Bayesian Design in Master Protocols

Kun He

R&G US INC

November 8, 2019

- **Bayesian versus Frequentist**
- **NCT02034110 Trial**
- **I-SPY 2**
- **GBM AGILE**

Bayesian versus Frequentist Approaches in Clinical Trials

Variable	Bayes	Frequentist
Differences		
Main goal of inference	Predict outcomes of future trials and absolute risk for future patients.	Estimate population average effects.
Assumptions	Requires explicit specification of prior distributions of unknown population parameters. Incorporates a priori knowledge and clinical judgment formally. May be sensitive to specification of prior distributions.	Does not require explicit specification of prior distributions of unknown population parameters. Incorporates a priori knowledge and clinical judgment informally.
Interim monitoring	Only the data actually obtained are relevant for final conclusions (e.g., a credible interval or predictive probability). Whether or not a clinician examines accumulating evidence with the possibility of stopping the trial does not affect inference.	Both the data actually obtained and the probabilities of data not obtained are relevant for final conclusions (e.g., a P value). Whether or not a clinician examines accumulating evidence with the possibility of stopping the trial does affect inference.
Ease of use	Often computationally complex; careful modeling often requires simulation-based calculations.	Often computationally simple, though careful modeling may require simulation-based calculations.

Bayesian versus Frequentist Approaches in Clinical Trials

Similarities	
Adaptation	Can incorporate adaptive designs, multistage trials, early stopping, and adaptive randomization.
Role of statistical judgment	Options for data-driven analyses are available. Skill and substance-area knowledge of the data analyst are important in drawing correct conclusions.
Compatibility	It is feasible to combine a Bayesian design with a frequentist analysis or a frequentist design with a Bayesian analysis.
Prior knowledge	Both approaches rely on prior knowledge and clinical judgment (though they incorporate them in different fashions).

Harrington et al, NEJM, 2016

Bayesian Method

- 2004 CDER Bayes Conference at NIH; Special issue of Clinical Trials (2005)
- CDER is committed to exploring the use of Bayes and other novel approaches for trial designs under PDUFA VI and 21st Century Cures → pilot program
- Bayesian methods have been accepted in Oncology for
  - Phase I dose finding
  - Phase II hypothesis generating
  - Phase III futility or exploratory analyses
  - Pediatric trials
  - Master Protocols

NCT02034110 – Dabrafenib + Trametinib

Study Type :	**Interventional (Clinical Trial)**
Actual Enrollment :	**206 participants**
Intervention Model:	**Parallel Assignment**
Intervention Model Description:	**9 indications**
Masking:	**None (Open Label)**
Primary Purpose:	**Treatment**
Official Title:	**A Phase II, Open-label, Study in Subjects With BRAF V600E-Mutated Rare Cancers With Several Histologies to Investigate the Clinical Efficacy and Safety of the Combination Therapy of Dabrafenib and Trametinib**
Actual Study Start Date :	**March 12, 2014**
Estimated Primary Completion Date :	**June 29, 2020**
Estimated Study Completion Date :	**June 29, 2020**

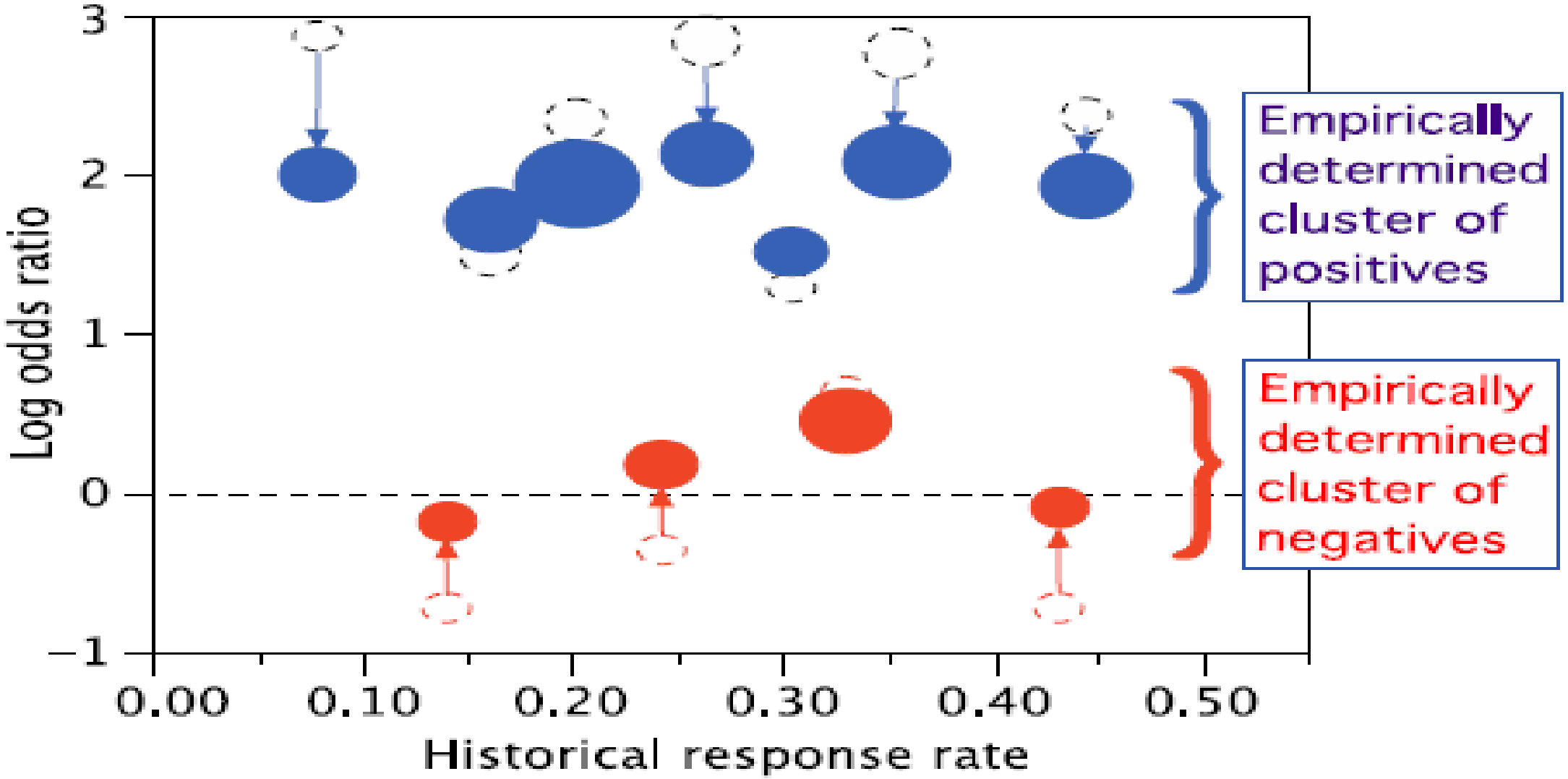
NCT02034110 Incidence Rates

Histology	Overall Incidence Rates in US (2011)	BRAF V^{600E} Mutation Rate
Anaplastic thyroid cancer (ATC)	0.10/100,000	24%
Biliary Tract Cancer (BTC)	0.6/100,000	7 - 30%
Diffuse Large B Cell Lymphoma (DLBCL)	9.17/100,000	4%
Gastrointestinal stromal tumor (GIST)	0.7 - 1.1/100,000	2 - 5%
Germ Cell Tumor (GCT) (~50% non-seminomatous)	6.31/100,000 (white males) 1.38/100,000 (black males)	3%
High-Grade Cerebral Glioma (HGG)	2 - 4/100,000	~3% (GBM)
Hairy Cell Leukemia (HCL)	0.33/100,000	90 - 100%
Multiple Myeloma (MM)	5.579/100,000	4%
Adenocarcinoma of Small Intestine	0.073/100,000	~10%

NCT02034110 Design

- Treatment: dabrafenib + tramentinib
- Single arm with 9 histology
- Primary endpoint: ORR
- Statistical method: Bayesian hierarchical modeling

NCT02034110 Bayesian Hierarchical Model



Berry et al CT 2013

Indication Finder (Tumor Agnostic)

- We observe multiple subtypes of a disease.
  - Likelihood within each subtype $F(y|\theta_g)$, where y is a vector $y_1,\ldots,y_{ng}$ for the subjects within the subtype.
- We relate the subtypes through a hierarchical model
  - Thus, $\theta_1,\ldots,\theta_G \sim H$.
  - The structure of H is key to the borrowing behavior of the model, can range from no borrowing to complete pooling.

Common choices for H

We have $\theta_1,\ldots,\theta_G \sim H$

- If H specifies independent draws from a fixed distribution, we have no borrowing, all subtypes are treated separately.
- if H specifies $\theta_1=\ldots=\theta_G=\theta$ (θ typically with a fixed prior distribution), then all subtypes are pooled.
- if $H\sim N(\mu,\tau)$ with nondegenerate priors on μ and τ, we acquire the model used in agnostic (degenerate priors revert a fixed distribution and separate analysis). Here τ is the key parameter for borrowing.

Pros/Cons of Agnostic

$\theta_1,\ldots,\theta_G \sim N(\mu,\tau)$ with priors on μ and τ.

- Advantages
  - allows dynamic borrowing between subgroups based on estimation of the hyper-parameters, particularly τ.
  - less type I error in certain cases, more power, smaller sample sizes, etc. compared to separate trials.
- Disadvantages
  - Like all models, it could be wrong. $\theta_1,\ldots,\theta_G \sim N(\mu,\tau)$ does not allow for outlying subtypes or clusters.
  - can increase misclassification error (compared to separate trials) in certain cases.

Clustering Models

- Typical situations where we see increased misclassification errors in Agnostic are “cluster” situations
  - drug works well in some subtypes, AND
  - drug doesn’t work at all in others.
- Note this is a HARD problem, pooling does far worse than Agnostic in this setting.

