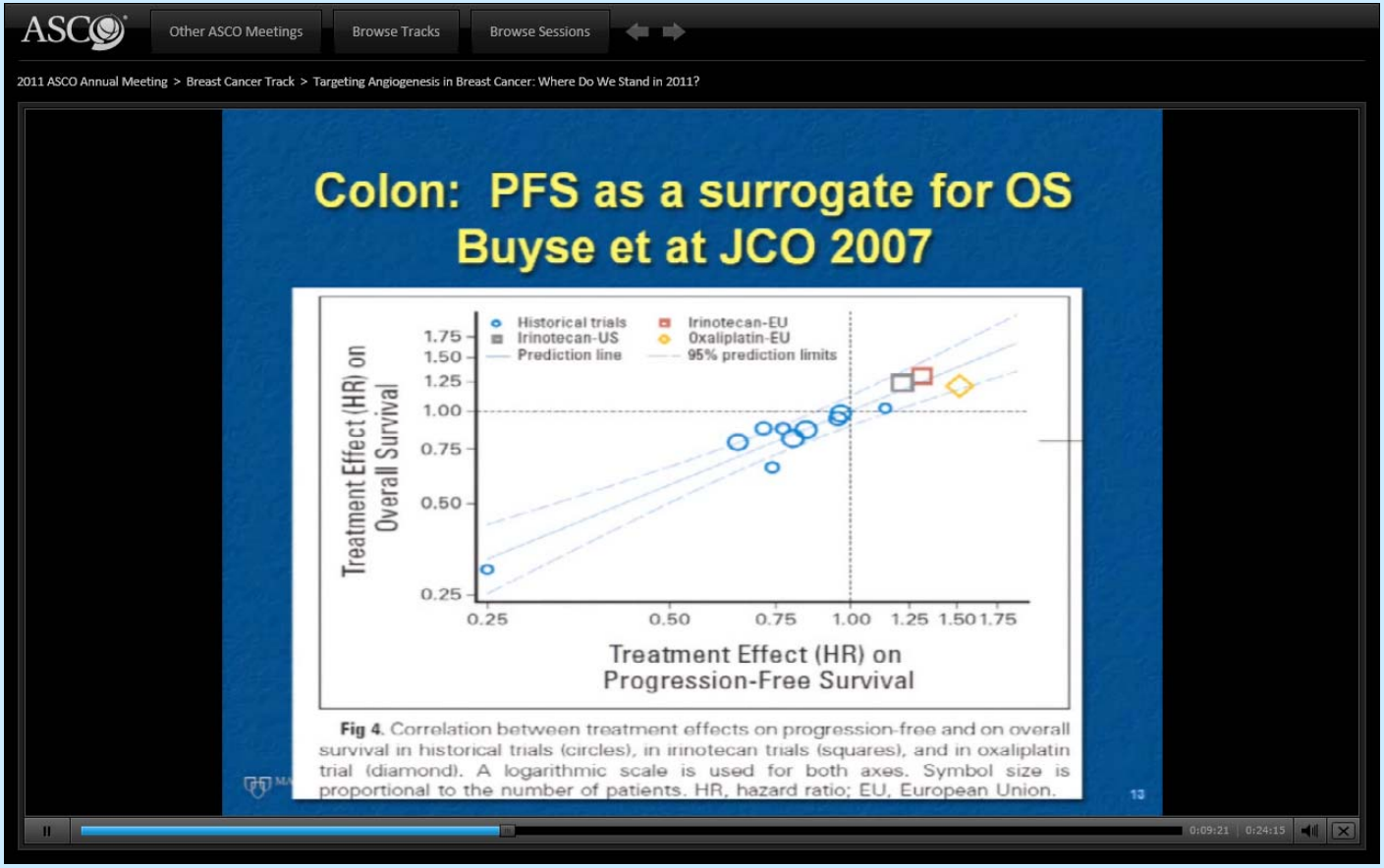
致死的疾患において便益が充分で副作用受容可能なら、PFS向上を受け入れない理由は無い



PFSの長所とOSに対する代替性

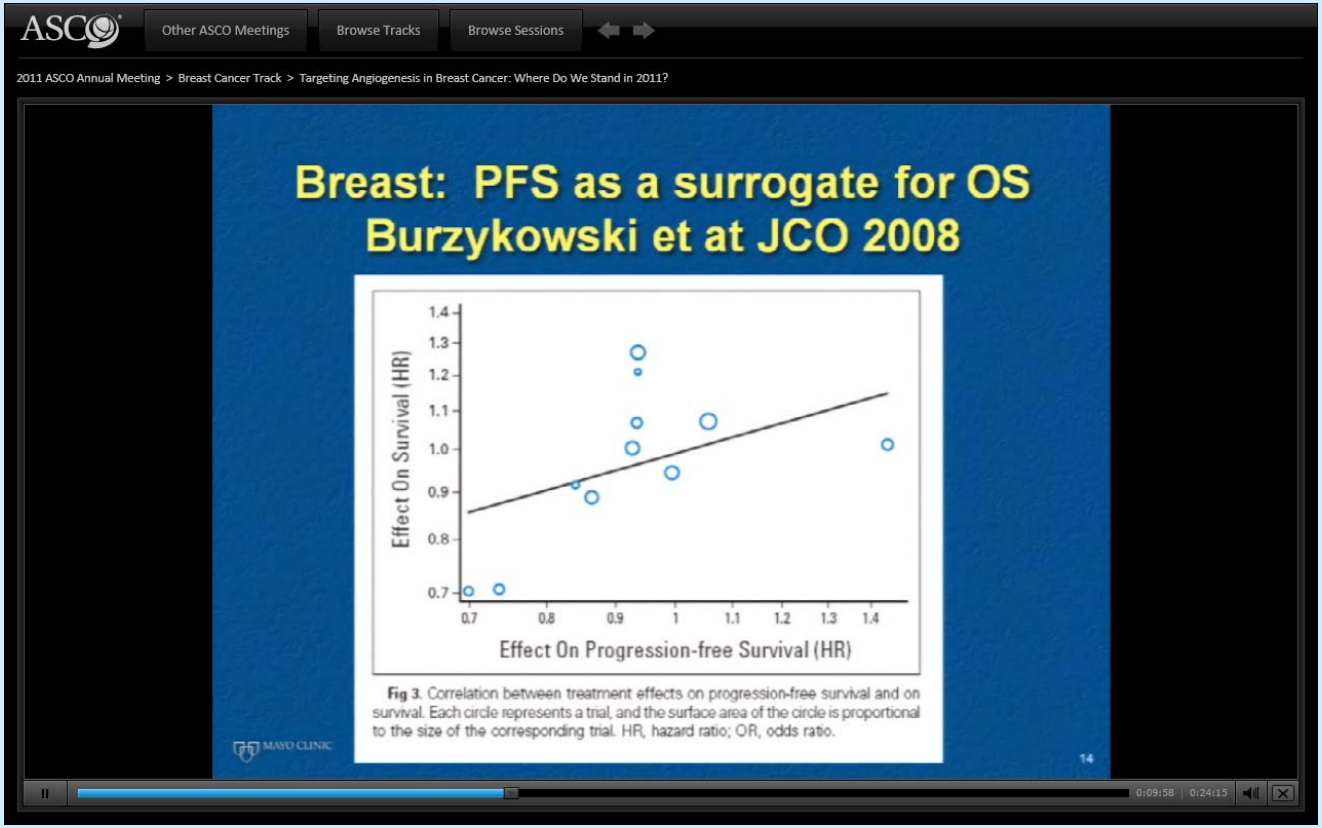


大腸癌では代替性高かった



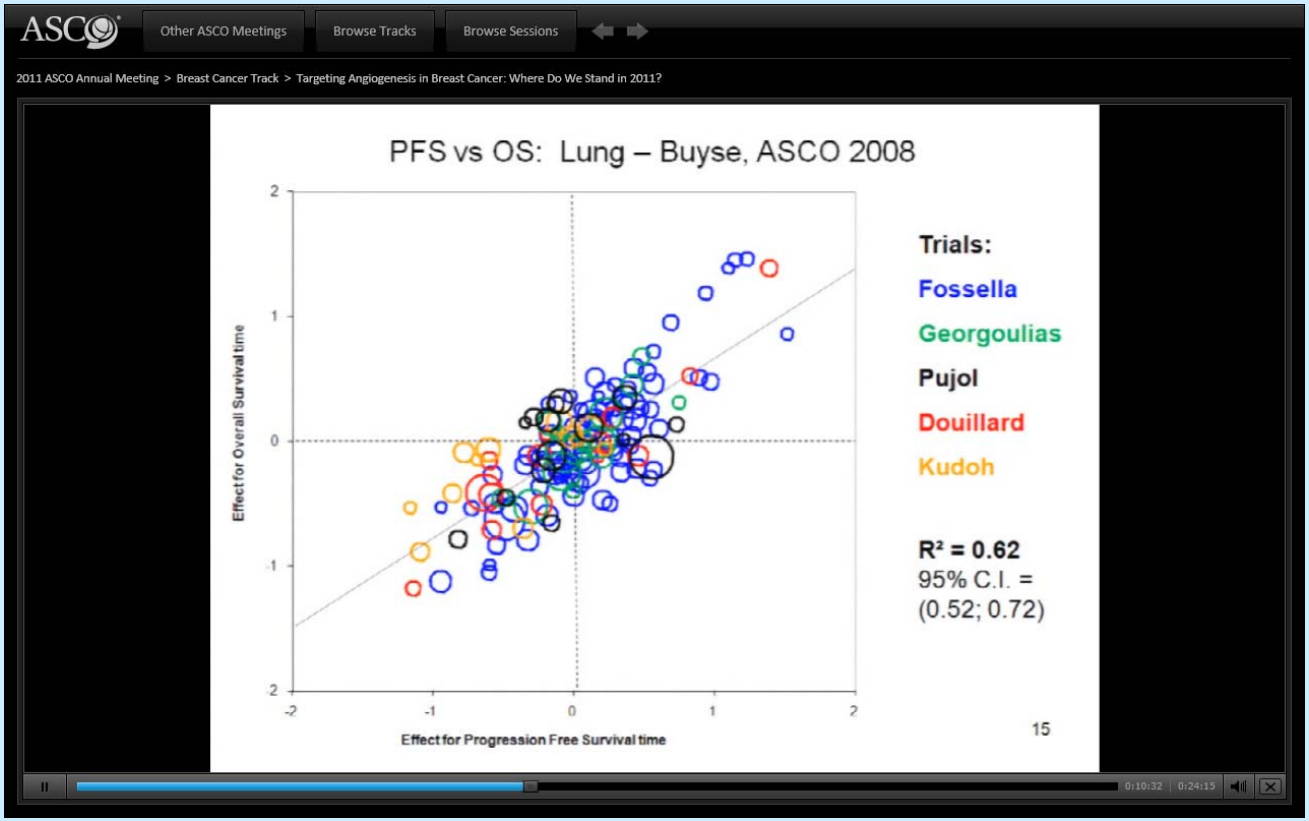
29

乳癌ではだめ


ASCO
Other ASCO Meetings
Browse Tracks
Browse Sessions
2011 ASCO Annual Meeting > Breast Cancer Track > Targeting Angiogenesis in Breast Cancer: Where Do We Stand in 2011?
Breast: PFS as a surrogate for OS
Burzykowski et at JCO 2008
Effect On Survival (HR)
1.4
1.3
1.2
1.1
1.0
0.9
0.8
0.7
0.7
0.8
0.9
1
1.1
1.2
1.3
1.4
Effect On Progression-free Survival (HR)
Fig 3. Correlation between treatment effects on progression-free survival and on survival. Each circle represents a trial, and the surface area of the circle is proportional to the size of the corresponding trial. HR, hazard ratio; OR, odds ratio.
MAYO CLINIC
14
0:09:58
0:24:15

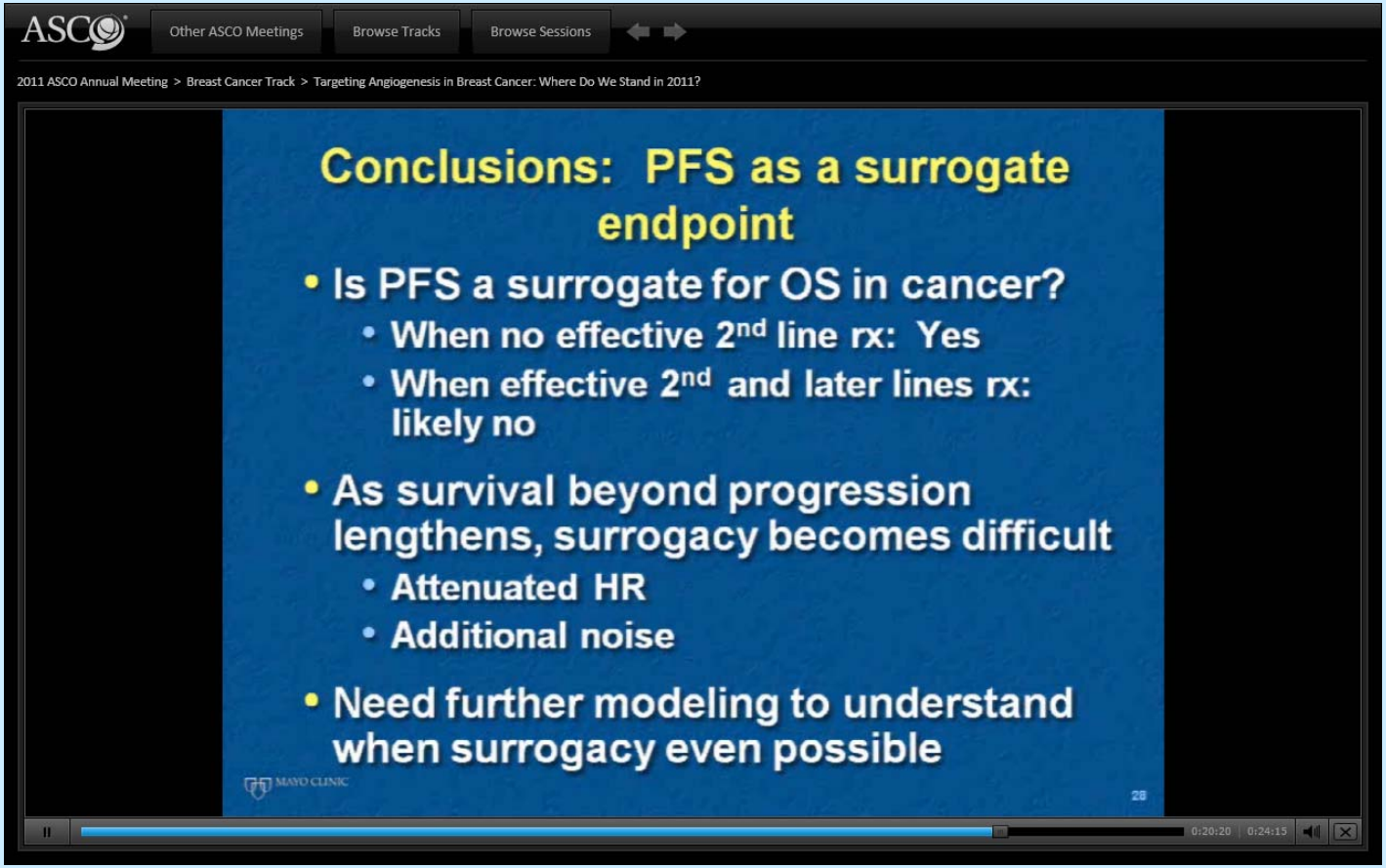
30

NSCLCではそれなり



31

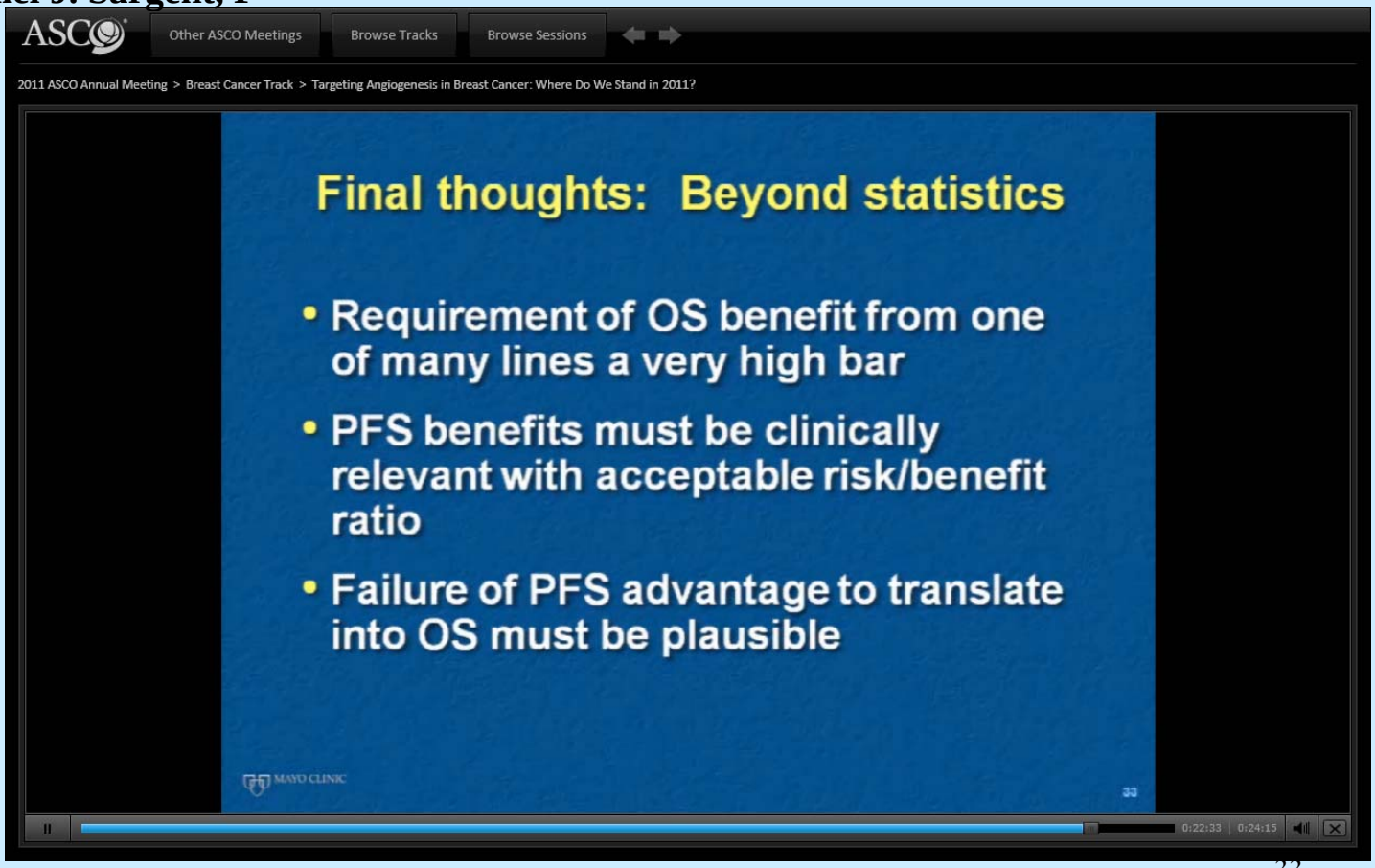
結論は：　代替性がどのような場合に成り立つか理解するために、さらに（モデリングの）検討必要



32

ASCO2011:Targeting Angiogenesis in Breast Cancer: Where Do We Stand in 2011? Appropriate Endpoints for Clinical Trials and Source of U.S. Food and Drug Administration/Oncologic Drugs Advisory Committee Controversy

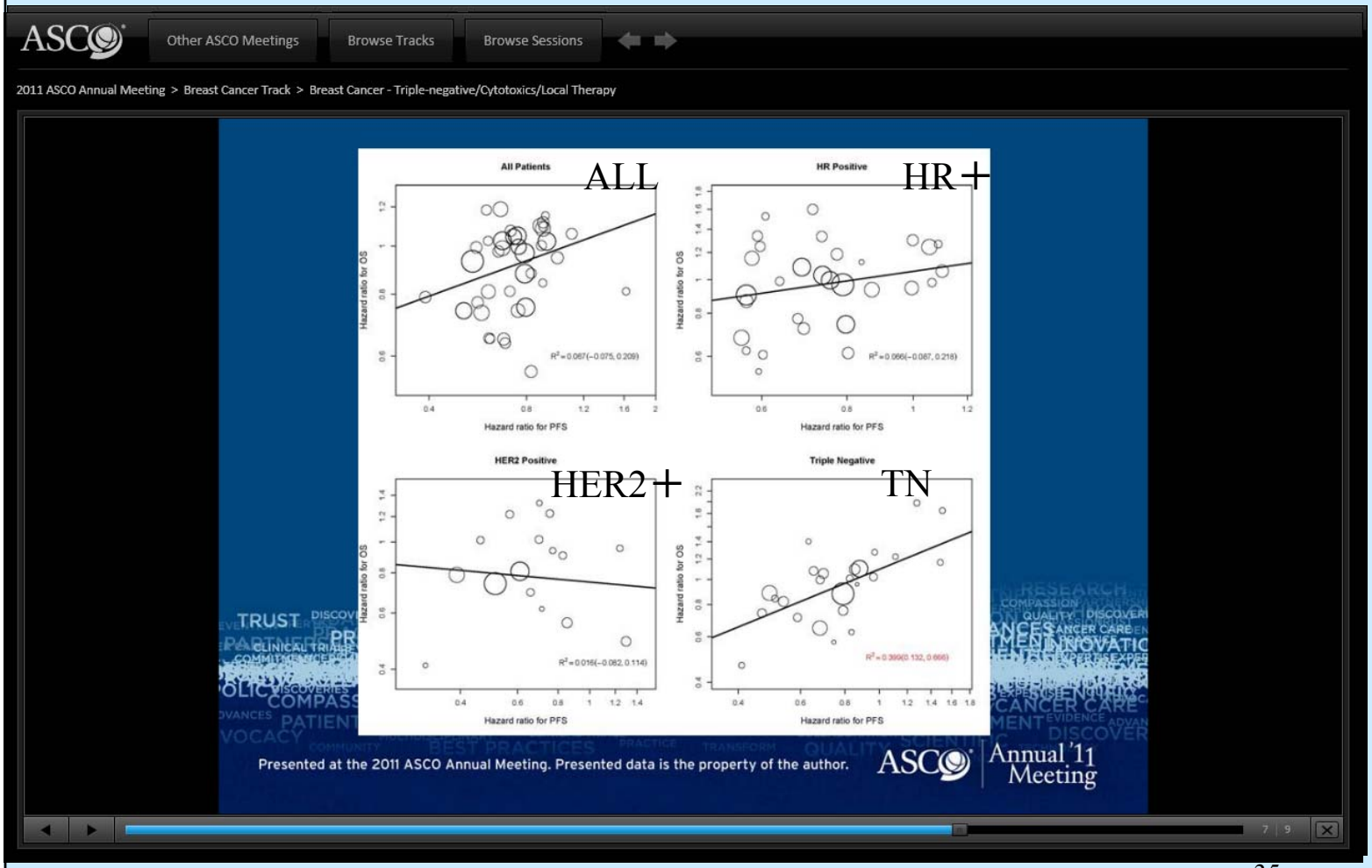
Daniel J. Sargent, P



33

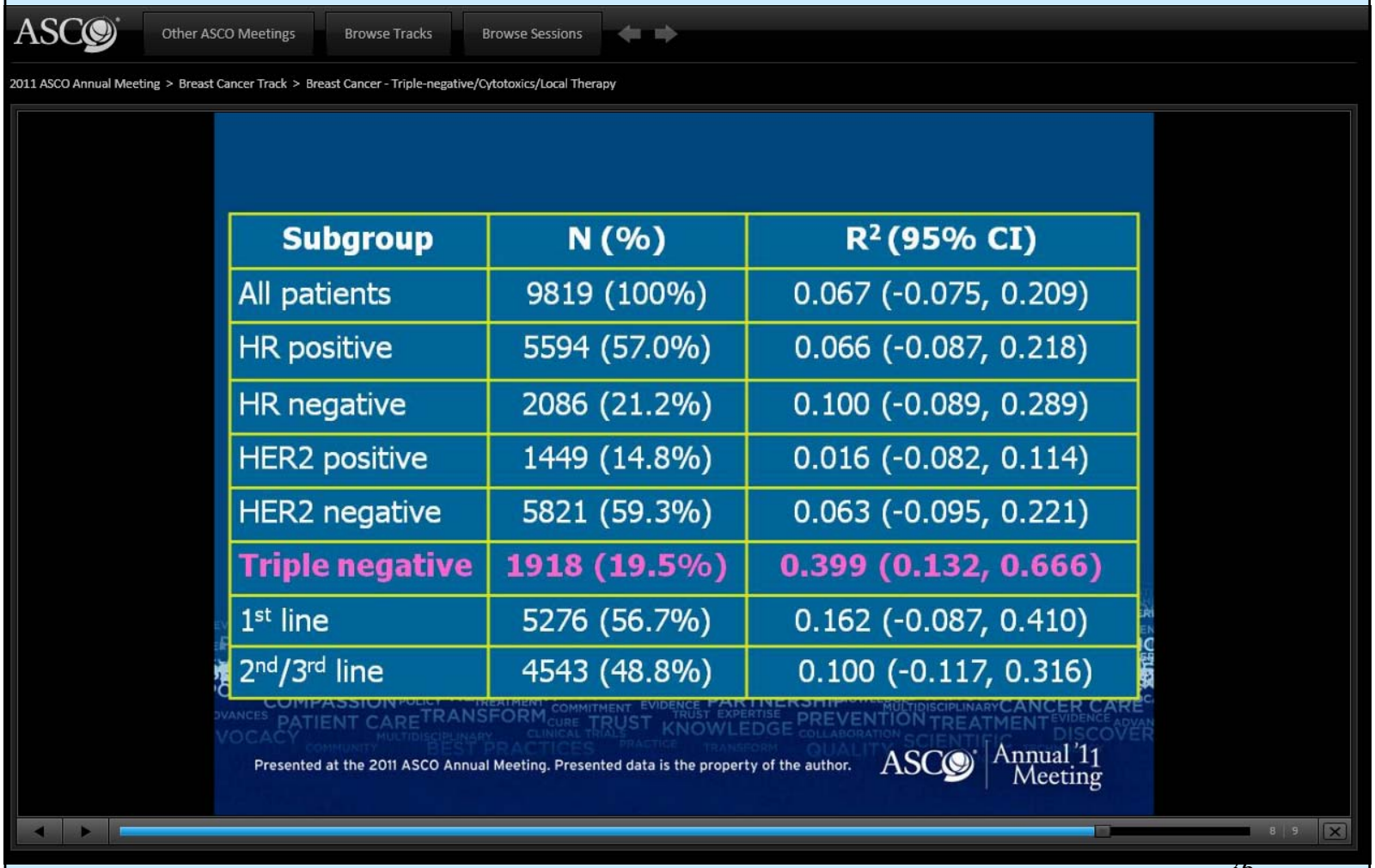
ASC02011- 1035
FDA申請データの乳癌データの解析　*J Clin Oncol* 2012;30:1705-1711

