Conference on Clinical Cancer Research

October 20, 2010 ~ Washington, DC

The Need for Adaptive Trials for Simultaneous Determination of Efficacy of a Therapeutic and a Diagnostic

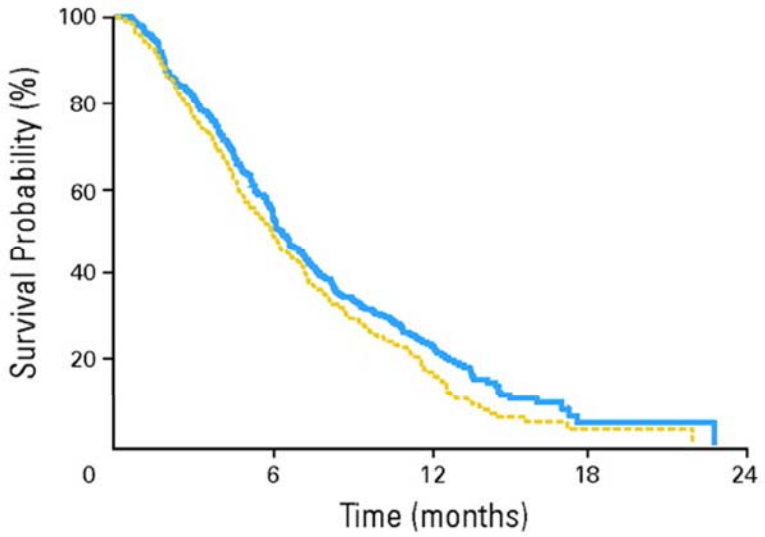
Eric H. Rubin

Merck

Problem Definition 1

- Most cancer drugs developed today are designed to inhibit specific cancer pathway targets
- Histology-based, “all comers” approaches to developing these drugs have typically led to failure in phase III studies, or demonstration of “success” based on statistically significant, but clinically questionable benefit in an “all comers” population

Statistically Significant but Clinically Relevant?



Recent Ph3 Failures of Targeted Drugs in Lung Cancer

Investigational Agent	Line of Therapy	Treatment regimen	Reason for Discontinuation
Cediranib	1L	combo with paclitaxel/carboplatin	Increased toxicity
Promune	1L	combo with paclitaxel/carboplatin	No improvement in OS
Sorafenib	1L	Combo with paclitaxel/carboplatin	Stopped for futility
Vadimezan	1L	combo with paclitaxel/carboplatin	Stopped for futility
Figitumumab	1L	combo with paclitaxel/carboplatin	Stopped for futility
Figitumumab	2L	combo with erlotinib	Stopped for futility;
Vandetanib	2L	combo with erlotinib	No improvement in OS
Vandetanib	2L	combo with docetaxel	No improvement in OS
Vorinostat	1L	combo with paclitaxel/carboplatin	Stopped for futility

Problem Definition 2

- Selection of a diagnostic test to identify patients who will benefit from treatment with a drug is difficult in early clinical trials and when this has been done, has often been incorrect
  - High EGFR protein expression was expected to predict responsiveness to EGFR-targeted antibodies, but this has not proven correct

VOLUME 22 · NUMBER 6 · MARCH 20 2004
JOURNAL OF CLINICAL ONCOLOGY ORIGINAL REPORT

Cetuximab Shows Activity in Colorectal Cancer Patients With Tumors That Do Not Express the Epidermal Growth Factor Receptor by Immunohistochemistry

Ei Young Chung, Jin-a Shin, Nancy E. Kemeny, Manish Shah, Gary K. Schwartz, Andrew Tse, Audrey Hamilton, Dorothy Pan, Deborah Schrag, Lawrence Schwartz, David S. Klimstra, Daniel Fridman, David P. Kelsen, and Leonard B. Saltz

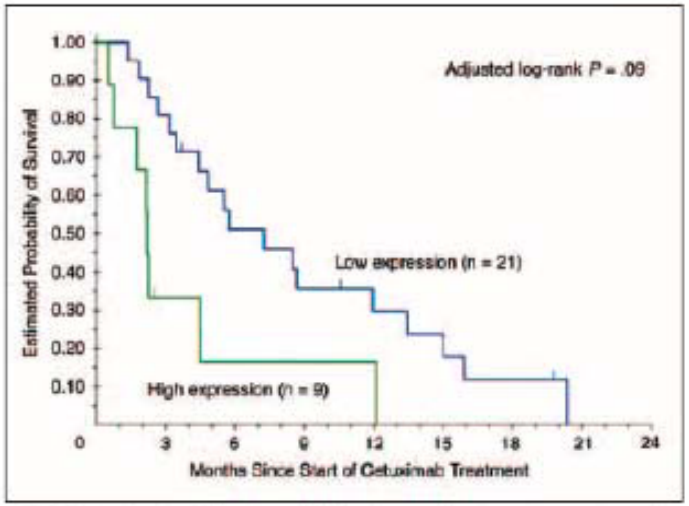


Fig 2. Plot of probability of survival for the study patients in relation to the epidermal growth factor receptor mRNA expression levels.

Molecular Determinants of Cetuximab Efficacy

Daniel Vallböhmer, Wu Zhang, Michael Gordon, Dong Yun Yang, Jim Yun, Oliver A. Press, Katrin E. Rhodes, Andy E. Sherrod, Syma Iqbal, Kathleen D. Danenberg, Susan Groshen, and Heinz-Josef Lenz

J Clin Oncol 23:3536-3544. © 2005 by American Society of Clinical Oncology

Problem Definition 3

- Conventional Phase I and II trials lack sufficient power to identify responsive subgroups
  - Also often lack a control group
  - BATTLE and I-SPY adaptive trials are an exception
- Progress in generating diagnostic tests that can be used to select responsive patients has been slow
  - Only 8 such tests currently listed in cancer drug labels
    - ER IHC, C-KIT IHC, 5q del chrom, EGFR IHC, HER2 IHC, RAS mutation, PML-RAR chrom, BCR-ABL chrom
  - Only 3 of the 8 tests are FDA approved
    - HER2 IHC, EGFR IHC, c-KIT IHC

Proposed Solution

- Design a pivotal phase III trial approach that adaptively identifies a responsive patient population and confirms the effectiveness of a new therapeutic in this population in a rigorous statistical manner
  - Designed to support simultaneous approval of a new therapeutic with an accompanying diagnostic test

Conference on Clinical Cancer Research

October 20, 2010 ~ Washington, DC

Targeting the Androgen Signaling Axis in Castration-Resistant Prostate Cancer

Howard I. Scher, MD

D. Wayne Calloway Chair in Urologic Oncology

Chief, Genitourinary Oncology Service

Memorial Sloan Kettering Cancer Center

Disclosures

I have the following financial relationships to disclose:

Consultant for: J&J Ortho Biotech (uncompensated)
Medivation (uncompensated)

Grant/Research support: J&J Ortho Biotech
Medivation
Aragon

I will discuss the following off label use and/or investigational use in my presentation:

Abiraterone acetate
MDV3100

A Clinical States Framework For The Prostate Cancer Disease Continuum Including Recent Approvals That Establish New *"Pre-"* and *"Post"* Chemotherapy Standards of Care

Diagnoses

↓

192, 280

Non-Castrate

Androgen depletion/blockade (bicalutamide)

Castration resistant

Deaths From Disease

↓

27,360

Clinically Localized Disease

→

Rising PSA

→

Clinical Metastases: Non-Castrate

→

2 Clinical Metastases: Castrate *Pre-* Provenge

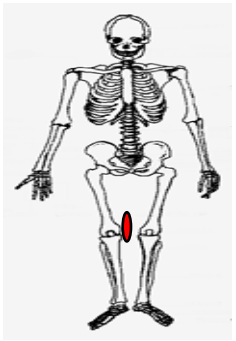
→

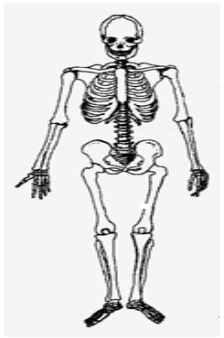
3 Clinical Metastases: Castrate 1st Line Docetaxel