NCT02034110 Two Level Models

- Top level clusters histology (could be one cluster, two, or many)
- Conditional on clustering, model borrows information within a cluster, but not across clusters
- Goal of model is to recognize which histology are similar and borrow between similar histology more than between dissimilar histology
- If the data for the histology within a cluster are quite similar, borrow extensively within the cluster. Otherwise adjust and borrow minimally.

NCT02034110 Bayesian Hierarchical Analysis

- Allows for the possibility that the response profile for the populations of histology may be heterogeneous or homogeneous. There may be a ‘cluster’ of histology in which the combination is effective
- Borrows information in a limited sense, especially from histology that demonstrate similar response rates
- Design is data-driven; the number of clusters used is based on the observed number of responses (and pre-specified model)
- Study is ongoing

NCT02034110 ATC Cohort

Table 16. Efficacy Results in the ATC Cohort Based on Independent Review of Study BRF117019

ATC Cohort Population (evaluable for response)	**n = 23**
Objective Response Rate (ORR)	
ORR (95% CI)[a]	61% (39%, 80%)
Complete Response Rate	4%
Partial Response Rate	57%
Duration of Response (DOR)	
% with DOR ≥6 months	64%

[a]CI = Confidence interval

[b] NE=Not estimable

I-SPY 2 Trial

Investigation of **S**erial Studies to **P**redict **Y**our **T**herapeutic **R**esponse with **I**maging **A**nd mo**L**ecular Analysis 2

I-SPY2 Background

- Breast cancer (BC) diagnosed in ~200,000 women annually in U.S.
- 45,000 women die annually of BC
- 10-20% of newly diagnosed BC present as locally advanced BC (LABC)
  - At high risk of recurrence
- Standard of care for women with LABC increasingly includes neoadjuvant therapy prior to surgical resection

I-SPY1 Trial

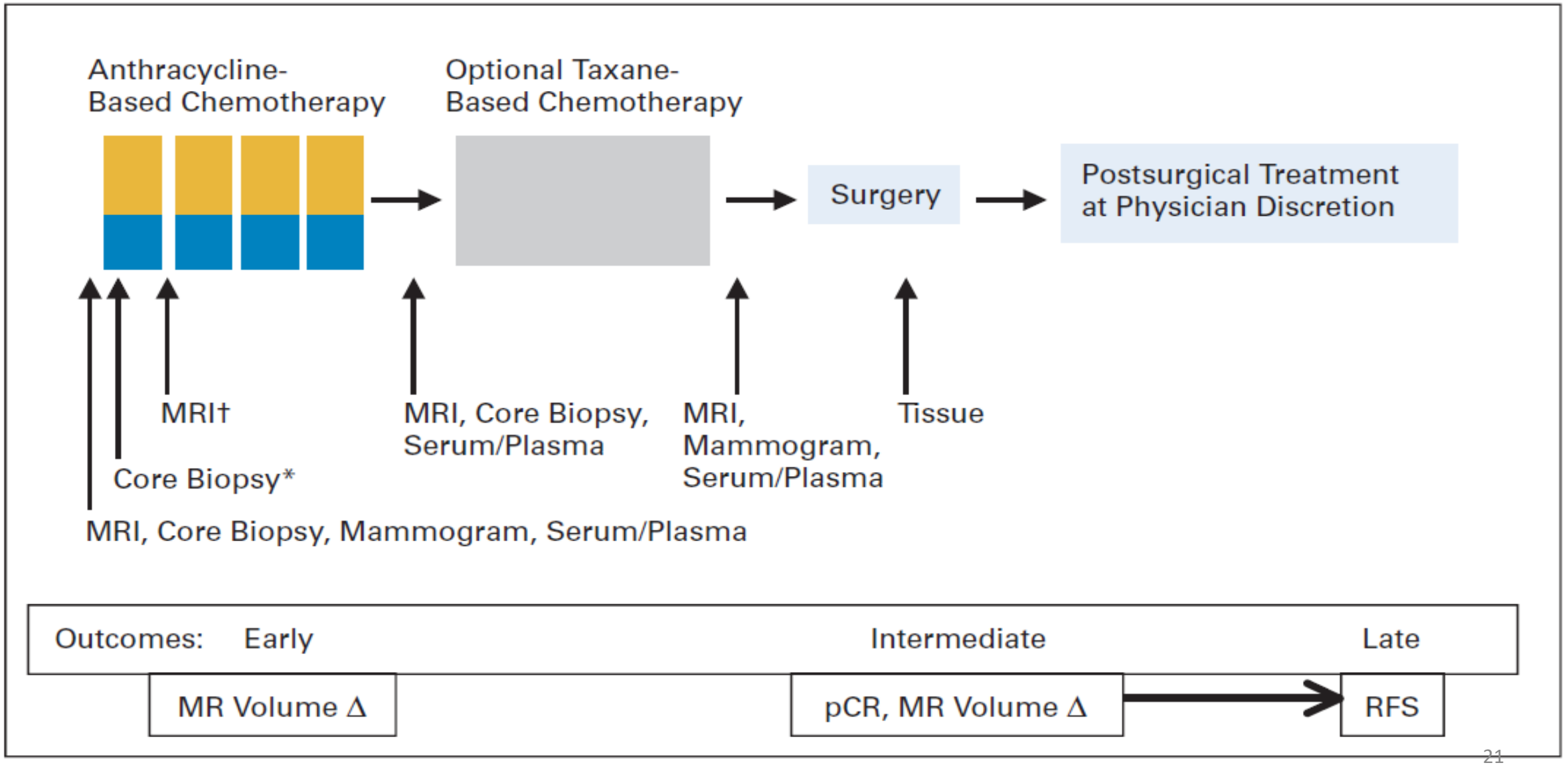
- Inter-SPORE collaboration (NCI Specialized Programs of Research Excellence):
  - American College of Radiology Imaging Network (ACRIN)
  - Cancer and Leukemia Group B (CALGB)
  - NCI Center for Biomedical Informatics and Information Technology (CBIIT)

Esserman et al JCO 2012

I-SPY1 Goal

- Intent was to evaluate and identify biomarkers of early response to standard chemotherapy
- All patients received neoadjuvant chemotherapy to test a comprehensive set of biomarkers for their ability to predict tumor response
- Predictors of 3-year survival:
  - Early endpoints: MRI changes; changes in gene expression
  - Intermediate endpoint: pCR rate
  - Longer term endpoint: 3-year RFS (relapse-free survival)

I-SPY1 Backbone



I-SPY1 Patient Population

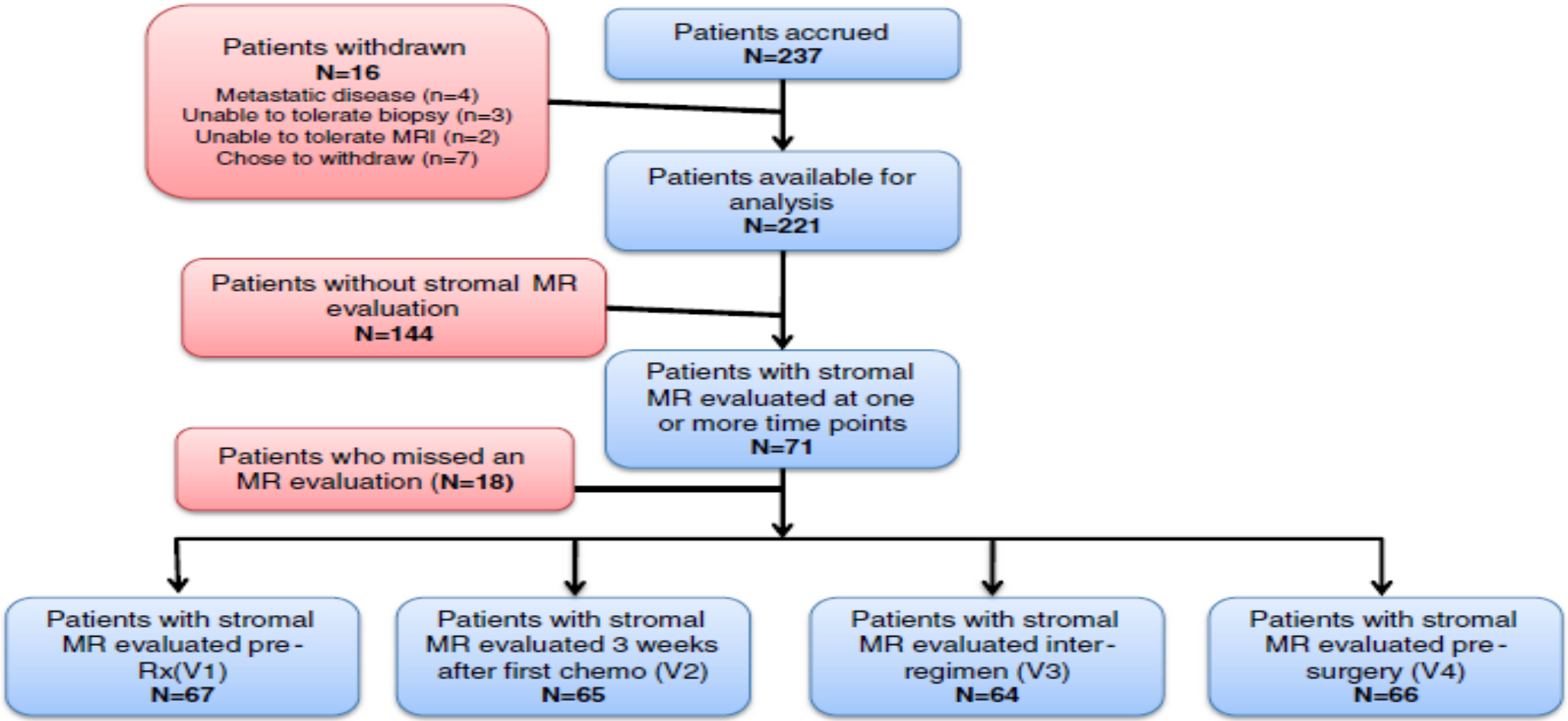


Fig. 3 Consort diagram for the study. Patients were accrued to ISPY 1 ($n = 237$). Of these, 221 were available for analysis, and 71 patients had MRIs that were assessable for stromal enhancement. The study protocol was to obtain four MRIs for each patient at V1 (prior to chemo); V2 (after first chemocycle); V3 (between AC and T chemotherapy); V4 (prior to surgery).

