ASA NJ Chapter Webinar, Dec 8, 11:00 am-12:30 pm EST

Randomized Trial Designs for Evaluating Predictive Biomarker Tests: What's the Estimand?

Gene Pennello, Ph.D.
Division of Biostatistics
FDA/CDRH/OSB

Acknowledgements: Qin Li, Ph.D., Jingjing Ye, Ph.D.

U.S. FOOD & DRUG
ADMINISTRATION

Outline

- Predictive Biomarker
- Companion Diagnostic Test (CDx)
- CDx Clinical Trial Design
  - Biomarker-Stratified
  - Biomarker-Strategy
  - Enrichment Design
  - Discordant Risk Randomization
- Estimands

Predictive Biomarkers

- **Predictive biomarker** informs on likely outcomes with specific treatments (e.g., relative sensitivity or resistance).
  - Other names: treatment selection biomarker, CDx
- **Prognostic biomarker** is biological characteristic indicating likelihood of disease progression in a homogeneous population of patients, either not receiving therapy (natural course) or on a standard therapy.
  - inform on outcomes independent of specific treatment (i.e. in oncology, ability of tumor to proliferate, invade, and/or spread)

Intended Uses / Claims

- Companion Diagnostic:
  - Provides information that is essential for the safe and effective use of a corresponding therapeutic product, allowing its benefits to exceed its risks.
  - EX. Defines the population for whom a therapeutic product is indicated.
- Complementary Diagnostic:
  - Provides clinically useful information about a therapeutic product yet is not a prerequisite for the therapeutic product's use (*not an official FDA definition*).

FDA Guidance, Predictive Markers

Beaver JA; Tzou A; Blumenthal GM; McKee AE; Kim G; Pazdur R; Philip R. An FDA Perspective on the Regulatory Implications of Complex Signatures to Predict Response to Targeted Therapies. Clin Cancer Res. **2017**, 23 (*6*), 1368-1372.

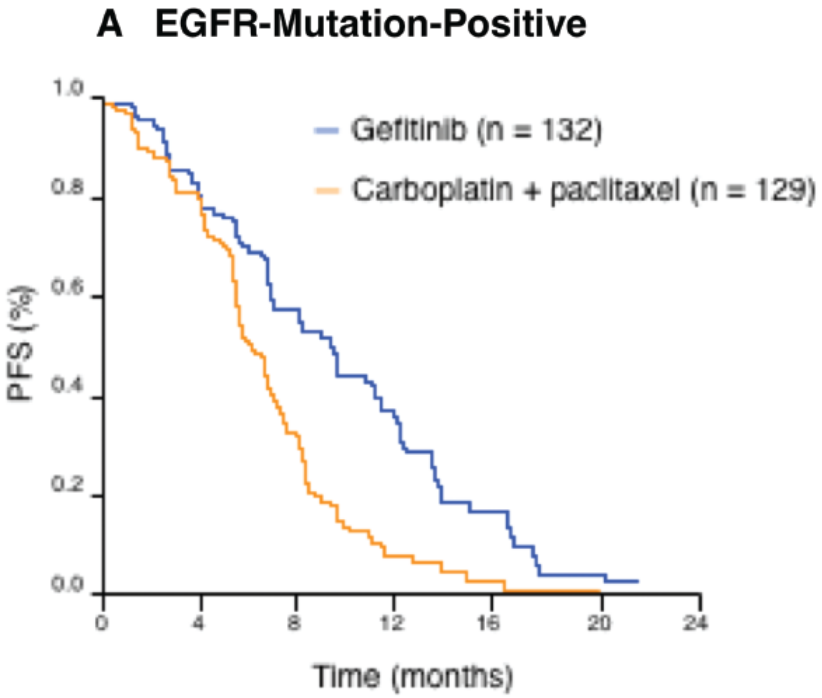
US FDA. Guidance on Enrichment Strategies for Clinical Trials to Support Approval of Human Drugs and Biological Products. US FDA: Silver Spring, MD, **2012**.

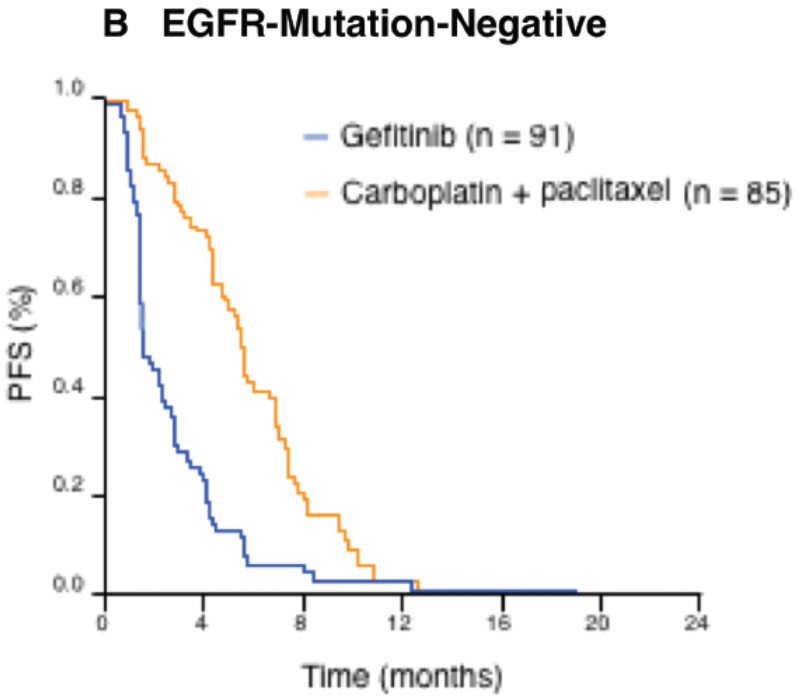
US FDA. In Vitro Companion Diagnostic Devices, US FDA: Silver Spring MD, **2014**.

US FDA. Principles for Codevelopment of an In Vitro Companion Diagnostic Device with a Therapeutic Product. US FDA: Silver Spring MD, **2016**.

US FDA-NIH Biomarker Working Group. BEST (Biomarkers, EndpointS, and other Tools) Resource [Internet]. Silver Spring (MD): Food and Drug Administration (US); **2016**.

Qualitative Interaction (NSCLC)





No. of patients at risk

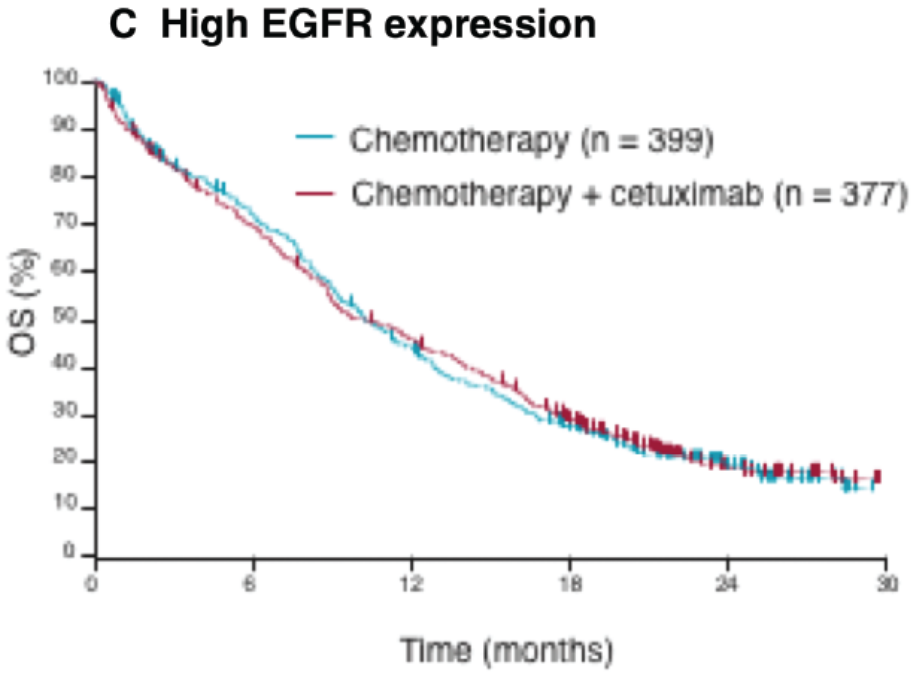
Time (months)	0	4	8	12	16	20	24
Gefitinib	132	108	71	31	11	3	0
Carboplatin plus paclitaxel	129	103	37	7	2	1	0

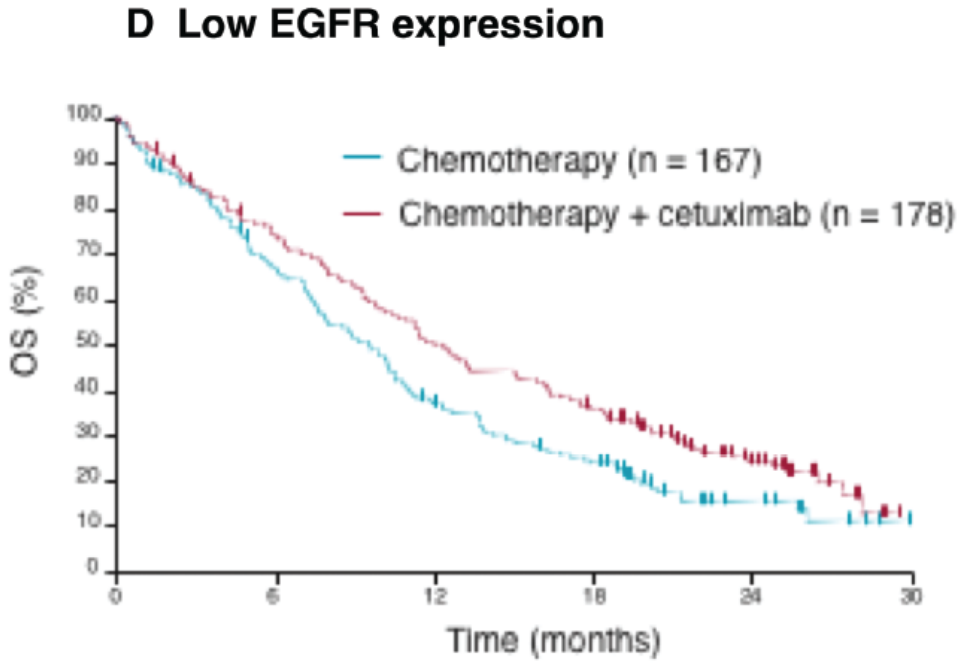
No. of patients at risk

Time (months)	0	4	8	12	16	20	24
Gefitinib	91	21	4	2	1	0	0
Carboplatin plus paclitaxel	85	58	14	1	0	0	0

Polley MC, Freidlin B, Korn EL, Conley BA, Abrams JS, McShane LM. Statistical and Practical Considerations for Clinical Evaluation of Predictive Biomarkers, J Natl Cancer Inst;2013;105:1677–1683

Qualitative Interaction (NSCLC)





