

# Adventures in Statistics

Donald A. Berry  
[dberry@mdanderson.org](mailto:dberry@mdanderson.org)



# **Adventures in Statistics**

**Clinical trials**

**Part 1: Bayesian bandits to platforms (today)**

**Part 2: Embracing adaptive Bayesian clinical trials (Oct 26)**

**Part 3: Regulators & “complex innovative designs” (Nov 30)**

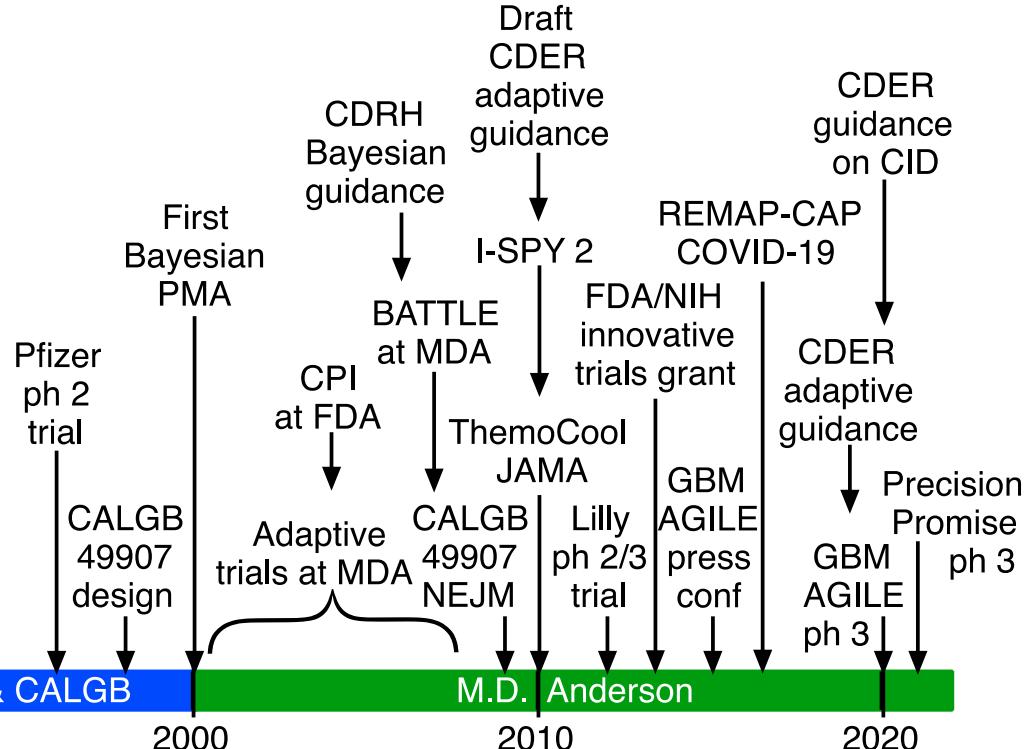
# Outline

- Why Bayes?
- What's a bandit strategy?
- Adaptive randomization
- Adaptive Bayesian platform trials, incl FDA views
- I-SPY 2 (neoadjuvant breast cancer)
- GBM AGILE (and Precision Promise)
- Need for simulations
- GBM AGILE ... the movie!

# What's different about Bayes?

Characteristic	Bayesian	Frequentist
Inferential unit	Patient (nested within trial)	Trial
Inferential measure	Probability of hypotheses and parameters	p-value: probability assuming null hypothesis
Probability applies to what?	All that's uncertain, including future results & hypotheses	Data: assuming particular hypothesis
Update probabilities	With each observation, or periodically	Not applicable
Predicting future results (including trial)	Based on current and prior information	Assumes particular hypothesis
Number of questions addressed	Many, perhaps scores	Preferably one
Modeling	Lots of modeling, including individual patients	Much less