I-SPY1 Results

- 215 / 237 (91%) patients had pathologic assessment available for analysis
- Mean tumor size = 6.0 cm; minimum = 3.0 cm
- pCR rate = 27%
- 36% had RCB (residual cancer burden) 0 or 1
  - RCB is a more complex and detailed pathologic evaluation; formula includes 6 variables
- pCR and RCB were predictive of 3-year RFS with 3.9 years mean follow-up (p=0.04 and 0.01, respectively)

I-SPY1 Conclusions

- Improvement on pCR or RCB may be a rapid way to screen for effectiveness of new targeting agents
- Most informative way to interpret results will be by combining pCR and RCB evaluations with molecular subgroup analysis
- MRI volume change is strong predictor of pCR
  - Hypothesize that MRI volume change can non-invasively determine response to new agents

I-SPY2 Goal

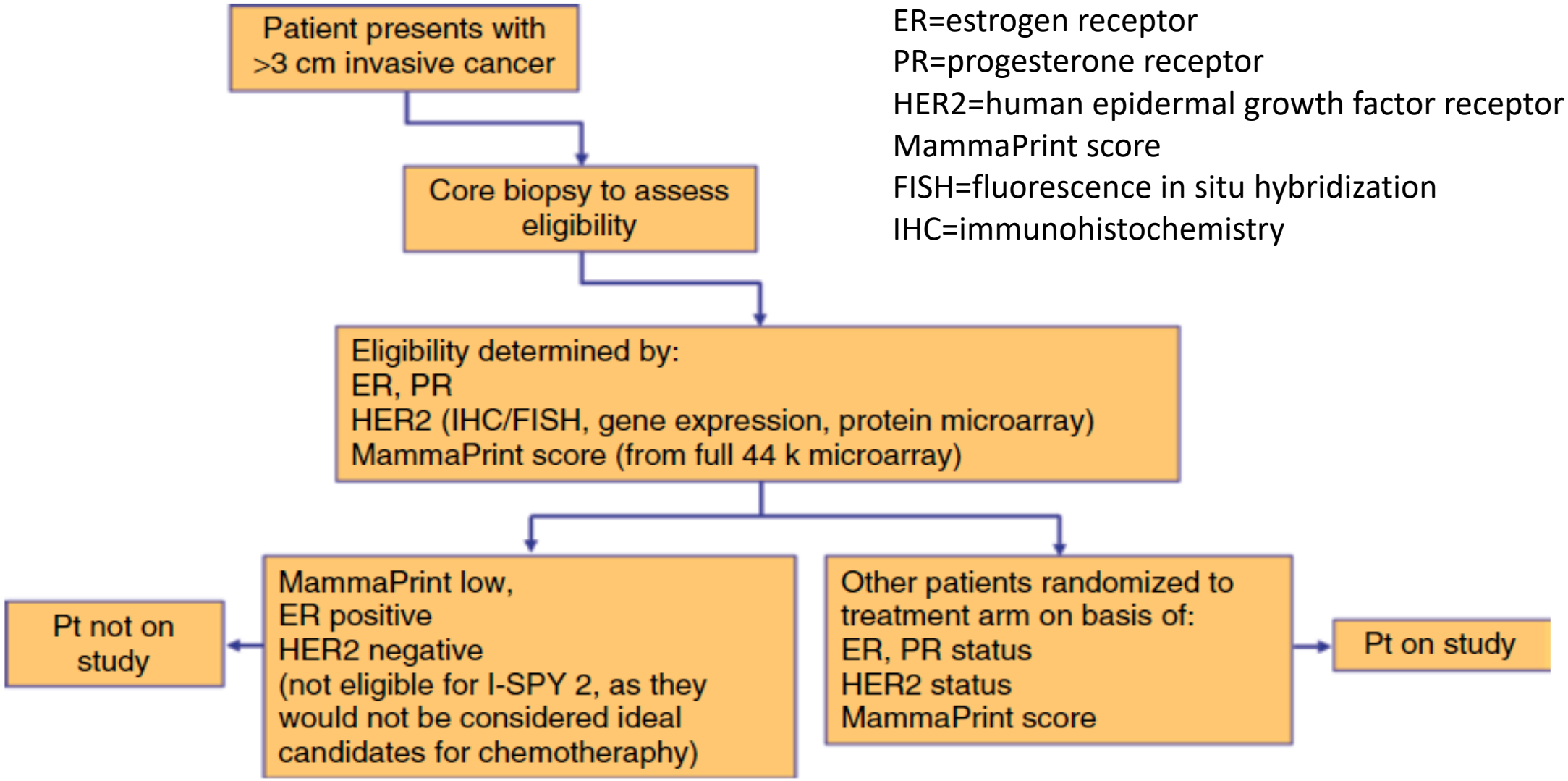
- Use adaptive design in neoadjuvant setting
  - Less patients for each signature
  - Faster throughput of agents by eliminating the need for new protocol each time an agent is added
- Biomarkers, imaging and pathology endpoints driven trial
- Validate, test and qualify biomarkers as new agents are tested
- Provides evidence for tailoring therapy

Barker et al CPT 2009

I-SPY2 Inclusion Criteria

- Screening phase
  - Histologically confirmed invasive BC
  - Clinically or radiologically measurable disease
  - No prior cytotoxic regimens
  - Age ≥ 18 years
  - ECOG PS 0-1
  - No ferromagnetic prostheses
- Treatment phase
  - Eligible tumors: stage II/III; T4, any N, M0; regional stage IV
  - Normal organ and marrow function
  - No uncontrolled/severe cardiac disease
  - No evidence of distant metastases
  - Specific tumor assay profile

I-SPY2 Patient Stratification



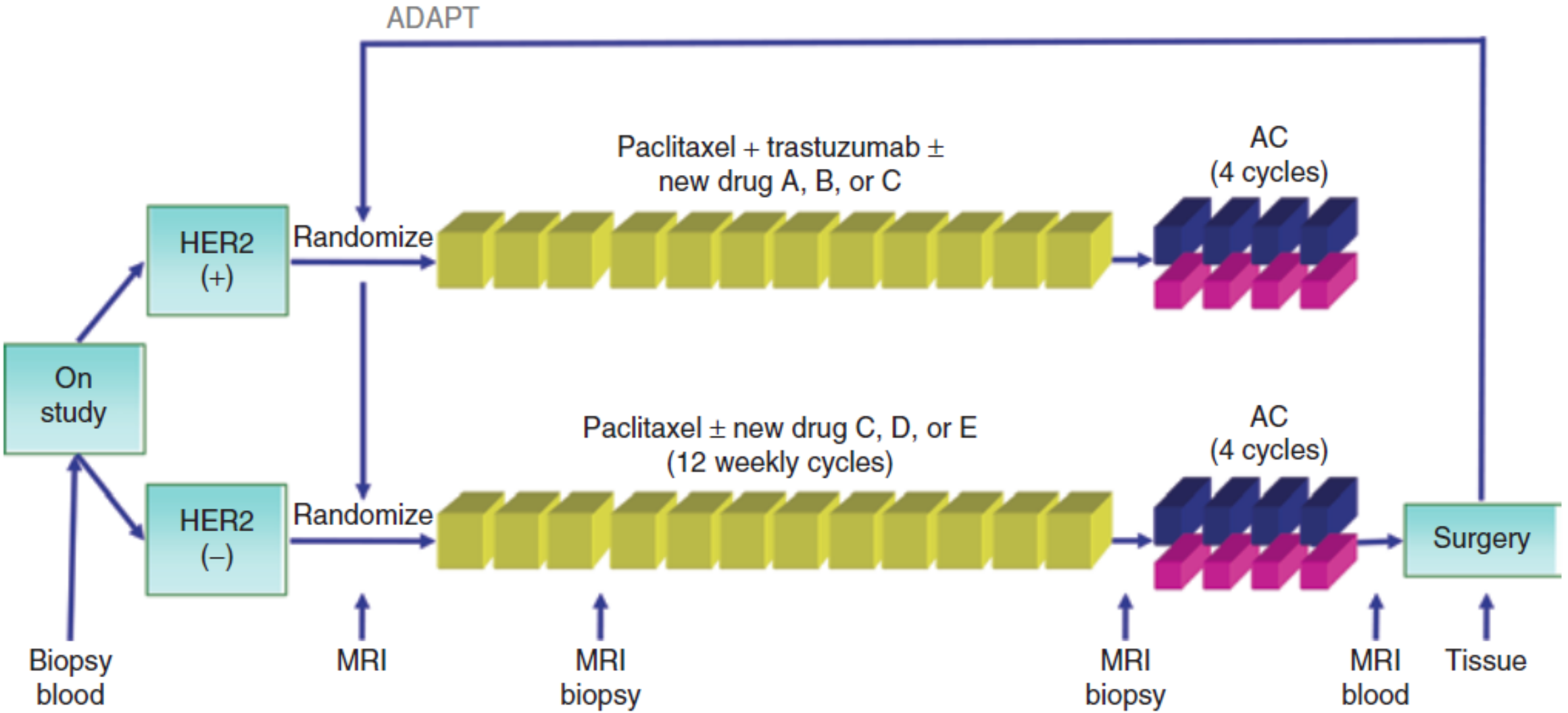
ER=estrogen receptor
PR=progesterone receptor
HER2=human epidermal growth factor receptor 2
MammaPrint score
FISH=fluorescence in situ hybridization
IHC=immunohistochemistry