No. of patients at risk

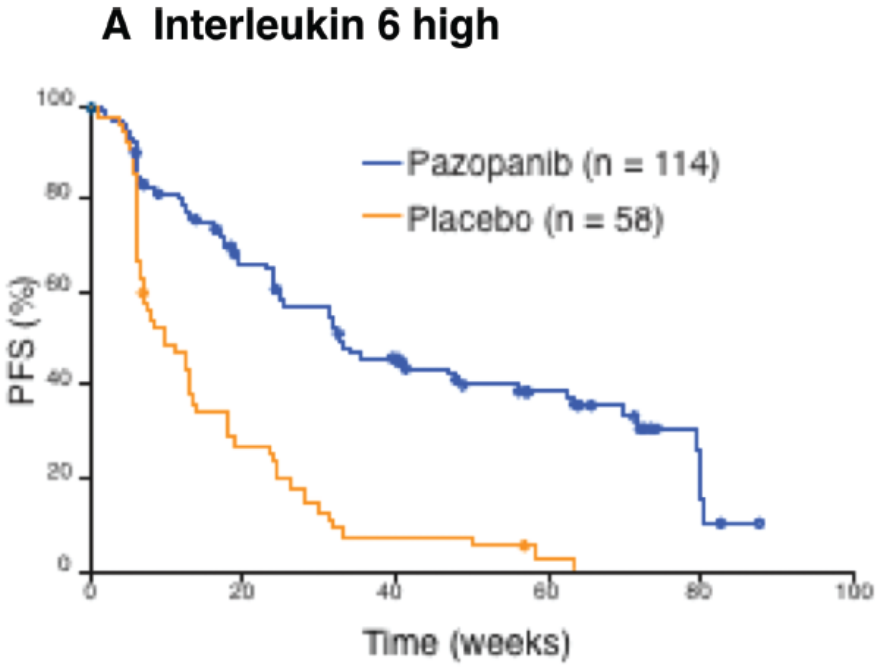
Time (months)	0	6	12	18	24	30
Chemotherapy	399	275	167	98	37	0
Chemotherapy plus cetuximab	377	254	164	93	29	3

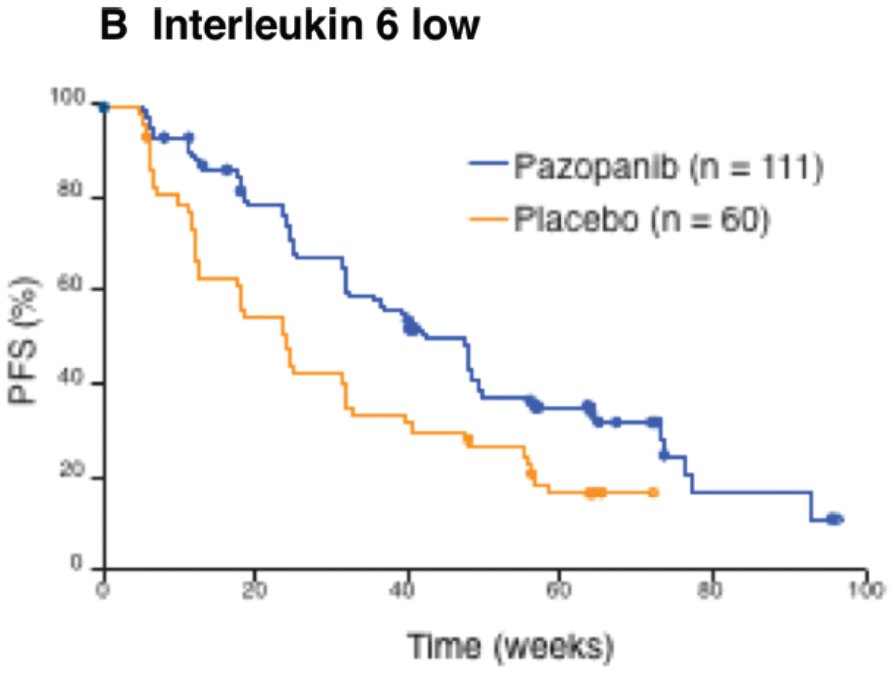
No. of patients at risk

Time (months)	0	6	12	18	24	30
Chemotherapy	167	108	58	36	11	0
Chemotherapy plus cetuximab	178	128	86	61	24	0

Figure 1. Examples of qualitative interactions. Gefitinib vs carboplatin + paclitaxel for first-line treatment of non–small cell lung cancer patients with *EGFR* mutation–positive tumors (**A**) and *EGFR* mutation–negative tumors (**B**) [adapted from Figure 2 of Mok et al (11). Reprinted with permission. Copyright 2009 Massachusetts Medical Society.]. Cetuximab + chempotherapy vs chemotherapy for first-line treatment of non–small cell lung cancer patients with high-expressing *EGFR* immunohistochemistry (IHC)–positive tumors (**C**) and low-expressing *EGFR* IHC–positive tumors (**D**) [adapted from Figure 4 of Pirker et al. (13). Reprinted with permission. Copyright 2012 Elsevier]. PFS = progression-free survival.

Quantitative Interaction (RCC)





No. of patients at risk

Time (weeks)	0	20	40	60	80	100
Pazopanib	114	64	42	25	4	0
Placebo	58	15	4	1	0	0

No. of patients at risk

Time (weeks)	0	20	40	60	80	100
Pazopanib	111	77	53	26	4	0
Placebo	60	31	19	8	0	0

Polley MC, Freidlin B, Korn EL, Conley BA, Abrams JS, McShane LM. Statistical and Practical Considerations for Clinical Evaluation of Predictive Biomarkers, J Natl Cancer Inst;2013;105:1677–1683

Quantitative Interaction (NSCLC)

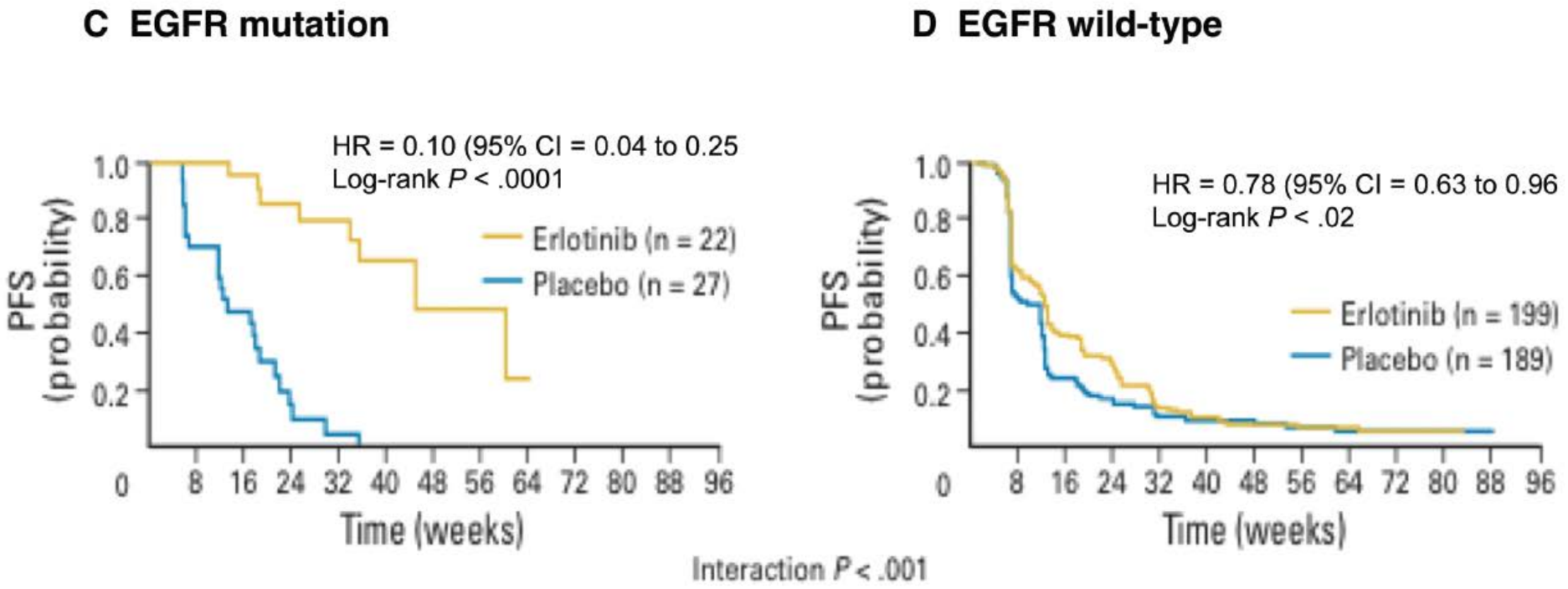


Figure 2. Examples of quantitative interaction: pazopanib vs placebo for locally advanced or metastatic renal cell carcinoma patients with high interleukin 6 (IL-6) values (**A**) and low IL-6 values (**B**) [adapted from Figure 2 of Tran et al. (14). Reprinted with permission. Copyright 2012 Elsevier]. Erlotinib maintenance therapy vs placebo for non–small cell lung cancer patients with *EGFR* mutation–positive tumors (**C**) and *EGFR* wild-type tumors (**D**) [adapted from Figure 3 of Brugger et al. (15). Reprinted with permission. Copyright 2011 American Society of Clinical Oncology]. Note that data were not available from Brugger et al. (15) to provide the number of patients at risk for (**C**) and (**D**). CI = confidence interval; HR = hazard ratio; PFS = progression-free survival.

References

Baker SG; Kramer BS; Sargent DJ; Bonetti M. Biomarkers, Subgroup Evaluation, and Clinical Trial Design. Discov Med. **2012**, 13 (*70*), 187-192.

Bossuyt PM; Lijmer JG; Mol BW. Randomised comparisons of medical tests: sometimes invalid, not always efficient. Lancet **2000**, 356 (*9244*), 1844-1847.

Buyse M; Michiels S. Omics-based clinical trial designs. Curr Opin Oncol. **2013**, 25 (*3*), 289–295.

Buyse M; Michiels S; Sargent DJ; Grothey A; Matheson A; de Gramont A. Integrating biomarkers in clinical trials. Expert Rev Mol Diagn. **2011**, 11(*2*), 171–182.

Freidlin; B.; McShane; L.M., Korn, E.L. Randomized Clinical Trials with Biomarkers: Design Issues. J Natl Cancer I. **2010**, 102 (*3*), 152-160.

Simon R. Stratification and partial ascertainment of biomarker value in biomarker driven clinical trials. J Biopharm Stat. **2014**, 24 (*5*), 1011-1021.

Simon R. Clinical trial designs for evaluating the medical utility of prognostic and predictive biomarkers in oncology. Personalized Medicine **2010**, 7 (*1*), 33-47.

Simon R; Maitournam A. Evaluating the Efficiency of Targeted Designs for Randomized Clinical Trials. Clin Cancer Res. **2004**, 10 (*20*), 6759–6763.

References

Pennello GA. Analytical and clinical evaluation of biomarkers assays: when are biomarkers ready for prime time? Clin Trials. **2013**, 10 (*5*), 666–676.

Pennello GA, Ye J. Companion Diagnostics. In *Encyclopedia of Biopharmaceutical Statistics 4th Ed.,* Ed. Shein-Chung Chow. CRC Press, **2017**.

Sharma A; Zhang G; Aslam S; Yu K; Chee M; Palma JF. Novel Approach for Clinical Validation of the cobas KRAS Mutation Test in Advanced Colorectal Cancer. Mol. Diagn. Ther. **2016**, 20 (*3*), 231–240.