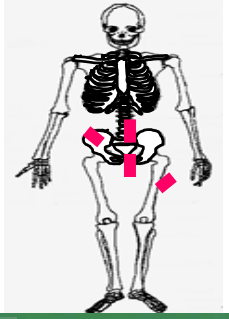
→

4 Clinical Metastases: Castrate *Post-* Carbazitaxel

1 Rising PSA: Castrate

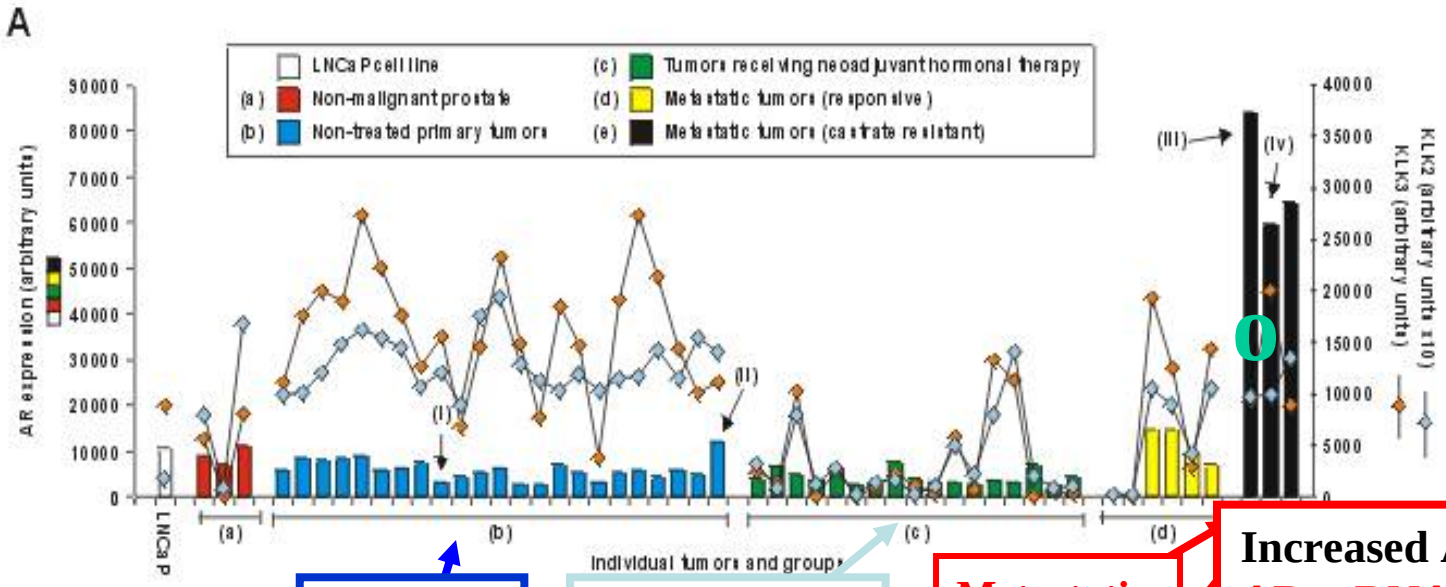


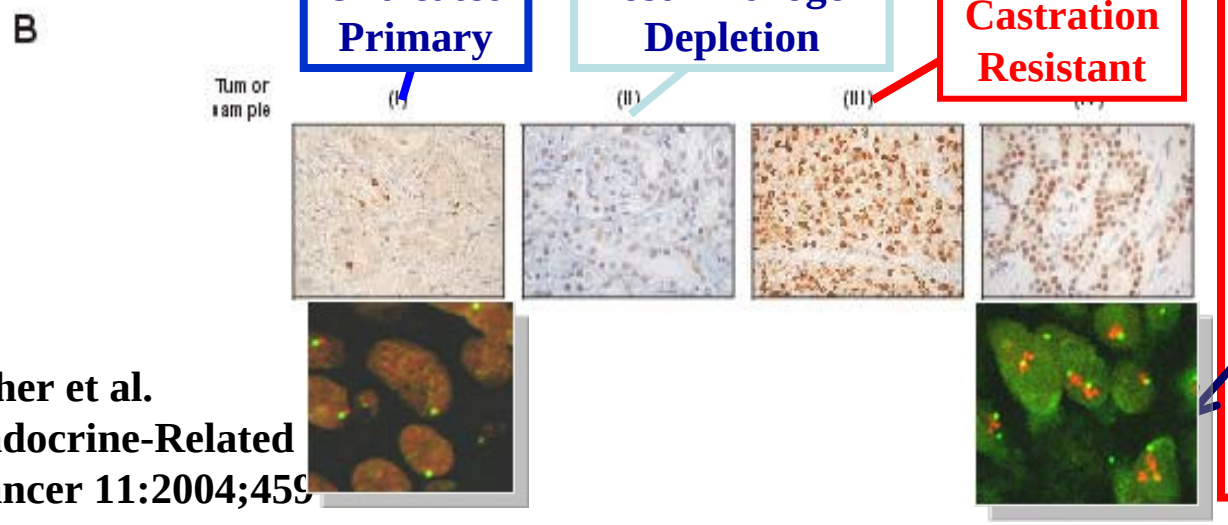




With detectable metastases: deaths from cancer exceed that from other causes

Oncogenic Alterations in Late State Prostate Cancers Are Targets for Therapy





Untreated Primary

Post-Androgen Depletion

Metastatic Castration Resistant

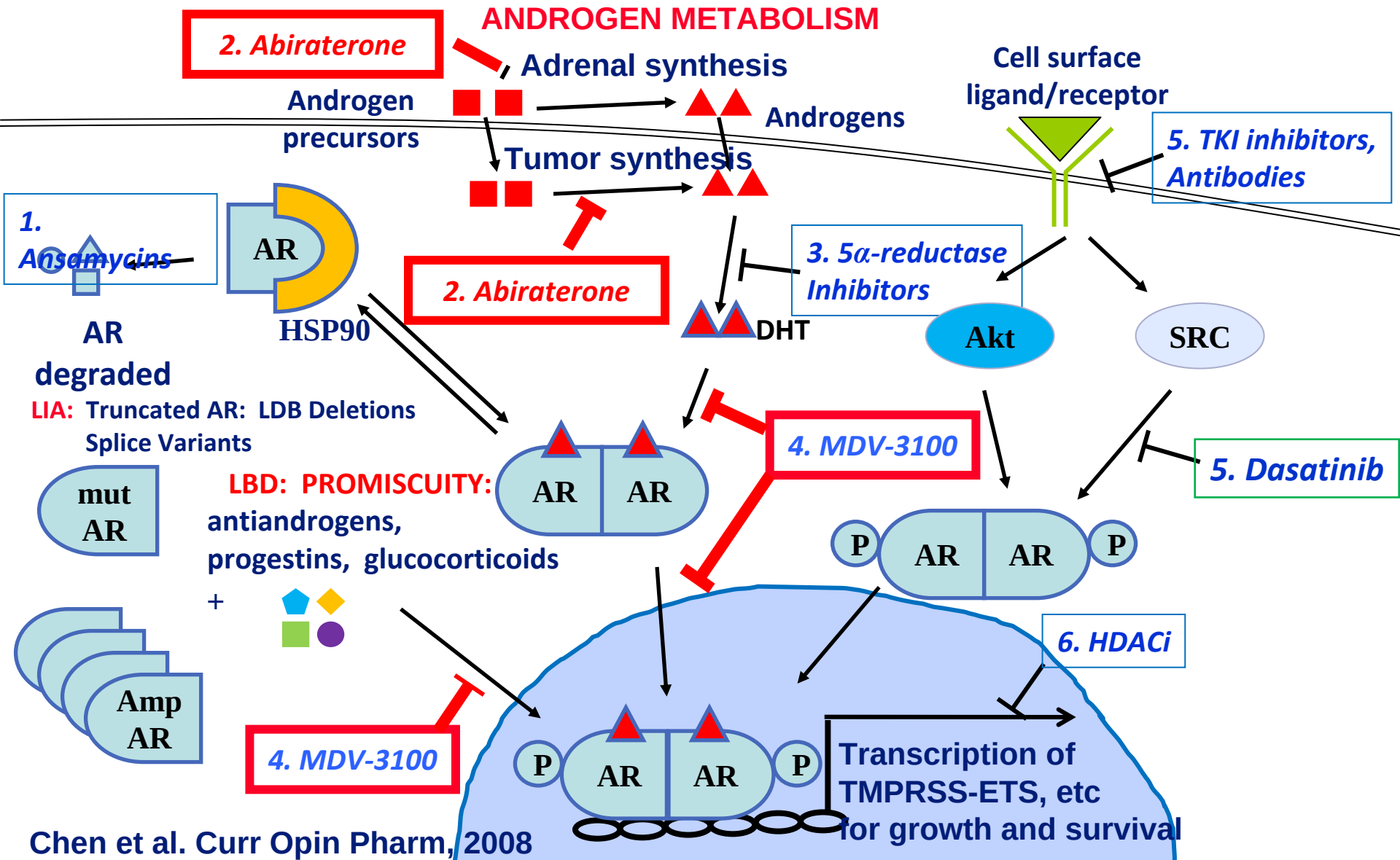
Increased AR protein
AR mRNA overexpression

AR amplification
AR mutations
Increased androgen
- biosynthesis: intracrine signaling