TNでは相関高い



35

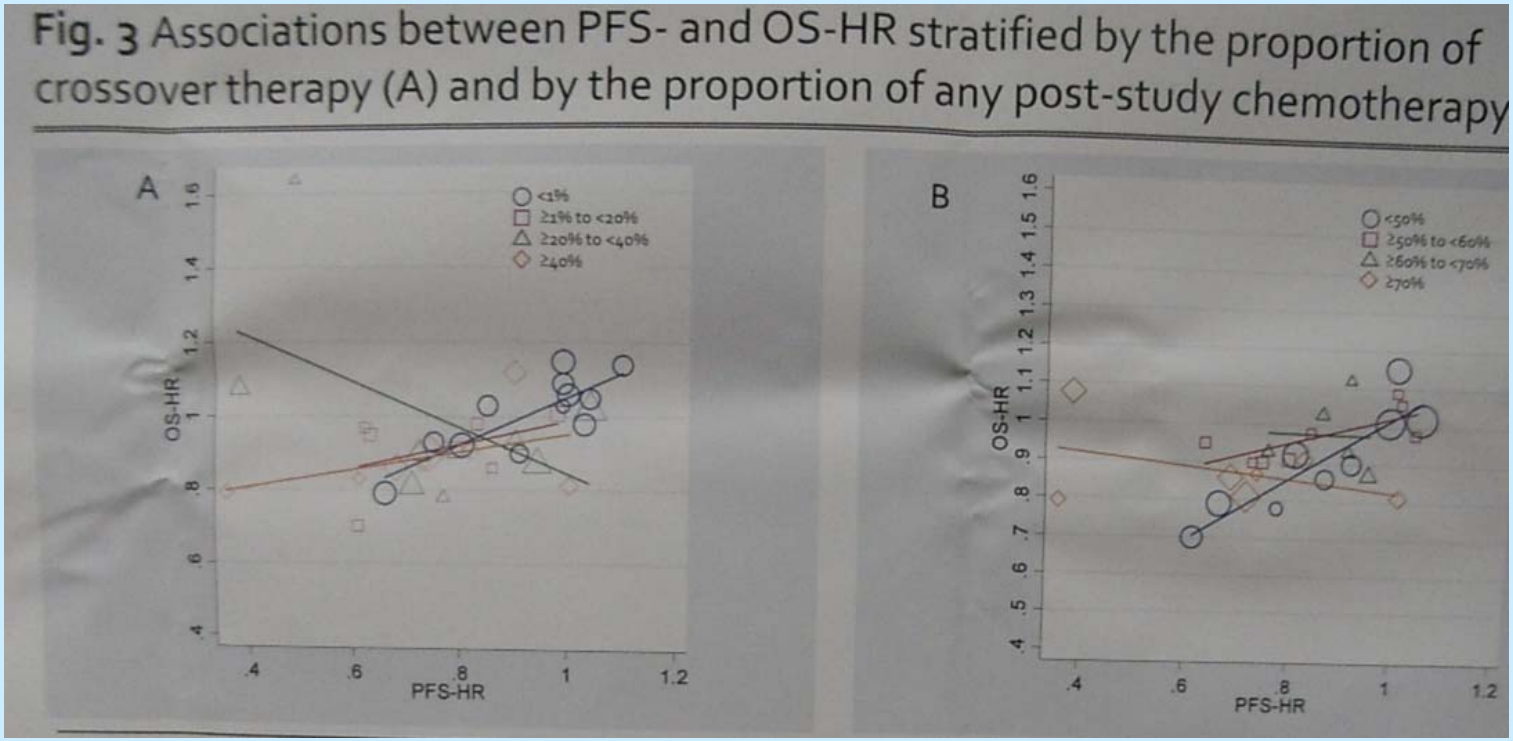
1035　TNは有効な2次治療が存在しない



Subgroup	N (%)	R^2 (95% CI)
All patients	9819 (100%)	0.067 (-0.075, 0.209)
HR positive	5594 (57.0%)	0.066 (-0.087, 0.218)
HR negative	2086 (21.2%)	0.100 (-0.089, 0.289)
HER2 positive	1449 (14.8%)	0.016 (-0.082, 0.114)
HER2 negative	5821 (59.3%)	0.063 (-0.095, 0.221)
Triple negative	**1918 (19.5%)**	**0.399 (0.132, 0.666)**
1st line	5276 (56.7%)	0.162 (-0.087, 0.410)
2nd/3rd line	4543 (48.8%)	0.100 (-0.117, 0.316)

36

ASCO2011-7540　Hotta et al.
肺癌の事例：クロスオーバーと2次治療の影響
2次治療へのクロスーオーバーによりPFS代替性失われる



37

胃がんでは･･･moderately correlated、検証はできなかった

R^2=0.61 (95% CI = 0.04 to 1.00).

Paoletti, Oba et al. JNCI 2013; 105: 1667-70.

BRIEF COMMUNICATION

Progression-Free Survival as a Surrogate for Overall Survival in Advanced/Recurrent Gastric Cancer Trials: A Meta-Analysis

Xavier Paoletti, Koji Oba, Yung-Jue Bang, Harry Bleiberg, Narikazu Boku, Olivier Bouché, Paul Catalano, Nozomu Fuse, Stefan Michiels, Markus Moehler, Satoshi Morita, Yasuo Ohashi, Atsushi Ohtsu, Arnaud Roth, Philippe Rougier, Junichi Sakamoto, Daniel Sargent, Mitsuru Sasako, Kohei Shitara, Peter Thuss-Patience, Eric Van Cutsem, Tomasz Burzykowski, Marc Buyse; on behalf of the GASTRIC group

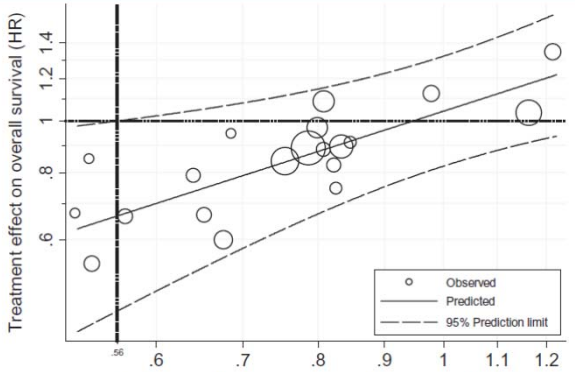
Manuscript received February 10, 2013; revised July 25, 2013; accepted July 25, 2013.

Correspondence to: Xavier Paoletti, PhD, Bisotatistics Dept / INSERM U900, Institut Curie, 26 rue d'Ulm, 75005 Paris, France (e-mail: xavier.paoletti@curie.net).

The traditional endpoint for assessing efficacy of chemotherapies for advanced/recurrent gastric cancer is overall survival (OS), but OS requires prolonged follow-up. We investigated whether progression-free survival (PFS) is a valid surrogate for OS. Using individual patient data from the GASTRIC meta-analysis, surrogacy of PFS was assessed through the correlation between the endpoints and through the correlation between the treatment effects on the endpoints. External validation of the prediction based on PFS was also evaluated. Individual data from 4069 patients in 20 randomized trials were analyzed. The rank correlation coefficient between PFS and OS was 0.853 (95% confidence interval [CI] = 0.852 to 0.854). The R^2 between treatment effects on PFS and on OS was 0.61 (95% CI = 0.04 to 1.00). Treatment effects on PFS and on OS were only moderately correlated, and we could not confirm the validity of PFS as a surrogate endpoint for OS in advanced/recurrent gastric cancer.

J Natl Cancer Inst;2013;105:1667–1670

GASTRIC グループ
20試験、4069患者



38

OSかPFSか？

- Survival Post Progression（SPP）が長くなる（たとえば12ヶ月以上）と、PFSのOSに対する代替性は弱くなる

  乳癌から、大腸癌そしてNSCLCへ

  Broglio and Berry, *JNCI* 2009; 101: 1642-9.

Detecting an Overall Survival Benefit that Is Derived From Progression-Free Survival

Kristine R. Broglio, Donald A. Berry

Background Whether progression-free survival (PFS) or overall survival (OS) is the more appropriate endpoint in clinical trials of metastatic cancer is controversial. In some disease and treatment settings, an improvement in PFS does not result in an improved OS.

Methods We partitioned OS into two parts and expressed it as the sum of PFS and survival postprogression (SPP). We simulated randomized clinical trials with two arms that had respective medians for PFS of 6 and 9 months. We assumed no treatment difference in median SPP. We found the probability of a statistically significant benefit in OS for various median SPP and observed *P* values for PFS. We compared the sample sizes required for PFS vs OS for various median SPP. We compare our results with the literature regarding surrogacy of PFS for OS by use of the correlation between hazard ratios for PFS and OS. All statistical tests were two-sided.

Results For a trial with observed *P* value for improvement in PFS of .001, there was a greater than 90% probability for statistical significance in OS if median SPP was 2 months but less than 20% if median SPP was 24 months. For a trial requiring 280 patients to detect a 3-month difference in PFS, 350 and 2440 patients, respectively, were required to have the same power for detecting a real difference in OS that is carried over from the 3-month benefit in PFS when the median SPP was 2 and 24 months.

Conclusions Addressing SPP is important in understanding treatment effects. For clinical trials with a PFS benefit, lack of statistical significance in OS does not imply lack of improvement in OS, especially for diseases with long median SPP. Although there may be no treatment effect on SPP, its variability so dilutes the OS comparison that statistical significance is likely lost. OS is a reasonable primary endpoint when median SPP is short but is too high a bar when median SPP is long, such as longer than 12 months.

J Natl Cancer Inst 2009;101:1642–1649

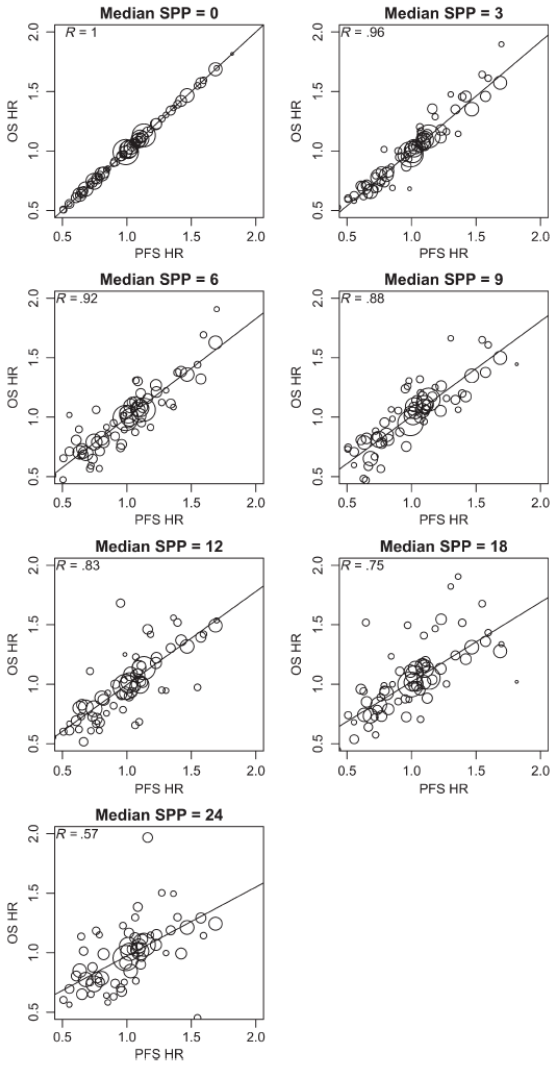


Figure 4. Association between progression-free survival (PFS) and overall survival (OS) for a single simulation of 67 trials. Each study had a randomly selected sample size and PFS hazard ratio, which remains fixed across scenarios while median survival postprogression (SPP) times were allowed to vary (0, 3, 6, 9, 12, 18, and 24 months). Hazard ratios (HRs) for PFS and OS were estimated with a proportional hazards model, and the correlation was estimated from a linear regression model weighted by the number of patients in each trial. The **size of the circle** is relative to the total sample size of the study. The diagonal line is the fitted weighted linear regression line.

シミュレーションによる仮想的試験でのPFSとOS

SPP(survival postprogression) が大きくなるとPFSのハザード比とOSのハザード比の相関低下

40

median PFS= 2:3 OS=PFS+SPP

Table 1. Summary of hazard ratios for overall survival (OS)*

Median SPP, mo	Median OS, HR (95% interval)
2	0.687 (0.514–0.909)
4	0.710 (0.517–0.966)
6	0.727 (0.511–1.023)
8	0.736 (0.502–1.068)
10	0.746 (0.491–1.100)
12	0.749 (0.479–1.140)
14	0.752 (0.470–1.174)
16	0.758 (0.462–1.207)
18	0.759 (0.448–1.241)
20	0.762 (0.440–1.277)
22	0.763 (0.428–1.304)
24	0.763 (0.416–1.333)

* The 95% interval of the OS hazard ratio values extends from the 2.5 percentile to the 97.5 percentile for 50 000 simulations. HR = hazard ratio; SPP = survival postprogression.

OSのハザード比メディアンは0.75前後、しかし信頼区間が広がり有意でない確率が上昇する

41

median PFS= 2:3 OS=PFS+SPP

Table 2. Summary of overall survival (OS) *P* values and power when the power for progression-free survival (PFS) is 80%, 85%, and 90%*

	80% power for PFS		85% power for PFS		90% power for PFS	
Median SPP, mo	Median *P* (95% interval)	OS power	Median *P* (95% interval)	OS power	Median *P* (95% interval)	OS power
2	.013 (<.001–.600)	0.696	.009 (<.001–.498)	0.748	.004 (<.001–.351)	0.822
4	.041 (<.001–.814)	0.535	.030 (<.001–.750)	0.589	.016 (<.001–.627)	0.672
6	.082 (.002–.888)	0.415	.063 (<.001–.867)	0.461	.040 (<.001–.811)	0.540
8	.125 (.001–.915)	0.335	.103 (.003–.905)	0.370	.069 (.002–.876)	0.441
10	.167 (.001–.935)	0.279	.141 (.001–.929)	0.317	.104 (.004–.901)	0.370
12	.199 (.001–.946)	0.244	.170 (.001–.937)	0.276	.132 (.001–.923)	0.322
14	.227 (.002–.951)	0.217	.202 (.001–.947)	0.238	.160 (.001–.936)	0.286
16	.257 (.002–.956)	0.194	.229 (.002–.949)	0.214	.190 (.001–.939)	0.258
18	.276 (.003–.957)	0.183	.253 (.002–.952)	0.195	.212 (.002–.946)	0.234
20	.298 (.003–.963)	0.167	.272 (.003–.956)	0.182	.235 (.002–.951)	0.214
22	.310 (.004–.966)	0.156	.289 (.003–.962)	0.169	.251 (.002–.952)	0.197
24	.322 (.005–.971)	0.146	.304 (.003–.966)	0.160	.265 (.003–.958)	0.188

* The 95% interval extends from the 2.5 percentile to the 97.5 percentile of the OS P values from 50 000 simulations. SPP = survival postprogression.

SPP(survival postprogression)が大きくなると、OSに対する検出力が減少

42

median PFS= 2:3 OS=PFS+SPP

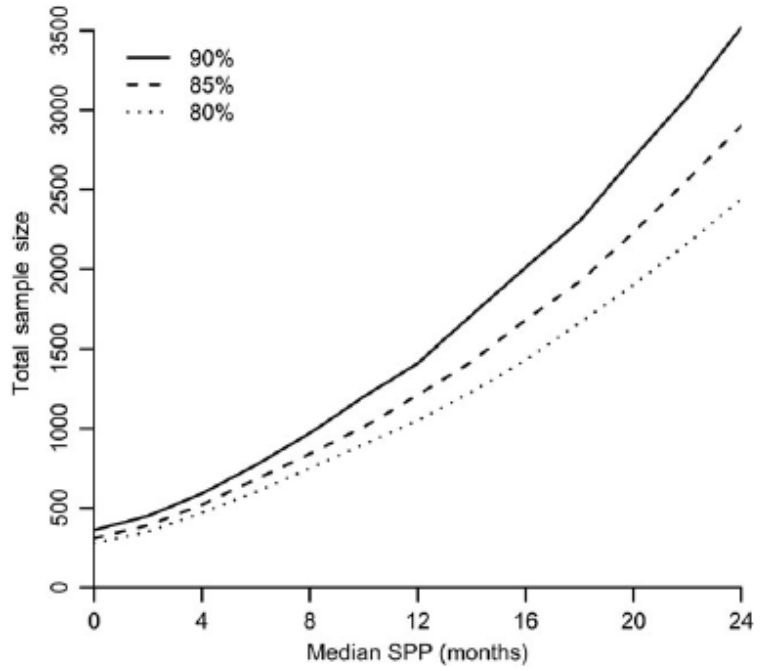


Figure 3. Sample sizes required for detecting a statistically significant difference in overall survival by median survival postprogression (SPP). The **three curves** were indexed by the power for overall survival (ie, powers of 90%, 85%, and 80%).

SPP(survival postprogression)が大きくなると必要サンプルサイズが増加

N=280 (検出力80%)
→
N=350(SPP=2)，2440(SPP=24)

43

OSかPFSか？

- Survival Post Progression（SPP）が長くなる（たとえば12ヶ月以上）と、PFSのOSに対する代替性は弱くなる

  乳癌から、大腸癌そしてNSCLCへ

  Broglio and Berry, *JNCI* 2009; 101: 1642-9.

- PFSには曖昧さ

  測定間隔・測定手段、予定していない検査の扱い

  主治医評価と中央評価の食い違い、QC

  盲検下でない場合の報告バイアスと脱落バイアス

  感度解析の必要性

  Freidlin et al. *JCO* 2007; 25: 2122-6.

  Bhattacharya et al. *JCO* 2009; 27: 5958-64.

  Dodd et al. *JCO* 2008; 3791-6.

- FDAの曖昧な態度も批判されている。十分なPFSの差とQOL向上が示されれば問題ないのであろうが基準設定は困難

盲検化されていない場合のPFS測定の問題と提案

Freidlin et al. *JCO* 2007; 25: 2122-6

- **報告バイアス**
  - 評価を行う医師が患者を試験治療にスイッチしたいと思う気持ちから、画像評価を過大に行い、対照群でより早くPFSを宣言してしまう（主観による評価バイアス）
  - 試験治療を受けたいと思う患者が対照治療群でより早く症状の進行を医師に報告し、その確認のための画像評価が早めになされてしまい、その結果早めにPFSが報告されることもありうる（評価時期のバイアス）。逆に毒性の強い試験治療で来院と画像評価が早めになり、試験治療群で早めにPFSが報告されることも起こりうる
- **患者脱落attritionバイアス**
  - 対照治療に割り付けられた患者に、新しい治療を受けたいという希望から脱落が多くなり、これがバイアスを引き起こす
- **正式の評価時点を2回に!**
  - 対照（標準）治療でのメディアンPFSとその2倍
  - 途中でのPD判定はその後の正式な時点で
  - 検出力はすべての時点を用いる方法と比べそれほど低下しない
  - 負担の軽減とバイアスの軽減

45

PFSに対する感度解析Sensitivity analysis

Bhattacharya et al. JCO 2009; 27: 5958-64

- PFSに曖昧さが伴うことは避けられない
- データの取り扱いを何通りかに変更し結論のrobustnessを検討
- 解析方法についてはプロトコルあるいは解析計画書に規定

- PFSに差がないという仮定のもとで判定時期の違いがどう影響するかシミュレーションを行う
- 両群を対等に扱う
- 試験群は保守的に、対照群はliberalに扱う
  - 予定していない判定時期、不完全情報なら遡らせる
  - 臨床情報に基づくPD判定を除く
  - 中央委員会で確認されなかったPD判定を打ち切り扱い

46

Sensitivity analysis 感度解析

Bhattacharya et al. JCO 2009; 27: 5958-64

Role of Sensitivity Analyses in Assessing Progression-Free Survival in Late-Stage Oncology Trials

Suman Bhattacharya, Gwen Fyfe, Robert J. Gray, and Daniel J. Sargent

A B S T R A C T

Sensitivity analysis is an important statistical technique that assesses whether the results of phase III trials are robust and likely to be generalizable. Until recently, sensitivity analyses were rarely included in phase III trials, and they remain poorly understood by many oncologists. Sensitivity analyses are critical to understanding the strength of conclusions made in the primary analysis of a late-stage clinical trial. They examine the influence of protocol design errors, unintended biases, deviations from assumptions underlying statistical models, and any unanticipated treatment delivery or practice patterns on trial results. In trials with complex or subjective end points, they also allow an understanding of the extent to which a positive outcome is driven by a single, possibly subjective, and therefore biased, element of an end point. The purposes of this article are to explain how sensitivity analyses are performed, to discuss areas of a clinical trial where sensitivity analyses should focus, and to illuminate the importance of this technique in the rigorous evaluation of late-stage clinical trial data, using specific examples. This article focuses on late-stage trials that use progression-free survival or time to progression as their primary end point, because sensitivity analyses are particularly important in these cases for which the end point is potentially subject to bias. Three sources of potential bias are explored: assessment time, symptomatic (ie, nonradiologic) disease progression, and missing data. For each source of potential bias, case studies are presented to highlight the role that sensitivity analyses play in determining whether the trial's conclusions are robust.