Biomarker-Drug Trials

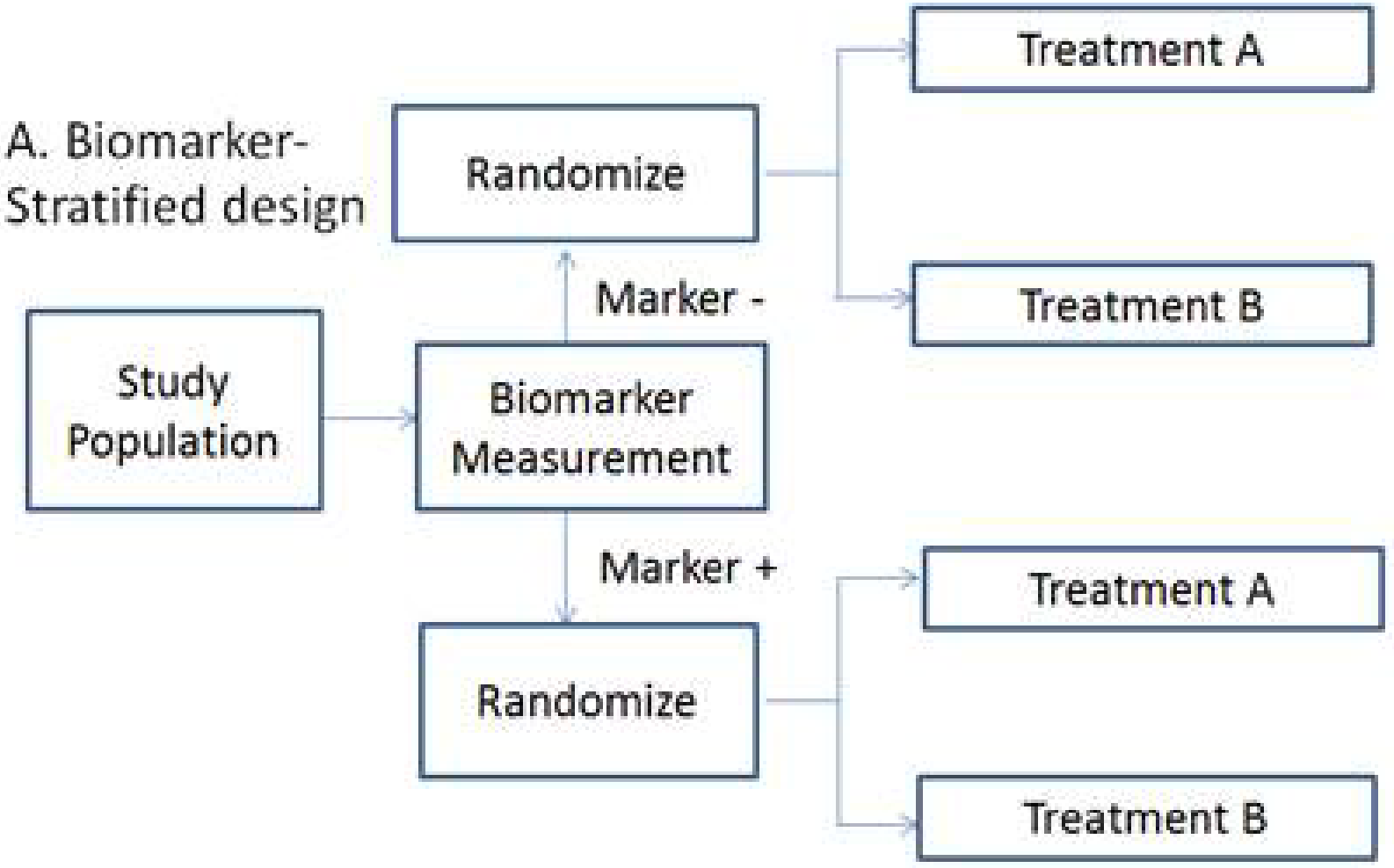
marker-positive, +

marker-negative, –

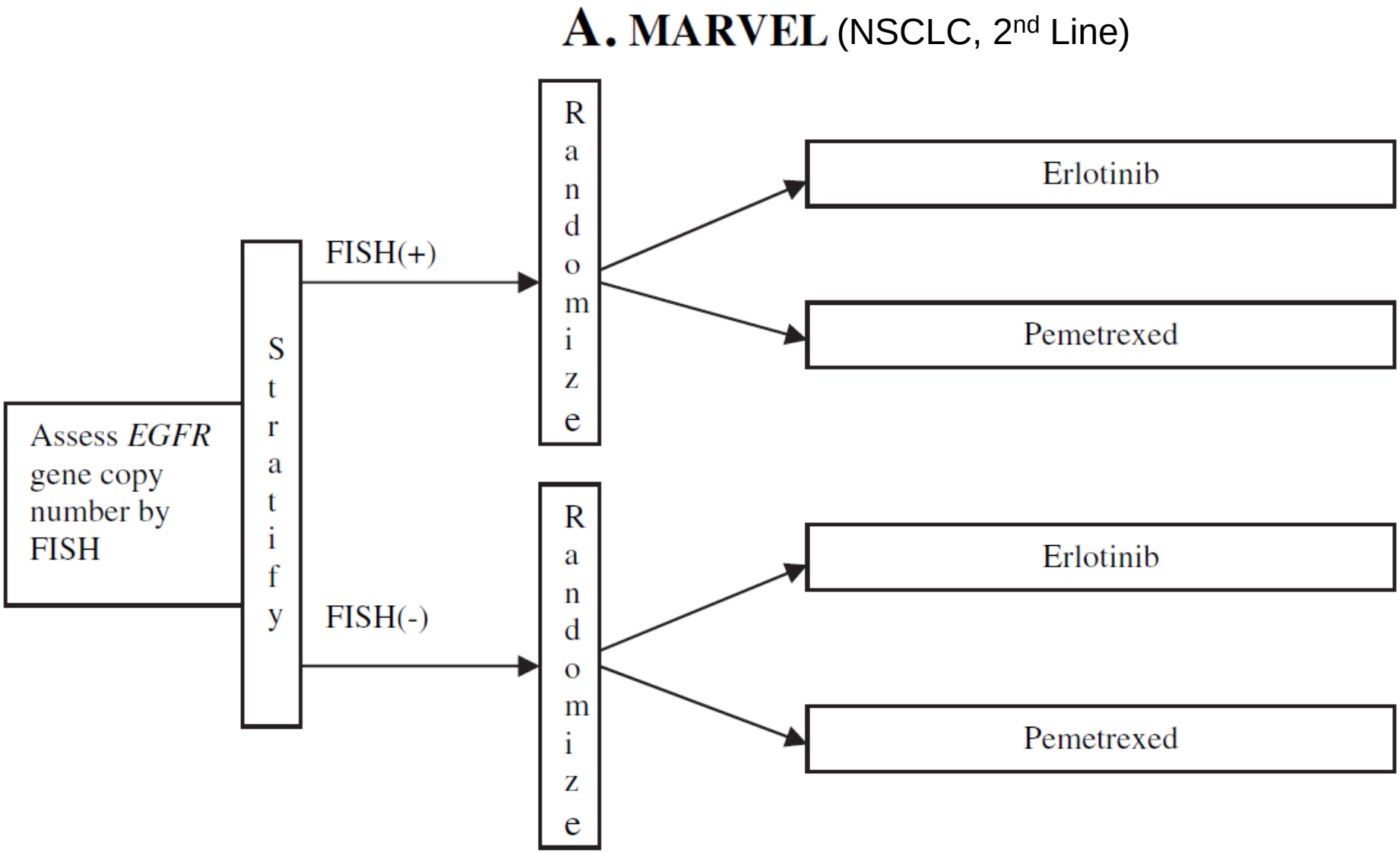
Treatment A is experimental arm.

Treatment B is control arm, typically standard-of-care or placebo.

Biomarker Stratified Design

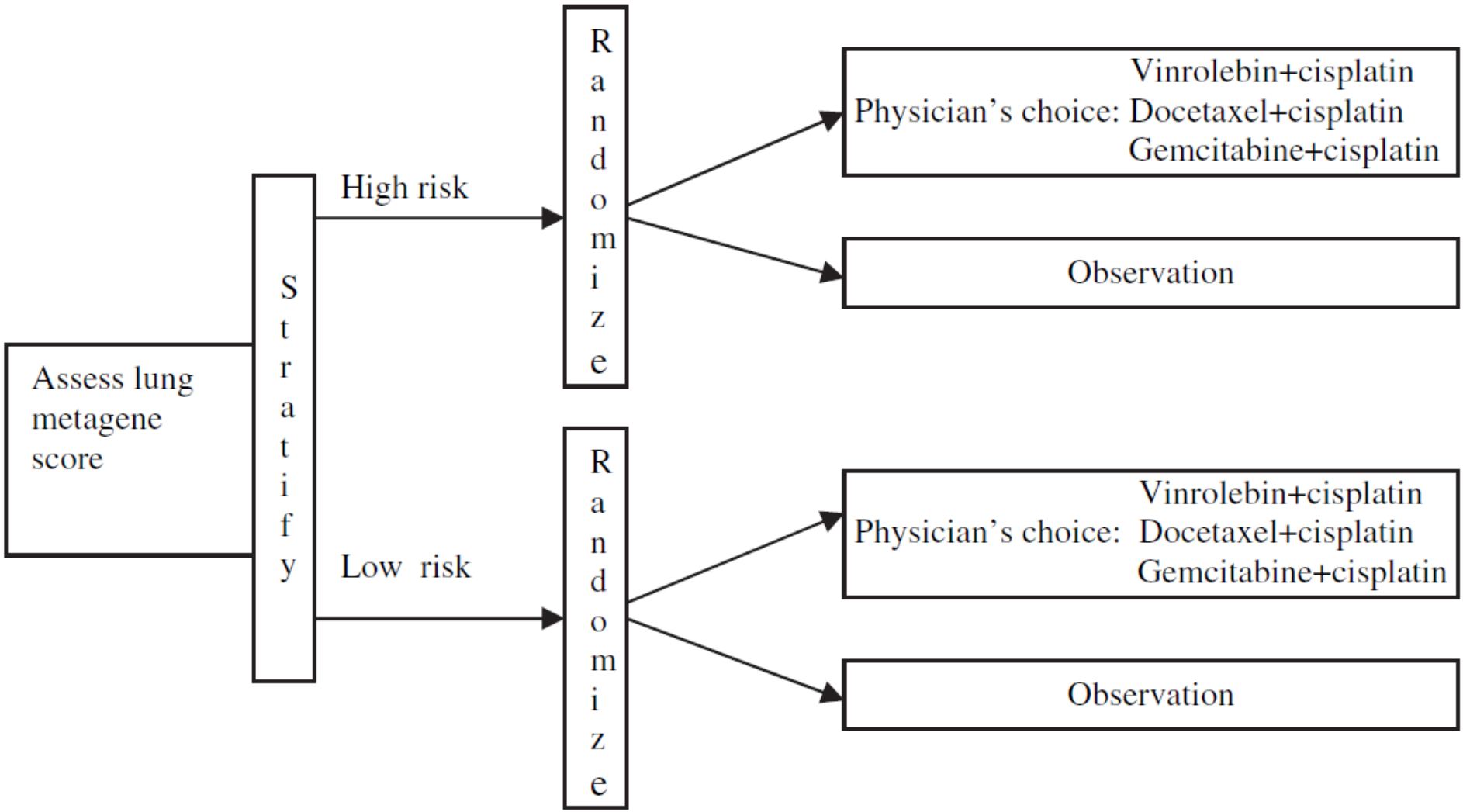


Freidlin; B.; McShane; L.M., Korn, E.L. Randomized Clinical Trials with Biomarkers: Design Issues. J Natl Cancer I. 2010, 102 (3), 152-160