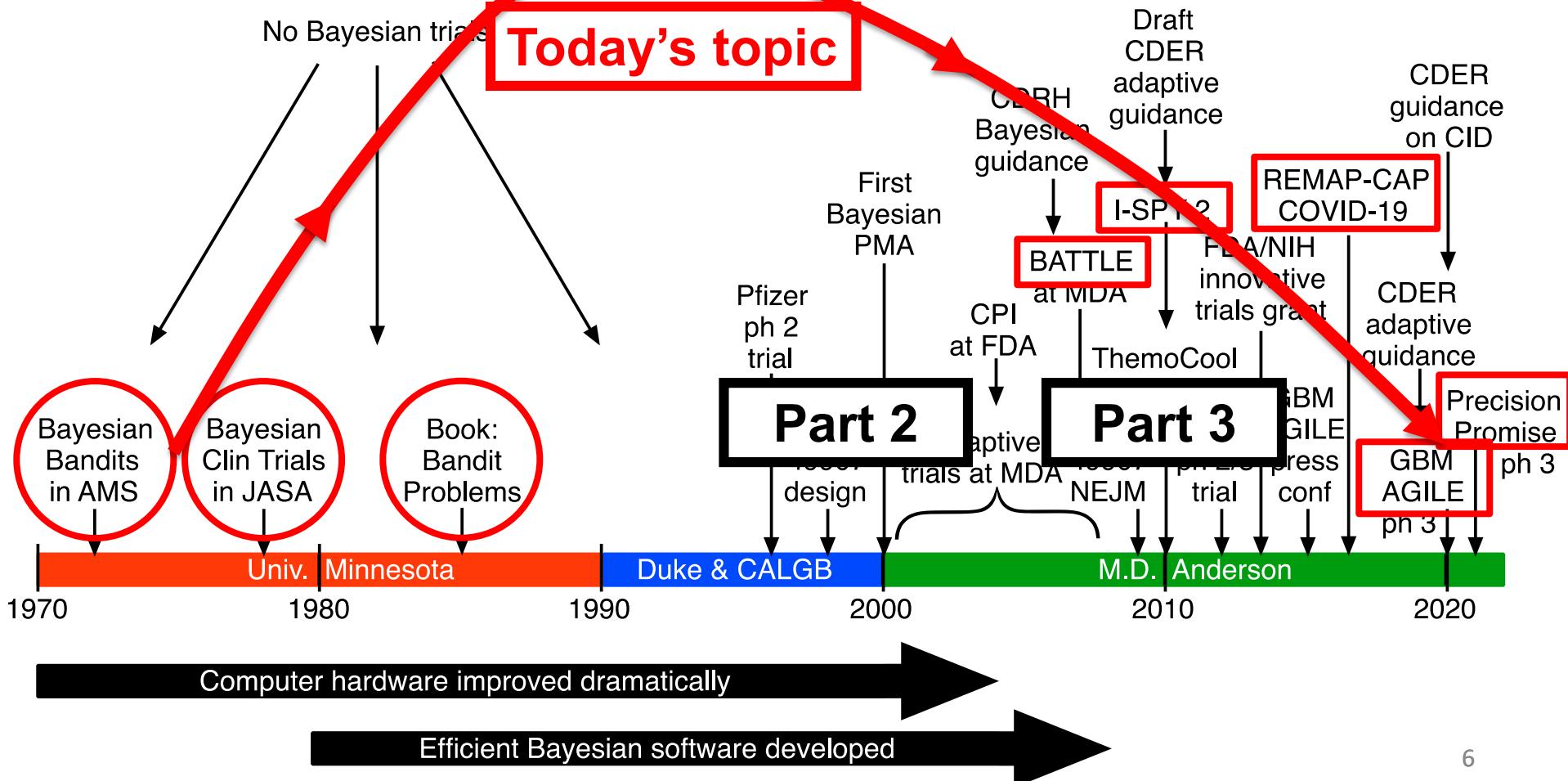
# Bayesian Adaptive Clinical Trials over Time