Alternative pathways: kinase activation

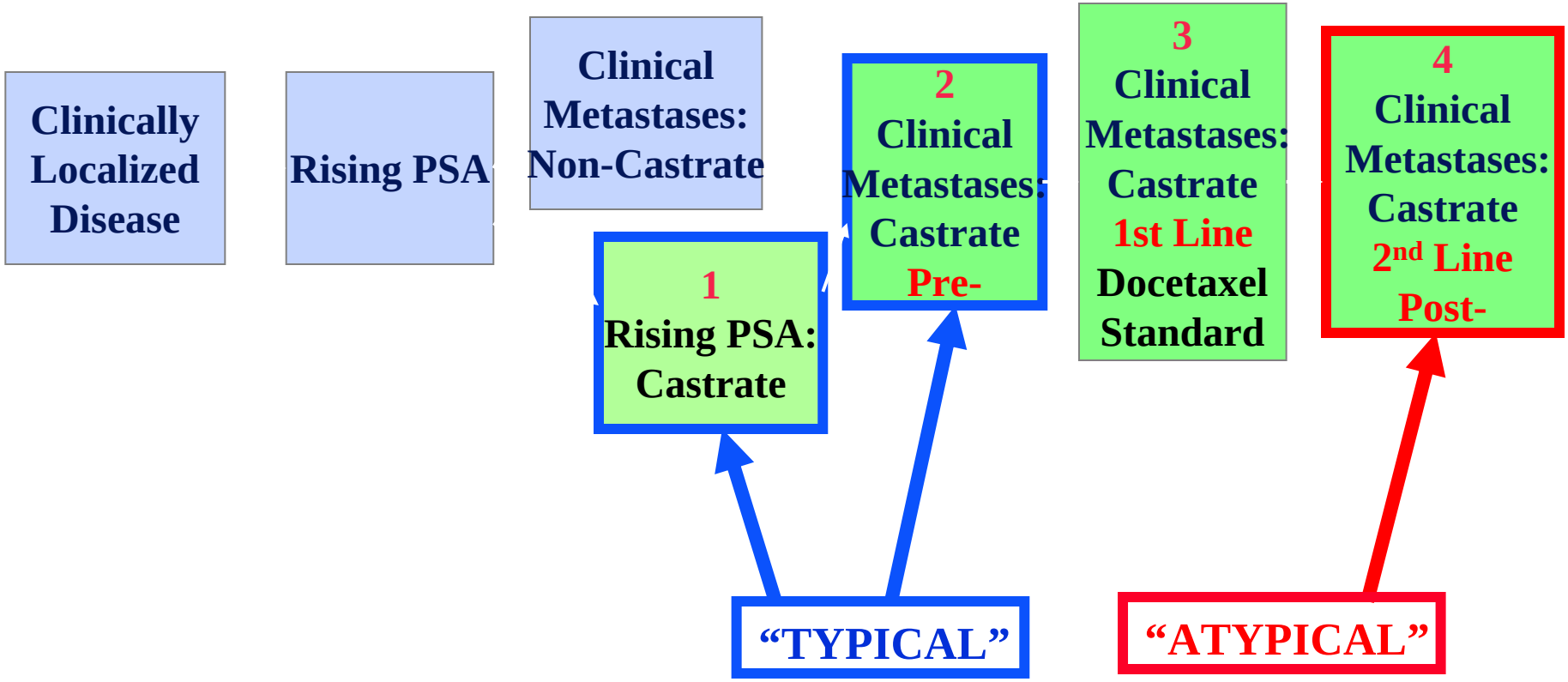
Scher et al.
Endocrine-Related
Cancer 11:2004;459

The Clinical Results With Abiraterone Acetate (Cyp 17 Inhibitor) and MDV3100 (A Next Gen Antiandrogen) Credential The Androgen Receptor Signaling Axis As A Relevant Therapeutic Target



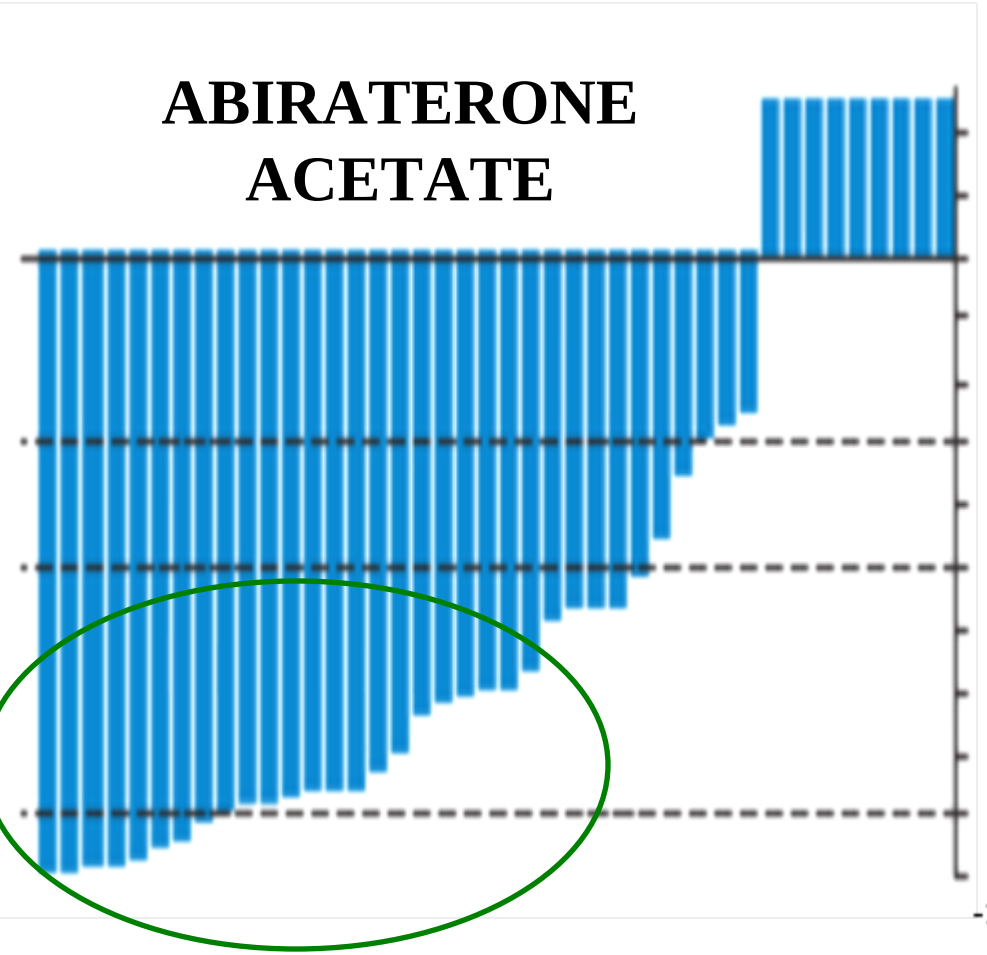
Chen et al. Curr Opin Pharm, 2008

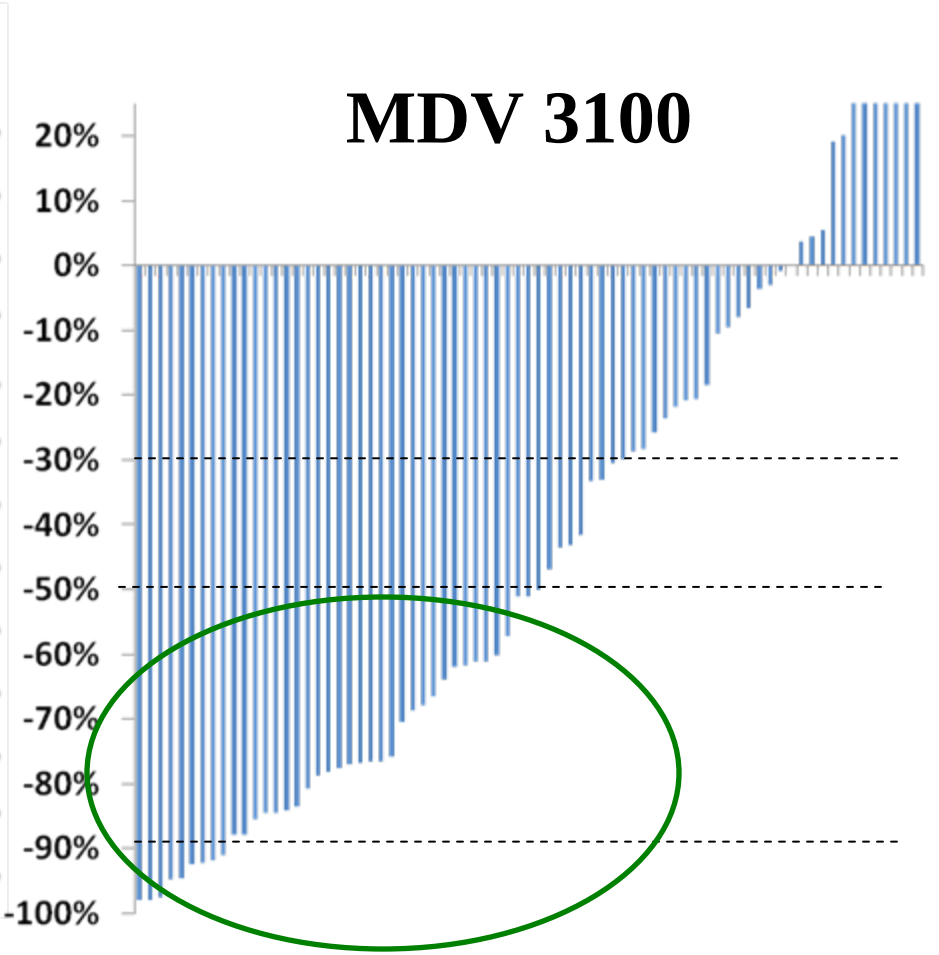
Hypothesizing that the Decision to Use Chemotherapy Would Not Change Disease Biology: Abiraterone Acetate and MDV3100 Were Studied in *Pre-* and *Post-* Chemotherapy mCRPC