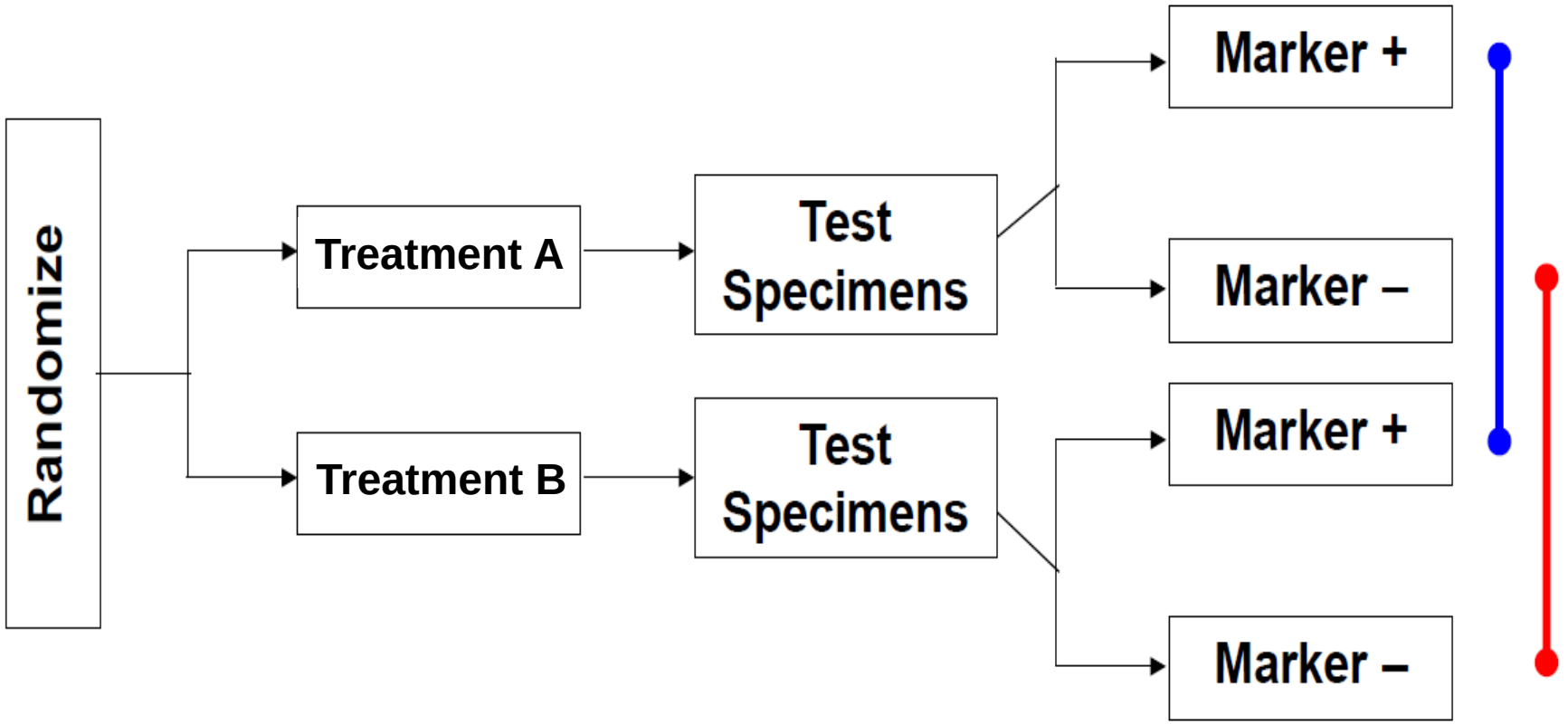


Freidlin; B.; McShane; L.M., Korn, E.L. Randomized Clinical Trials with Biomarkers: Design Issues. J Natl Cancer I. 2010, 102 (3), 152-160

B. CALGB-30506 (NSCLC, Stage 1)



Biomarker-Stratified Retrospective analysis of RCT



- Apply CDx biomarker test to stored specimens
- Evaluate treatment effect in CDx subgroups (– , +)

Prospective-Retrospective Biomarker-Stratified Design

Mack GS. FDA holds court on *post hoc* data linking *KRAS* status to drug response. Nature Biotech. **2009**; 27 (*2*) 110-112.

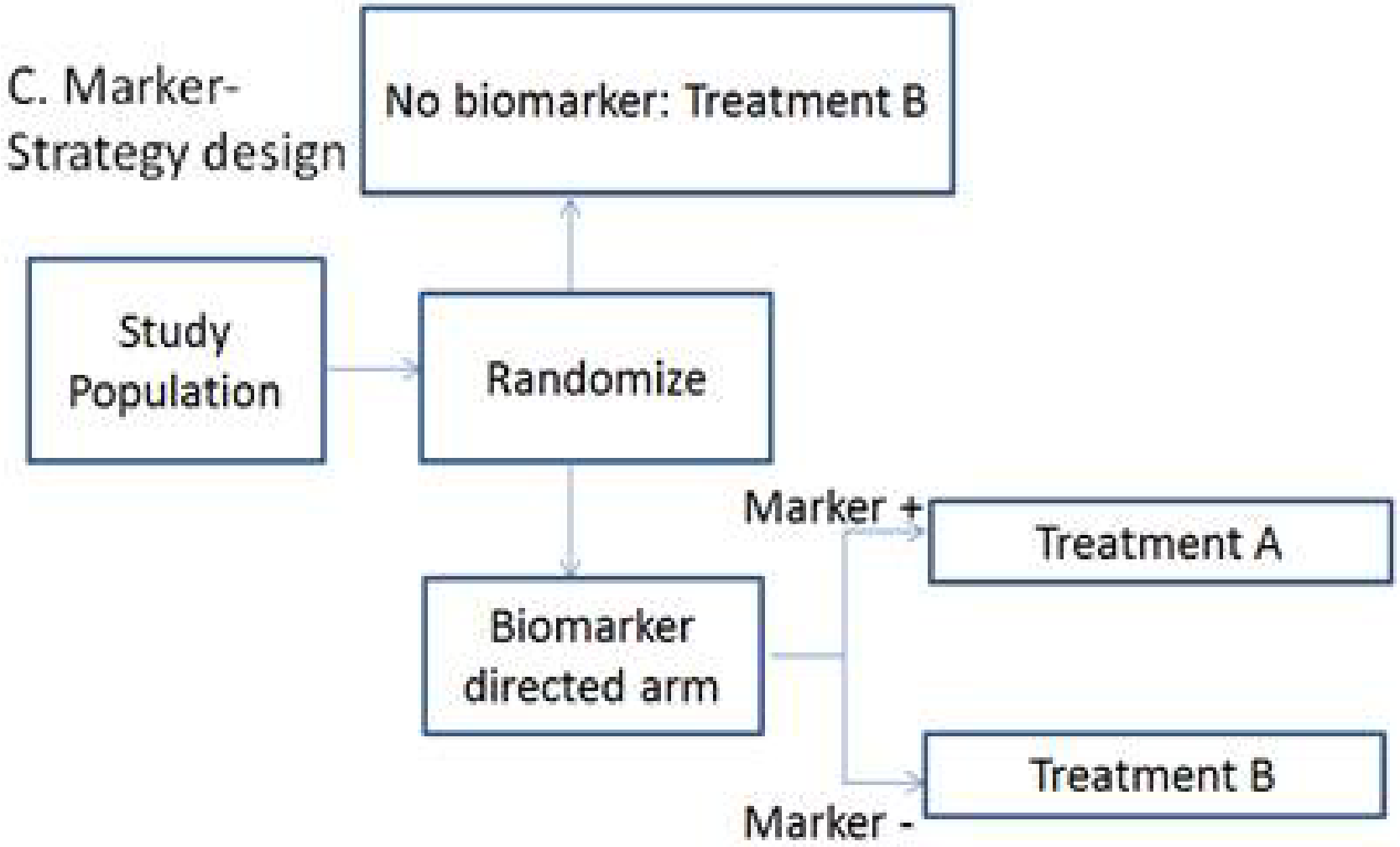
Ransohoff DF; Gourlay ML. Sources of bias in specimens for research about molecular markers for cancer. J Clin Oncol. **2010**, 28 (*4*), 698-704.

Simon RM; Paik S; Hayes DF. Use of archived specimens in evaluation of prognostic and predictive biomarkers. J Natl Cancer I. **2009**, 101 (*21*), 1446-1452.

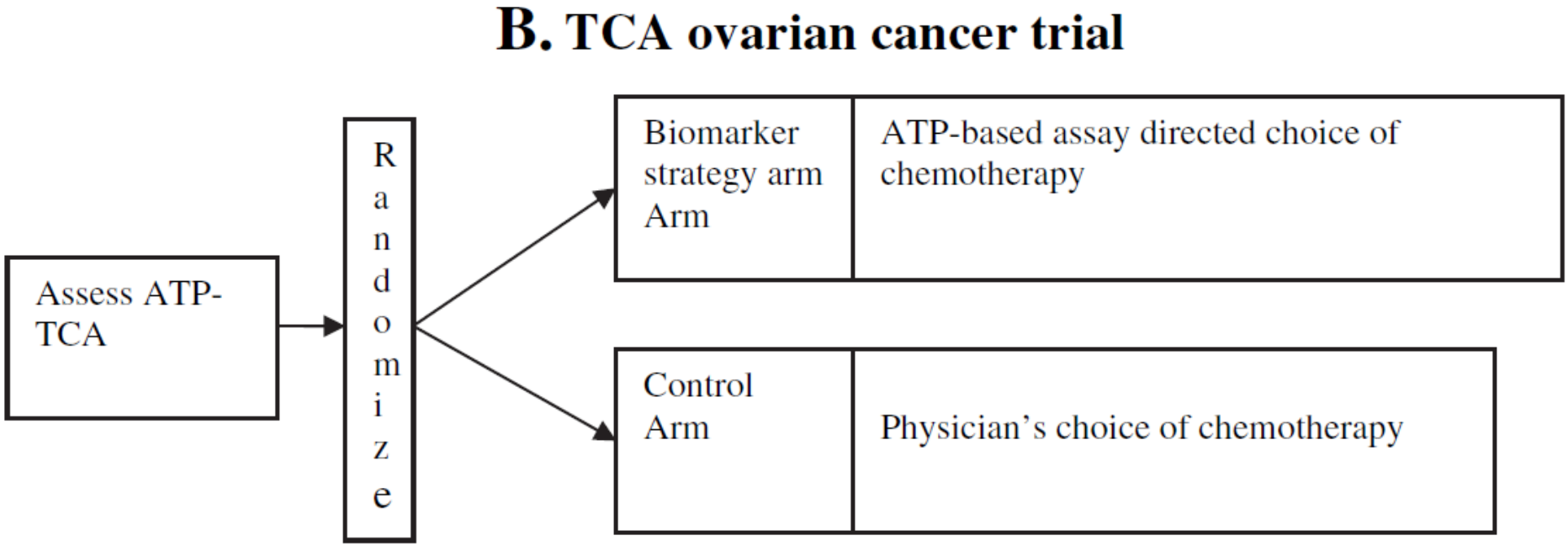
Simon R. Stratification and partial ascertainment of biomarker value in biomarker driven clinical trials. J Biopharm Stat. **2014**, 24 (*5*), 1011-1021.