J Clin Oncol 27:5958-5964. © 2009 by American Society of Clinical Oncology

Sensitivity analysis （Example of E2100）
robustness of the results is confirmed

Table 4. Sensitivity Analyses for Missing Data Bias

Type of Analysis	Experimental Arm	Control Arm	HR	(95% CI)
	Median PFS (months)			
	Bevacizumab + Paclitaxel (n = 368)	Paclitaxel (n = 354)		
Study E2100 (bevacizumab in metastatic breast cancer)[9]				
Primary analysis of PFS	11.3	5.8	0.48	0.39 to 0.61
PFS analysis of the patients (N = 649) with at least one scan submitted for the IRF review	11.3	6.0	0.50	0.40 to 0.63
The PD date was backdated to the date of the first missed tumor assessment if one or more tumor assessments were missing immediately preceding PD	11.2	5.6	0.48	0.38 to 0.61
The PD date for patients whose investigator-reported PD was not confirmed by the IRF was set to the last tumor assessment + 1 day	9.2	5.0	0.46	0.37 to 0.56
Worst-case analysis: PD date was set to the last tumor assessment + 1 day in the bevacizumab arm and was censored in the paclitaxel-only arm for cases in which the investigator-reported PD was not confirmed by the IRF	9.2	5.8	0.60	0.48 to 0.74
Worst-case analysis: Censoring because of nonprotocol cancer therapy and early discontinuation were considered PD events in bevacizumab arm and censored in the paclitaxel-only arm	8.2	5.8	0.78	0.64 to 0.95

オリジナル解析

両群を同様に保守的に扱う

試験群を保守的に対照群をliberalに扱う

48

再掲：最近までの、とりあえずの結論は

- 2次治療以降に有効な治療が登場すれば
- SPP(survival post progression)が延長すれば
  PFSのOSに対する代替性は薄まる（ハザードは希釈される）

- PFSの曖昧さとそれに対する対処の必要性
- 大きなPFSの改善と優れたrisk/benefit profileなら（さらに
  経済的に許容できるなら）PFSによる承認は当然ありうる

FDA: Guidance for Industry Clinical Trial Endpoints for the Approval of Non-Small Cell Lung Cancer Drugs and Biologics, 2011 June

EMA: Appendix 1 to the guideline on the evaluation of anticancer medicinal products in man. EMA/CHMP/27994/2008/Rev.1, 2012, December

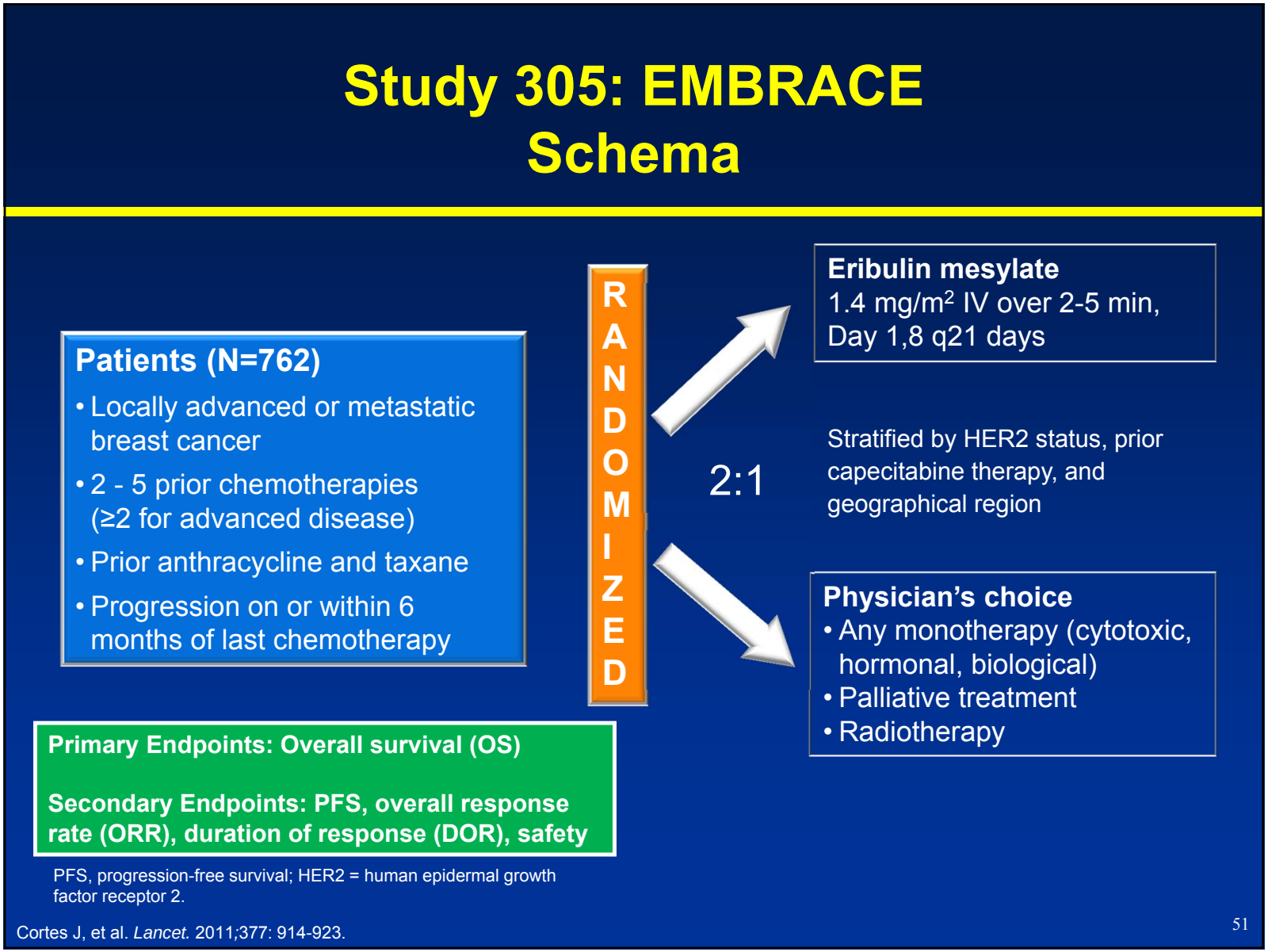
49

さて・・・

- **PFSで同等なのにOSが優れる例の登場**
  - **MBC: Eribulin 305, 301**
  - **NSCLC: Pemetrexed JMDB**
  - **Colon：Cetuximab　vs　Bevacizumab**
  - **Gastric;（Next ESMO）**
- **その理由は？**

50

Study 305: EMBRACE Schema

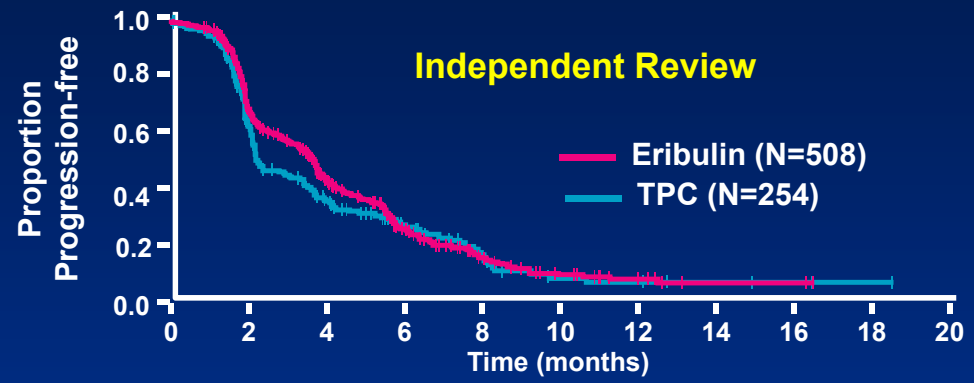


PFS, progression-free survival; HER2 = human epidermal growth factor receptor 2.

Cortes J, et al. *Lancet.* 2011;377: 914-923.

51

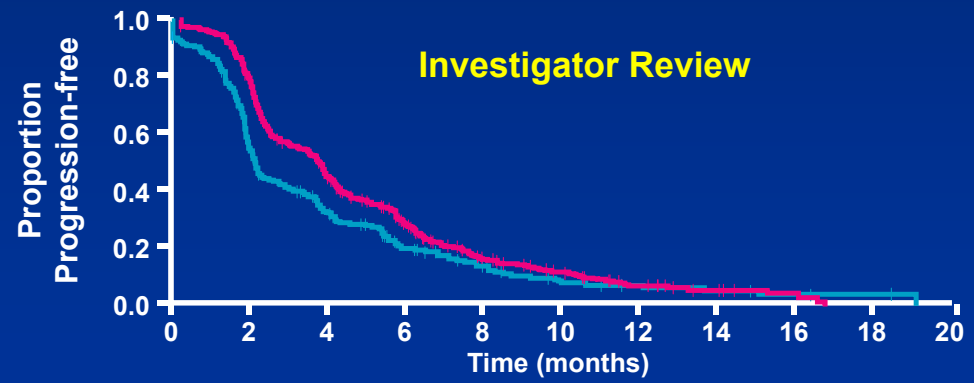
Study 305: EMBRACE
Progression-Free Survival (ITT Population)



Median (months)

Eribulin	3.7
TPC	2.2

HR 0.87 (95%CI 0.71, 1.05; *P*=0.137)



Median (months)

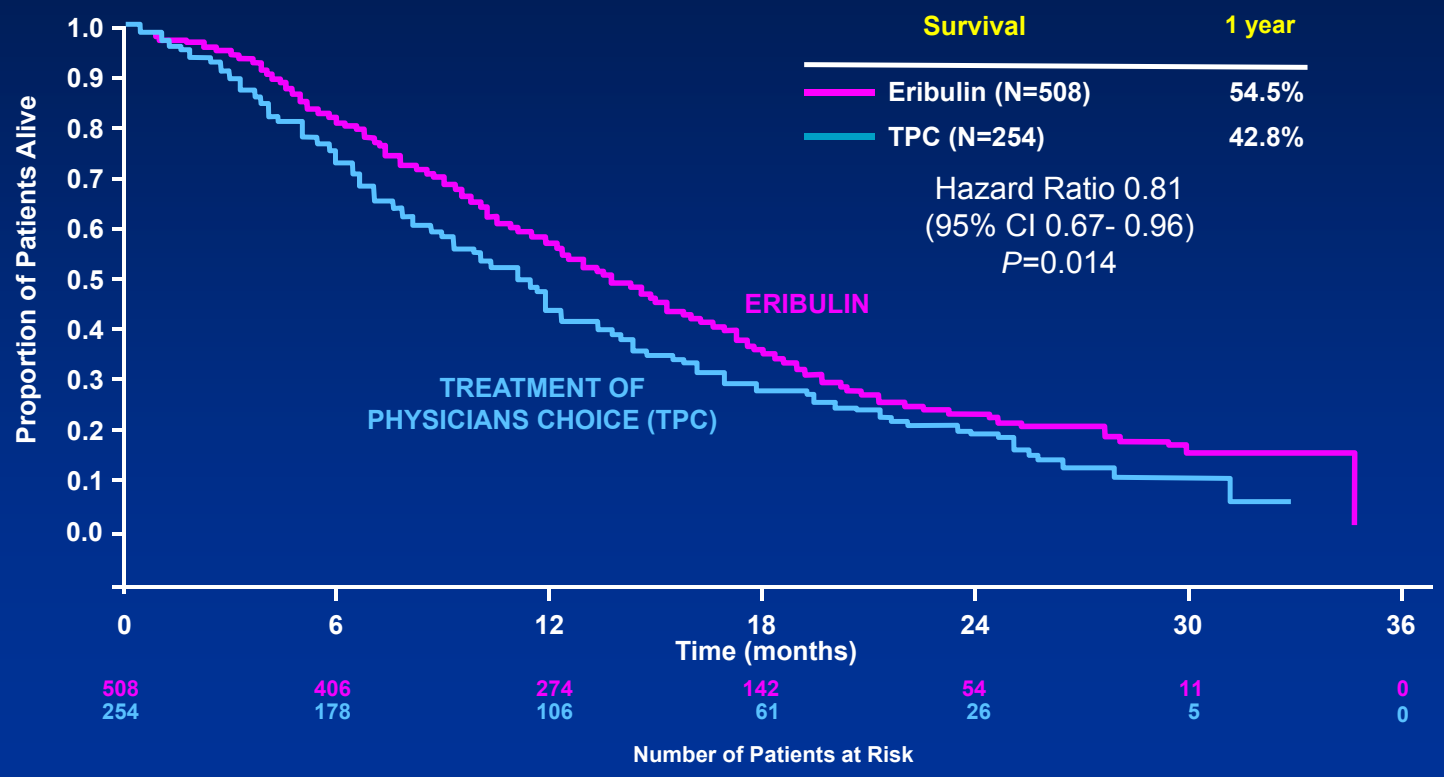
Eribulin	3.6
TPC	2.2

HR 0.76 (95%CI 0.64, 0.90; *P*=0.002)

Cortes J, et al. *Lancet.* 2011;377: 914-923.

52

Study 305: Kaplan-Meier Analysis of Overall Survival
Overall Survival Update Analysis*

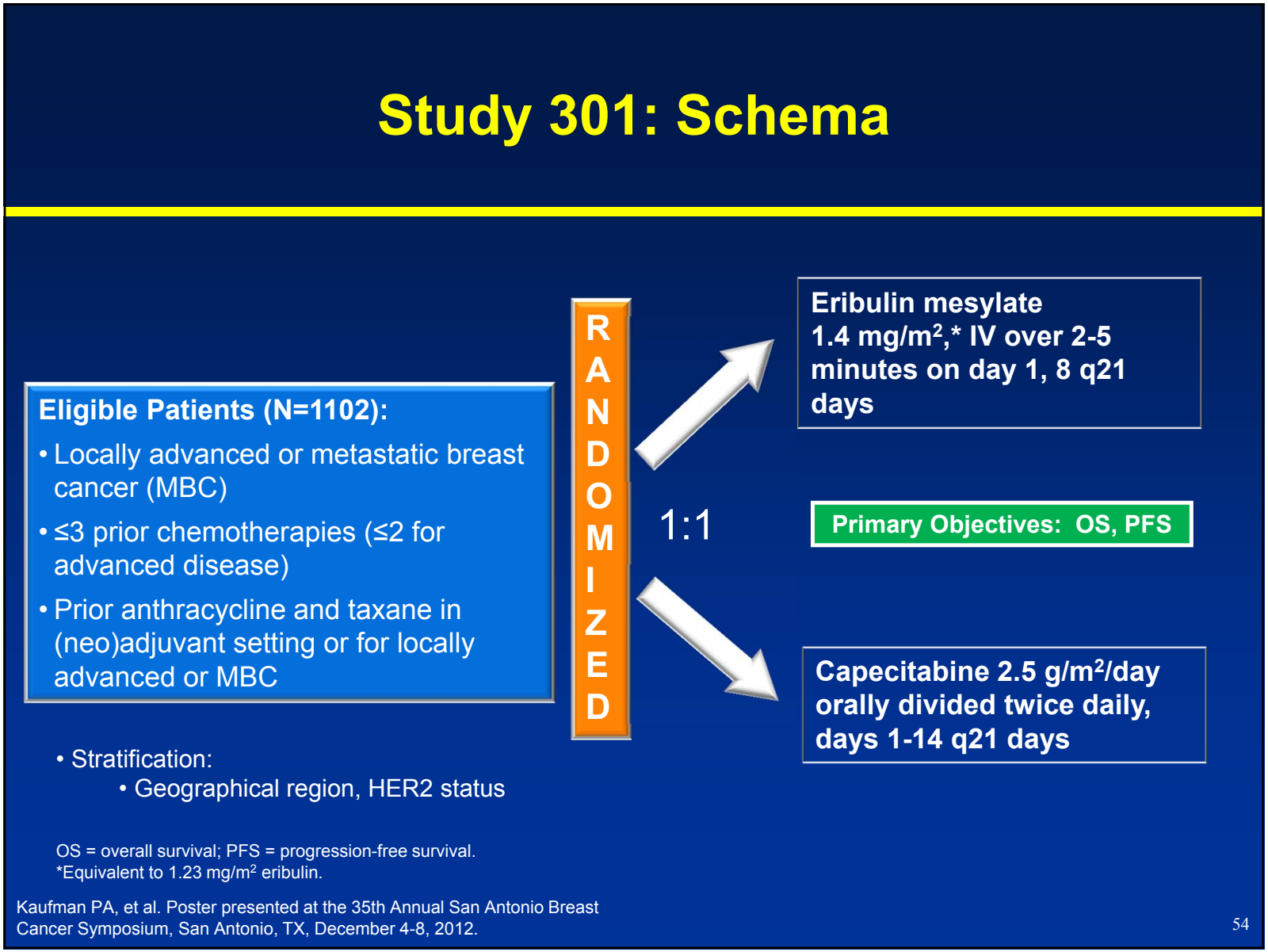


*Intent-to-treat population analysis was not protocol prespecified.

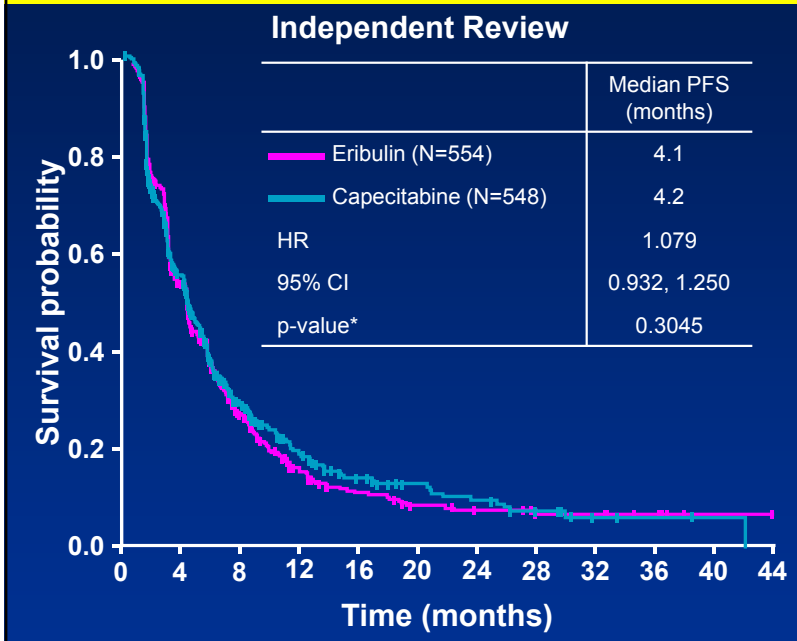
Cortes J, et al. *Lancet.* 2011;377: 914-923.

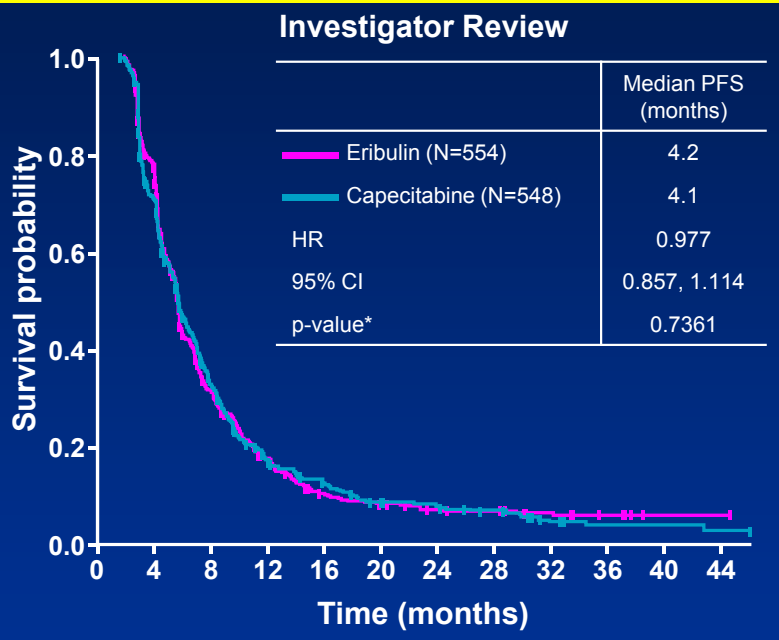
53

Study 301: Schema



Study 301: Progression-Free Survival



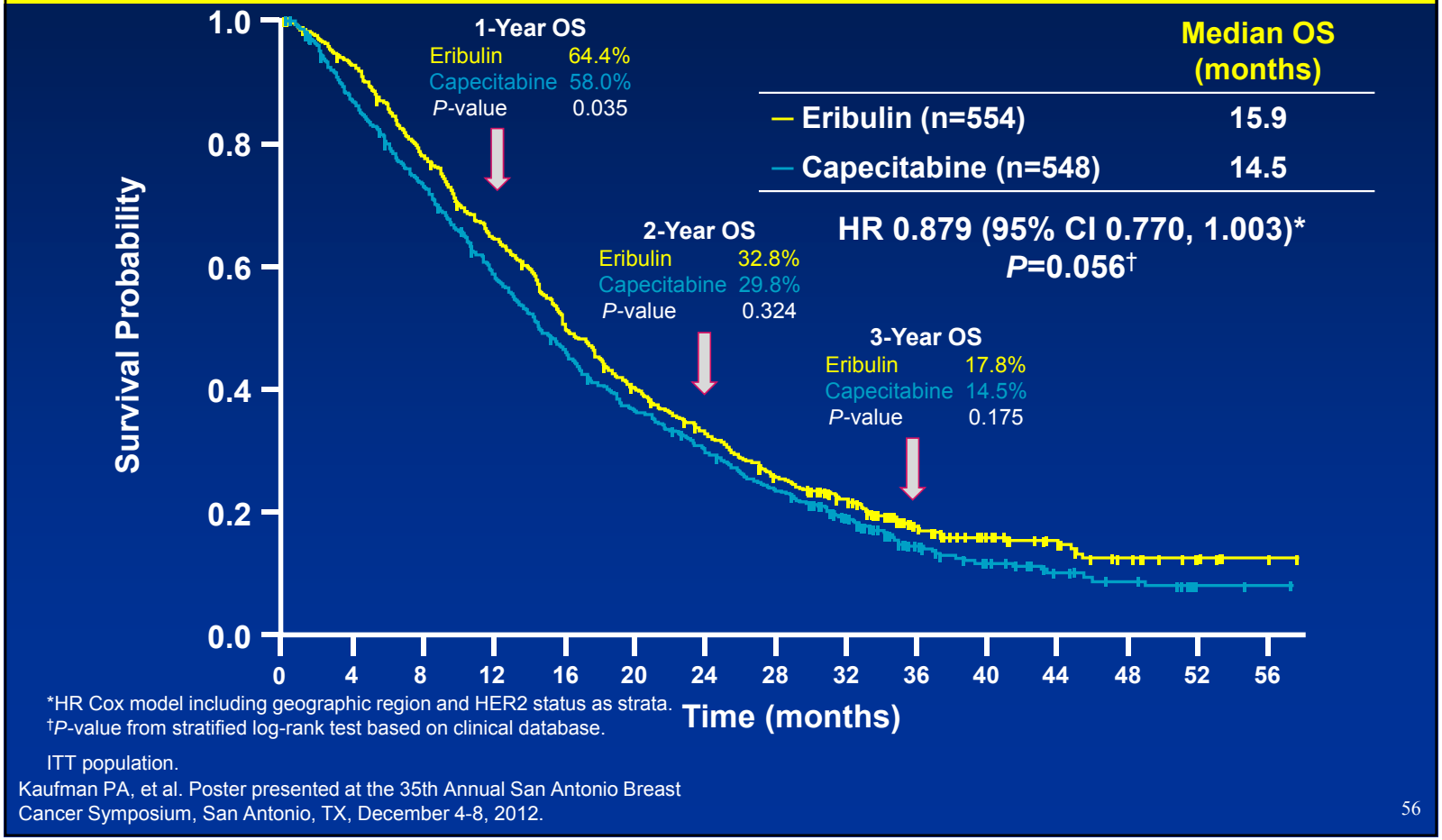


*Stratified log-rank test based on IVRS data for Region and Clinical Database for HER2 status.
ITT Population

Kaufman PA, et al. Poster presented at the 35th Annual San Antonio Breast Cancer Symposium, San Antonio, TX, December 4-8, 2012.