Survival based endpoints are assessable earlier in the *post* chemotherapy setting: a shorter track to the "goal line"

Significant Activity Was Observed in Post-Chemotherapy Treated Patients Leading to Phase 3 Registration Trials in Unselected (All Comers) Patient Populations

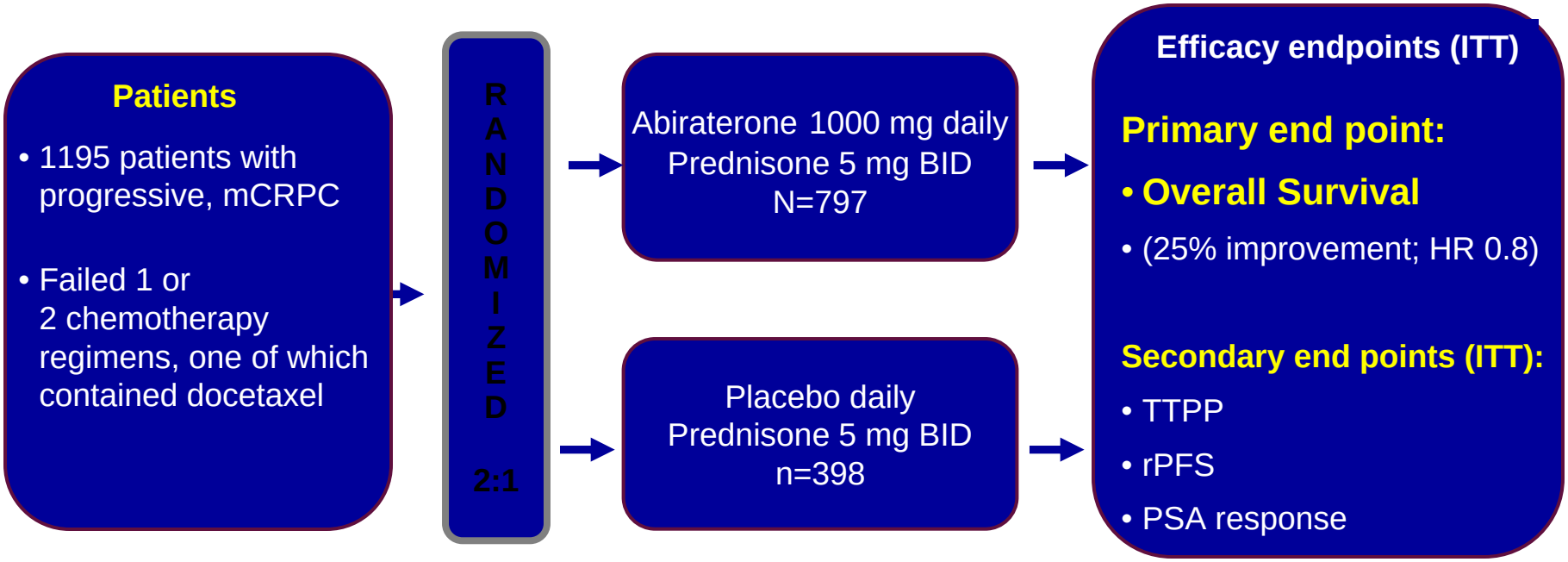




Danila et al. JCO 27:1496, 2010
Also: Reid et al. JCO 27: 1489. 2010

Scher et al. Lancet 75:1437, 2010

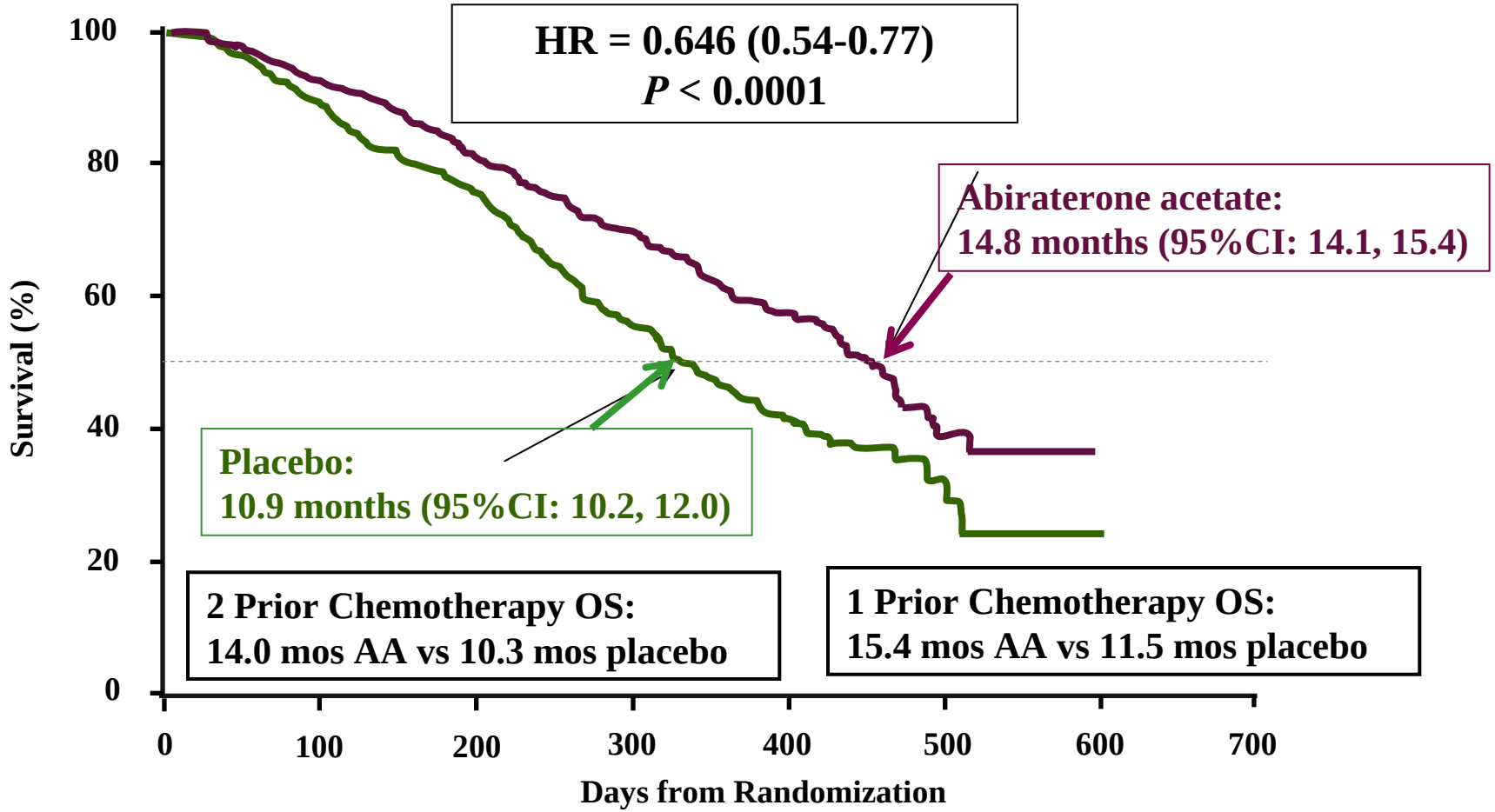
COU-AA-301, A Randomized Double-Blind Placebo-Controlled Phase 3 Trial of Abiraterone Acetate In Post-Chemotherapy Treated CRPC