US FDA. Principles for Codevelopment of an In Vitro Companion Diagnostic Device with a Therapeutic Product. US FDA: Silver Spring MD, **2016**.

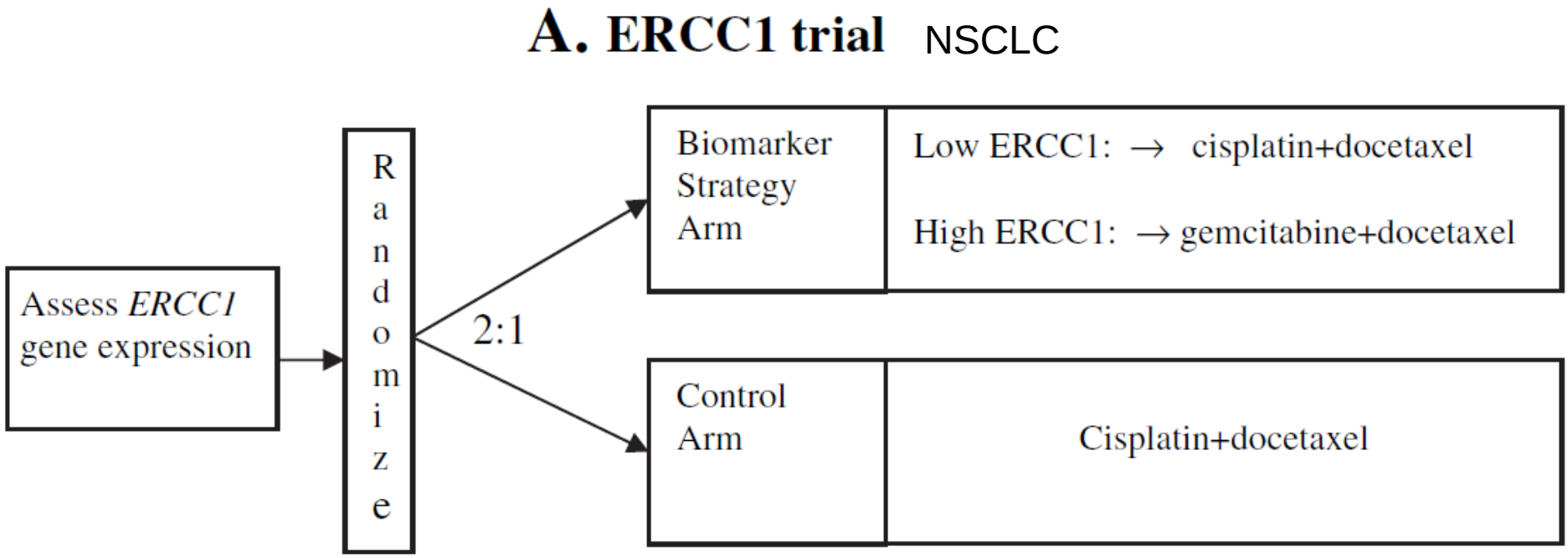
Biomarker Strategy Design



Freidlin; B.; McShane; L.M., Korn, E.L. Randomized Clinical Trials with Biomarkers: Design Issues. J Natl Cancer I. 2010, 102 (3), 152-160

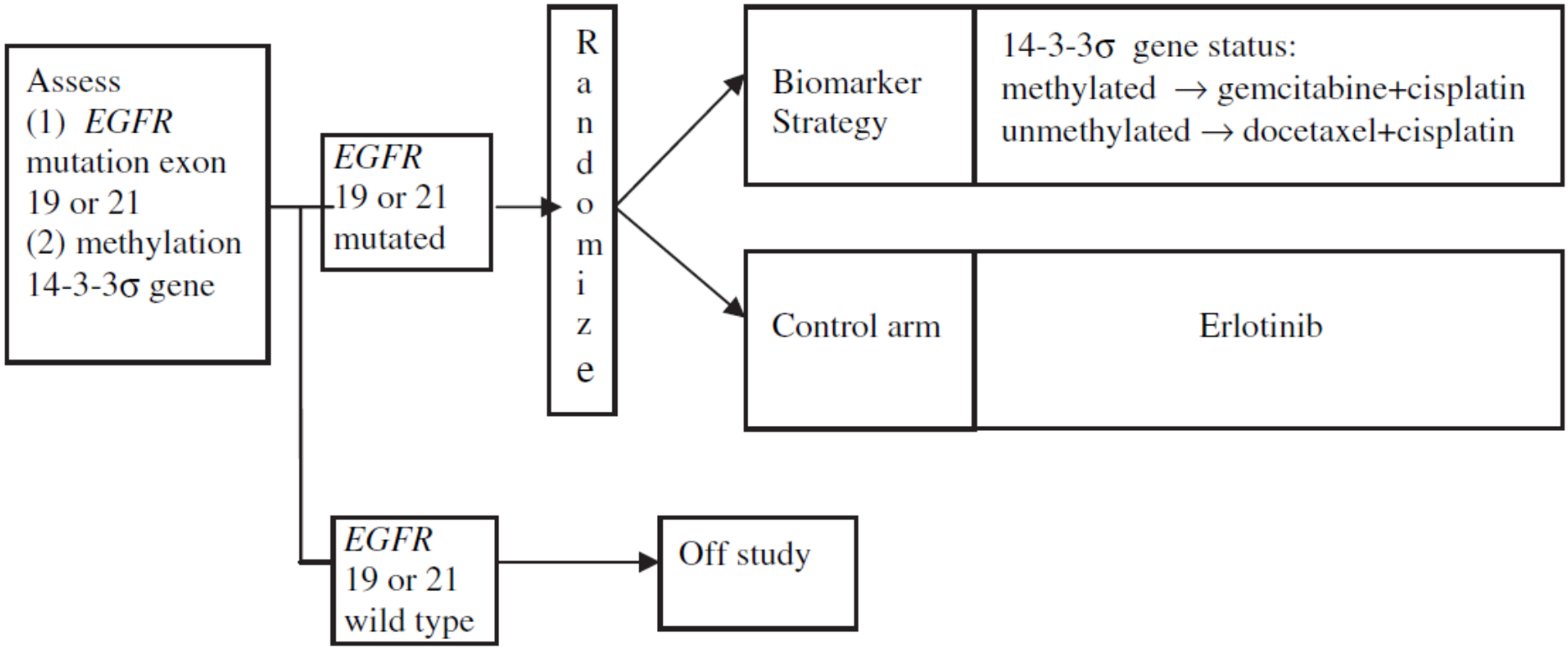


Freidlin; B.; McShane; L.M., Korn, E.L. Randomized Clinical Trials with Biomarkers: Design Issues. J Natl Cancer I. 2010, 102 (3), 152-160

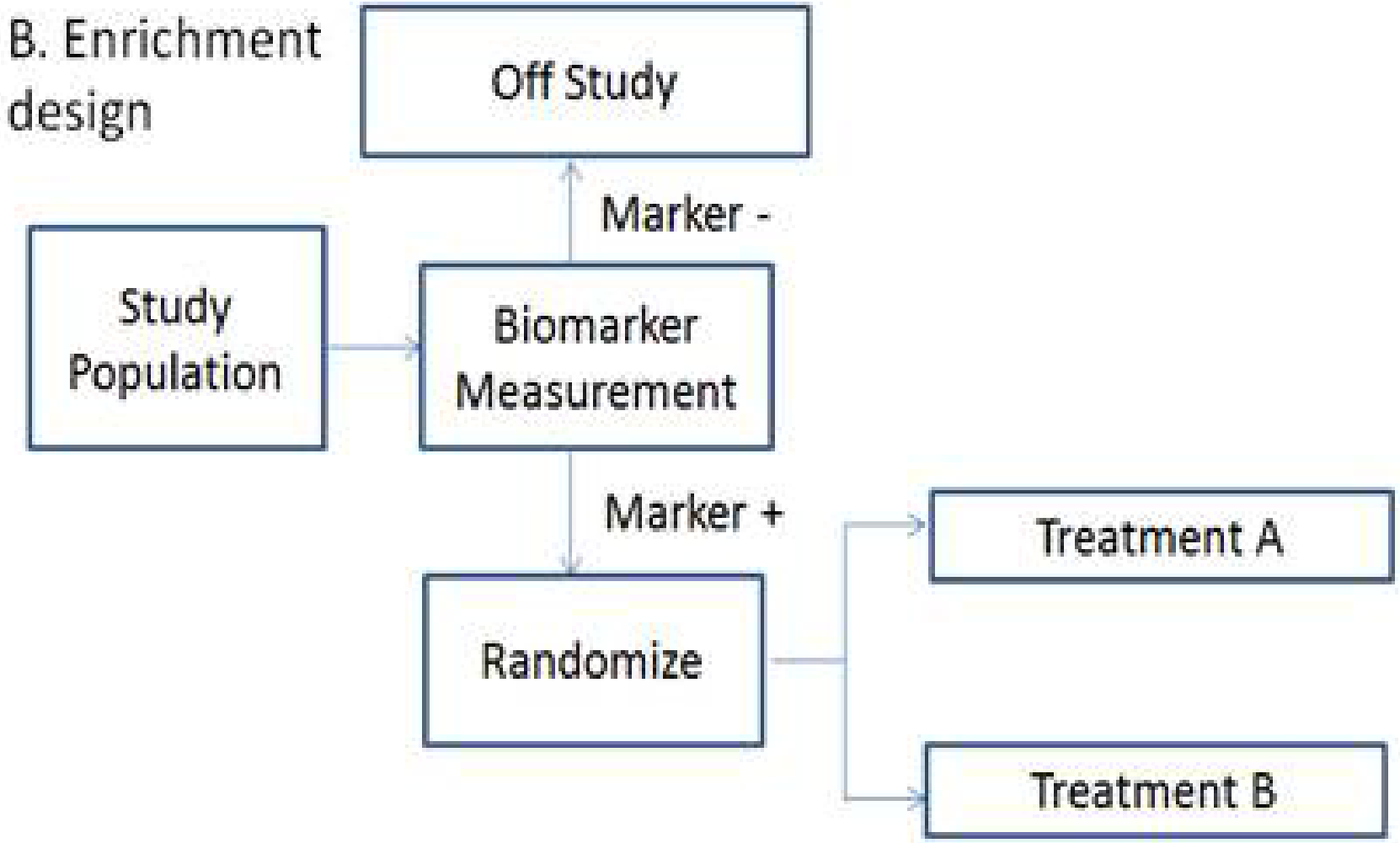


Freidlin; B.; McShane; L.M., Korn, E.L. Randomized Clinical Trials with Biomarkers: Design Issues. J Natl Cancer I. 2010, 102 (3), 152-160