55

Study 301: Overall Survival



*HR Cox model including geographic region and HER2 status as strata.
†*P*-value from stratified log-rank test based on clinical database.

ITT population.

Kaufman PA, et al. Poster presented at the 35th Annual San Antonio Breast Cancer Symposium, San Antonio, TX, December 4-8, 2012.

56

Scagliotti et al.
***JCO* 2008; 26: 3543-51.**

Phase III Study Comparing Cisplatin Plus Gemcitabine With Cisplatin Plus Pemetrexed in Chemotherapy-Naive Patients With Advanced-Stage Non–Small-Cell Lung Cancer

Giorgio Vittorio Scagliotti, Purvish Parikh, Joachim von Pawel, Bonne Biesma, Johan Vansteenkiste, Christian Manegold, Piotr Serwatowski, Ulrich Gatzemeier, Raghunadharao Digumarti, Mauro Zukin, Jin S. Lee, Anders Mellemgaard, Keunchil Park, Shehkar Patil, Janusz Rolski, Tuncay Goksel, Filippo de Marinis, Lorinda Simms, Katherine P. Sugarman, and David Gandara

A B S T R A C T

Purpose
Cisplatin plus gemcitabine is a standard regimen for first-line treatment of advanced non–small-cell lung cancer (NSCLC). Phase II studies of pemetrexed plus platinum compounds have also shown activity in this setting.

Patients and Methods
This noninferiority, phase III, randomized study compared the overall survival between treatment arms using a fixed margin method (hazard ratio [HR] < 1.176) in 1,725 chemotherapy-naive patients with stage IIIB or IV NSCLC and an Eastern Cooperative Oncology Group performance status of 0 to 1. Patients received cisplatin 75 mg/m^2 on day 1 and gemcitabine 1,250 mg/m^2 on days 1 and 8 (n = 863) or cisplatin 75 mg/m^2 and pemetrexed 500 mg/m^2 on day 1 (n = 862) every 3 weeks for up to six cycles.

Results
Overall survival for cisplatin/pemetrexed was noninferior to cisplatin/gemcitabine (median survival, 10.3 *v* 10.3 months, respectively; HR = 0.94; 95% CI, 0.84 to 1.05). Overall survival was statistically superior for cisplatin/pemetrexed versus cisplatin/gemcitabine in patients with adenocarcinoma (n = 847; 12.6 *v* 10.9 months, respectively) and large-cell carcinoma histology (n = 153; 10.4 *v* 6.7 months, respectively). In contrast, in patients with squamous cell histology, there was a significant improvement in survival with cisplatin/gemcitabine versus cisplatin/pemetrexed (n = 473; 10.8 *v* 9.4 months, respectively). For cisplatin/pemetrexed, rates of grade 3 or 4 neutropenia, anemia, and thrombocytopenia ($P \leq .001$); febrile neutropenia ($P = .002$); and alopecia ($P < .001$) were significantly lower, whereas grade 3 or 4 nausea ($P = .004$) was more common.

Conclusion
In advanced NSCLC, cisplatin/pemetrexed provides similar efficacy with better tolerability and more convenient administration than cisplatin/gemcitabine. This is the first prospective phase III study in NSCLC to show survival differences based on histologic type.

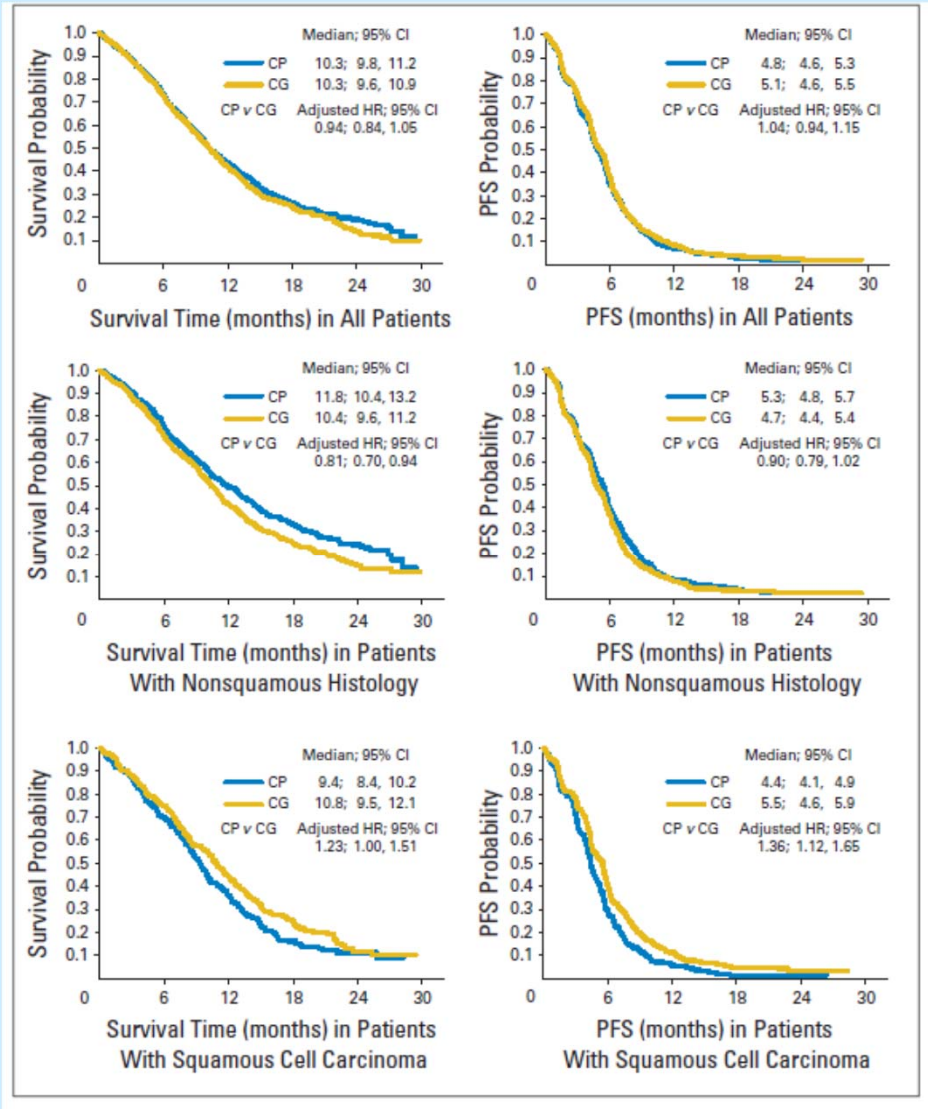
J Clin Oncol 26:3543-3551. © 2008 by American Society of Clinical Oncology

57

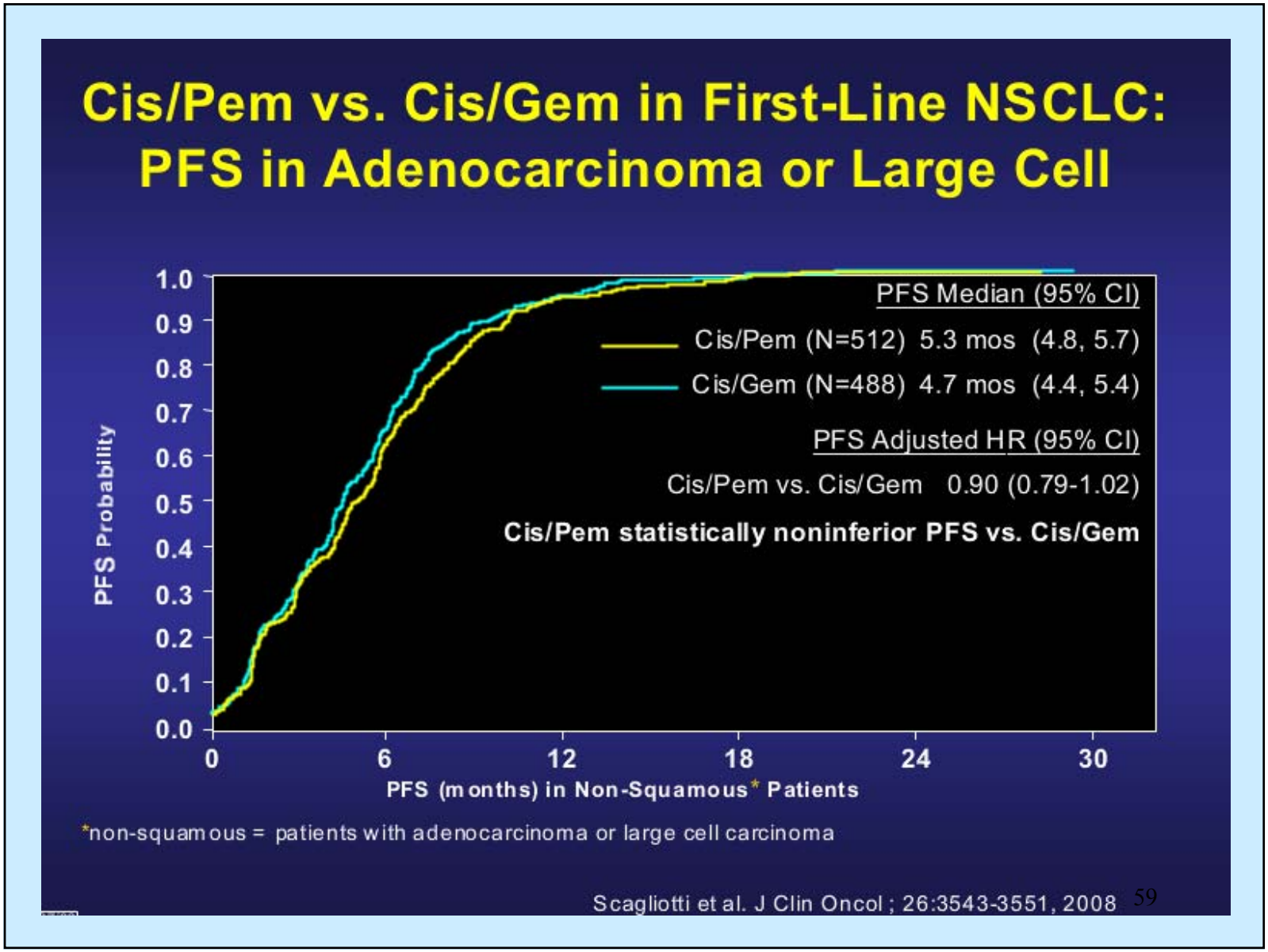
組織型による
効果の違い

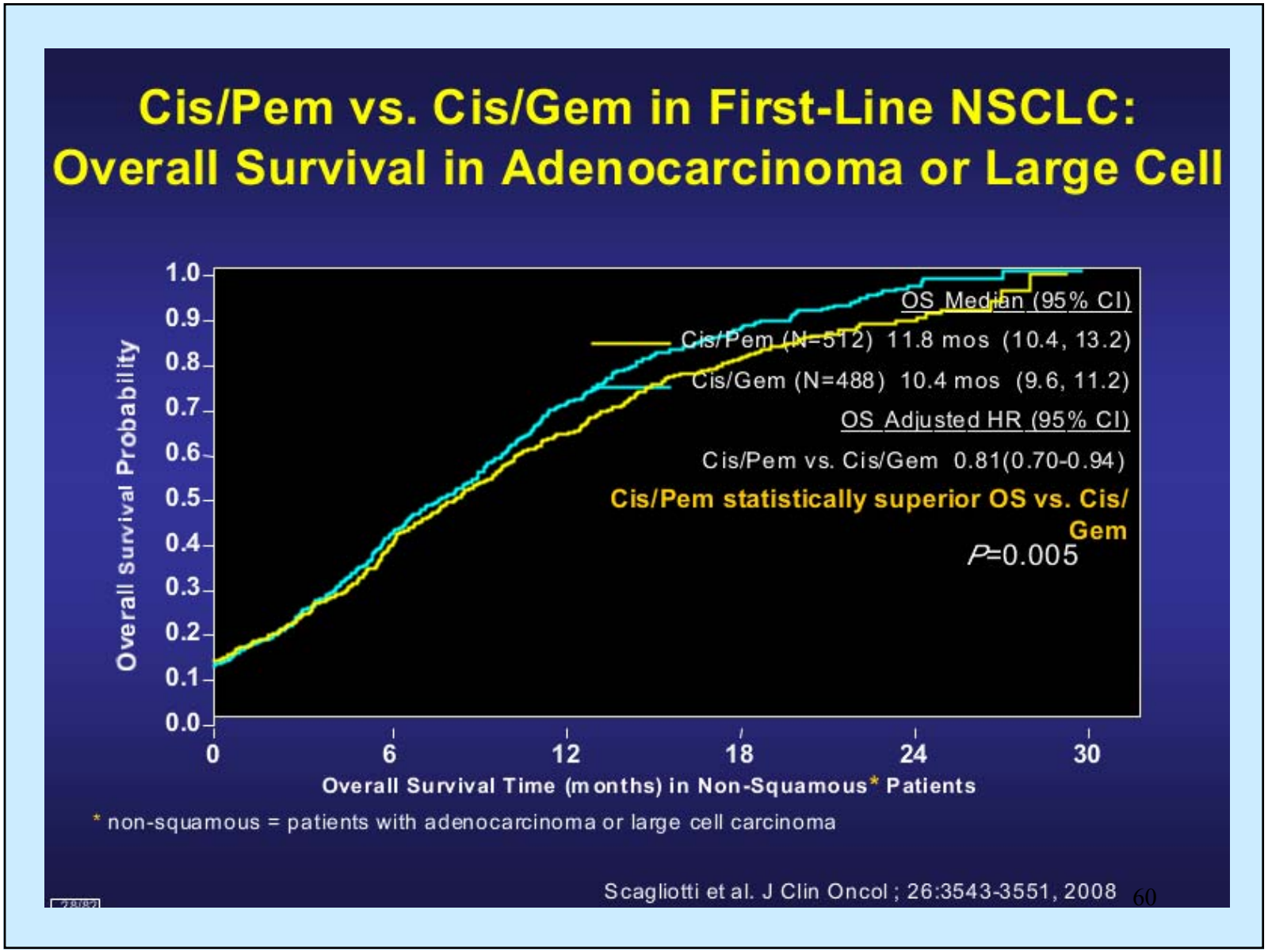
そして

OSに強い効果



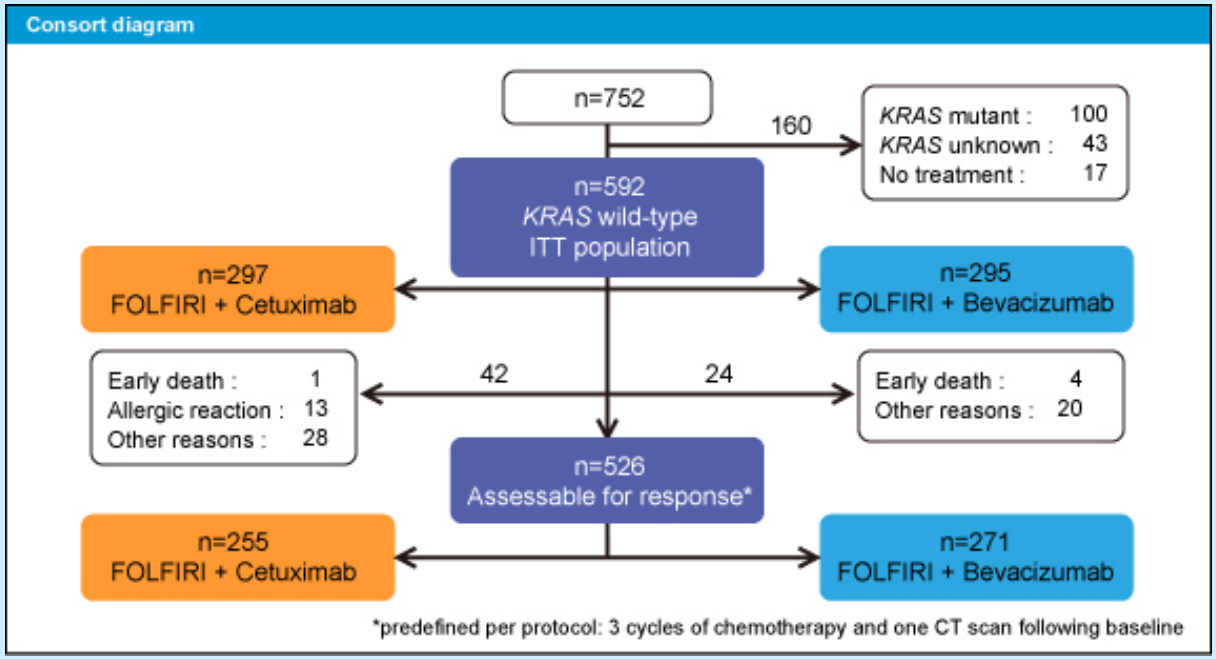
58





Abstract #LBA3506 (ASCO2013)

Randomized Comparison of FOLFIRI Plus Cetuximab versus FOLFIRI Plus Bevacizumab as 1st-line Treatment of *KRAS*Wild-type Metastatic Colorectal Cancer: German AIO Study KRK-0306 (FIRE-3).

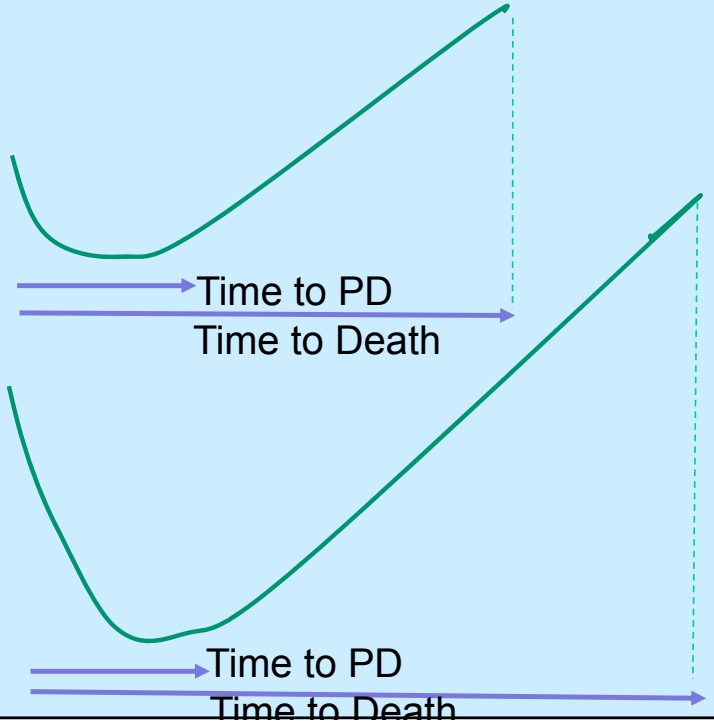


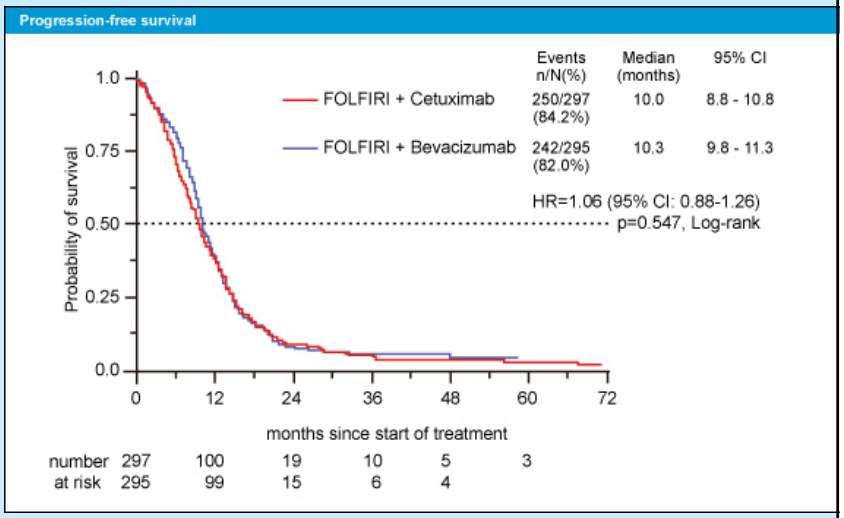
61

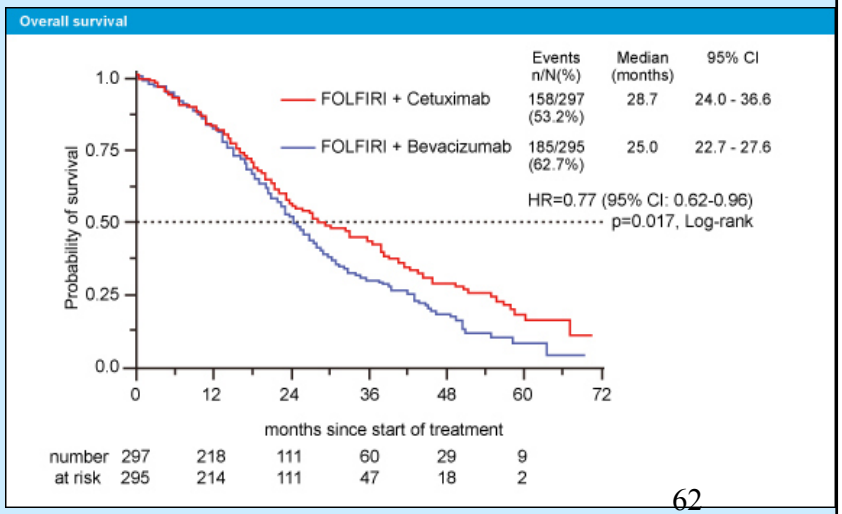
一つの説明

Cetuximabによる腫瘍縮小の方が
程度が大きい。

PDと判定されても元の大きさに戻る
あるいは致死的な大きさになるまで
には、Cetuximabの方が時間がかかる。







62

OSとPFSの乖離の理由

- これまでに言われてきたような、
  - 後治療の影響と初期治療効果の希釈
  - PFSの曖昧さとバイアス
- 効果の異なる部分集団の存在
- PFSは複合事象
  - PDイベントの発生パターンが治療群で異なる
  - イベントの種類によって死亡リスクが異なる

63

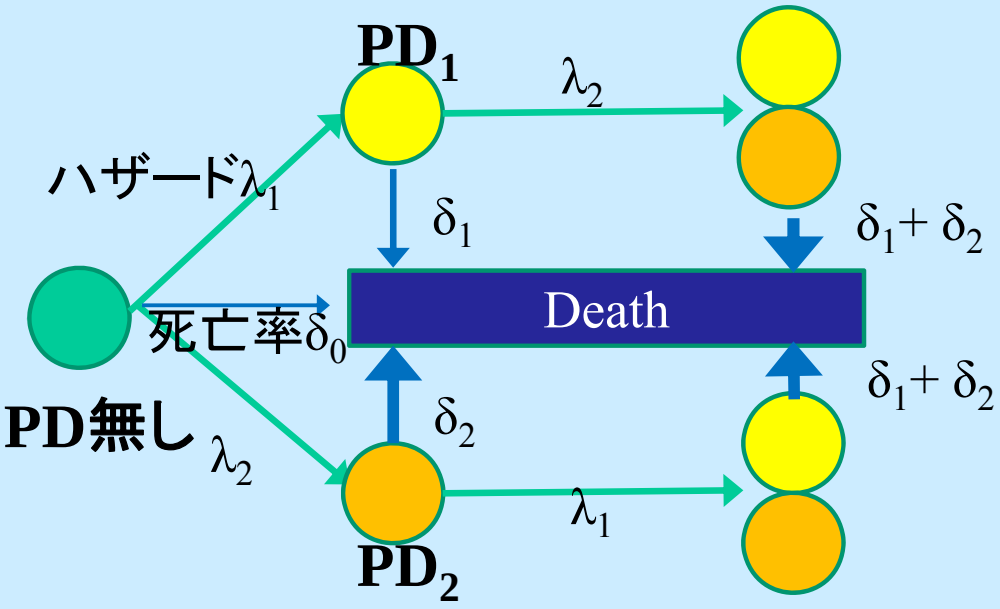
EMAガイダンスの示唆

PFSは複合事象であるので、さまざまなタイプのイベントに対する治療効果の違いを探索するために、タイプ別集計や競合危険（competing risk) 解析を行い、タイプ別の解析を行うことが適切である。

EMA: Appendix 1 to the guideline on the evaluation of anticancer medicinal products in man. EMA/CHMP/27994/2008/Rev.1, 2012, December

64

PFSとOSの乖離を説明する簡単な数学モデル



ステージ移行にはマルコフ性の仮定

$\delta_0 \sim 0$, $\delta_1 < \delta_2$

PFSは近似的に　$(\lambda_1+\lambda_2)t$　OSは　$(\delta_1\lambda_1+\delta_2\lambda_2)t^2/2$

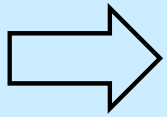
65

PFSは近似的に　$(\lambda_1+\lambda_2)t$　OSは$(\delta_1\lambda_1+\delta_2\lambda_2)t^2/2$

(本当はもっと複雑だが省略)

例

	λ_1	λ_2
治療1	3	2
治療2	2	3



δ_1　δ_2

3:1　なら

PFSのHR=1

OSの　1.22

66

Table 1. Overall survival for patients whose disease progressed due to a new metastasis or due to an increase in the size of a pre-existing lesions

	Progression due to new metastasis		Progression due to pre-existing lesion	
	Eribulin n=271	Capecitabine n=285	Eribulin n=147	Capecitabine n=129
Median OS, months (95% CI)	15.5 (14.2, 17.5)	12.9 (11.3, 14.5)	17.4 (14.4, 19.7)	17.4 (15.3, 20.9)
HR (95% CI)	0.81 (0.68, 0.97)		1.13 (0.87, 1.46)	
P-value	0.02		0.35	

CI, confidence interval; OS, overall survival

Median, HRの解釈？は困難

67

Results (cont.)