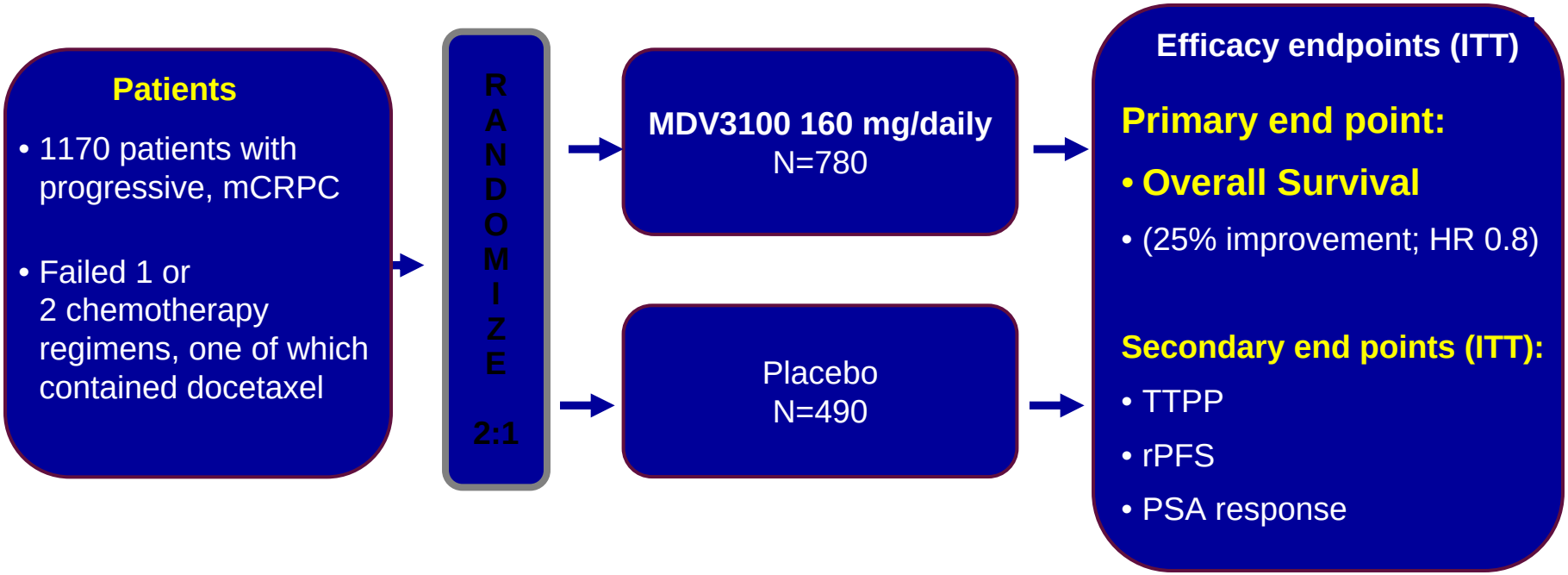
ESMO, October, 2010

J. DeBono, H. Scher, Co-PI

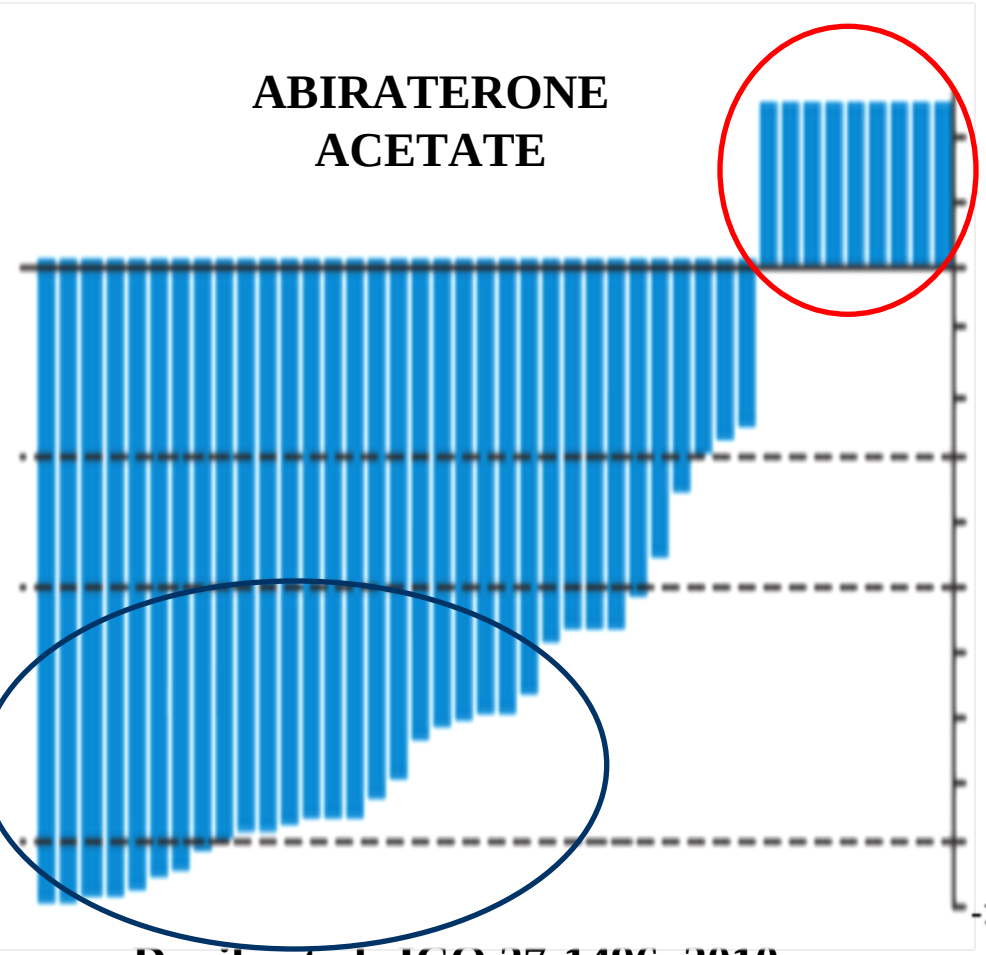
Abiraterone Acetate Prolongs Overall Survival in Patients With mCRPC Who Have Progressed After Docetaxel-Based Chemotherapy



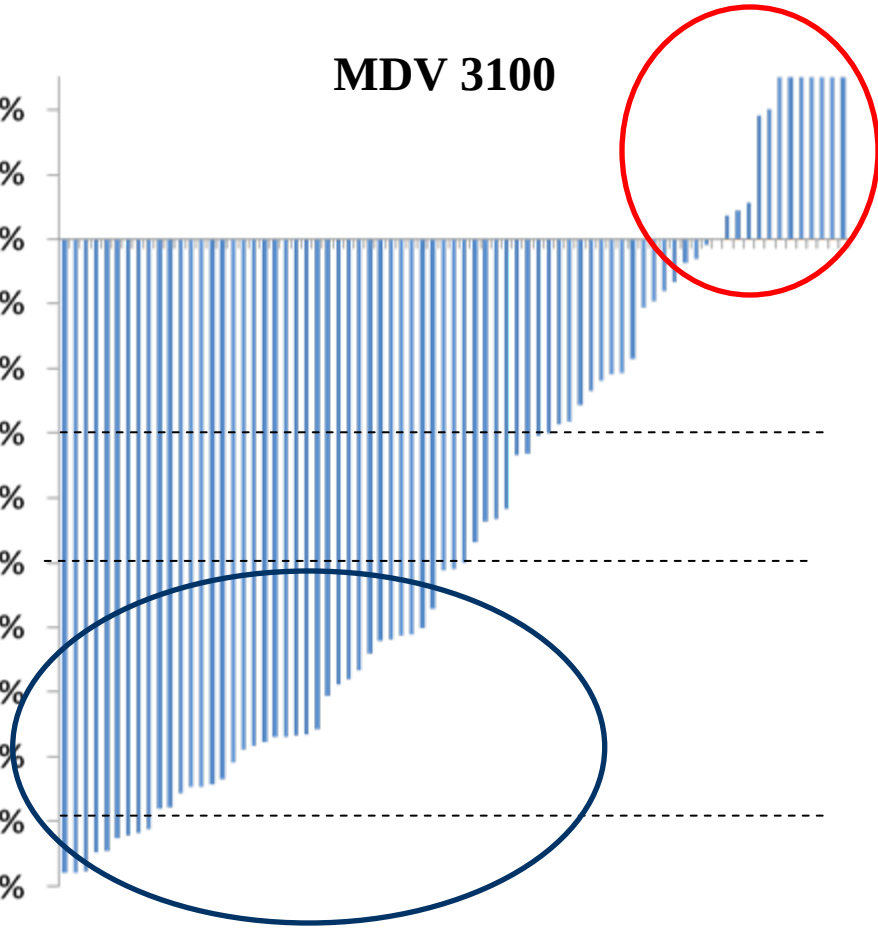
AFFIRM, A Randomized Double-Blind Placebo Controlled Phase 3 Trial of MDV3100 In Post-Chemotherapy Treated CRPC is Anticipated to Fully Accrue in December, 2010



For Both Drugs, The Pattern of Change in PSA Suggests the Presence of Biological Predictors of Sensitivity and Resistance: Are The Predictors The Same Or Different For These Two Agents?



Danila et al. JCO 27:1496, 2010
Also: Reid et al. JCO 27: 1489. 2010



Scher et al. Lancet 75:1437, 2010

Addressing Sensitivity to AR Signaling Inhibitors

1. **Discovery**:

   Profile GEMM's, cell lines and/or xenografts that replicate human prostate cancer.