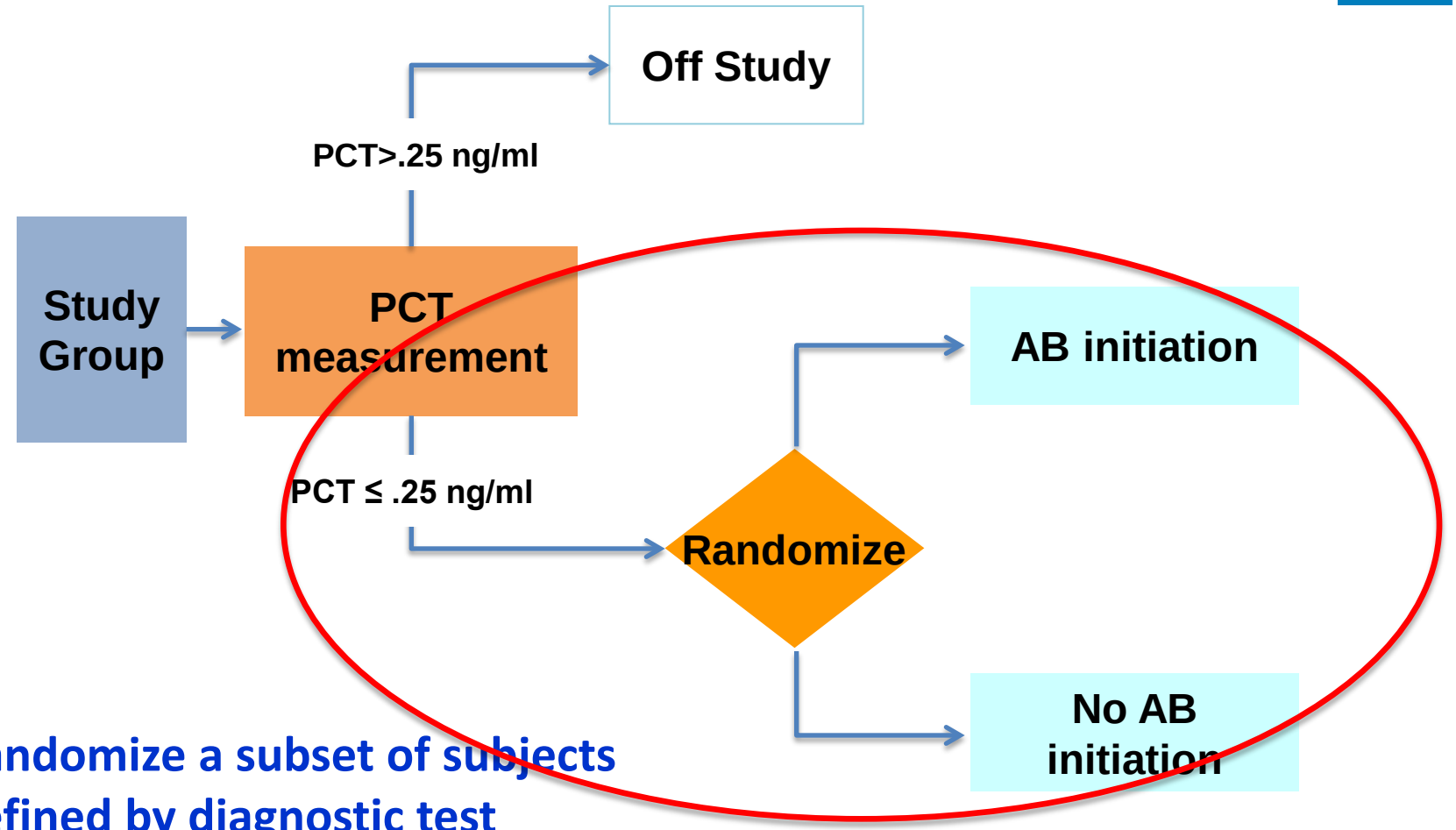
A. SLCG0601 NSCLC, Stage IV



Enrichment Design



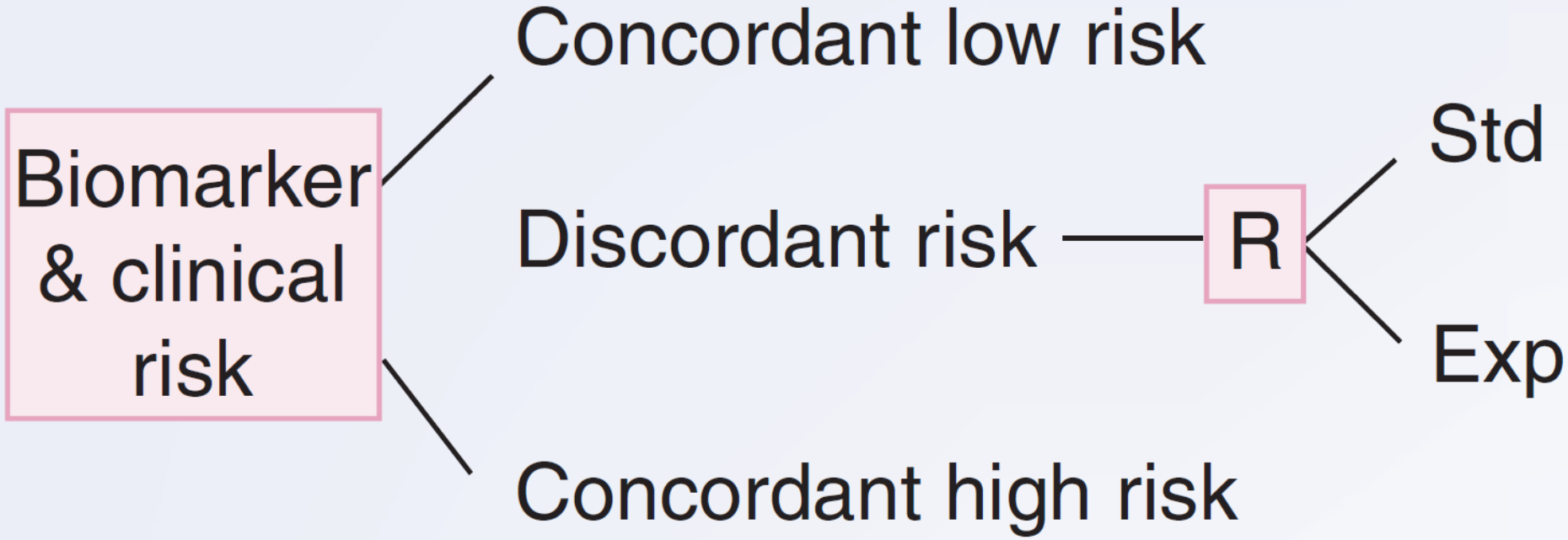
Enrichment Design



Randomize a subset of subjects defined by diagnostic test value (TRAP-LRTI on PCT ≤.1)

Targeted Reduction of Antibiotics using Procalcitonin in outpatient adults with suspect lower respiratory tract infection (TRAP-LRTI), VTEU PI Geeta K Swamy, MD, MPH; ARLG PI Ebbing Lautenbach, Vance Fowler, Henry “Chip” Chambers.

Discordant Risk Randomization



Buyse M; Michiels S; Sargent DJ; Grothey A; Matheson A; de Gramont A. Integrating biomarkers in clinical trials. Expert Rev Mol Diagn. **2011**, 11(*2*), 171–182.

Discordant Risk Randomization

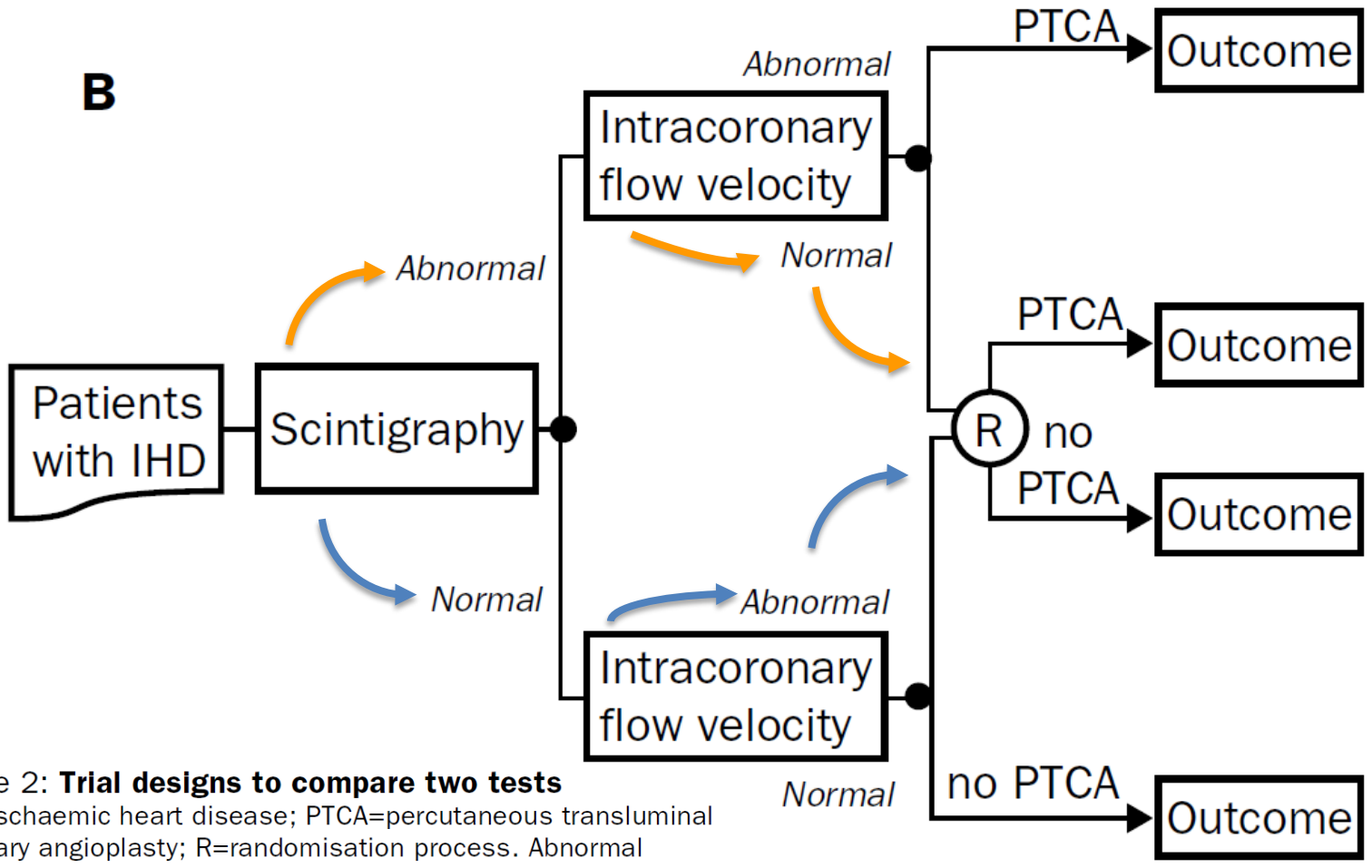


Figure 2: **Trial designs to compare two tests**
IHD=ischaemic heart disease; PTCA=percutaneous transluminal coronary angioplasty; R=randomisation process. Abnormal scintigraphy=reversible perfusion defect; abnormal intracoronary flow velocity=insufficient reserve.