I-SPY2 Signatures

Biomarker signature	Types (Hormone Receptor, HER2, MP)							
	+++	++-	+-+	+--	-++	-+-	--+	---
(All)	x	x	x	x	x	x	x	x
(HR+)	x	x	x	x				
(HR-)					x	x	x	x
(HER2+)	x	x			x	x		
(HER2-)			x	x			x	x
(MammaPrint+)	x		x		x		x	
(--)							x	x
(-+)					x	x		
(+-)			x	x				

HR (either ER+ or PgR+; both ER- and PgR-); HER2 (positive (+); normal (-)); MammaPrint status (High2 (+), High1 (-))

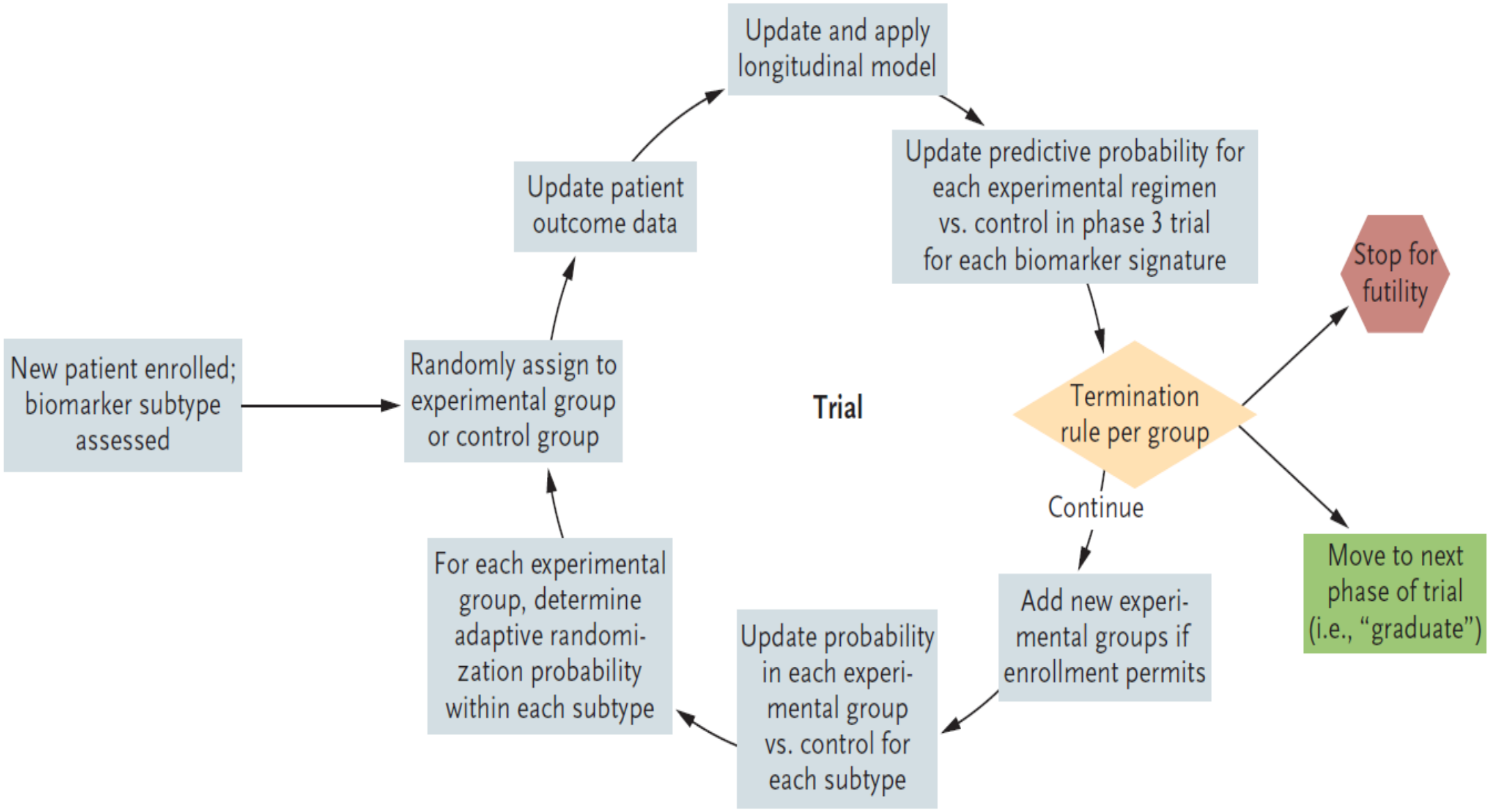
I-SPY2 Basic Schema



I-SPY2 Adaptive Design

- New agents are assigned to all signatures for which they may be effective
- Control arm applies to all signatures
- Randomization probabilities determined based on accumulating data (0.2 for control; min of 0.1 for exp. arms)
  - Regimens that are performing better for patient's biomarker type will have greater assignment probability
- Continuously, throughout the trial, each agent's probability of success in phase 3 will be calculated for each signature
  - Decisions: graduate, drop, or continue?

I-SPY2 Adaptive Design



I-SPY2 Endpoints

- Primary – pCR
  - Defined as no residual cancer in the breast (at time of definitive surgical resection) or in lymph nodes (no invasive tumor by H&E)
- Secondary – change in MRI volume from baseline to completion of paclitaxel base therapy; RCB at time of pathologic assessment of residual disease
- Others – 3 and 5 year RFS and OS
  - RFS: local/regional invasive recurrence, invasive ipsilateral breast tumor recurrence, distant recurrence, inoperable due to progression and/or death from BC
  - OS: death from BC, non-BC or unknown causes

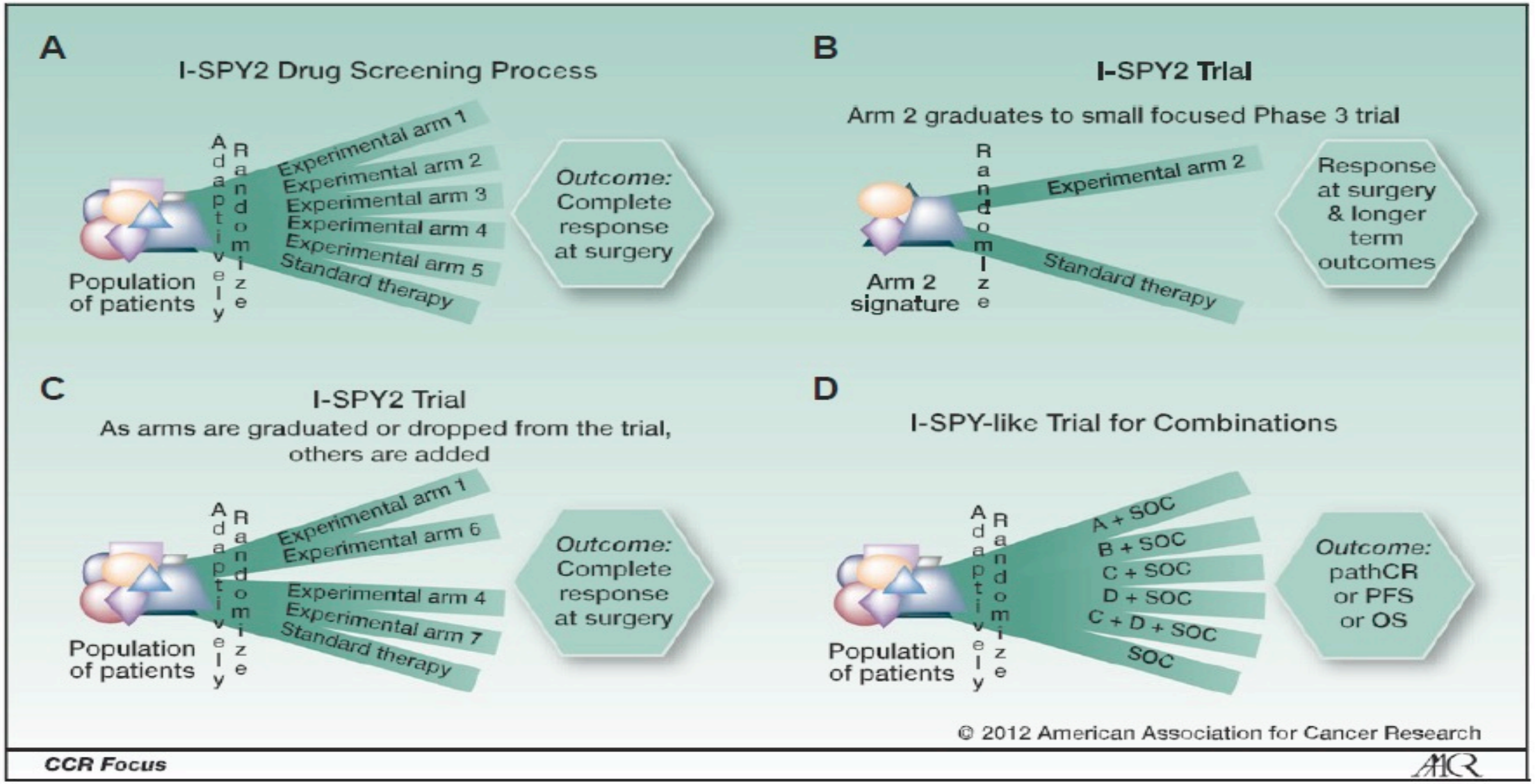
I-SPY2 Trial Decisions

- Regimen will be dropped for futility if its predictive probability (PP) drops sufficiently low for all biomarker signatures
  - Minimum of 20 pts enrolled before dropping
  - Once dropped, patients will revert to control regimen; however, their outcomes will remain with originally assigned arm
- Regimen will graduate to phase III if PP(s) for ≥1 biomarker signatures reaches sufficiently high level
  - Minimum of 60 pts enrolled before graduating
- If maximum sample size of 120 patients/regimen (over all biomarker types) reached, no more assignment to that regimen
- All patients expected to have surgery to assess pCR after regimen graduates or max N reached