   Genotype tumors for common genomic alterations:

TMPRSS2-ERG fusion	**50%**
PTEN deletion	**40%**
AR amplification	**30%**
MYC amplification	**20%**

2. **Assay development and validation:**

3. **Clinical qualification:**

   Study a biomarker across clinical states: primary vs. metastatic disease
   castration sensitive vs. resistant

4. **Prospective trials to explore associations with clinical outcome(s):**

Biomarker Development

Assay validation:

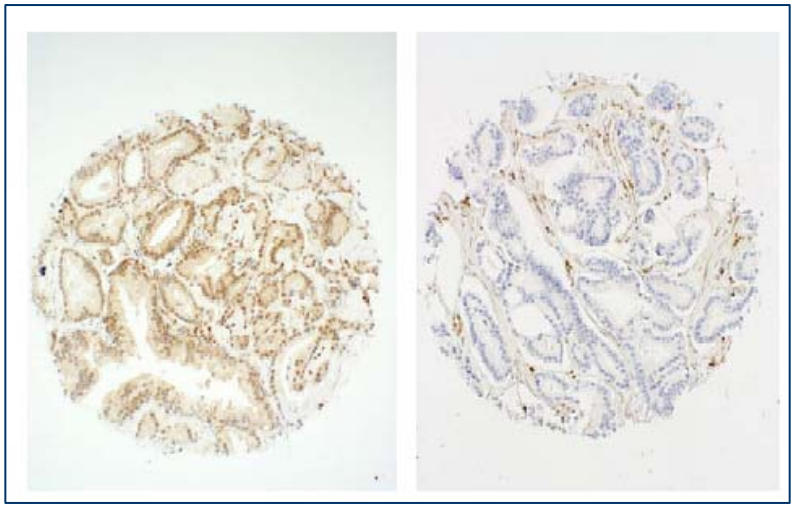
- Establish minimum performance characteristics for an assay to justify clinical testing.
- Achieve analytical validity (short of full CLIA) across laboratories/centers.

Clinical qualification:

3. To develop performance metrics in the clinic that are unrelated to an intervention to justify further testing.
- To design trials in a sequence to qualify a “biomarker” for a specific “context of use” (label) that will affect/impact/guide medical decision making.

A Clinical Trial is Under Development That Includes an Analytically Valid PTEN IHC Assay: "Null" or "Any"

PTEN expression in tumors from two different patients on a tissue microarray



Present in tumor and stroma

"Null" in tumor Present in stroma

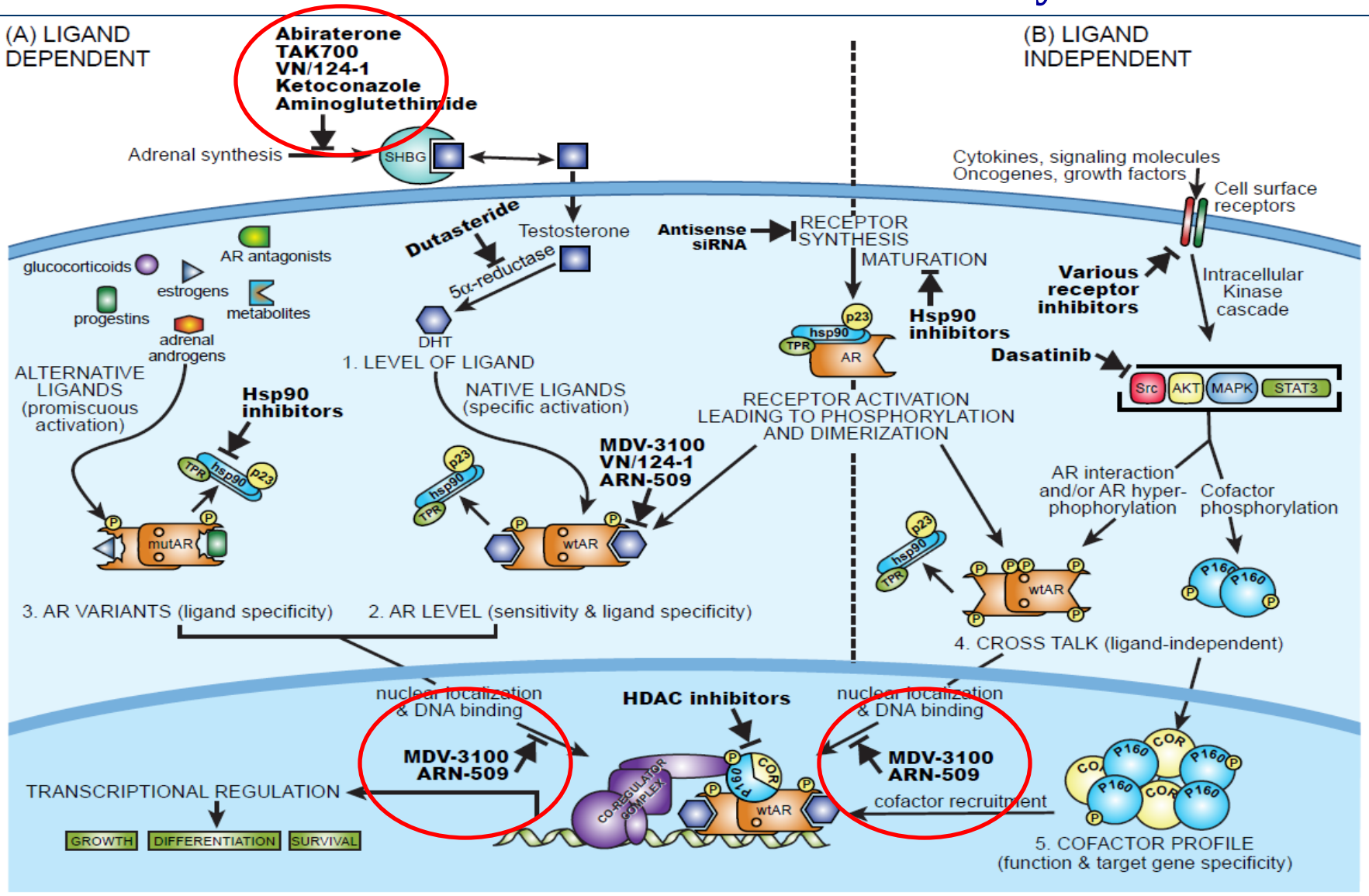
Courtesy of V. Reuter, MSKCC

MSKCC, JHU and DFCI SPORE

1. Validation in cell lines.
2. **IHC in primary and metastatic tissue: "Null" or any** H-score
3. qPCR.
4. CNA.

DeMarzo (JHU) Loda (DFCI) and Reuter (MSKCC).

There are a Range of Agents in Clinical Development That Target The Androgen Receptor Signaling Axis Which Are Anticipated to Show Similar Patterns of Sensitivity



Demonstrating Clinical Benefit for These Agents Will Be More Difficult

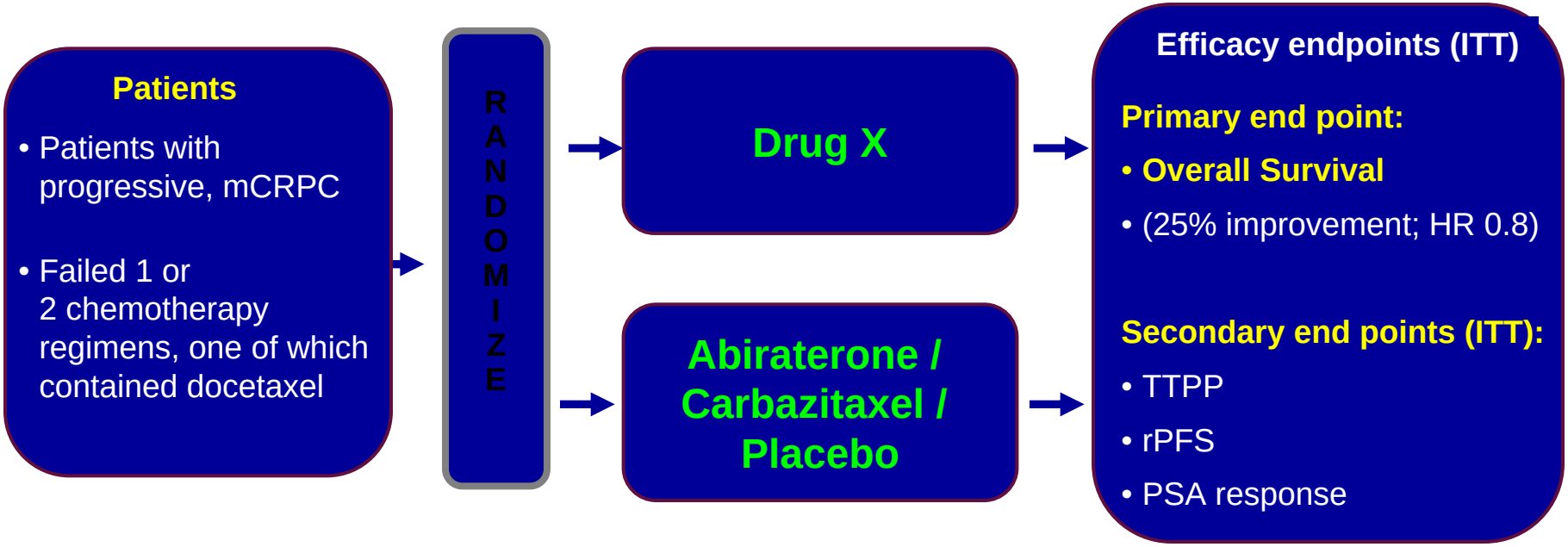
1. Changing standards of care.

2. The availability of more treatments that are effective.

3. “All comers” eligibility ultimately dilutes the treatment effect.