タイプによる
死亡ハザードの違い
おそらく過小評価

- Once patients were deemed to have tumour progression due to a new metastasis they were at higher risk of death (HR 2.1; 95% CI 1.8, 2.4; Wald nominal *P*<0.0001; versus tumour progression with no new metastasis, whether stratified by treatment group or not)
- There was a non-significant trend in new metastasis-free survival in favour of eribulin with a median difference of 0.6 months (HR 0.90; 95% CI 0.77, 1.05) (Figure 1)
  - data from independent review were largely consistent with investigator review, with a non-significant trend in new metastasis-free survival in favour of eribulin (median difference 0.3 months)
- The incidence of new metastases by site is shown in Table 2. The incidence of new metastasis in the CNS or lungs were lower in eribulin-treated patients compared with patients who received capecitabine. In contrast, there were more new metastases in the lymph nodes of patients in the eribulin group compared with capecitabine-treated patients
- The time to a new metastasis observed in the CNS, lungs or liver is shown in Figure 2, with a trend favouring eribulin over capecitabine

68

課題

- PFSの再検討は必要か？
- 少なくとも今後は、OSに与えるPDタイプ別の影響解析は必要
- となると、OSは当然、PD後のイベントの情報もある程度必要

- 解析方法は
  - PFSについてはタイプ別にcompeting riskとして解析
  - タイプ別の群間比較も必要（Grayの方法など）
  - タイプ別の死亡に対するリスク評価（時間依存性共変量）

PFSに替わる新しいOSの代替エンドポイントが作れるか？

69

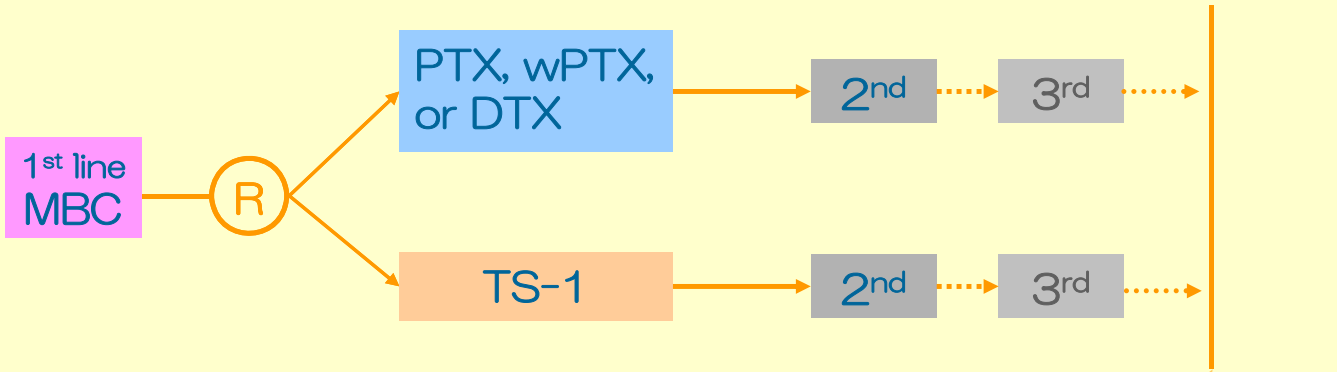
課題（続き）

- PFSに替わる新しいOSの代替エンドポイントが作れるか？
  いわれてみればEBCTCGでは遠隔転移を重視している
- このような解析に耐えるデータがあるか？

70

SELECT BC：PI　国立がん研究センター東病院　向井　博文　先生

SELection of Effective ChemoTherapy for advanced Breast Cancer



登録症例数：618例(2006.10～2010.7)
primary endpoint　：OS
secondary endpoint　：PFS、TTF、有害事象、QOL、医療経済評価

【付随研究（疫学研究）】

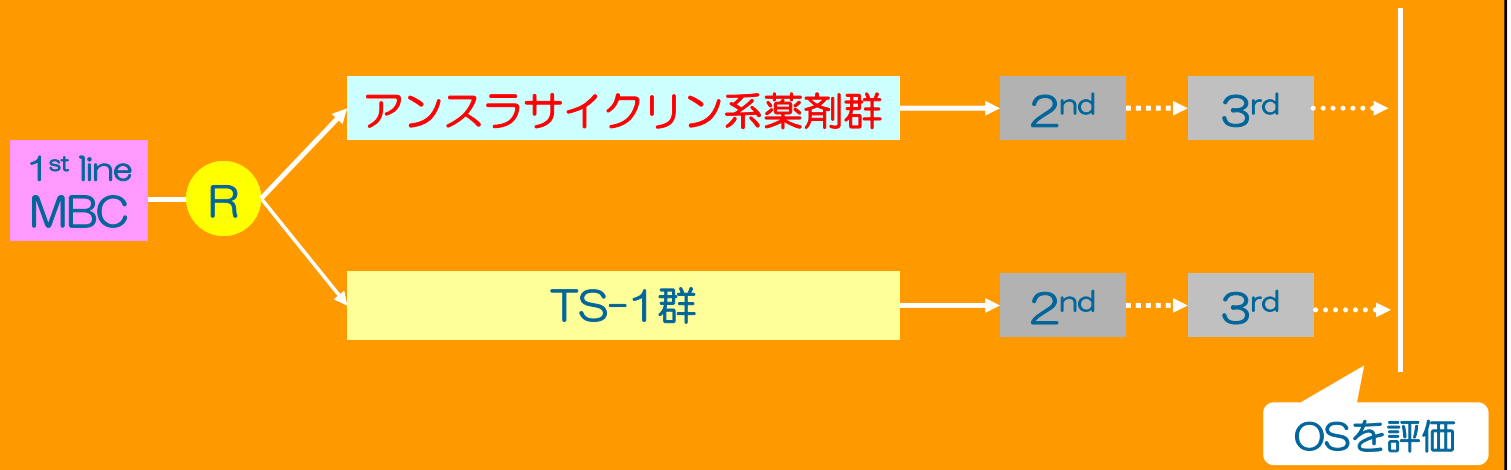
◆SELECT BC-ECO　：PI　九州がんセンター　大野　真司　先生
- ◆乳癌患者における臨床試験参加・辞退の生存予後調査

◆SELECT BC-FEEL　：PI　国立がん研究センタC東病院　石原　幹也　先生
- ◆乳癌患者に於ける臨床試験参加・辞退に影響する要因の質問紙調査

71

新規研究

SELECT BC-CONFIRM ：PI　国立がんC東病院　向井　博文　先生



目標症例数：200例
primary endpoint　：OS
secondary endpoint　：PFS、TTF、有害事象、QOL

【付随研究（疫学研究）】SELECT BCの付随研究　2研究を継続予定

◆ECO　乳癌患者における臨床試験参加・辞退の生存予後調査

◆FEEL 乳癌患者に於ける臨床試験参加・辞退に影響する要因の質問紙調査

72

課題（続き）

- PFSに替わる新しいOSの代替エンドポイントが作れるか？
  いわれてみればEBCTCGでは遠隔転移を重視している

- このような解析に耐えるデータがあるか？
  CSPOR, JCOG, others
- 今後の進行がん臨床試験のデザインは？
  OSの追跡は必須
  PD後の必要情報？
  PDのタイプ分け？
  QOL（PRO）の評価

73

Moving Beyond the Hazard Ratio in Quantifying the Between-Group Difference in Survival Analysis

Hajime Uno, Brian Claggett, Lu Tian, Eisuke Inoue, Paul Gallo, Toshio Miyata, Deborah Schrag, Masahiro Takeuchi, Yoshiaki Uyama, Lihui Zhao, Hicham Skali, Scott Solomon, Susanna Jacobus, Michael Hughes, Milton Packer, and Lee-Jen Wei

A B S T R A C T

In a longitudinal clinical study to compare two groups, the primary end point is often the time to [illegible] event (eg, disease progression or death). The hazard ratio is routinely used to empirically quantify the between-group difference under the assumption that the ratio of the two hazard functions is approximately constant over time. When this assumption is plausible, such a ratio estimate may capture the relative difference between two survival curves. However, the clinical meaning of such a ratio estimate is unclear (if not impossible) to interpret when the underlying proportional hazards assumption is violated (ie, the hazard ratio is not constant over time). Although [illegible] studied extensively and [illegible] the hazard ratio estimator have been discussed in the statistical literature, such crucial information does not seem to have reached their counterparts in oncology practice researchers. In this article, we summarize several critical concerns regarding this conventional practice and discuss various well-known alternatives for quantifying the underlying differences between two arms with respect to a time-to-event endpoint. The data from [illegible] clinical trials, [illegible] a variety of scenarios, are used throughout to illustrate our discussions. When there is not sufficient information about the profile of the between-group difference at the design stage of the study, we encourage practitioners to consider a prespecified, clinically meaningful, model-free measure for quantifying the difference and to use robust estimation procedures to draw primary inferences.

定番のハザード比による要約は適切か？

比例ハザード性が成立しなければ誤った解釈

（追跡時間によって結果が変わる）

十分な追跡・症例数が存在しても、イベント数が

少ないと、情報不足と言う誤解につながる

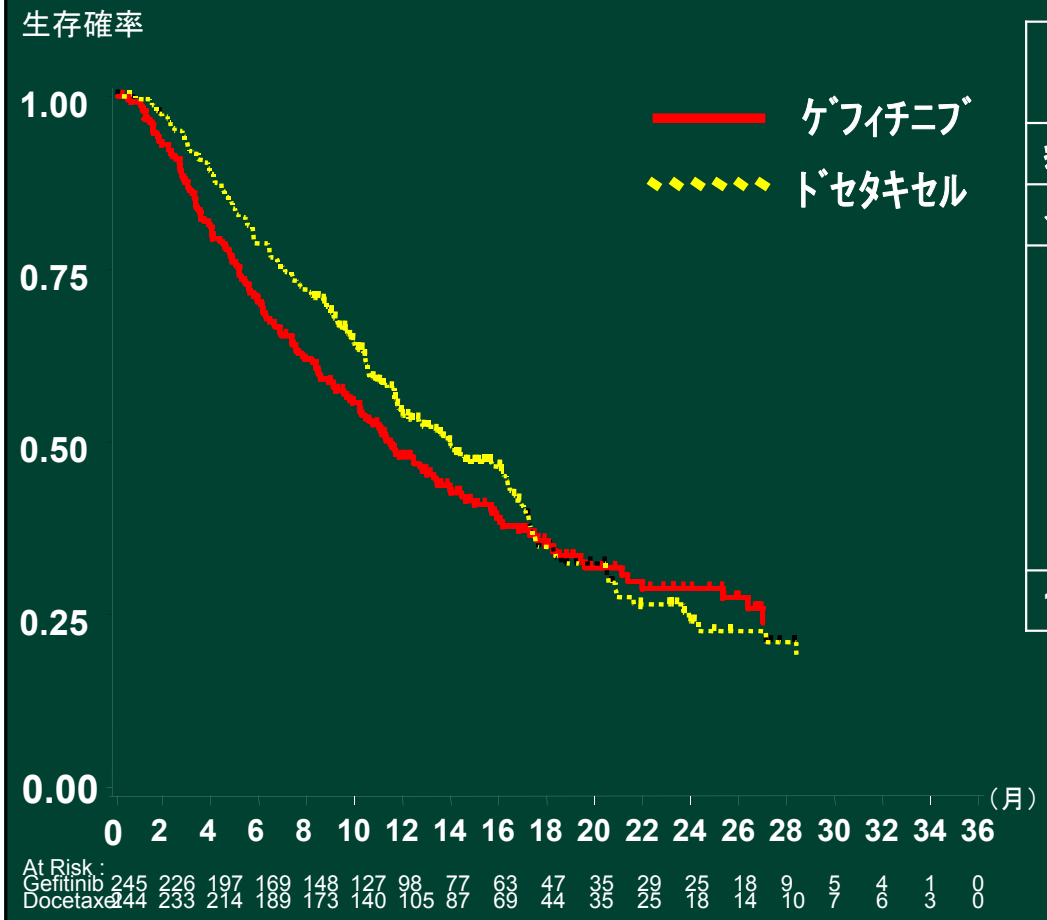
著者：宇野一（ダナファーバー）、宇山佳明氏（医薬品医療機器総合機構）、
宮田俊男氏（元厚生労働省医薬食品局審査管理課）が参加

74

イレッサ公式サイト
http://www.iressa.jp/ekigaku/eki_index.asp

75

主要評価項目ー　全生存期間（ITT*）



	ｹﾞﾌｨﾁﾆﾌﾞ割付群	ﾄﾞｾﾀｷｾﾙ割付群
症例数	245	244
イベント数	156	150
プロトコールで規定された主解析：共変量を考慮しないCox回帰分析 ハザード比（95.24%信頼区間）= 1.12（0.89, 1.40）、p=0.330 統計学的に非劣性は証明されなかった		
1 年生存率	48%	54%

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

76
13

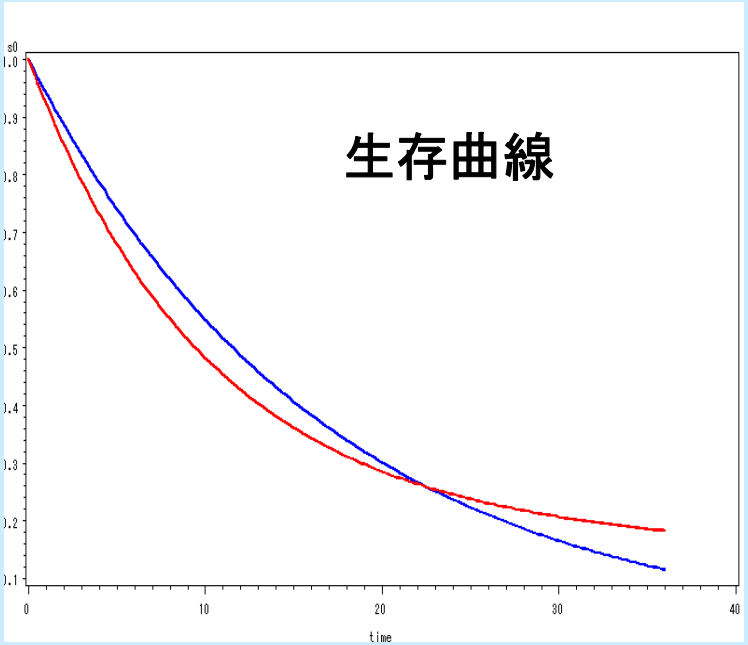
イレッサ公式サイト:結果の概要

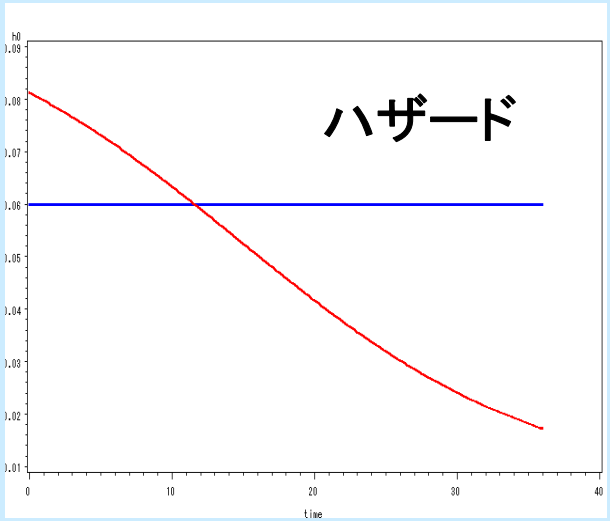
本試験の主要目的は、試験実施計画書で規定した統計学的な判定基準（Cox回帰分析を行い、ハザード比の信頼区間の上限が1.25以下であること）に従い、全生存期間におけるイレッサのドセタキセルに対する非劣性を証明することです。
この規定された解析を行った結果、主要目的であるイレッサのドセタキセルに対する非劣性の証明には至りませんでした（ハザード比 1.12、95.24%信頼区間 0.89-1.40、p= 0.330）。
なお、この解析は、全生存期間におけるドセタキセル群に対するイレッサ群の効果が時間の経過にかかわらず一定であることを前提としていますが、第三者統計専門家が行った追加の解析では、当該効果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？

77

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



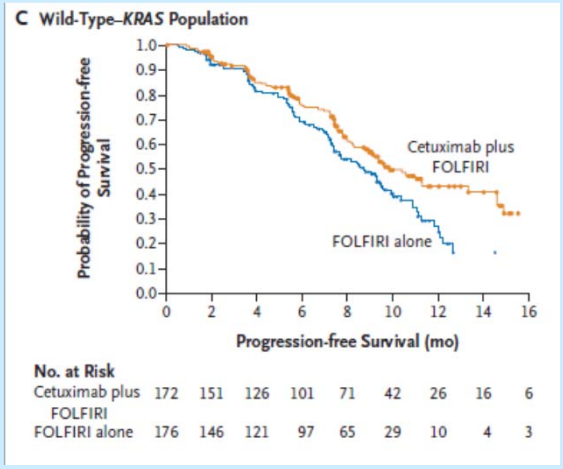


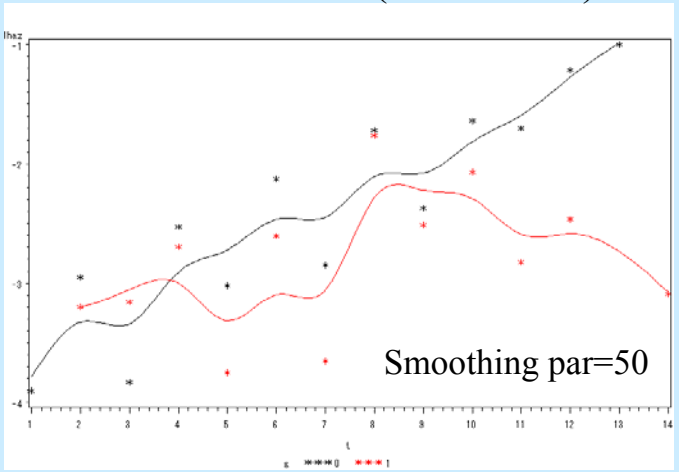
赤：奏効集団20%（MST 10倍）残り80%（MST 0.6倍）

78

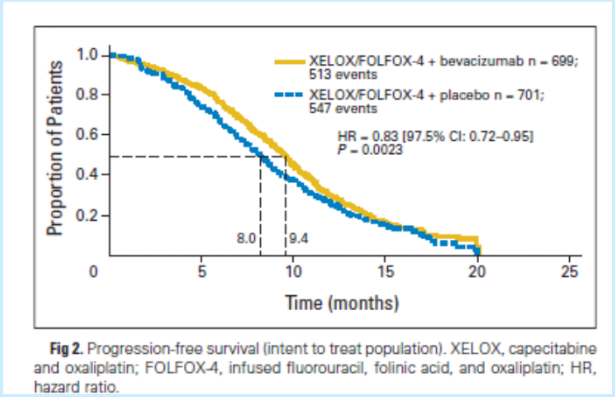
進行大腸癌PFS　ハザードの薬剤による違い

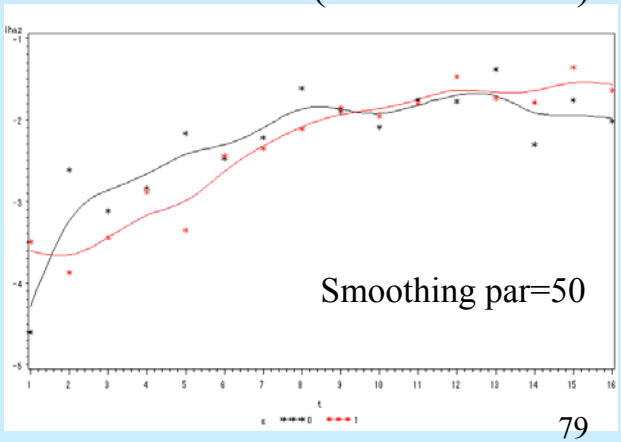
Cetuximab NEJM 2009 (CRYSTAL)





Bevacizumab JCO2008 (FOLFOXbase)