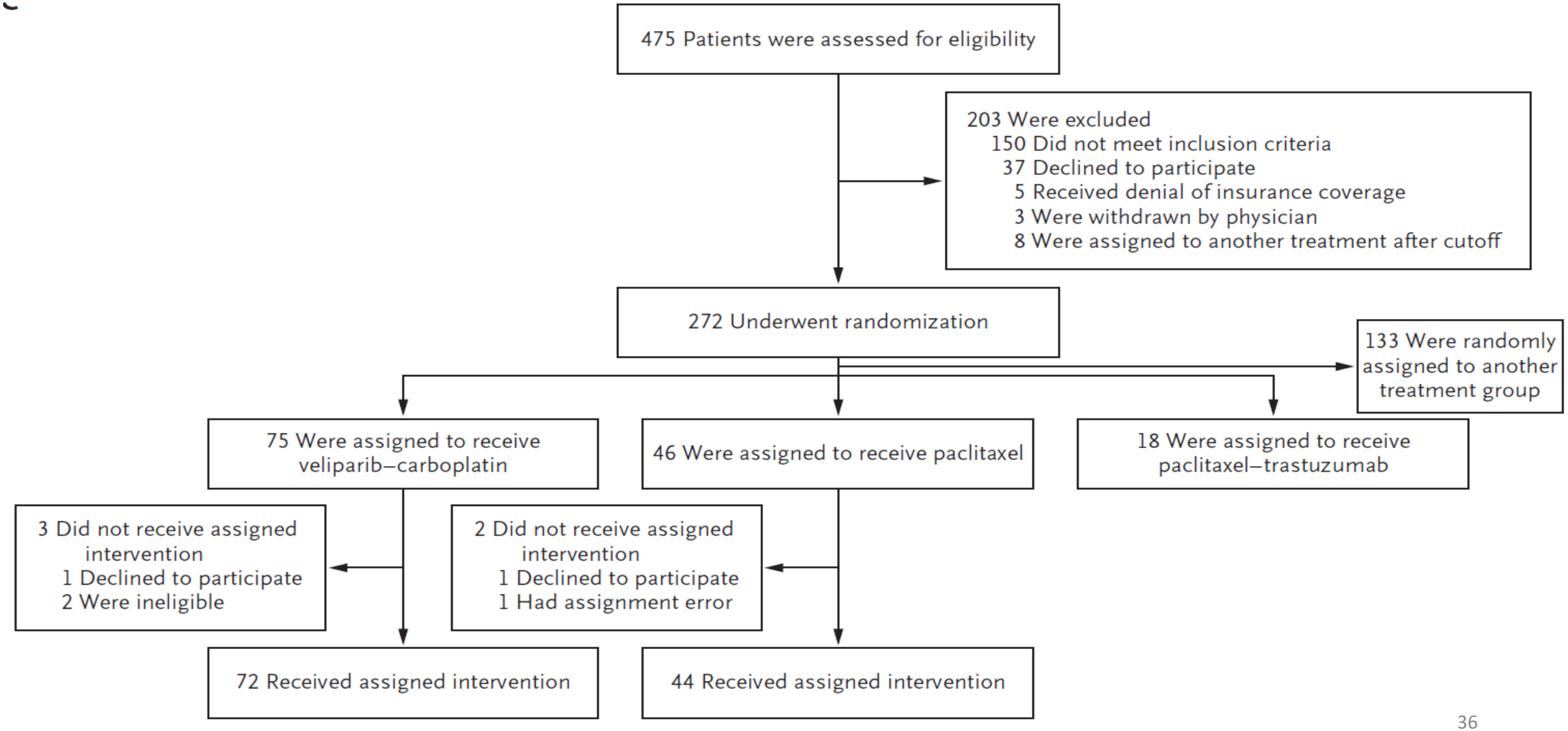
I-SPY2 Prediction

- Predicted probability of experimental treatment being successful in Phase 3 compared to control is calculated for:
  - Every biomarker signature
  - Every experimental treatment
  - Updated weekly
- Assume Phase 3 will equally randomize 300 patients between experimental and control arms
  - Trial will conclude in favor of experimental arm if:
    - $\Pr [\pi(\mathbf{R},T=1) > \pi(\mathbf{R},T=0) \mid Y_{F,0}, Y_{F,1}] > 0.85$
      - $\pi(\mathbf{R},T=t)$ = probability of pCR for signature **R** and treatment T
      - $Y_{F,0}$ and $Y_{F,1}$ = future # of responses for T=0 and T=1

I-SPY2


A
I-SPY2 Drug Screening Process
Population of patients
Adaptively
Randomize
Experimental arm 1
Experimental arm 2
Experimental arm 3
Experimental arm 4
Experimental arm 5
Standard therapy
Outcome: Complete response at surgery
B
I-SPY2 Trial
Arm 2 graduates to small focused Phase 3 trial
Arm 2 signature
Randomize
Experimental arm 2
Standard therapy
Response at surgery & longer term outcomes
C
I-SPY2 Trial
As arms are graduated or dropped from the trial, others are added
Population of patients
Adaptively
Randomize
Experimental arm 1
Experimental arm 6
Experimental arm 4
Experimental arm 7
Standard therapy
Outcome: Complete response at surgery
D
I-SPY-like Trial for Combinations
Population of patients
Adaptively
Randomize
A + SOC
B + SOC
C + SOC
D + SOC
C + D + SOC
SOC
Outcome: pathCR or PFS or OS
© 2012 American Association for Cancer Research
CCR Focus
AACR

I-SPY2 Veliparib-Carboplatin



I-SPY2 Veliparib-Carboplatin

Table 2. Final Predictive Probabilities.*

Biomarker Signature	Estimated Rate of Pathological Complete Response (95% PI)		Probability of Veliparib–Carboplatin Being Superior to Control	Predictive Probability of Success in Phase 3 Trial
	Veliparib–Carboplatin	Control		
			percent	
All HER2 negative	33 (23–43)	22 (10–35)	91	53
Hormone-receptor positive and HER2 negative	14 (3–25)	19 (5–33)	28	8
Triple negative	51 (36–66)	26 (9–43)	99	88

* HER2 denotes human epidermal growth factor receptor 2, and PI probability interval.

Rugo et al NEJM 2016

I-SPY2 Veliparib-Carboplatin

RESULTS

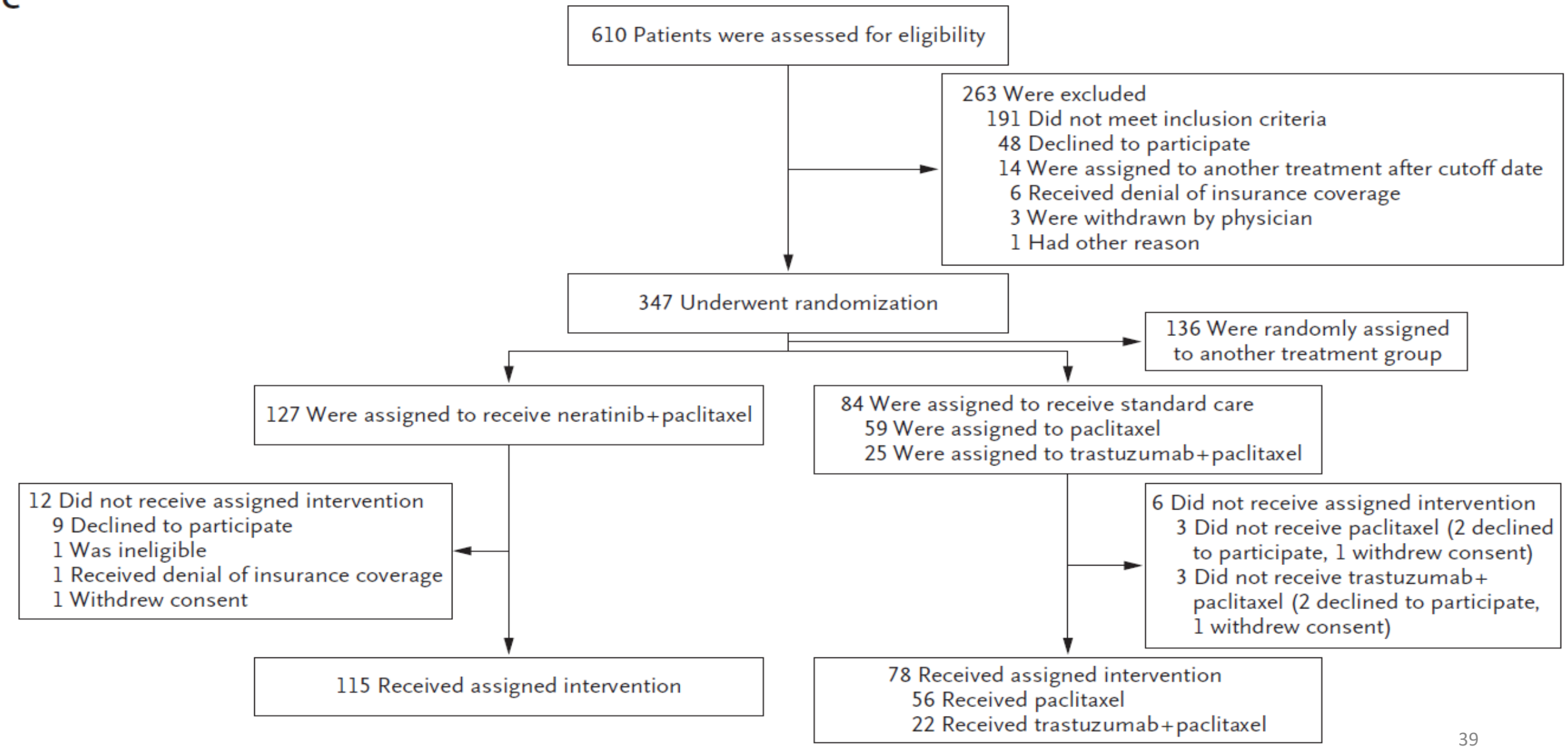
With regard to triple-negative breast cancer, veliparib–carboplatin had an 88% predicted probability of success in a phase 3 trial. A total of 72 patients were randomly assigned to receive veliparib–carboplatin, and 44 patients were concurrently assigned to receive control therapy; at the completion of chemotherapy, the estimated rates of pathological complete response in the triple-negative population were 51% (95% Bayesian probability interval [PI], 36 to 66%) in the veliparib–carboplatin group versus 26% (95% PI, 9 to 43%) in the control group. The toxicity of veliparib–carboplatin was greater than that of the control.