4. Concurrent development of companion diagnostics is essential: using an adaptive design.

A Randomized Double-Blind Phase 3 Trial of Drug "X" vs. A "Standard" In Post-Chemotherapy Treated CRPC: Biomarker Discovery and Validation Proceed in Parallel



1. **As predictive biomarkers are unknown at present, trial eligibility should require collection of primary prostate tissue.**
2. **Practically, assays should be performed on formalin fixed paraffin embedded tissue.**
3. **For selected determinants, a metastatic tumor sample may be necessary.**
4. **Blood based biomarkers can also be considered.**

Conference on Clinical Cancer Research

October 20, 2010 ~ Washington, DC

Adaptive Signature Design for Clinical Trial of Advanced Prostate Cancer

Richard Simon, D.Sc.

Chief, Biometric Research Branch, National Cancer Institute

http://brb.nci.nih.gov

- Cancers of a primary site often represent a heterogeneous group of diverse molecular diseases which vary fundamentally with regard to
  - the oncogenic mutations that cause them
  - their responsiveness to specific drugs

- How can we develop new drugs in a manner more consistent with modern tumor biology and obtain reliable information about what regimens work for what kinds of patients?

- Developing a drug with a companion test increases complexity and cost of development but should improve chance of success and have substantial benefits for patients and for the economics of medical care

- Although the randomized clinical trial remains of fundamental importance for predictive genomic medicine, some of the conventional wisdom of how to design and analyze rct's requires re-examination
- The concept of doing an rct of thousands of patients to answer a single question about average treatment effect for a target population presumed homogeneous with regard to treatment efficacy in many cases no longer has an adequate scientific basis

- Predictive biomarker
  - Measured made before treatment to identify who is likely to benefit from a particular treatment
- Classifier
  - Decision tool based on one or more predictive biomarkers

In Ideal Settings

1. Develop a completely specified classifier identifying the patients most likely to benefit from a new drug
   - Based on biology, pre-clinical data and phase I-II studies
2. Establish analytical validity of the classifier
3. Design and analyze a focused clinical trial to evaluate effectiveness of the new treatment and how it relates to the classifier

- Cancer biology is complex and it is not always possible to have the right single completely defined predictive classifier identified and analytically validated by the time the pivotal trial of a new drug is ready to start accrual
  - Adaptive methods for the refinement and evaluation of predictive biomarkers in the pivotal trials in a non-exploratory manner
  - Use of archived tissues in focused “prospective-retrospective” re-analysis of previously conducted randomized pivotal trials
    - Simon, Paik, Hayes; JNCI 101:1-7, 2009

Conference on Clinical Cancer Research

October 20, 2010 ~ Washington, DC

Adaptive Signature Design

Boris Freidlin and Richard Simon

Clinical Cancer Research 11:7872-8, 2005

End of Trial Analysis

- Compare outcomes X to C for **all patients** using significance threshold 0.01
  - If overall H_0 is rejected, then claim effectiveness of X for eligible patients
  - Otherwise
    - Compare outcomes X to C in **one** adaptively defined subset of patients using threshold of significance 0.04

- Divide the patients randomly into a training set T and a validation set V. The training set will contain one-third of the patients.
- Using the biomarker information, treatment and outcome for the patients in T, develop a binary classifier that identifies the subset of patients who appear most likely to benefit from the new treatment X compared to control C
  - f(B1,B2,B3,B4) = log hazard ratio of death for X relative to C as a function of biomarker values
  - If f(B1,B2,B3,B4)/ser <c then Classifier(B1,B2,B3,B4)=X
  - If f(B1,B2,B3,B4)/ser >c then Classifier(B1,B2,B3,B4)=C
  - Cutpoint c optimized

- Use the classifier developed in training set T to classify the patients in the validation set V.
- Let V_X denote the subset of patients in V who are classified as likely to benefit from X
- Compare survivals of patients who received T to survivals of those who received C for patients in V_X
  - If the difference in survival is significant at level 0.04, then the new treatment is more effective than the control for patients with biomarker values for which Classifier(B1,B2,B3,B4) =X
  - The classifier identifies the indication for use of X for future patients

- This approach can also be used to identify the subset of patients who don't benefit from X in cases where X is superior to C overall at the 0.01 level. The patients in $V_C = V - V_X$ are predicted not to benefit from X. Survivals of X vs C can be examined for patients in that subset and a confidence interval for the hazard ratio calculated.

- This design has improved statistical power for identifying treatments that benefit a subset of patients in molecularly heterogeneous diseases
- It has greater specificity than the standard approach for identifying which patients are not likely to benefit from a new treatment
- The standard approach results in treatment with approved drugs of many patients who do not benefit from them

Sample Size Planning for Advanced Prostate Cancer Trial

- Survival endpoint
- Final analysis when there are 700 deaths total
  - 90% power for detecting a 25% overall reduction in hazard at two-sided 0.01 significance level (increase in median from 12 months to 9 months)
- 80% power for detecting 37% reduction in hazard for subset consisting of 33% of patients
  - 700 * (2/3) * (1/3) = 157
  - 157 deaths required for 80% power to detect 37% reduction in hazard at two-sided 0.04 significance level.
  - To have 700 deaths at final analysis, 935 patients will be accrued and followed till the event rate is 75%

Sample Size Planning

- The number of required patients can be substantially reduced by
  - Targeting larger treatment effects
  - Targeting treatment benefits that apply to more than 33% of the patients
  - Refining the simple interim analysis for futility described for this example

- Tumor specimen at entry as condition for eligibility
- Specimen preserved for later assay
- Assays will be performed prior to analysis using analytically validated tests
  - Reproducible, robust and accurate for use with archived tissue
  - No cut-point required
  - Additional markers could be included prior to using specimens

Interim Futility Analysis

- Interim futility analysis conducted when there are approximately 340 patients who have been followed for 6 months after randomization
- The analysis will use 6-month progression-free survival as intermediate endpoint.
- If difference between X group and C group is not significant at one-sided 0.20 level, then accrual will be terminated
- Power 90% for detecting 12 percentage point increase in proportion free of recurrence at 6 months from baseline of 40%

Interim Futility Analysis

- Interim futility analysis does not utilize any of the 5% type I error of the study
- Using 6 month PFS as endpoint for interim futility analysis does not assume that PFS is a valid surrogate of survival; only that it is plausible to not expect a survival benefit if there is no PFS benefit
- The one-sided 0.20 significance level is used because the overall effect may be weak if the treatment benefits only a 33% subset of the patients.

If the Markers Were Measured at Randomization

- Analytically validated tests would be required by the start of accrual
- The interim analysis could involve marker-defined subsets of patients
- Restricting accrual based on interim evaluation of marker specific treatment effects could substantially reduce sample size but would introduce issues not addressed in the current design

Key Features

- Trial-wise type I error limited to 0.05
  - Chance of any false positive conclusion of treatment benefit limited to 0.05
- Randomized treatment assignment
- Regulatory endpoint
- Sample size sufficient for
  - evaluating treatment effect in 33% subset
- Biomarkers measured using analytically validated tests
- Analysis algorithm pre-defined, and specific analysis plan defined prior to any assaying of tumors or data analysis

- This approach is as sound statistically as the conventional “one treatment fits all” design
- In settings where a single conventional “average effect” trial would be the basis for drug approval, this design should be the basis for approval either overall or for the identified subset
- This approach is more science based and consistent with tumor biology

Conference on Clinical Cancer Research

October 20, 2010 ~ Washington, DC

Adaptive Clinical Trial Designs

Dr. R. Sridhara

Director, Division of Biometrics V

CDER, FDA

The Term Adaptive

- Change in eligibility criteria
- Change in planned sample size
- Change in choice of test statistic – analytic methods
- Change in choice of hypothesis
- Change in choice of primary endpoint
- Change in choice of dose groups/drop or add treatment arms
- Change in allocation to treatment to achieve balance or assign fewer subjects to the inferior treatment – randomization procedure
- Change to enrich subpopulation

Potential Advantages of ADs

- Increase of efficiency to collect same information
- Increase likelihood of success on the study objective
- Improved understanding of the treatment's effect

Types of ADs

- Exploratory:
  - Less restrictive
  - Explore without adjusting for multiple looks, multiple adaptations to generate hypothesis to be tested
- Hypotheses Testing or Confirmatory:
  - Adequate and well controlled (A & WC) studies
  - Pre-planned, type I error rate (false positive rate) well controlled
  - Decision rules specified for each adaptation

Important to Remember

- Confirmatory studies are not exploratory studies
  - Prospectively Planned
  - Study Integrity Maintained
- Not considered as AD:
  - Revisions after unplanned findings from IA
- Reactive revisions difficult to interpret – judgmental

‘Targeted’ Therapy

- Target? – generally biomarker guided
- Biomarker measurement – accuracy, reliability, etc.
- Biomarker cut-off threshold – are positive and negative distinct from each other – generally measured on a continuous scale
  - Target treatment effect large in marker positive group
  - Target treatment have quantitatively less effect in marker negative group
  - Target treatment have harmful effect in marker negative group