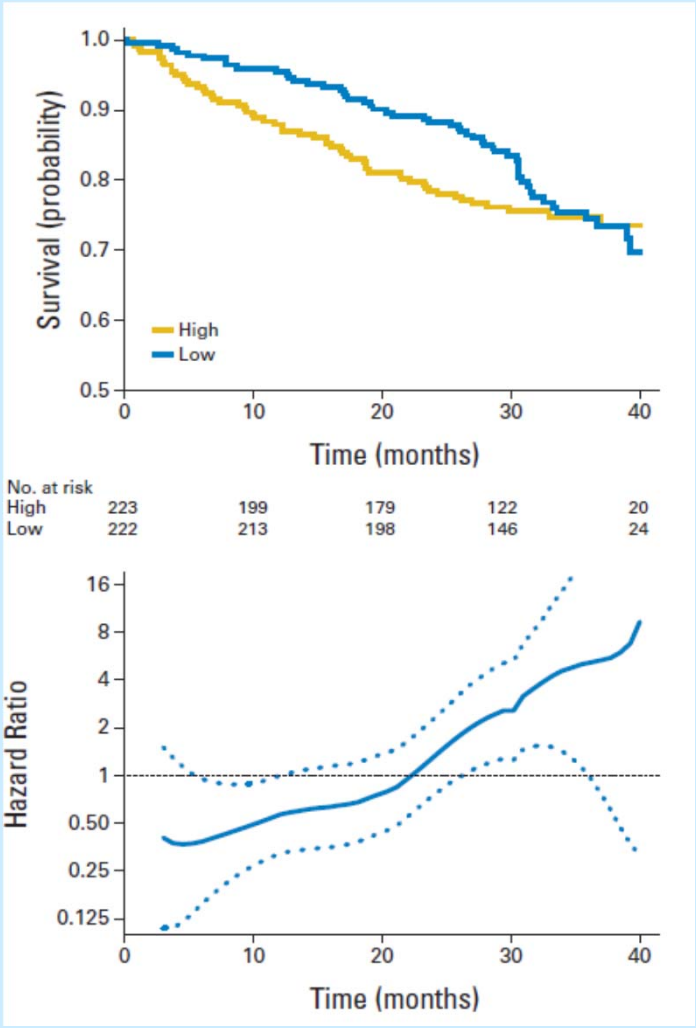
79

ハザード比にかわる要約指標

- ある定まった時点での生存確率の差、比
- メディアンなどパーセント点の差、比
- ある追跡上限までの生存時間平均、あるいはその上限からの差（失われた生存）

80

多発性骨髄腫に対するデキサメタゾンの用量比較



ハザードと時間の質的交互作用

ハザード推定値は

0.87（0.60~1.27），p=0.47

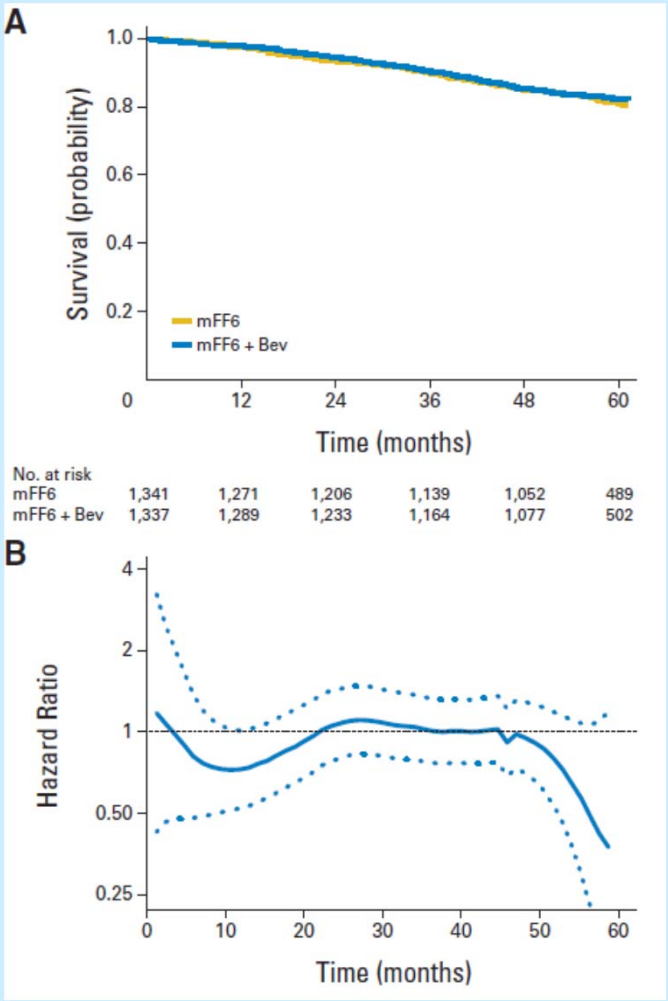
24ヶ月の生存割合の比

1.13（1.03~1.23）

10%点は20.3ヶ月と9.5ヶ月で比は有意

81

結腸癌術後補助療法：FOLFOXへのベバシズマブの上乗せ



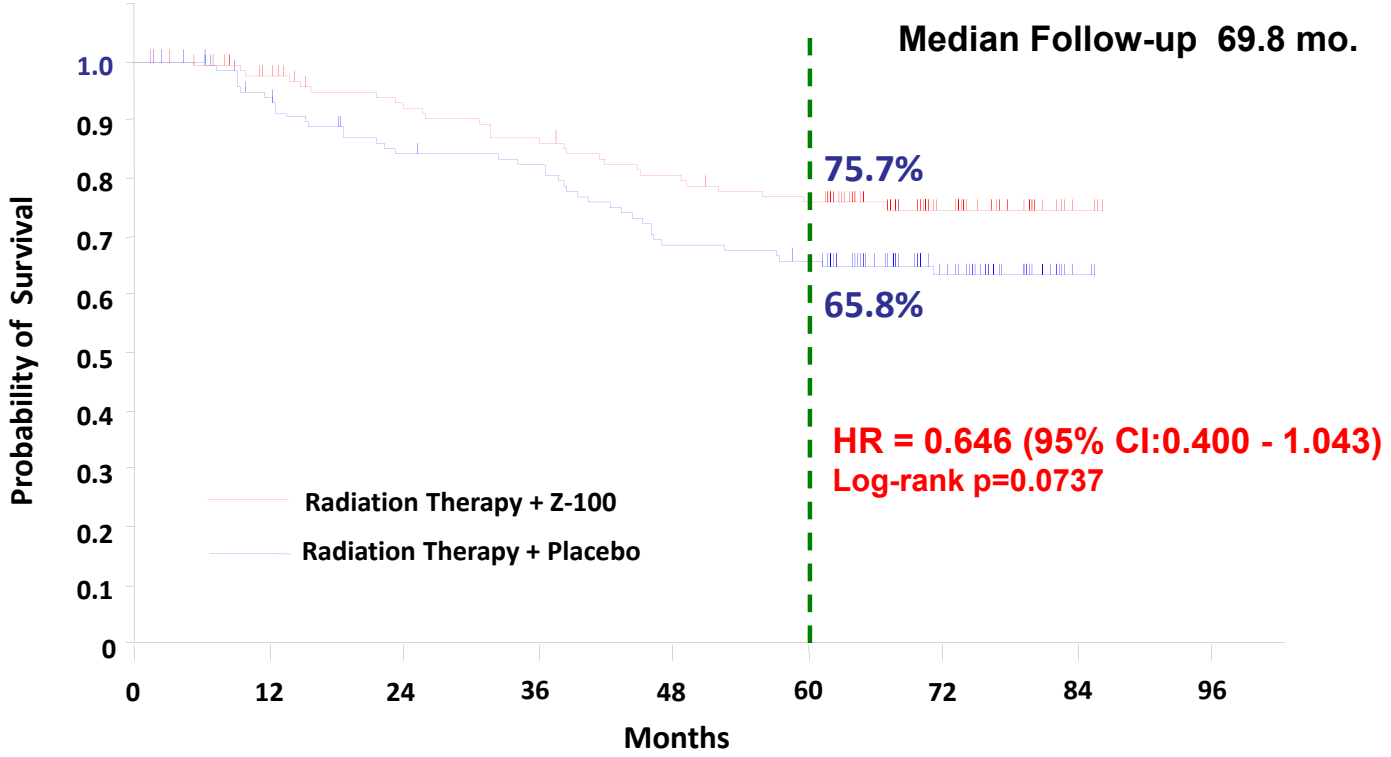
ハザード比
　0.95（0.79~1.13）, p=0.56
情報不足の解釈は正しいか？

60ヶ月上限での平均値は
　55.2ヶ月対54.9ヶ月
　差は0.3（-0.7~1.2）

82

Primary ： Overall Survival

Sugiyama et al. Ann Oncology 2014; Feb 25



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

「OSかPFSか」論争を受けての感想とまとめ

最近のがん臨床試験のデザイン

- がん領域の特徴とその試験デザインへの反映
- 新しいアプローチの必要性（ベイズ流など）
- Time-to-eventをめぐる問題（OSかPFSか）
- より効率的・患者視点の臨床試験に向けて

ASCO2014でのValue, Precision Medicineの強調

分子標的を意識した新薬の開発・免疫療法ががん領域では盛んである。試験方法論の提案とその応用例も増えている。奏効率・PFSの見直し、さらにハザード比による要約など、これまで常識と思われていた方法論の見直しさえ行なわれている。

全ての試験関係者にとって、「何のために、なぜ」という本質的な問いを常に心がけることがますます重要となろう。

84

BACK-UP

85

PRO

Basch et al. JNCI 2011; 24: 1808-10.

EDITORIALS

Use of Patient-Reported Outcomes to Improve the Predictive Accuracy of Clinician-Reported Adverse Events

Ethan Basch, Antonia Bennett, M. Catherine Pietanza

Correspondence to: Ethan Basch, MD, MSc, Department of Medicine, Health Outcomes Research Group, and Center for Health Policy and Outcomes, Memorial Sloan-Kettering Cancer Center, 307 East 63 St, New York, NY 10065 (e-mail: ebasch@mskcc.org).

Abundant research has now demonstrated that patient and clinician reports of symptoms—and particularly symptomatic toxicities (ie, adverse events) during cancer treatment—provide discrepant yet complementary data (1–3).

How can this be? Can't only the patient *or* the clinician be "right"? The more patient-centered among us might state that the patient is always right by definition because nobody (not even the most sensitive clinician) can truly know another person's subjective experience. But the more traditional among us might assert that clinicians should be considered right because they have an "objective" perspective based on experience and training, which prevents them from exaggerating or understating what they observe.

In fact, it appears that both the patient and clinician provide information of value, which when combined provides a more accurate understanding of the patient's symptoms. This finding is good news for those of us who are interested in improving the measurement of symptoms in clinical trials and practice. The optimistic NCI intergroup trial N9741 (7), in which an abundance of life-threatening gastrointestinal serious adverse events was ultimately detected (8). Therefore, availability of PRO data not only enhances the accuracy of clinician CTCAE reports but also may improve safety.

So, operationally how might this work? There are three potential approaches:

1) "Independent reporting," in which patient and clinician toxicity data are collected, analyzed, and reported completely separately from each other;
2) "Merged reporting," in which patient and clinician data are collected separately and then merged analytically into a single metric; and
3) "Collaborative reporting," in which patients directly report symptomatic toxicity information, which is then provided to clinicians to inform their CTCAE reporting.

個々の主観的経験を報告するのに最適であるのは患者、これを疾患の観点から説明するのに最適なのは医師。両者は相補的・・・、時期のCTCAEv5は症状の正確性を向上させるためPROを組み込む予定

86

San Antonio Breast Cancer Symposium 36 2013DEC10-14

87

サンアントニオ乳癌シンポジウム雑感
（2013DEC10-14)

- アナストロゾールの乳癌予防効果（S3-01)
- 検診の過剰診断（PL1)
- I-SPY2試験（adaptive　PII）初めての「卒業レジメ」（S5-02)
- 原発と転移でのゲノム情報の違い（S6-06,07)
- 予後因子としてのTIL（S1-05,06,07)と免疫治療（多数のポスター）
- 医療経済的分析（S3-04、多数のポスター）
- Survivorの社会心理的側面/QOL（PL3)

治療から予防へ、しかし検診の有用性についてはまだまだ議論
分子標的からパスウエイ抑制に、しかしがん多様性と標的自体の変化
免疫治療の時代到来か
経済的負担、そしてQOL

88

「OSかPFSか」論争を受けての感想とまとめ

最近のがん臨床試験のデザイン

- がん領域の特徴とその試験デザインへの反映
- 新しいアプローチの必要性
- Time-to-eventをめぐる問題（OSかPFSか）
- より効率的・患者視点の臨床試験に向けて

分子標的を意識した新薬の開発ががん領域では盛んである。試験方法論の提案とその応用例も増えている。奏効率・PFSの見直しなど、これまで常識と思われていた方法論の見直しさえ行なわれている。

全ての試験関係者にとって、「何のために、なぜ」という本質的な問いを常に心がけることがますます重要となろう。

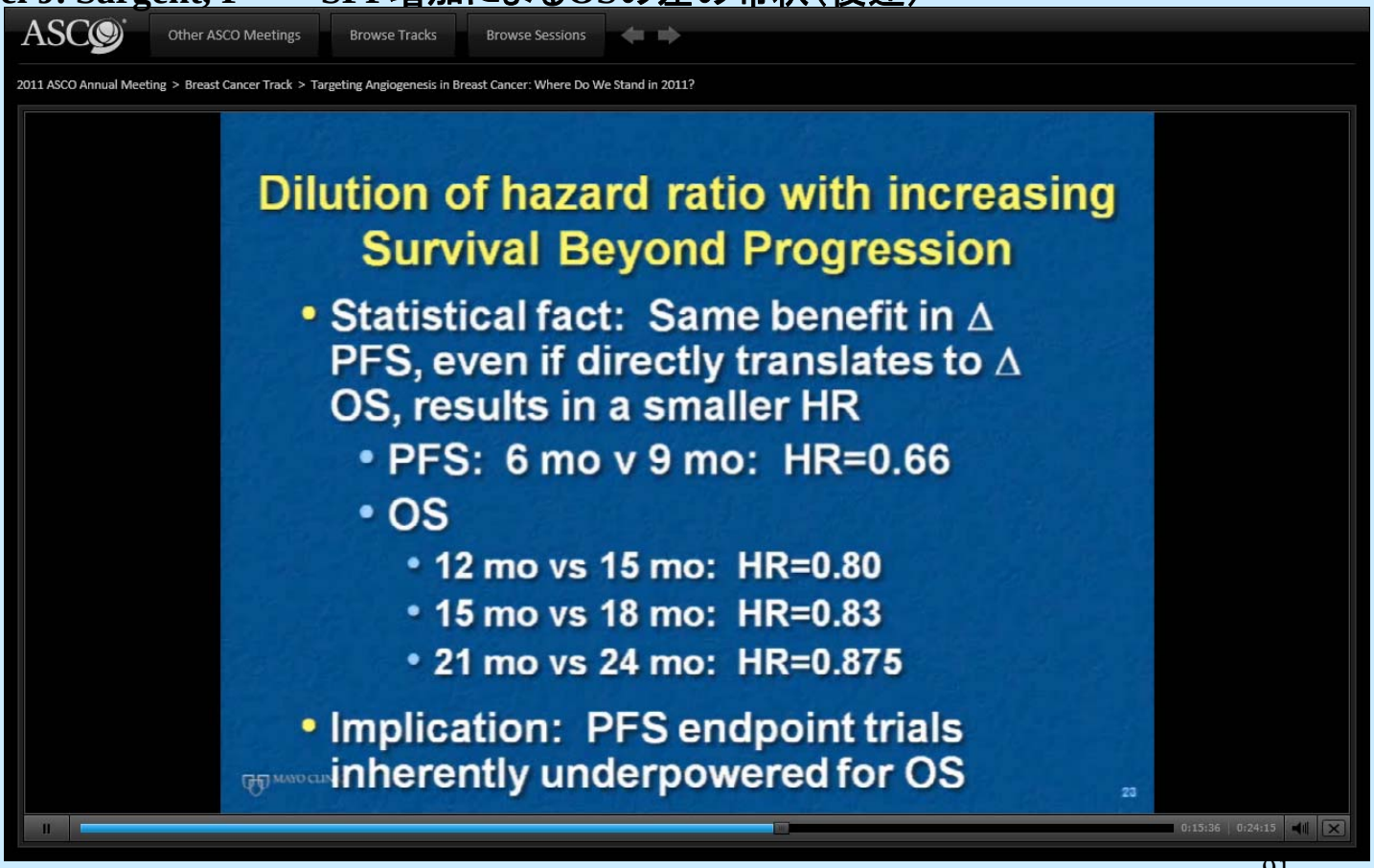
89

何が必要か？　具体的には・・・

- ヒストリカルデータの活用（症例の登録・追跡、これを可能とする施設の体制、検定バンクと標準化された手法によるマーカー測定、データ利用規定とデータ管理・解析を行う組織、病理医の協力、適切な倫理審査が必要となるが、わが国ではこのようなインフラストラクチャが脆弱である）
- 患者による毒性・QOLのPROによる評価（とくに第III相試験あるいは市販後の試験）
- 大規模試験あるいは市販後試験において試験実施の負担を軽減するための、毒性・併用治療に関するデータ収集適正化
- 新薬の承認や適応拡大・標準治療確立のために必要なデータは何か、そのためにどのような臨床試験（かつ市販後の監視）が必要かつ可能か、に関する当局あるいは統計家との協議。それを可能にする当局・統計家の能力とレギュラタリーサイエンスの発展

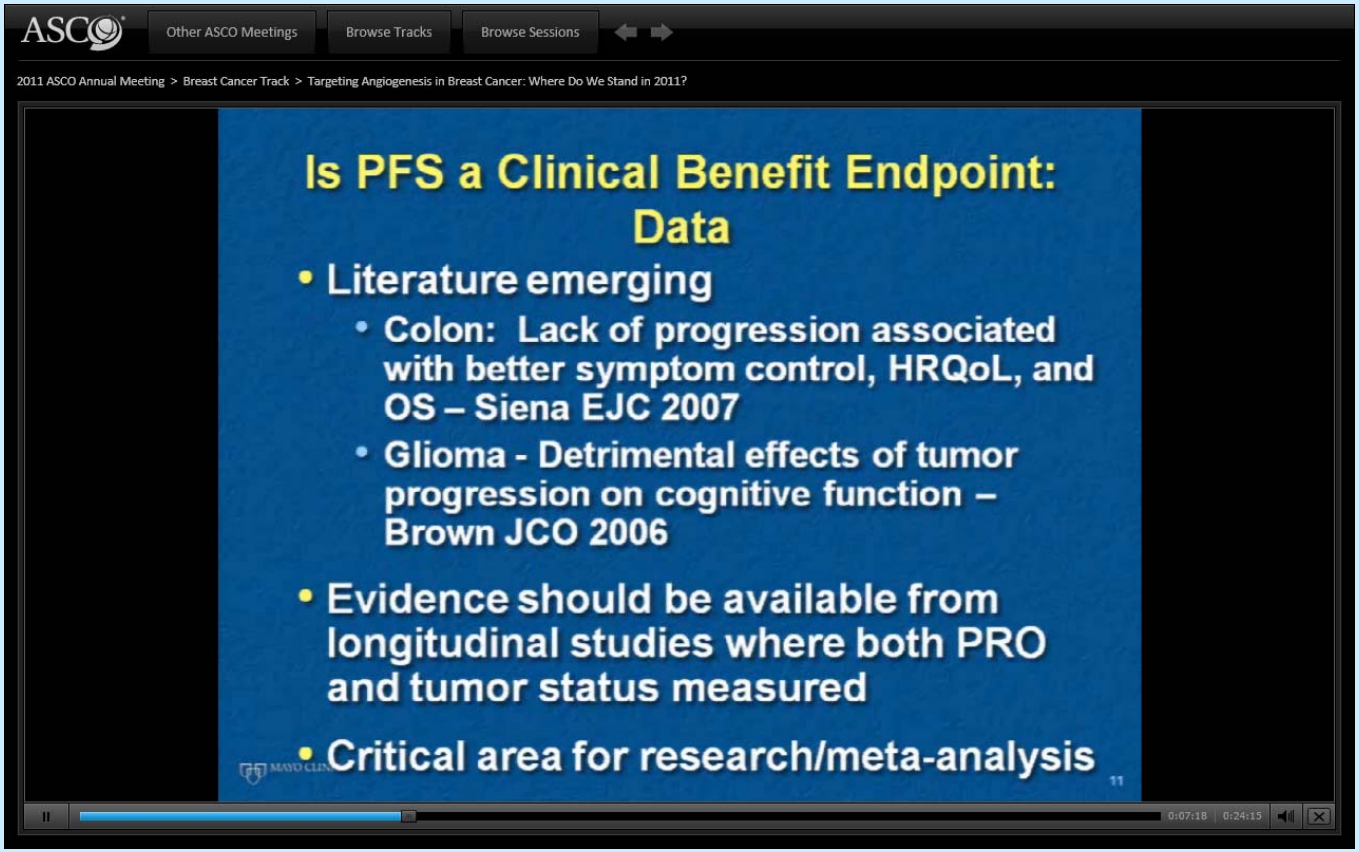
90

ASCO2011:Targeting Angiogenesis in Breast Cancer: Where Do We Stand in 2011? Appropriate Endpoints for Clinical Trials and Source of U.S. Food and Drug Administration/Oncologic Drugs Advisory Committee Controversy
Daniel J. Sargent, P　　SPP増加によるOSの差の希釈（後述）



91

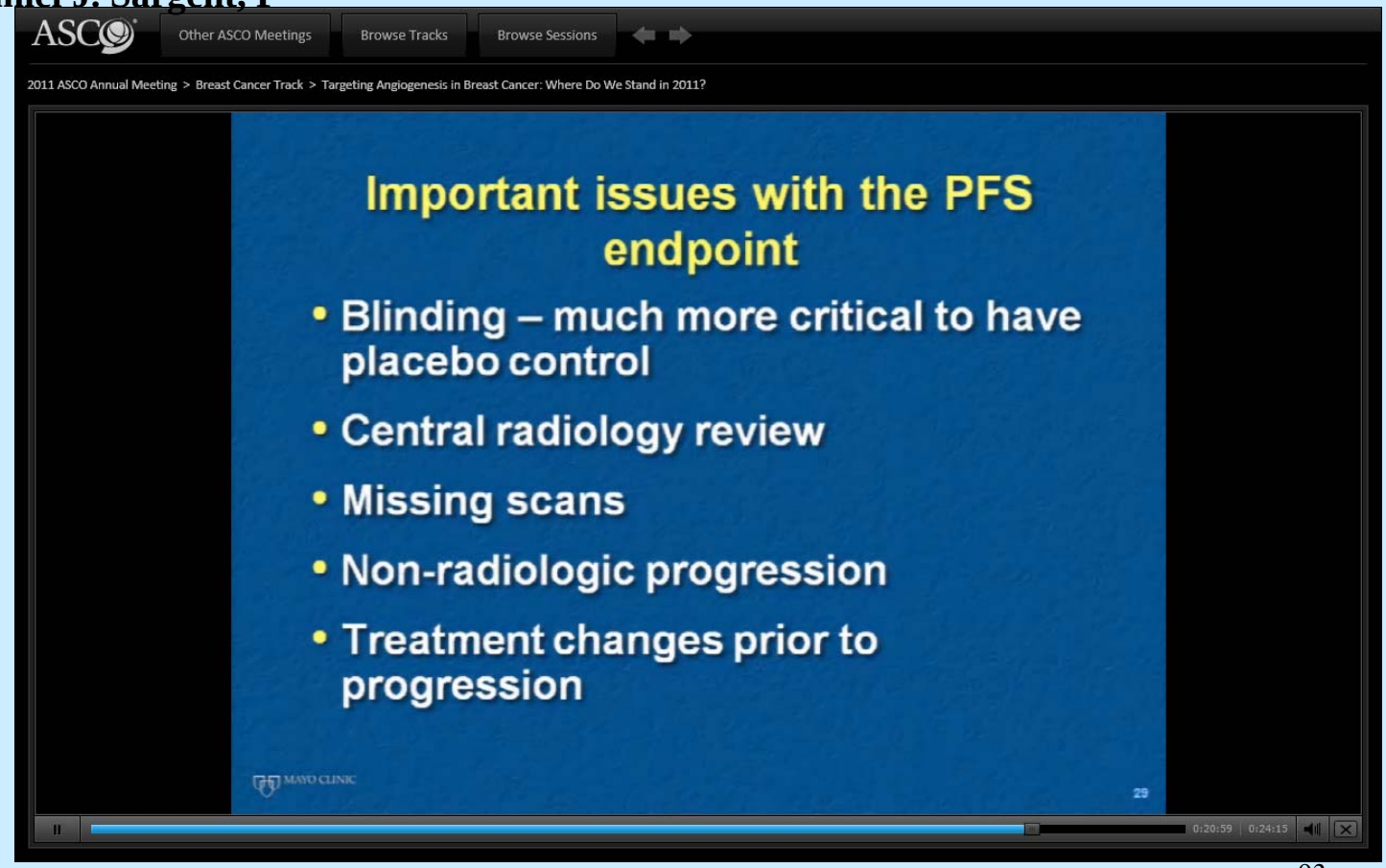
Targeting Angiogenesis in Breast Cancer: Where Do We Stand in 2011? Appropriate Endpoints for Clinical Trials and Source of U.S. Food and Drug Administration/Oncologic Drugs Advisory Committee Controversy
Daniel J. Sargent, P