Bossuyt PMM, Lijmer JG, Mol BWJ. Randomised comparisons of medical tests: sometimes invalid, not always efficient. *Lancet* 2000; 356: 1844–47

Notation

- $\theta_{ab} = E_{ab}(Y) =$ expectation of Y for treatment $A = a$, biomarker status $B = b$ $(A, B = 0,1)$.
  - objective response (0,1), event-free survival time
- $\theta^*_{at} = E_{at}(Y) =$ expectation of Y for treatment $A = a$, biomarker test result $T = t$ $(A, T = 0,1)$

$$\theta^*_{at} \overset{NDME}{=} \Sigma_{b=0}^{1} \theta_{ab} \Pr(B = b | T = t)$$

$$= \theta_{a0}(1 - p_t) + \theta_{a1} p_t,$$

$$p_t = \Pr(B = 1 | T = t)$$

Notation

- $\delta_b = \theta_{1b} - \theta_{0b} =$ treatment effect (difference) in parameter value between treatment arms $a = 0,1$ given biomarker status $B = b\ (= 0,1)$
- $\delta_t^* = \theta_{1t}^* - \theta_{0t}^* =$ treatment effect (difference) in parameter value between treatment arms $a = 0,1$ given test result $T = t\ (= 0,1)$
- $\Delta_{A.B} = \delta_1 - \delta_0 =$ predictive biomarker *capacity*.

Biomarker Stratified Design

$$\delta_t^* = \theta_{1t}^* - \theta_{0t}^*$$

$$\delta_b = \theta_{1b} - \theta_{0b}$$

Estimand $\Delta_{A.T}^* = \delta_1^* - \delta_0^*$

$$= (p_1 - p_0)(\delta_1 - \delta_0)$$

$$= (PPV + NPV - 1)\Delta_{A.B},$$

= treatment arm by biomarker interaction

$\Delta_{A.B}$ attenuated by the factor $PPV + NPV - 1$.

Biomarker Strategy Design

- Δ^*_{T-S}= difference between test-strategy arm and SoC arm in outcome Y. Let $\tau = \Pr(T = 1)$.

Estimand $\Delta^*_{T-S} = \tau\delta^*_1$

$$= \tau[\delta_0 + p_1(\delta_1 - \delta_0)]$$

$$\overset{\delta_1 = \delta_0 = \delta}{= \tau\delta}$$

$$\overset{\delta_0 = 0}{= \tau p_1 \delta_1 = \Pr(TP)\,\delta_1} \overset{p_1 = p_0 = p}{= \tau p \delta_1}$$

Remarks. Inefficient: δ^*_1 diluted by τ .

Invalid: Δ^*_{T-S} can be > 0 wo HTE or if test random

Enrichment Design

- $\Delta^*_{T+} = \delta^*_1 =$ treatment effect given test result $T = 1$:

Estimand $\Delta^*_{T+} = \delta^*_1$

$$= \delta_0 + p_1(\delta_1 - \delta_0)$$

$$\overset{\delta_1 = \delta_0 = \delta}{=} \delta$$

$$\overset{\delta_0 = 0}{=} p_1\delta_1 = PPV\,\delta_1 \overset{p_1 = p_0 = p}{=} p\delta_1$$

Remarks. Efficient: δ^*_1 not diluted by τ.

Invalid?: Δ^*_{T+} can be > 0 wo HTE or if test random

Discordant Risk Design

- $\Delta^*_{A.R}$ = difference in treatment effect between two discordant risk arms

Estimand $\Delta^*_{A.R} = \delta^*_{10} - \delta^*_{01}$

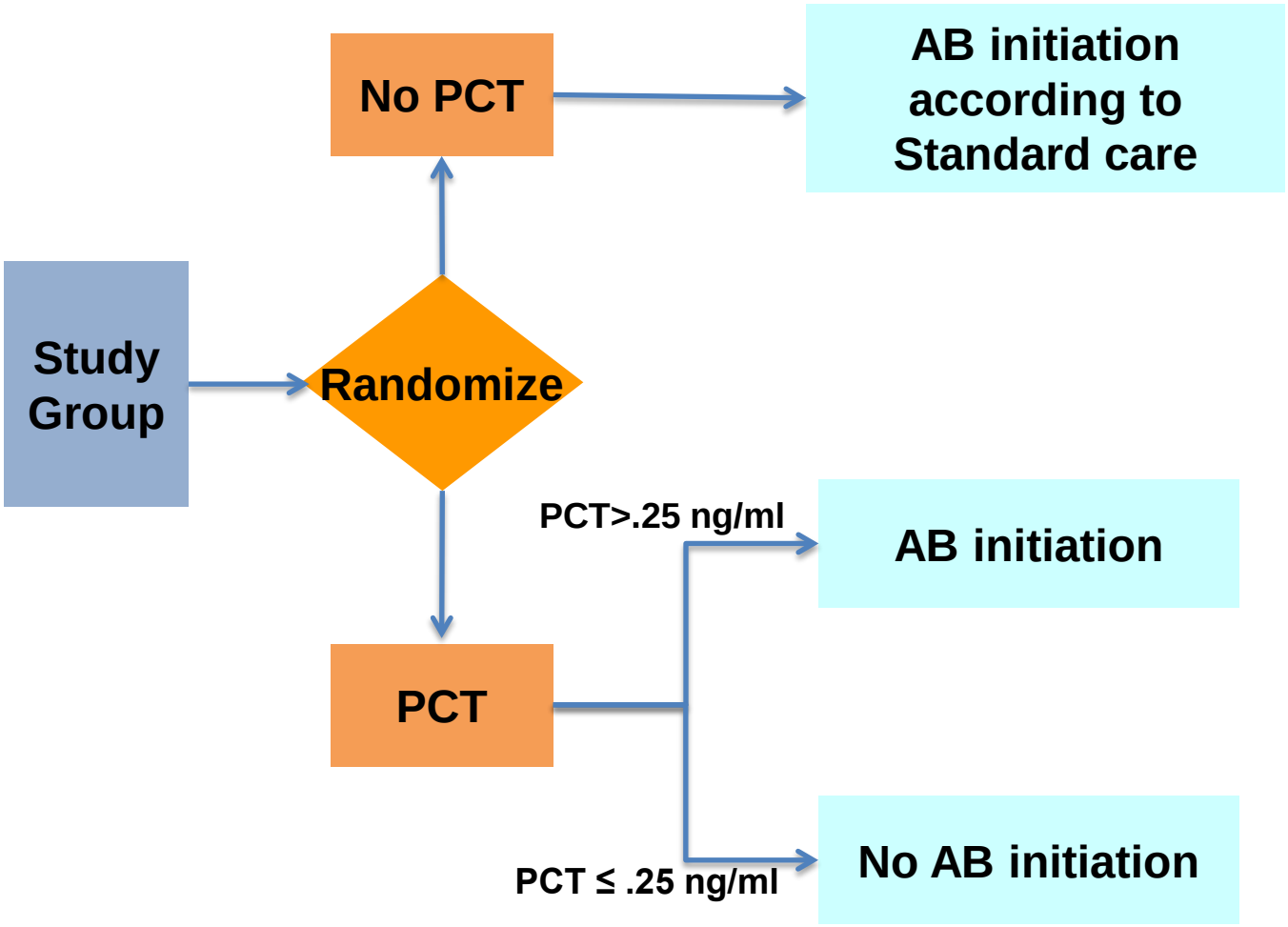
$$= (p_{10} - p_{01})(\delta_1 - \delta_0)$$

δ^*_{ts} = treatment effect given new and standard test results $T = t\ (= 0,1)$ and $S = s\ (= 0,1)$,

$p_{ts} = \Pr(B = 1 | S = s, T = t)$

Remarks. If T is not associated with B given S, then $p_{10} - p_{01} = p_{\bullet 0} - p_{\bullet 1} < 0$

Study Design of RCTs in Literature Marker Strategy Design



Ex. Length of Hospital Stay (Days)

	Infection Status		
PCT	Not Bact	Bact	Pr
$\leq 0.25(-)$	TN	FN	**0.75**
$> 0.25(+)$	FP	TP	**0.25**
Prev	**0.9**	**0.1**	
No AB	**6**	**18**	
AB	**6**	**6**	

Random Test

$1 - Sp = Se = 0.25$

$p_0 = p_1 = 0.10$

PCT Value	PCT Directed	AB Always
$\leq 0.25(-)$	TN(6) + FN (18) = 5.4	0.9(6)
$> 0.25(+)$	FP(6) + TP (6) = 1.5	0.1(6)
	6.9	6

Difference in mean length of stay between PCT-directed and AB always groups is $\Delta^*_{Dir-ABI} = \mathbf{6.9 - 6.0 = 0.9}$ days.

Ex. Length of Hospital Stay (Days)

	Infection Status		
PCT	Not Bact	Bact	Pr
$\leq 0.25(-)$	0.675	0.075	0.75
$> 0.25(+)$	0.225	0.025	0.25
Prev	0.9	0.1	
No AB	6	18	
AB	6	6	

Random Test

$1 - Sp = Se = 0.25$

$p_0 = p_1 = 0.10$

PCT Value	PCT Directed	AB Always
$\leq 0.25(-)$	0.675(6) + 0.075(18) = 5.4	0.9(6)
$> 0.25(+)$	0.225(6) + 0.025(6) = 1.5	0.1(6)
	6.9	6

Difference in mean length of stay between PCT-directed and AB always groups is $\Delta^*_{Dir-ABI} = \mathbf{6.9 - 6.0 = 0.9}$ days.

Ex. Length of Hospital Stay (Days)

	Infection Status		
PCT	Not Bact	Bact	Pr
$\leq 0.25(-)$	0.675	0.010	0.685
$> 0.25(+)$	0.225	0.090	0.315
Prev	0.9	0.1	
No AB	6	18	
AB	6	6	

Informative Test

$Sp = .25\ Se = 0.90$

$p_0 = .0146\ p_1 = .2857$

PCT Value	PCT Directed	AB Always
$\leq 0.25(-)$	0.675(6) + 0.010(18) = 4.23	0.9(6)
$> 0.25(+)$	0.225(6) + 0.090(6) = 1.89	0.1(6)
	6.12	6