Analytical Validation

- Analytical performance
  - Precision (repeatability, reproducibility)
  - Accuracy
  - Sensitivity, Limit of Detection
  - Specificity (interference, cross-reactivity)
  - Sample type / matrix
  - Sample preparation / conditions
  - Performance around the cut-off
  - Potential for carryover, cross-hybridization

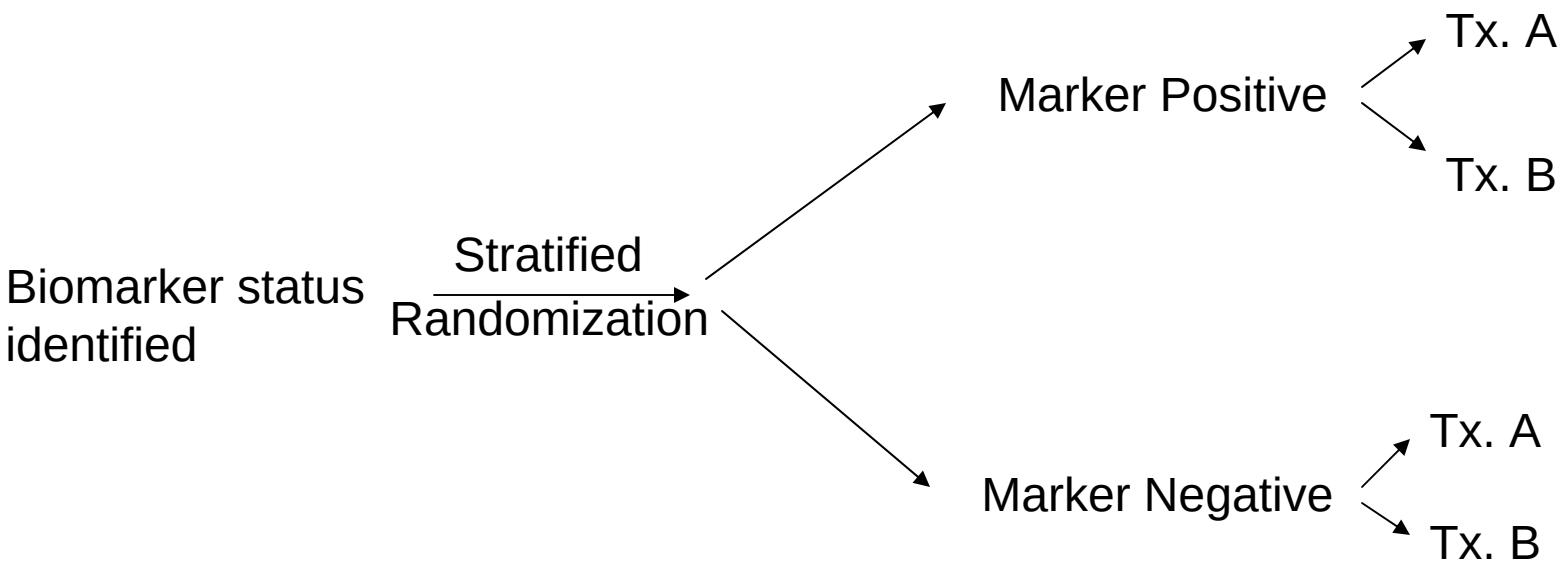
Clinical Validation Steps

- Training Set(s)
- Develop classifier and/or cut-offs
  - Fully specified device
- Test Set (s)
- Independent Validation on intended use population

Proposed ‘Signature’ Design

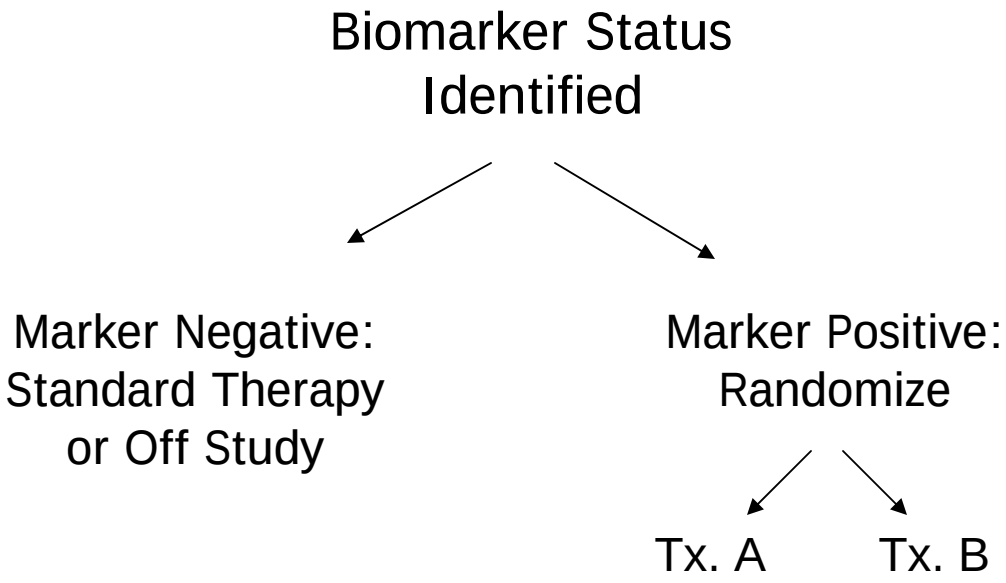
- Assumes: Biomarker assay validated and performance characteristics known; Scientific rationale for biomarker choice known; Reasonable prevalence of marker positive patients and no imbalance in prognostic factors between treatment arms in the marker positive subgroup.
- No adaptations during the study. RCT conducted in the overall patient population. After study completion, test if the treatment is effective in a biomarker defined subgroup.
- Pros: False positive rate is controlled; Pre-specified algorithm & analyses; Uses all the available information.
- Cons: Even if the Tx works in the overall population, chances of winning is low; Potentially a good number of marker negative patients may be treated who do not benefit and may be harmed.

Alternative Design

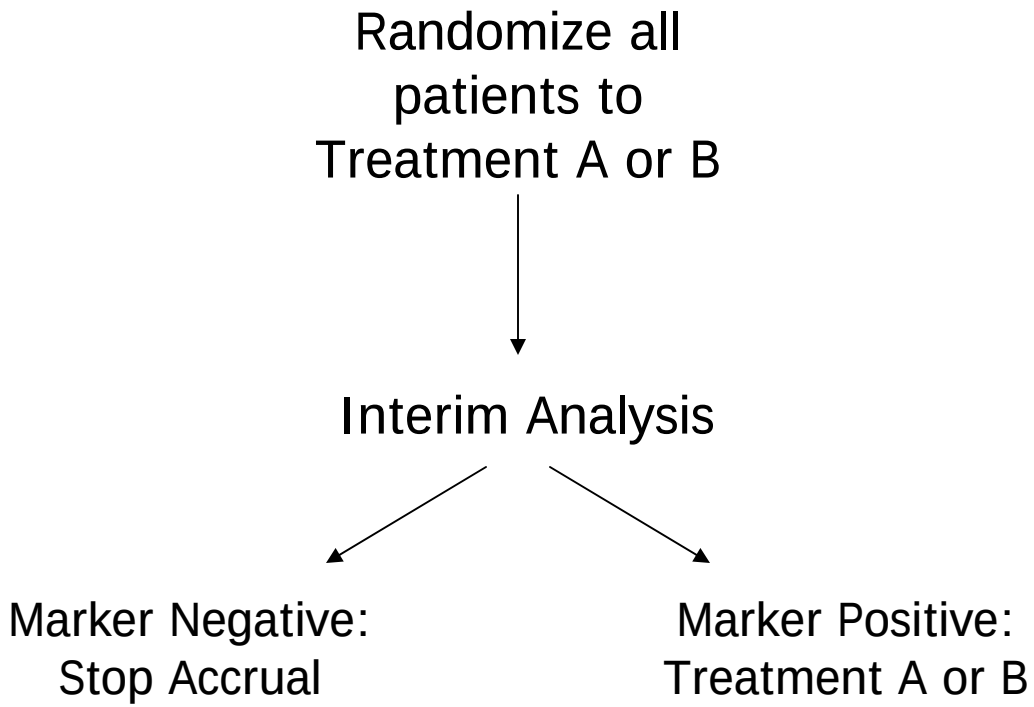


Targeted Enrichment

Adaptive Enrichment Design 1



Adaptive Enrichment Design 2



Examples of Success Stories

- Herceptin, Gleevec
  - Good scientific rationale
  - Pre-clinical data available
  - Enrichment designs
  - **Big** superior treatment clinical benefit
- Under development: PARP inhibitors, BRAF inhibitors, ALK inhibitors
  - Good scientific rationale
  - Pre-clinical data available
  - Enrichment designs
  - Expected **big** superior treatment clinical benefit

Summary

- FDA: Treat patients who are likely to experience substantial clinical benefit and do not treat patients in whom the treatment may cause harm
- Approval based on Risk - Benefit ratio
- Targeted drug development can achieve this goal
  - Key lessons for success: Good pre-clinical data, identification of target, a drug that actually hits the target, substantial treatment clinical benefit
- Many statistical enrichment adaptive RCT designs which control false positive rate are available to achieve this goal