92

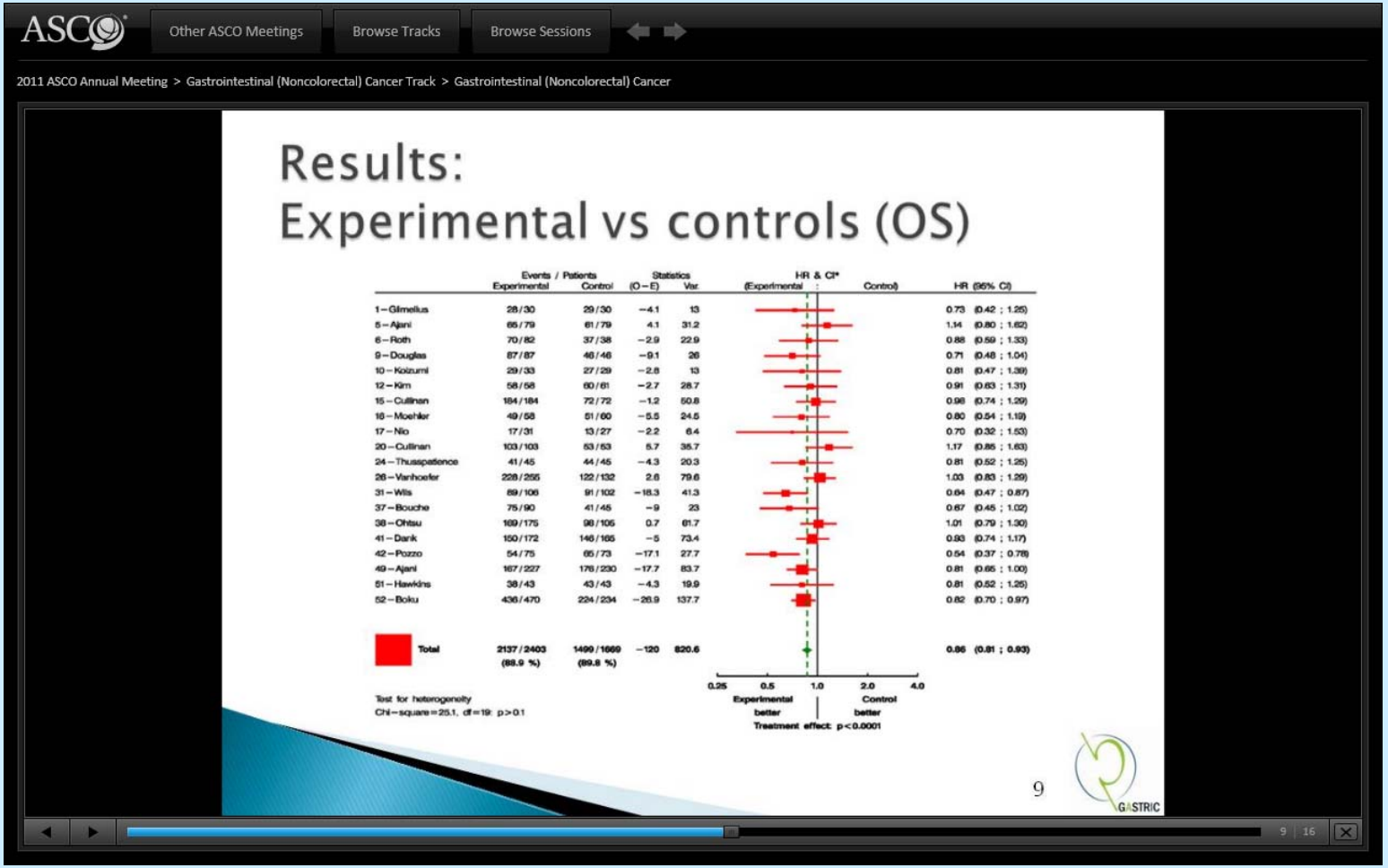
ASCO2011:Targeting Angiogenesis in Breast Cancer: Where Do We Stand in 2011? Appropriate Endpoints for Clinical Trials and Source of U.S. Food and Drug Administration/Oncologic Drugs Advisory Committee Controversy

Daniel J. Sargent, P



93

4095


ASCO
Other ASCO Meetings
Browse Tracks
Browse Sessions
2011 ASCO Annual Meeting > Gastrointestinal (Noncolorectal) Cancer Track > Gastrointestinal (Noncolorectal) Cancer
Results:
Experimental vs controls (OS)
Events / Patients
Experimental
Control
Statistics
(O – E)
Var.
HR & CI*
(Experimental : Control)
HR (95% CI)
1 – Gilmellus 28/30 29/30 −4.1 13 0.73 (0.42 ; 1.25)
5 – Ajani 65/79 61/79 4.1 31.2 1.14 (0.80 ; 1.62)
6 – Roth 70/82 37/38 −2.9 22.9 0.88 (0.59 ; 1.33)
9 – Douglas 87/87 46/46 −9.1 26 0.71 (0.48 ; 1.04)
10 – Koizumi 29/33 27/29 −2.8 13 0.81 (0.47 ; 1.39)
12 – Kim 58/58 60/61 −2.7 28.7 0.91 (0.63 ; 1.31)
15 – Cullinan 184/184 72/72 −1.2 50.8 0.96 (0.74 ; 1.29)
16 – Moehler 49/58 51/60 −5.5 24.5 0.80 (0.54 ; 1.19)
17 – Nio 17/31 13/27 −2.2 6.4 0.70 (0.32 ; 1.53)
20 – Cullinan 103/103 53/53 5.7 35.7 1.17 (0.85 ; 1.63)
24 – Thusspatience 41/45 44/45 −4.3 20.3 0.81 (0.52 ; 1.25)
26 – Vanhoefer 228/255 122/132 2.6 79.6 1.03 (0.83 ; 1.29)
31 – Wils 89/106 91/102 −18.3 41.3 0.64 (0.47 ; 0.87)
37 – Bouche 75/90 41/45 −9 23 0.67 (0.45 ; 1.02)
38 – Ohtsu 169/175 98/105 0.7 61.7 1.01 (0.79 ; 1.30)
41 – Dank 150/172 146/165 −5 73.4 0.93 (0.74 ; 1.17)
42 – Pozzo 54/75 65/73 −17.1 27.7 0.54 (0.37 ; 0.78)
49 – Ajani 167/227 176/230 −17.7 83.7 0.81 (0.65 ; 1.00)
51 – Hawkins 38/43 43/43 −4.3 19.9 0.81 (0.52 ; 1.25)
52 – Boku 436/470 224/234 −26.9 137.7 0.82 (0.70 ; 0.97)
Total 2137/2403 (88.9 %) 1499/1669 (89.8 %) −120 820.6 0.86 (0.81 ; 0.93)
Test for heterogeneity
Chi−square=25.1, df=19 p>0.1
0.25 0.5 1.0 2.0 4.0
Experimental better
Control better
Treatment effect: p<0.0001
9
GASTRIC
9 16

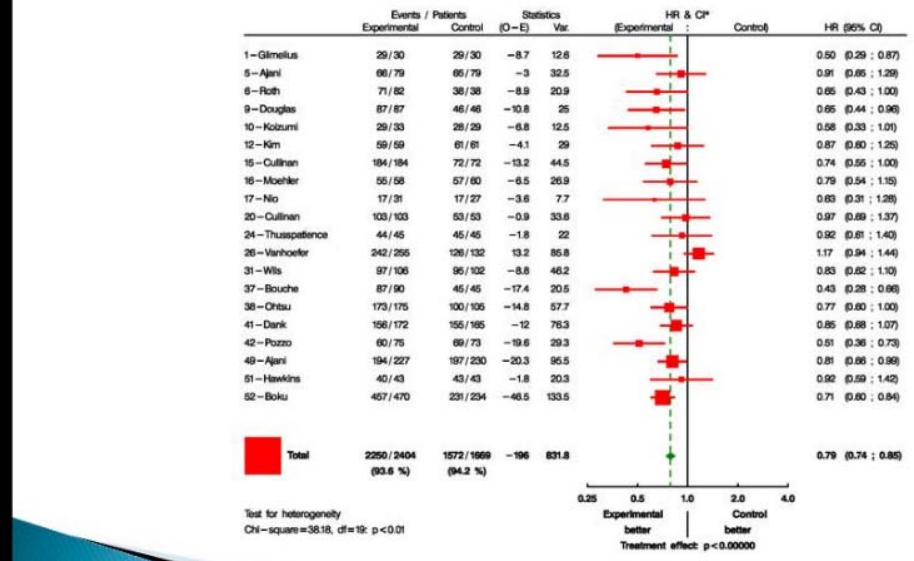
94

4095

ASCO | Other ASCO Meetings | Browse Tracks | Browse Sessions

2011 ASCO Annual Meeting > Gastrointestinal (Noncolorectal) Cancer Track > Gastrointestinal (Noncolorectal) Cancer

Results: Experimental vs controls (PFS)



10

GASTRIC

10 | 16

95

それではPFS（Progresson Free Survival）はreliabileか？

- 進行・再発の乳癌、卵巣癌、NSCLC、大腸癌*では次第にPFSがプライマリエンドポイントになりつつあるが・・

- どこを「時点」とするのか？　検査オーダー日、診断日？
- 測定間隔はどう設定するか？
- 測定手段はどこまでを必須とするか？　臨床症状は？
- 予定していない（患者の訴え等による）検査の扱いは？
- 主治医判定と中央判定、判定のQCは？

*Yothers JCO 2007; 25: 5153-4.　進行大腸癌に対する薬物療法の目的は治癒、患者負担の軽減、症状の安定化と進行の抑制であり、単なる生存の延長は二義的な意味しか持たない。よってPFSはOSの代替エンドポイントというより、それ自体で臨床的意義を有するプライマリエンドポイントになりうる。OSの測定は明確かつ信頼性は高いものの、セカンダリー以降の治療の影響を被る。一方、PFSの測定には、画像評価の時期設定と判定など、バイアスや誤差の入りうる余地がOSより大きい。その一方で（同じ時点で比較すれば）情報量は多く、かつセカンダリー以降の治療の影響を受けることはない。今後PFSをプライマリエンドポイントに採用する試験が増えることを予想。

Difference between central review and local evaluation

- 施設判定と中央判定結果の食い違いは大きいもの（不一致率は23-36%）の、有効性に関する試験結果は両者で食い違わない

Table 1. Trials That Have Used Retrospective Blinded Independent Central Reviews

Trial	Sample Size	Discrepancy Rate Between Central and Local Progression Times	Per Central Review HR	Per Central Review 95% CI	Per Local Review HR	Per Local Review 95% CI	P
Renal cell carcinoma							
Sorafenib *v* placebo*[17]	903	NR	0.44	0.35 to 0.55	0.51	0.43 to 0.60	
Sunitinib *v* interferon alpha[18]	750	NR	0.42	0.32 to 0.54	0.42	0.33 to 0.52	
Colorectal cancer							
Panitumumab plus best supportive care *v* best supportive care[19]	463	NR	0.54	0.44 to 0.66	0.39	0.32 to 0.48	
Breast cancer							
Lapatinib plus capecitabine *v* capecitabine[20]	324	Lapatinib plus capecitabine, 40 of 163 = 25%; capecitabine, 40 of 161 = 25%	0.49	0.34 to 0.71	0.59	0.42 to 0.84	
Bevacizumab plus capecitabine *v* capecitabine[21]	462	105 of 462 = 23%†	0.98	0.77 to 1.25	NR, value not significantly different from 1		
Bevacizumab plus paclitaxel *v* paclitaxel[10,11]	722	232 of 649 = 36%‡	0.42	0.34 to 0.52	0.48	0.39 to 0.61	
Ixabeplone plus capecitabine *v* capecitabine,[22] months	752	NR	NR		NR		
Median PFS			5.8 *v* 4.2		5.3 *v* 3.8		< .0011
95% CI			5.45 to 6.97 *v* 3.91 to 4.50		NR		

NOTE. We reviewed the literature and searched PubMed for studies in breast cancer, colorectal cancer, lung cancer and renal cell carcinoma. Search terms included, "progression- free survival" or "time to progression," with filters of "randomized controlled trial" and "published in last five years." This revealed 209 manuscripts, of which only six reported having a central review of progression. The bevacizumab plus paclitaxel trial in breast cancer (last row) was included separately because it generated much discussion during an US Food and Drug Administration Oncologic Drug Advisory Committee meeting on December 5, 2007. All of these trials implemented a retrospective blinded independent central review (option 2). The panitumimab trial[11] allowed cross-over at the time of locally determined progression among patients receiving the control treatment. As a result, patients for whom progression was not confirmed centrally continued to be evaluated centrally for progression.

Abbreviations: HR, hazard ratio; NR, not reported, PFS, progression-free survival.

*Double-blinded trial.

†Seventy-nine of the 105 discrepancies were investigator to determined progressions unconfirmed by central review. The investigators state that the investigator progressions were most commonly called by physical exam. The remaining 26 progressions were called by the independent review before the locally determined time.

‡The discrepancy rate was estimated amongst the 649 patients for whom images were available for central review. Due to limited data availability, an agreement was counted if dates were within 6 weeks of one another.

Difference between central review and local evaluation

- 珍しく施設判定の進行後にも追跡が遂行されたベバシズマブの腎癌に対する2重盲検試験の例
- 116患者を対象としたプラセボ、低用量、高用量の3群試験
- 42例で判定が不一致であるが大半（33/42）が1時点の食い違い
- ハザード比に推定値間に本質的な違いは生じていない

Table 2. Agreement/Disagreement in Progression Times by Central and Local Reviewers from Yang et al[13]

Progression Time	No. of Patients		
	Placebo	Low Dose	High Dose
Local progression called more than 60 days after central	3	0	3
Local progression called 1 to 60* days after central review	6	6	8
Local and central progression times equal	27	22	16
Central progression called 1 to 60 days after local review	2	5	6
Central progression called more than 60 days after local	2	0	1
Total	40	33	34

*After a first evaluation, patients were evaluated every 2 months (approximately 60 days) for progression the first year then every 3 months thereafter.

Table 3. Hazard Ratios from Yang et al[13]

	Low Dose *v* Placebo		High Dose *v* Placebo	
	HR for PFS	95% CI	HR for PFS	95% CI
Central review	0.68	0.42 to 1.1	0.37	0.22 to 0.67
Local review	0.82	0.51 to 1.31	0.47	0.29 to 0.77
Central review with censoring at unconfirmed local progression*	0.68	0.40 to 1.14	0.38	0.21 to 0.68

NOTE. The original manuscript reports hazard ratios (HRs) based on adjudicated progression times and reports time to progression, not progression-free survival (PFS).
*Subjects for whom local progressions were called before the centrally determined time were artificially censored.

98

進行乳癌におけるSPP （レビュー）

Overall Survival and Post-Progression Survival in Advanced Breast Cancer: A Review of Recent Randomized Clinical Trials

Everardo D. Saad, Artur Katz, and Marc Buyse

A B S T R A C T

With the availability of several lines of therapy, overall survival (OS) has been progressively substituted by progression-free survival (PFS) and other tumor-based assessments as the primary efficacy end point in advanced breast cancer trials. We investigated the frequency and determinants of OS gain in the recent literature and the duration of post-progression survival (PPS) according to treatment type and line. We used PubMed to search for phase III trials on systemic antineoplastic therapies published between January 1998 and December 2007 in 11 leading journals. The primary end point was the one stated explicitly, used for N calculation, or listed first. Significant gain was considered as reported $P < .05$ for superiority trials or proven non-inferiority or equivalence otherwise. We retrieved 76 trials, and gain in OS was reported in 15 cases (19.7%). The median gain in OS was 4.7 months, and such gain was more frequent when there was significant gain in PFS and in second-line and third-line trials. The average median OS was 20.7 months in trials assessing first-line chemotherapy and 31.1 months with first-line hormone therapy. The median proportion of OS accounted for by PPS was significantly longer in hormone therapy trials than in chemotherapy trials, but varied little across treatment lines. A statistically significant gain in OS has been reported in about one in five recent phase III trials in advanced breast cancer, despite the fact that OS has seldom been used as the primary end point. PPS represents nearly two thirds of patient survival after on-trial disease progression.

J Clin Oncol 28:1958-1962. © 2010 by American Society of Clinical Oncology

99

進行乳癌におけるSPP （レビュー）

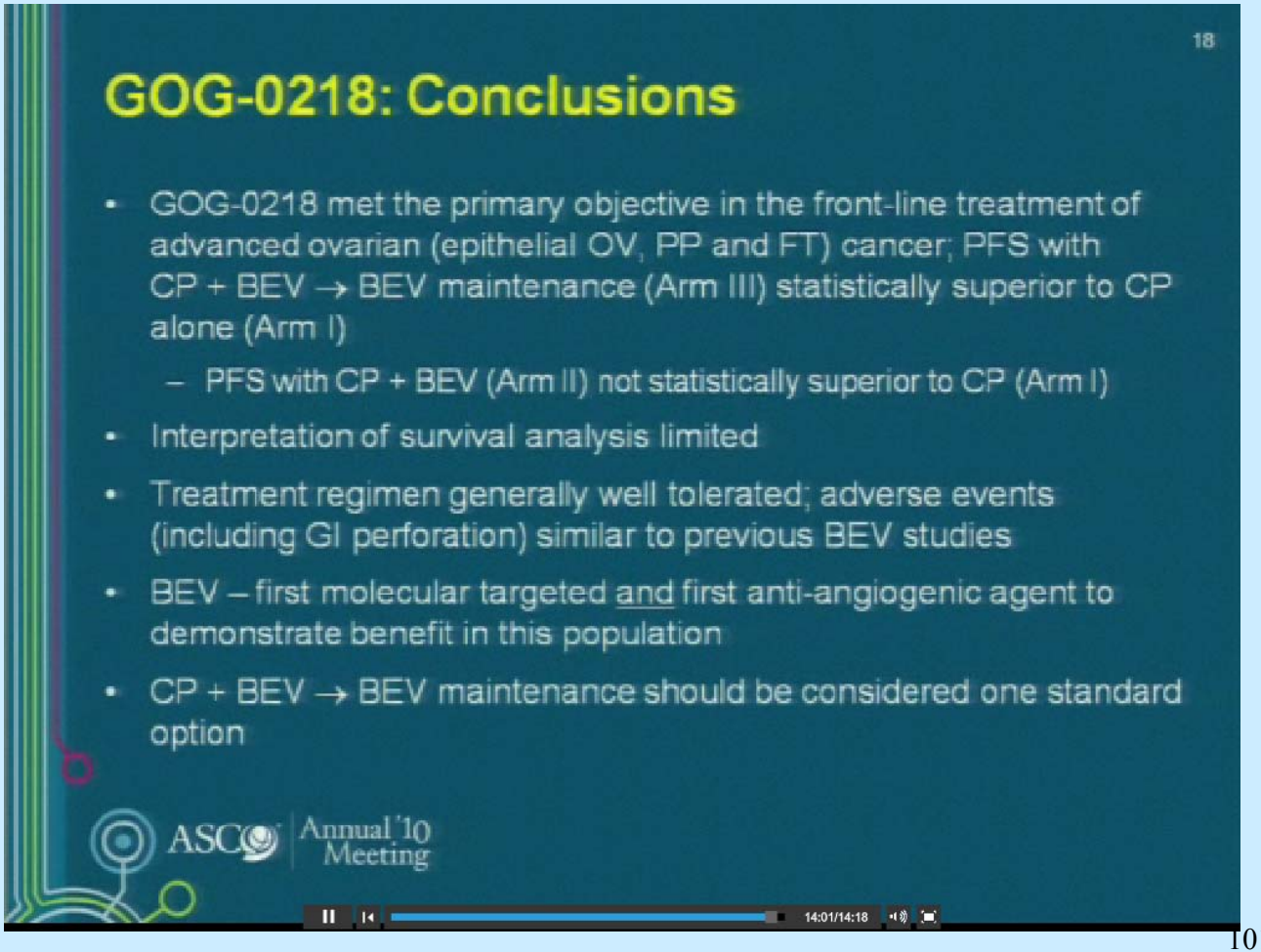
Table 3. Average Median PFS, OS, and PPS, and the Median Proportion of OS Accounted for by PPS, for Trial Arms Overall and According to Treatment Type and Line

Trial Type	No. of Arms	Average Median (months)			Median Proportion of OS Accounted for by PPS (%)
		PFS	OS	PPS	
Overall	127	6.9	20.5	13.6	65.7
Chemotherapy ± targeted therapy	96	7.0	18.3	11.4	63.7
Only first-line	59	7.8	20.7	12.9	64.4
Only second-line	14	6.7	15.2	8.5	59.8
Other (mixed lines)	23	5.0	14.1	9.1	64.4
Hormone therapy ± chemotherapy	29	6.9	27.4	20.5	74.4
Only first-line	13	9.3	31.1	21.8	69.0
Only second-line	13	5.0	23.2	18.1	77.9

Abbreviations: PFS, progression-free survival; OS, overall survival; PPS, post-progression survival.

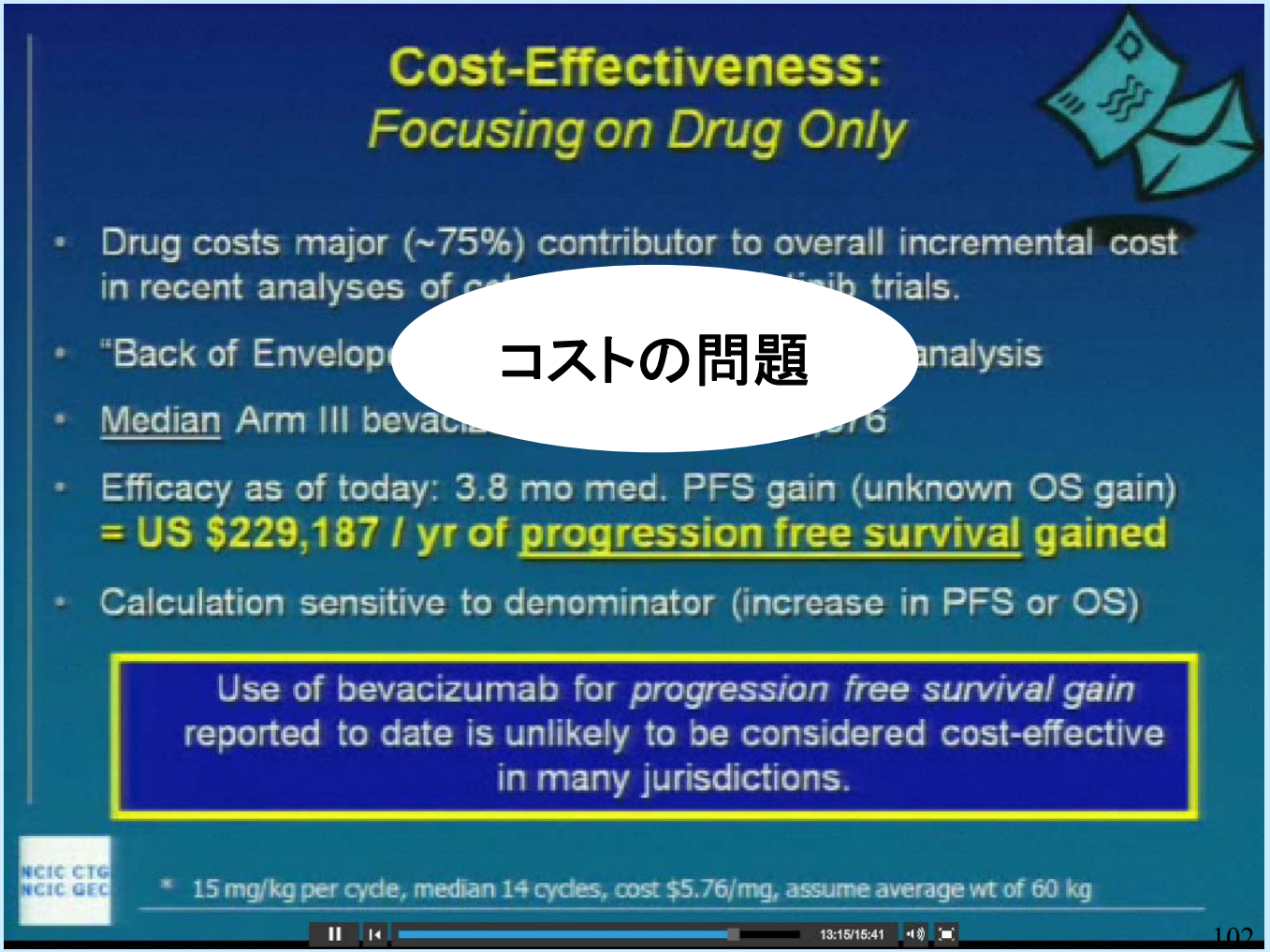
100

ASCO2010　LBA1　Burger et al.