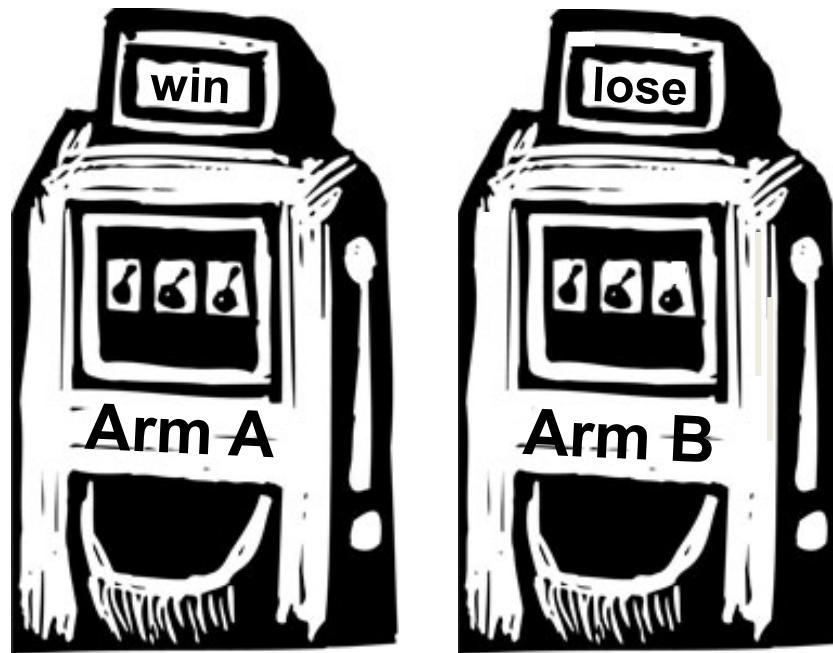
Computer hardware improved dramatically

Efficient Bayesian software developed

# Bayesian Adaptive Clinical Trials over Time



# Two-armed bandits



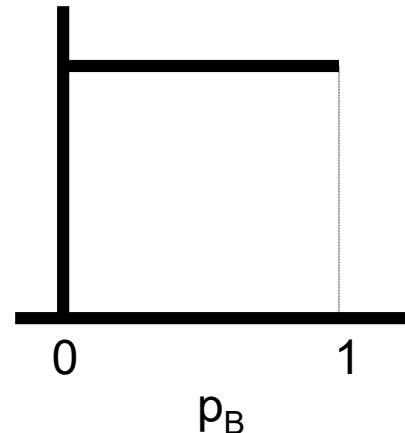
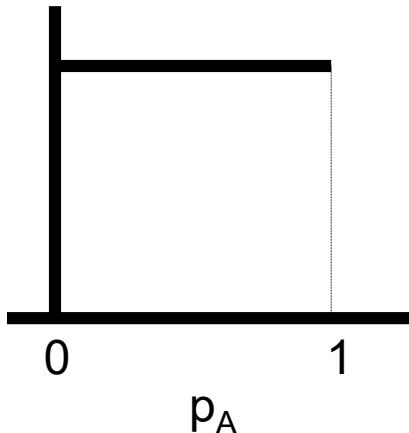
**Goal:**

**Maximize expected winnings (\$)**

**or ...**

**Maximize expected number  
of patients treated successfully**

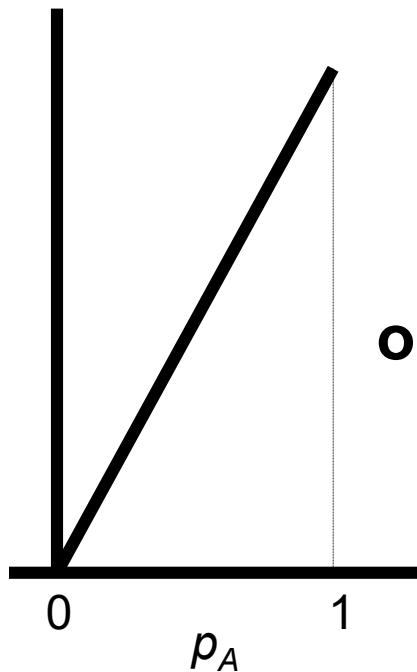
**Faced with 2 arms having win rates  
 $p_A$  and  $p_B$  and these prior distributions:**