CONCLUSIONS

The process used in our trial showed that veliparib–carboplatin added to standard therapy resulted in higher rates of pathological complete response than standard therapy alone specifically in triple-negative breast cancer. (Funded by the QuantumLeap Healthcare Collaborative and others; I-SPY 2 TRIAL ClinicalTrials.gov number, NCT01042379.)

Rugo et al NEJM 2016

I-SPY2 Neratinib



I-SPY2 Neratinib

Table 2. Final Posterior and Predictive Probabilities of Neratinib Efficacy with Regard to 10 Biomarker Signatures.

Biomarker Signature	Estimated Rate of Pathological Complete Response (95% Probability Interval)		Probability of Neratinib Being Superior to Control	Predictive Probability of Success in Phase 3 Trial
	Neratinib	Control		
		percent		
Any	33 (24–40)	23 (14–33)	93	48
Hormone-receptor positive	23 (13–33)	16 (6–28)	81	40
Hormone-receptor negative	44 (30–55)	31 (17–45)	92	58
HER2 positive	39 (28–51)	23 (8–38)	95	73
HER2 negative	28 (15–37)	24 (13–35)	69	25
High-risk category 2 on 70-gene profile*	48 (30–60)	29 (11–48)	93	72
HER2 positive, hormone-receptor positive	30 (18–44)	17 (3–32)	91	65
HER2 positive, hormone-receptor negative	56 (37–73)	33 (11–54)	95	79
HER2 negative, hormone-receptor positive	14 (3–25)	16 (5–27)	42	14
HER2 negative, hormone-receptor negative	38 (22–50)	31 (15–46)	77	40

Park et al NEJM 2016

I-SPY2 Neratinib

Table 2. Final Posterior and Predictive Probabilities of Neratinib Efficacy with Regard to 10 Biomarker Signatures.

Biomarker Signature	Estimated Rate of Pathological Complete Response (95% Probability Interval)		Probability of Neratinib Being Superior to Control	Predictive Probability of Success in Phase 3 Trial
	Neratinib	Control		
		percent		
Any	33 (24–40)	23 (14–33)	93	48
Hormone-receptor positive	23 (13–33)	16 (6–28)	81	40
Hormone-receptor negative	44 (30–55)	31 (17–45)	92	58
HER2 positive	39 (28–51)	23 (8–38)	95	73
HER2 negative	28 (15–37)	24 (13–35)	69	25
High-risk category 2 on 70-gene profile*	48 (30–60)	29 (11–48)	93	72
HER2 positive, hormone-receptor positive	30 (18–44)	17 (3–32)	91	65
HER2 positive, hormone-receptor negative	56 (37–73)	33 (11–54)	95	79
HER2 negative, hormone-receptor positive	14 (3–25)	16 (5–27)	42	14
HER2 negative, hormone-receptor negative	38 (22–50)	31 (15–46)	77	40

Park et al NEJM 2016

I-SPY2 Neratinib

RESULTS

Neratinib reached the prespecified efficacy threshold with regard to the HER2-positive, hormone-receptor–negative signature. Among patients with HER2-positive, hormone-receptor–negative cancer, the mean estimated rate of pathological complete response was 56% (95% Bayesian probability interval [PI], 37 to 73%) among 115 patients in the neratinib group, as compared with 33% among 78 controls (95% PI, 11 to 54%). The final predictive probability of success in phase 3 testing was 79%.

CONCLUSIONS

Neratinib added to standard therapy was highly likely to result in higher rates of pathological complete response than standard chemotherapy with trastuzumab among patients with HER2-positive, hormone-receptor–negative breast cancer. (Funded by QuantumLeap Healthcare Collaborative and others; I-SPY 2 TRIAL ClinicalTrials.gov number, NCT01042379.)

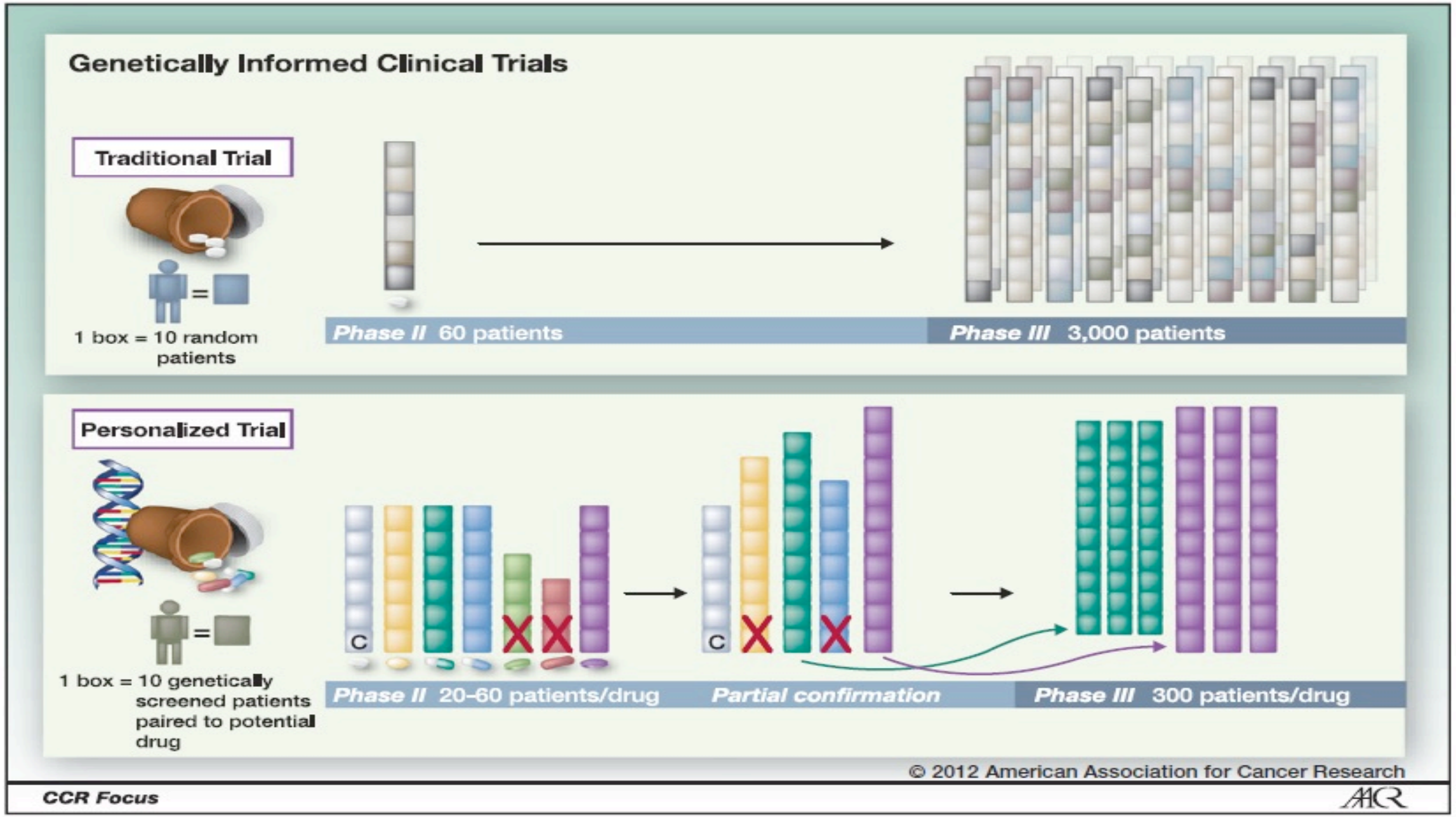
Park et al NEJM 2016

Table 2. Bayesian vs. Hypothetical Standard Frequentist Design.

Variable	I-SPY 2 Design	Standard Frequentist Design for I-SPY 2
Main goal of inference	Posterior distributions of rates of pathological complete response for the investigational drug (neratinib or veliparib) and the control. Predicted probability of success in a subsequent phase 3 trial.	Odds ratio or relative risk of response, investigational drug vs. control, with confidence interval and P value
Assumptions	Specification of prior distributions of response rates for investigational drug and control; specification of model for adapting randomization fraction as information becomes available, including model for imputation of pathological complete response based on imaging in previous patients	Specification of anticipated rates of pathological complete response in the control group and of clinically relevant target differences; specification of prior stratification for randomization of subtypes of breast cancer; distributions of unknown population parameters
Randomization	Adaptive randomization increases likelihood of participant receiving treatment assignment that may be of benefit. Estimates of pathological complete response rates must be model-based because of lack of balance of patients' baseline characteristics across treatments.	Constant randomization probabilities do not preferentially target patients who may benefit from a treatment; heterogeneity of patient groups receiving a treatment may dilute estimates of treatment effects. Constant randomization probabilities ensure approximate balance of baseline characteristics across treatments and allow direct comparisons.