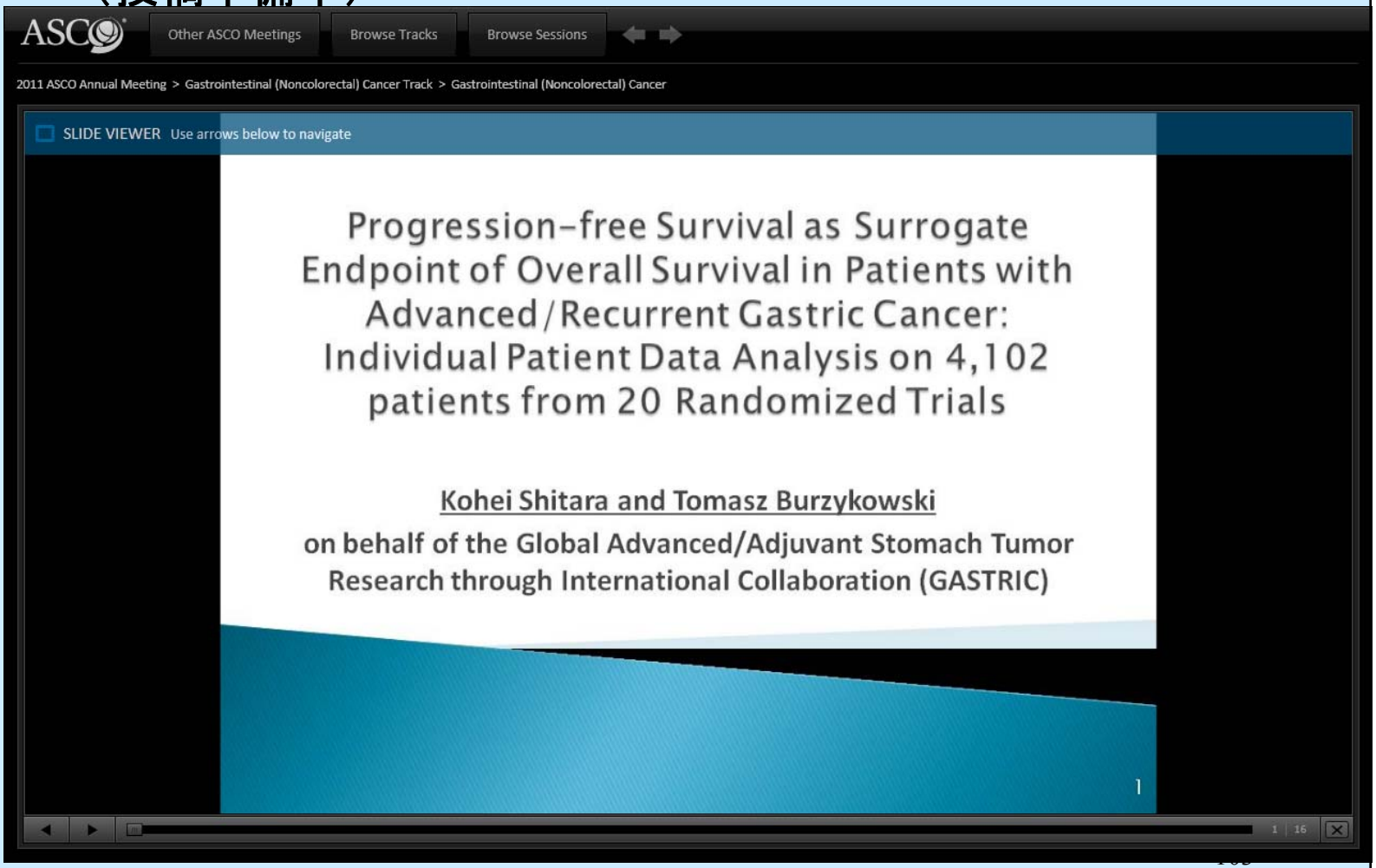
101

ASCO2010　LBA1　Discussion to Burger et al.

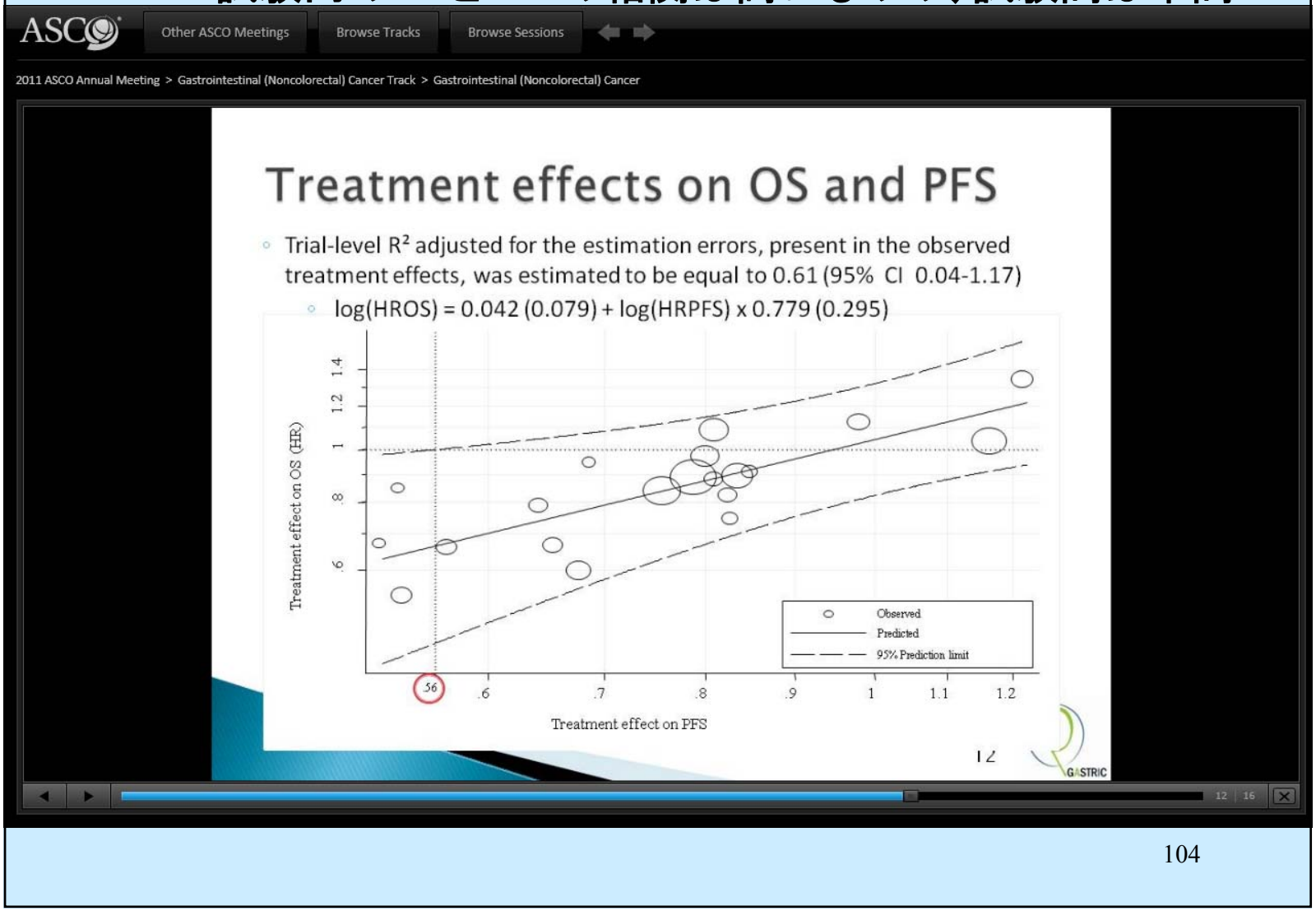


コストの問題

ASCO2011- 4095　GASTRICの胃癌メタアナリシスグループ（投稿準備中）

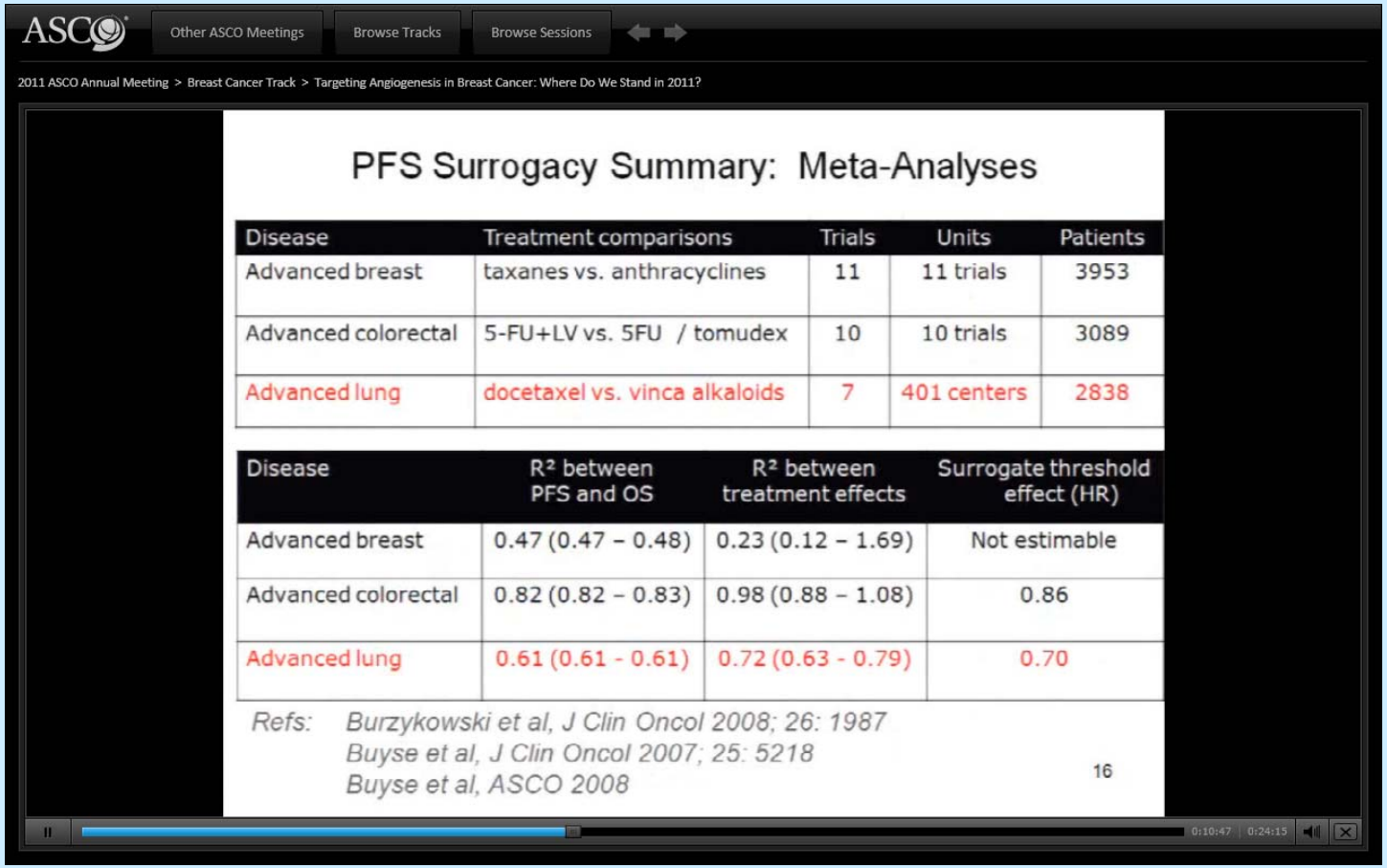


4095　試験内のOSとPFSの相関は高いものの、試験間は中間

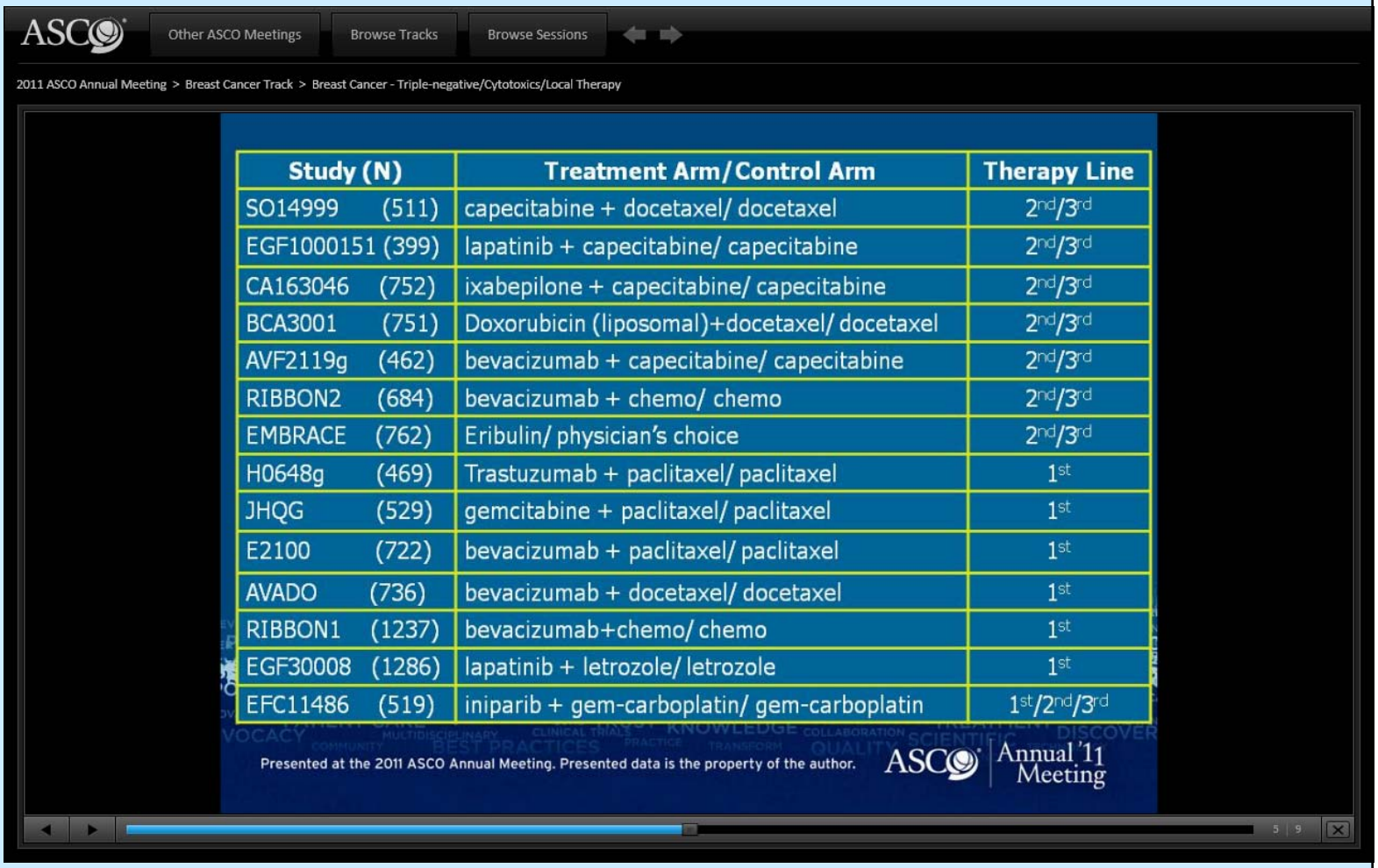


104

ASCO2011:Targeting Angiogenesis in Breast Cancer: Where Do We Stand in 2011? Appropriate Endpoints for Clinical Trials and Source of U.S. Food and Drug Administration/Oncologic Drugs Advisory Committee Controversy
Daniel J. Sargent, P



1035



Study (N)	Treatment Arm/Control Arm	Therapy Line
SO14999 (511)	capecitabine + docetaxel/ docetaxel	2nd/3rd
EGF1000151 (399)	lapatinib + capecitabine/ capecitabine	2nd/3rd
CA163046 (752)	ixabepilone + capecitabine/ capecitabine	2nd/3rd
BCA3001 (751)	Doxorubicin (liposomal)+docetaxel/ docetaxel	2nd/3rd
AVF2119g (462)	bevacizumab + capecitabine/ capecitabine	2nd/3rd
RIBBON2 (684)	bevacizumab + chemo/ chemo	2nd/3rd
EMBRACE (762)	Eribulin/ physician's choice	2nd/3rd
H0648g (469)	Trastuzumab + paclitaxel/ paclitaxel	1st
JHQG (529)	gemcitabine + paclitaxel/ paclitaxel	1st
E2100 (722)	bevacizumab + paclitaxel/ paclitaxel	1st
AVADO (736)	bevacizumab + docetaxel/ docetaxel	1st
RIBBON1 (1237)	bevacizumab+chemo/ chemo	1st
EGF30008 (1286)	lapatinib + letrozole/ letrozole	1st
EFC11486 (519)	iniparib + gem-carboplatin/ gem-carboplatin	1st/2nd/3rd

106

PFSのブラインド下での中央判定は必要か?

Dodd et al. JCO 2008; 3791-6

VOLUME 26 · NUMBER 22 · AUGUST 1 2008

JOURNAL OF CLINICAL ONCOLOGY　STATISTICS IN ONCOLOGY

Blinded Independent Central Review of Progression-Free Survival in Phase III Clinical Trials: Important Design Element or Unnecessary Expense?

Lori E. Dodd, Edward L. Korn, Boris Freidlin, C. Carl Jaffe, Lawrence V. Rubinstein, Janet Dancey, and Margaret M. Mooney

From the Branches of Biometric Research, Investigational Drug, Cancer Investigations, and Diagnostic Imaging, Division of Cancer Treatment and Diagnosis, National Cancer Institute, Rockville, MD.

Submitted January 14, 2008; accepted May 21, 2008.

Authors' disclosures of potential conflicts of interest and author contributions are found at the end of this article.

Corresponding author: Lori Dodd, PhD, National Cancer Institute, 6130 Executive Blvd, MSC 743, EPN Rm 8140, Rockville, MD 20892; e-mail: doddl@mail.nih.gov.

Published by the American Society of Clinical Oncology

0732-183X/08/2622-3791/$20.00

A B S T R A C T

Progression-free survival is an important end point in advanced disease settings. Blinded independent central review (BICR) of progression in randomized clinical trials has been advocated to control bias that might result from errors in progression assessments. However, although BICR lessens some potential biases, it does not remove all biases from evaluations of treatment effectiveness. In fact, as typically conducted, BICRs may introduce bias because of informative censoring, which results from having to censor unconfirmed locally determined progressions. In this article, we discuss the rationale for BICR and different ways of implementing independent review. We discuss the limitations of these approaches and review published trials that report implementing BICR. We demonstrate the existence of informative censoring using data from a randomized phase II trial. We conclude that double-blinded trials with consistent application of measurement criteria are the best means of ensuring unbiased trial results. When such designs are not practical, BICR is not recommended as a general strategy for reducing bias. However, BICR may be useful as an auditing tool to assess the reliability of marginally positive results.

J Clin Oncol 26:3791-3796. Published by the American Society of Clinical Oncology

PFSのブラインド下での中央判定は必要か?
結論

- PFSが望ましいエンドポイントである場合、2重盲検試験がバイアス最小化のための最良の方法
- これが不可能な場合、バイアス軽減の一般的な戦略としてはブラインド中央判定を推奨できない
- 施設での進行判定後の追加画像撮影はinformative censoringの問題を軽減する点で推奨できるが、実装は難しいであろう。BICRを測定バラツキ軽減の手段とすることも考えられるが、informative censoringバイアスとのバランスを考慮すべき
- BICRを最終的な解析手法として推奨しないものの、施設判定のバイアスをチェック(audit)する仕組みとしては有効であり、臨床的有効性の観点からはぎりぎりの試験結果の信憑性を高めることにはつながる
- そもそもPFSをエンドポイントとする臨床試験は、臨床的にも重要な意義のある、大きな治療効果を目指すべきである。このような状況下では、本稿で議論したようなバイアスに対し結論は頑健であろう

Informative censoring:ローカルで進行と判断され中央で進行せずとされた例を打ち切り扱いにすると、生存率を上げる方にバイアス

108

Informative censoring

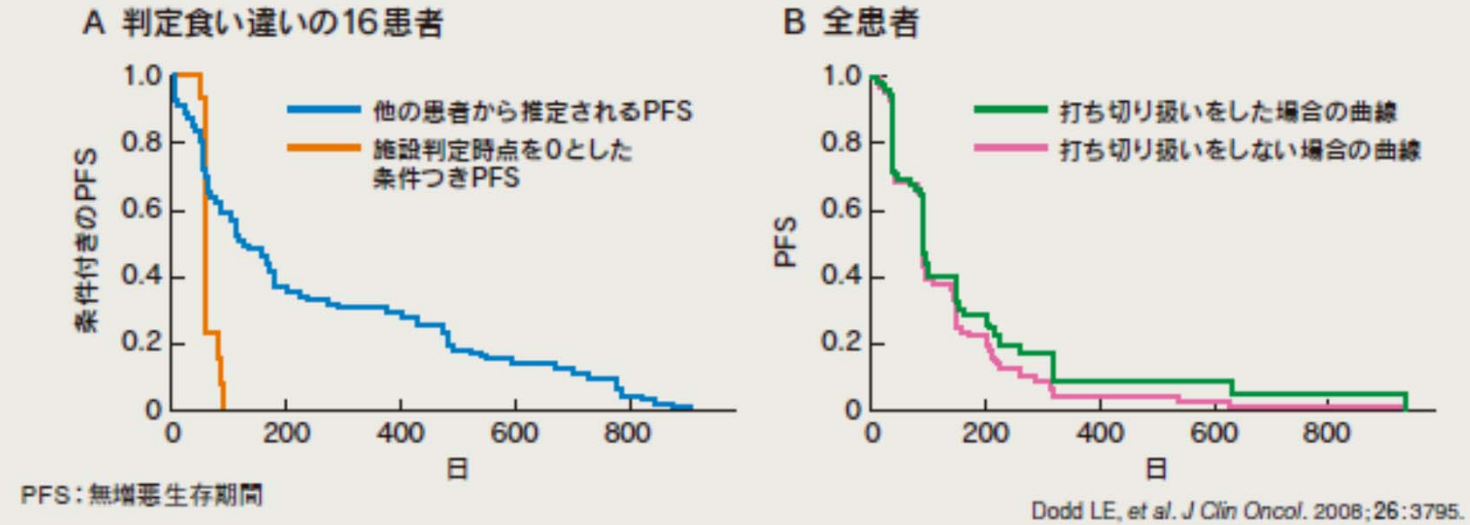
- 中央判定で進行とする前に施設で進行を判定
- 施設で進行が判定されれば治療法も変更され、それ以降の画像評価も行われないことが多い。つまり中央判定が不能となる。（わが国においても同様の状況）
- このような場合にFDAのガイドラインは打ち切り判定を推奨
- しかし、この種の打ち切りはinformativeであり評価にバイアス、みかけの成績を上げる効果
- もし対照治療群にこの打ち切りが多く発生すれば、治療効果が薄まる方向のバイアスが生ずる

109

Bias due to informative censoring

- 16患者においてBICR判定に先立って施設で進行が判定
- 16患者を打ち切り扱いすることにより上側のバイアス

図1 無増悪曲線の対比



PFS：無増悪生存期間

Dodd LE, *et al. J Clin Oncol*. 2008;26:3795.

110

進行・再発（肺）がん1次治療 OSに替わる評価？

- 2次治療以降は無視して大きなPFSの違いをめざす
- 治療戦略としての評価

  PFS、OS

  そしてQOL評価（Quality Adjusted Life Year)と

  経済評価

例：SELECT-BC

進行・再発乳癌1次治療 600例の登録完了

Taxane　VS　TS1

2次以降の治療は主治医選択

OSでのTS1の非劣性試験、ただしQALYを死亡まで評価

111

$$PFS = 1 - \exp(-(\lambda_1 + \lambda_2)t)$$

$$event1 : \frac{\lambda_1}{\lambda_1 + \lambda_2}(1 - \exp(-(\lambda_1 + \lambda_2)t) \quad event2 : \frac{\lambda_2}{\lambda_1 + \lambda_2}(1 - \exp(-(\lambda_1 + \lambda_2)t)$$

$$densityOS = \delta_0 \exp(-(\lambda_1 + \lambda_2)t) + \delta_1(1 - \exp(-\lambda_1 t)) + \delta_2(1 - \exp(-\lambda_2 t))$$

$$OS = \frac{\delta_0}{\lambda_1 + \lambda_2}(1 - \exp(-(\lambda_1 + \lambda_2)t) + \delta_1(t - \frac{1}{\lambda_1}(1 - \exp(-\lambda_1 t))) + \delta_2(t - \frac{1}{\lambda_2}(1 - \exp(-\lambda_2 t)))$$

112