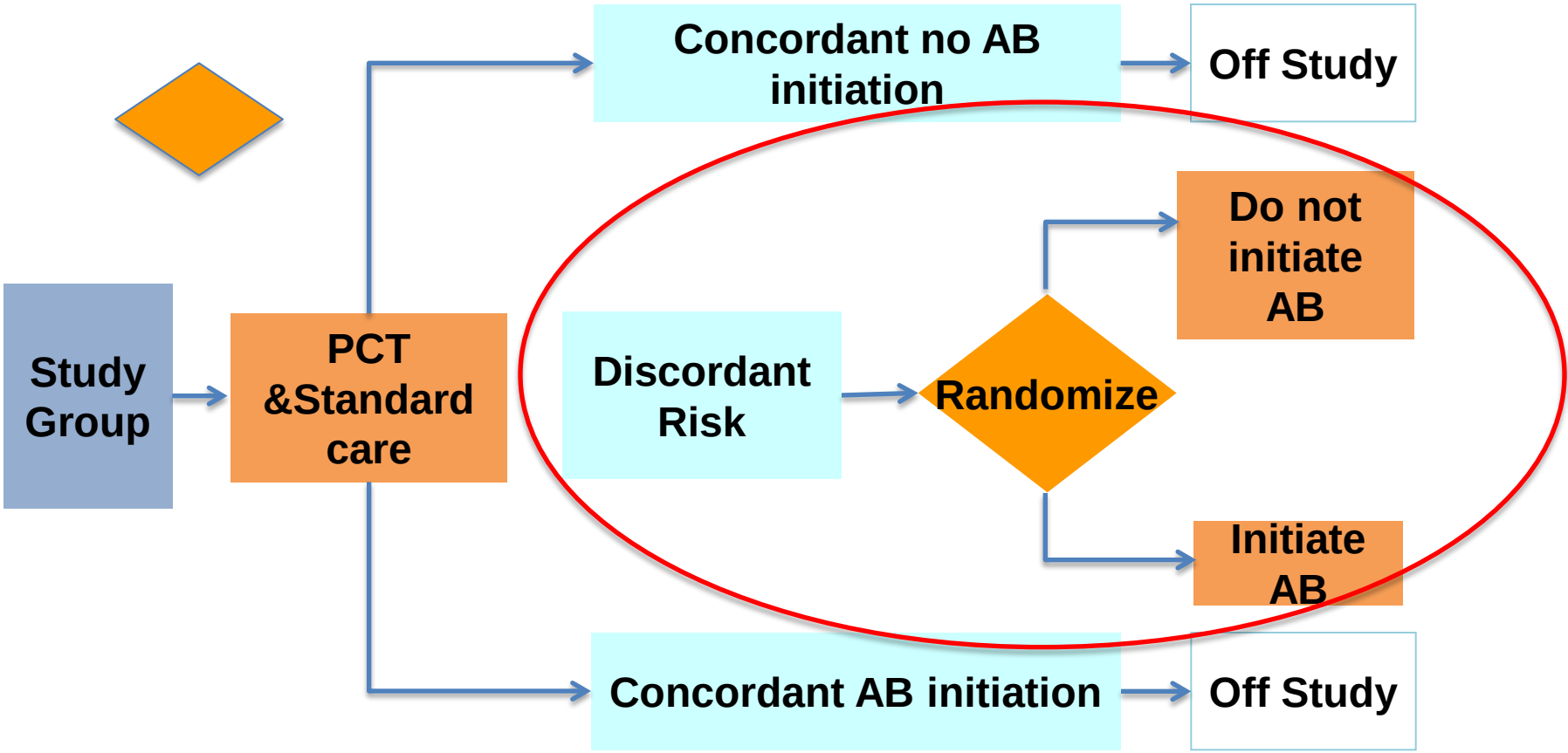
Difference in mean length of stay between PCT-directed and AB always groups is $\Delta^*_{Dir-ABI} = \mathbf{6.12 - 6.0 = 0.12}$ days.

Estimand Values, PCT EX.

Trial Design	Estimand	Perfect Test	Random Test	Decent Test
		$\tau_0 = 0$ $\tau_1 = 1$	$\tau_0 = 0.25$ $\tau_1 = 0.25$	$\tau_0 = 0.25$ $\tau_1 = 0.90$
Biomarker Capacity	$\delta_1' - \delta_0'$	$p_0 = 0$ $p_1 = 1$	$p_0 = 0.1$ $p_1 = 0.1$	$p_0 = .0146$ $p_1 = .2857$
Biomarker Stratified	$\Delta_{A.T}^{*\prime} = \delta_1^{*\prime} - \delta_0^{*\prime} =$ $(p_1 - p_0)(\delta_1' - \delta_0')$	12	0	3.25
Biomarker Strategy	$\Delta_{Dir-ABI}^{*} = (1-\tau)\delta_0^{*\prime}$ $= (1-\tau)[\delta_0' + p_0(\delta_1' - \delta_0')]$ $\overset{\delta_0'=0}{=} (1-\tau)p_0\delta_1' = \Pr(FN)\delta_1'$	0	0.9	0.12
Enrichment	$\Delta_{T-}^{*} = \delta_0^{*\prime} = \delta_0' + p_0(\delta_1' - \delta_0')$ $\overset{\delta_0'=0}{=} p_0\delta_1' = (1-NPV)\delta_1'$	0	1.2	0.1752

$\delta_t' = -\delta_t,\ \delta_t^{*\prime} = -\delta_t^*, \tau_b = \Pr(T = 1|B = b)$

Discordant Risk Randomization Design



Randomize subjects for whom test result and clinician disagree on treatment decision

Meeting Materials of the FDA's Microbiology Devices Panel, 10 November 2016.
https://www.fda.gov/advisorycommittees/committeesmeetingmaterials/medicaldevices/medicaldevicesadvisorycommittee/microbiologydevicespanel/ucm515517.htm

Key Subgroups for Adjunctive Tests

- Marker-strategy design compares PCT + SoC and SoC groups on whole population.
- Alternatively, the comparison can be restricted to those subgroups for whom PCT mattered (**changed** the treatment decision):

	SoC + PCT	
SoC	**no ABI**	**ABI**
no ABI	**No Change**	**Change**
ABI	**Change**	**No Change**

ABI = antibiotic initiation

Estimand Values, PCT EX.

Test	Test Positive Fractions	Predictive Value	Correlation
S	$\tau_0 = 0.40$ $\tau_1 = 0.90$	$p_0 = .0182$ $p_1 = .2000$	$\rho_0 = .0000$ $\rho_1 = .5556$
T	$\tau_0 = 0.25$ $\tau_1 = 0.90$	$p_0 = .0146$ $p_1 = .2857$	

Estimand	Value
p_{10}	0.0288
p_{01}	0.0146
$p_{10} - p_{01}$	0.0142
$\Delta^{*\prime}_{A.R} = (p_{10} - p_{01})(\delta'_1 - \delta'_0)$	0.1701

$\delta'_t = -\delta_t,\ \delta^{*\prime}_t = -\delta^*_t, \tau_b = \Pr(T = 1|B = b),\ \delta'_1 - \delta'_0 = 12$

Discussion

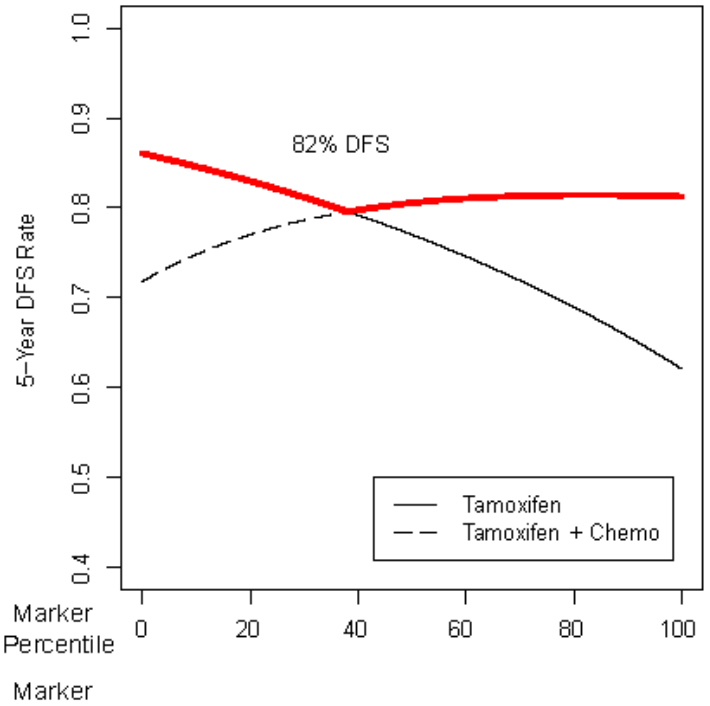
- Clinical trials are conducted in efforts to translate trial results to clinical practice.
- A predictive biomarker test has direct clinical consequences because it is essential for the safe and effective use of a corresponding therapeutic product.
- A predictive biomarker test for a therapeutic product should be evaluated using an estimand that provides a clear link between test results and the outcomes of treatment decisions.
- Unfortunately, some trial designs do not separate predictive accuracy of the test from treatment effect, leading to an inefficient, misleading, or otherwise uninterpretable estimand for evaluating the test / treatment combination.
- Some new performance measures (estimands) have similar interpretability problems.

Global measure of marker performance

Change in response rate under marker-based treatment:

$$\Theta = P(R = 1 | T = 1, \Delta(Y) > 0)\, P(\Delta(Y) > 0) + P(R = 1 | T = 0, \Delta(Y) < 0)\, P(\Delta(Y) < 0) - P(R = 1 | T = 1)$$
$$= 0.82 - 0.79 = 0.03$$

- 3% increase in 5-year DFS rate under marker-based treatment



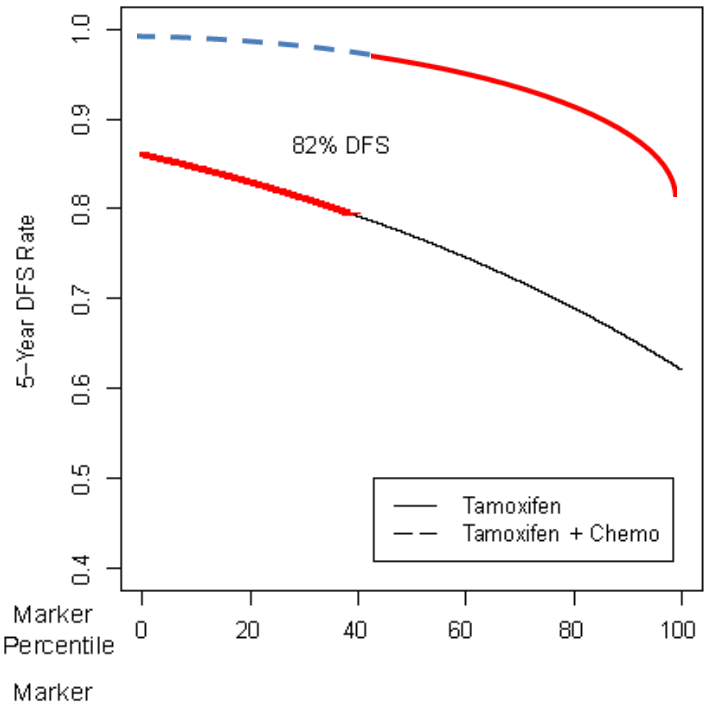
Janes, Brown, Pepe, Huang. Statistical Methods for Evaluating and Comparing Biomarkers for Patient Treatment Selection, UW Biostatistics Working Paper Series, 1-9-2013, Section 4.3.
http://biostats.bepress.com/uwbiostat/paper389/

Global measure of marker performance

Change in response rate under marker-based treatment:

$$\Theta = P(R = 1 | T = 1, \Delta(Y) > 0)\, P(\Delta(Y) > 0) + P(R = 1 | T = 0, \Delta(Y) < 0)\, P(\Delta(Y) < 0) - P(R = 1 | T = 1)$$
$$= 0.82 - 0.79 = 0.03$$

- 3% increase in 5-year DFS rate under marker-based treatment



Janes, Brown, Pepe, Huang. Statistical Methods for Evaluating and Comparing Biomarkers for Patient Treatment Selection, UW Biostatistics Working Paper Series, 1-9-2013, Section 4.3.
http://biostats.bepress.com/uwbiostat/paper389/