Interim monitoring	A treatment is declared potentially successful if predicted probability of success in phase 3 trial is at least 85%. Predicted probability of success is evaluated frequently during the trial. Experimental treatment is dropped for futility if predictive probability of success in a phase 3 trial is <10% in all 10 signatures.	Summary test statistics calculated a small number of times (typically 3 or 4) with P values checked against interim monitoring boundaries for futility and efficacy. Treatment effect estimates can be used for future trials, but groups are not selected on the basis of predicted success rates of future phase 3 trial.
Ease of use	Software for calculation of posterior distribution of pathological complete response or predictive probability of success not generally available. Accruing information must be updated frequently and accurately for adaptive randomization.	Summary and test statistics based on ratios or differences of proportions of pathological complete response. Open-source or other software for design and analysis widely available. Few software packages available for adjusting estimates for treatment effects after early stopping. Accurate data updates required for interim monitoring.

Harrington et al NEJM 2016



Berry et al CCR 2012

GBM AGILE

Adaptive Global Innovative Learning Environment for Glioblastoma

GBM AGILE Background

- Brain tumor diagnosed in ~23,000 annually in U.S.
- Glioblastoma (GBM) is the most common type, accounts for 45%
- Prognosis is poor, with 1- and 5-year survival rates of only 35% and 4.7%.
- Standard of care for newly diagnosed is temozolomide and radiation, and not clear for recurrent as no therapies showing survival vs. lomustine

Alexander et al CCR 2017

GBM AGILE Stratification Factors

Line of Rx:

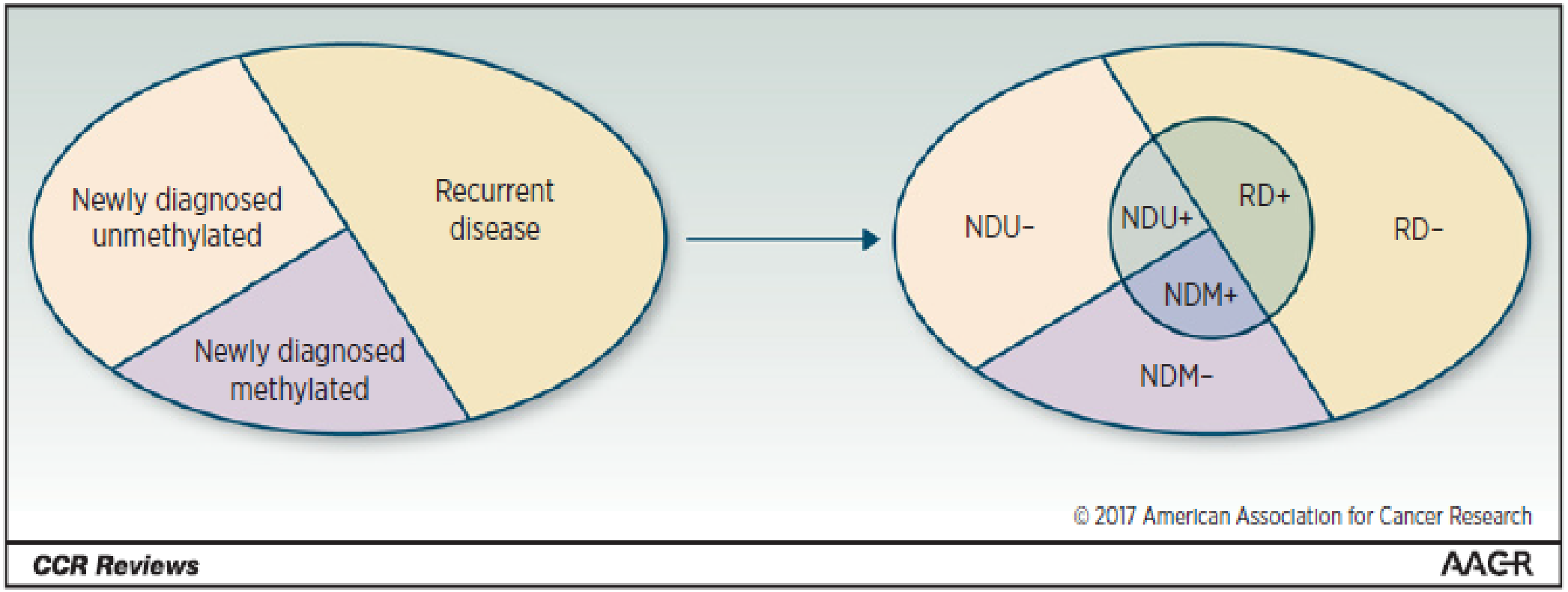
newly diagnosed vs recurrent

In newly diagnosed:

Methylation (M) vs unmethylated (N)

GBM AGILE Therapy w/ Enrichment Biomarker A

Doubles # of signatures



GBM AGILE Signatures

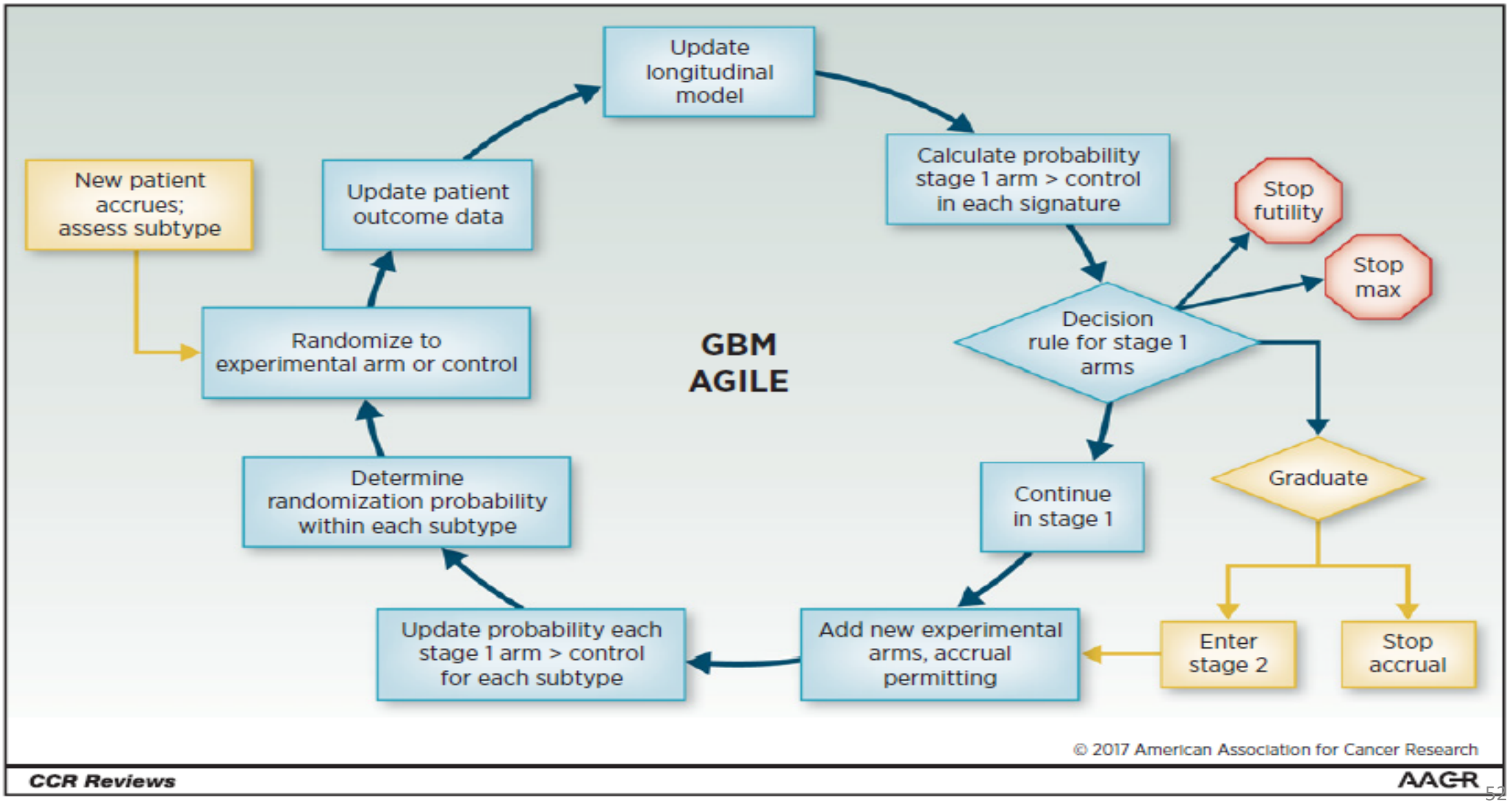
Possible signatures	Therapy's subtypes, including enrichment biomarker + or -					
	NDM+	NDM-	NDU+	NDU-	RD+	RD-
1: All	X	X	X	X	X	X
2: ND	X	X	X	X		
3: NDM	X	X				
4: NDU			X	X		
5: RD					X	X
6: All+	X		X		X	
7: ND+	X		X			
8: NDM+	X					
9: NDU+			X			
10: RD+					X	

All = All patients who are eligible for GBM AGILE; ND = Newly diagnosed; NDM = Newly diagnosed methylated; NDU = Newly diagnosed unmethylated; RD = Recurrent disease; + = enrichment biomarker positive, if the therapy has an enrichment biomarker

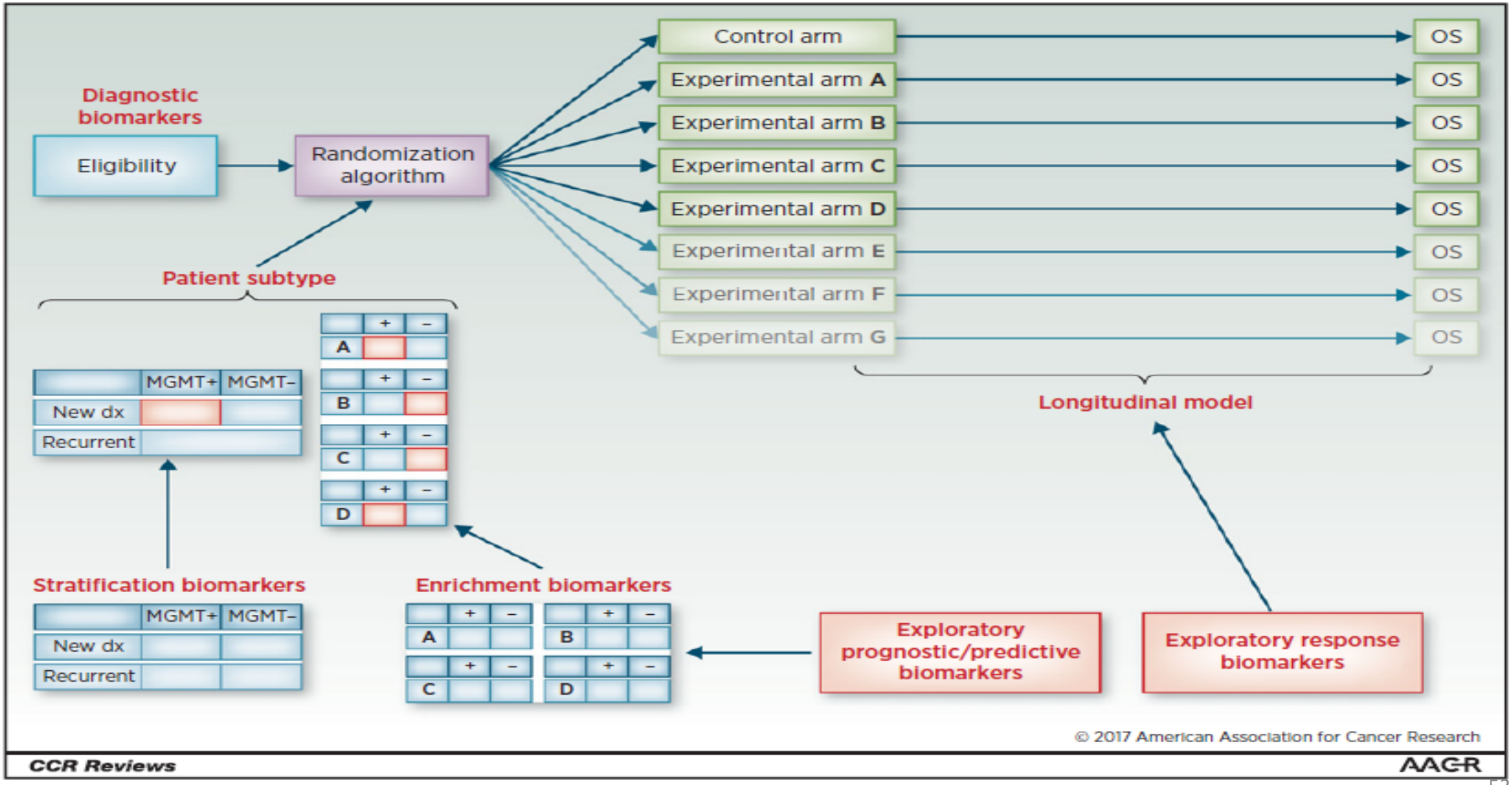
GBM AGILE Goal

- Multi-arm randomized platform trials enable solid links for bridging time periods
- Primary endpoint: OS
- Inform OS using longitudinal model

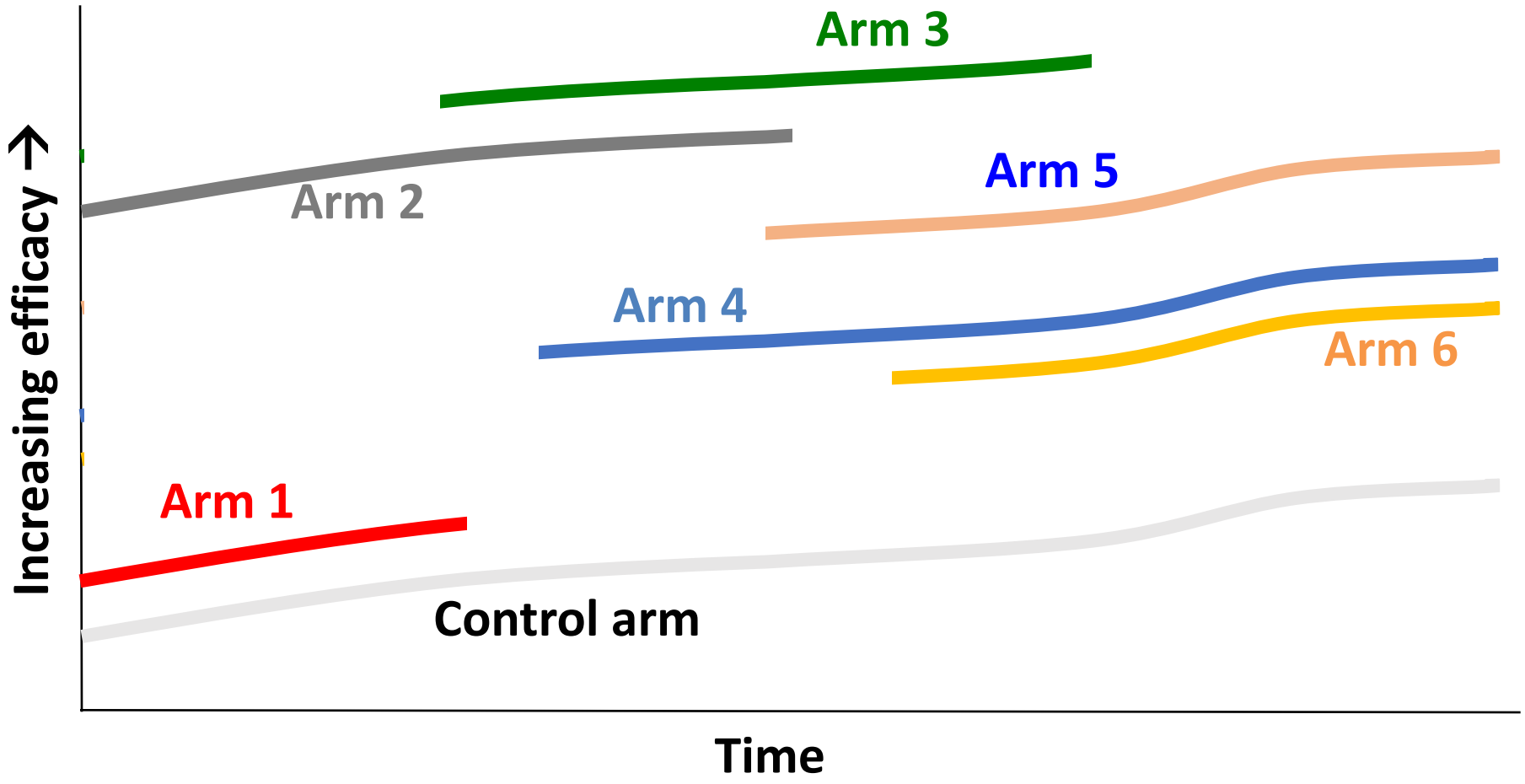
GBM AGILE Schema



GBM AGILE Schema



GBM AGILE The Time Machine in Platform Trials



GBM AGILE Modeling Time Aspects

- Drives adaptive randomization, graduation, futility, final primary analysis
- Allows simultaneous estimation time effects and treatment arm effects
- Concurrent data critical for relative effects of any two arms, but estimates of treatment effects and their precision is enhanced using all available information

GBM AGILE Traditional Two-armed Trials

- Fixed randomization (1:1, 2:1, etc.) (Adaptive randomization may introduce bias; may not compensate for slight improvement in efficiency and ethics)
- Absolute requirement for concurrently randomized controls
- One critical consequence: trial ends if control arm ends
- Traditional Principles Do Not Apply in Multi-armed Platform Trials!

GBM AGILE

Reference

1. Alexander BM, Ba S, Berger M et al (2017) Adaptive Global Innovative Learning Environment for Glioblastoma: GBM AGILE, CCR, DOI: 10.1158/1078-0432.CCR-17-0764
2. Barker AD, Sigman CC, Kelloff GJ, et al (2009) I-SPY 2: An Adaptive Breast Cancer Trial Design in the Setting of Neoadjuvant Chemotherapy, Clinical Pharmacology & Therapeutics 86 (1): 96-100
3. Berry DA (2015) The Brave New World of clinical cancer research: adaptive biomarker-driven trials integrating clinical practice with clinical research, Molecular Oncol 9:951-959
4. Berry S, Broglio K, Groshen S, Berry DA (2013) Bayesian hierarchical modeling of patient subpopulations: efficient designs of Phase II oncology clinical trials, Clinical Trials 10:720-734
5. Esserman LJ, Berry DA, DeMichele A et al (2012) Pathologic Complete Response Predicts Recurrence-Fee Survival More Effectively by Cancer Subset: Results From the I-SPY 1 TRIAL-CALGB 150007/150012, ACRIN 6657, J Clin Oncol 30:3242-3249
6. Harrington D, Parmigiani G (2016) I-SPY 2 – A Glimpse of the Future of Phase 2 Drug Development? NEJM 375(1): 7-9
7. Olshen A, Wolf D, Jones EF et al (2018) Features of MRI stromal enhancement with neoadjuvant chemotherapy: a subgroup analysis of the ACRIN 6657/I-SPY TRIAL, J Medical Imaging, 5(1)
8. Park JW, Liu MC, Yee D et al (2016) Adaptive Randomization of Neratinib in Early Breat Cancer, NEJM 375(1): 11-22
9. Rugo HS, Olopade OI, DeMichele A (2016) Adaptive Randomization of Veliparib-Carboplatin Treatment in Breast Cancer, NEJM 375(1): 23-34

Thank You!