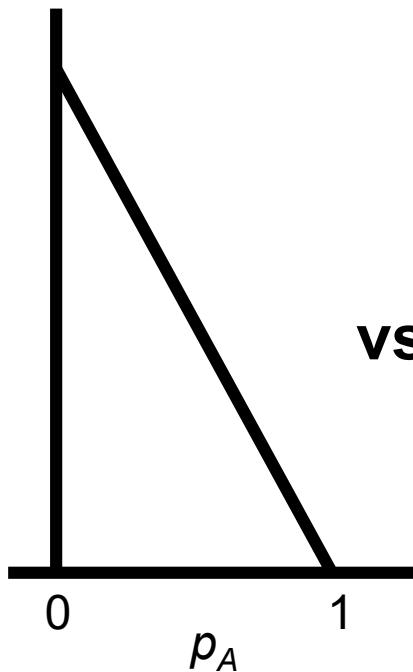
**Which do you pick?**

**Suppose pick Arm A.**  
**Win \$1 w/prob 1/2, but you learn something too ...**

If win (1/2)

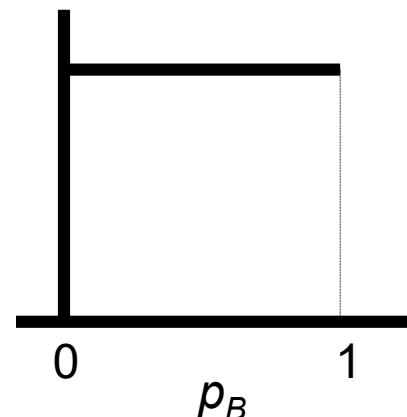


If lose (1/2)



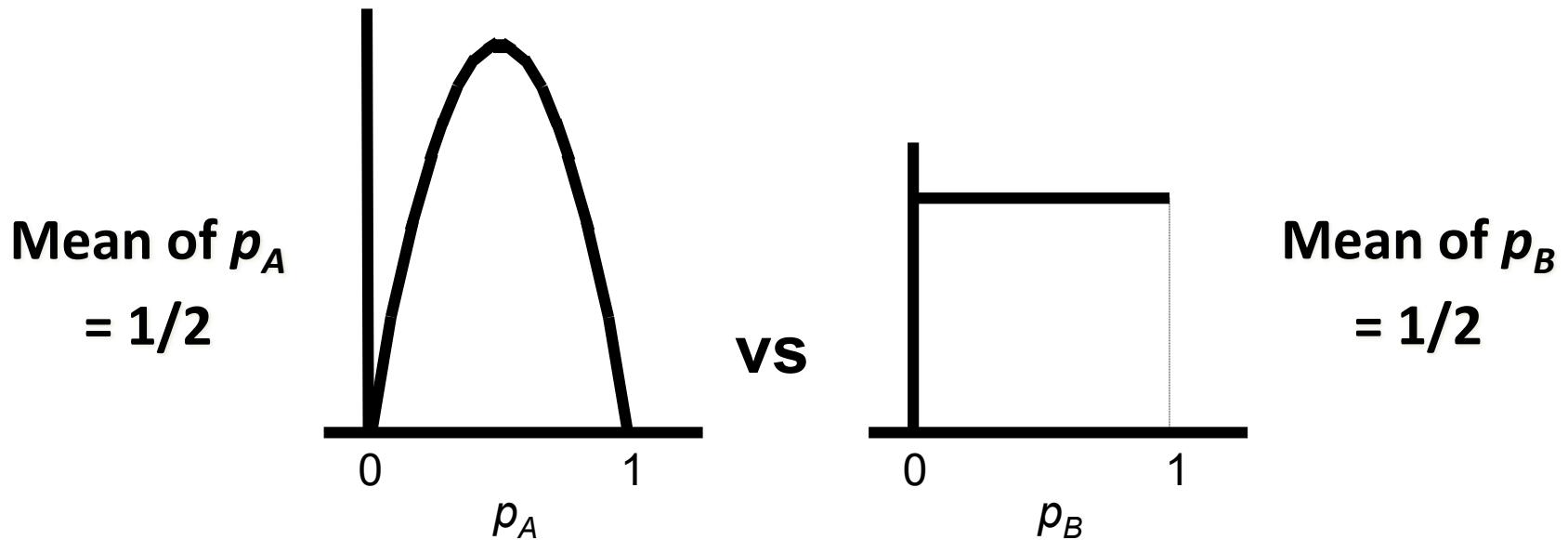
or

vs



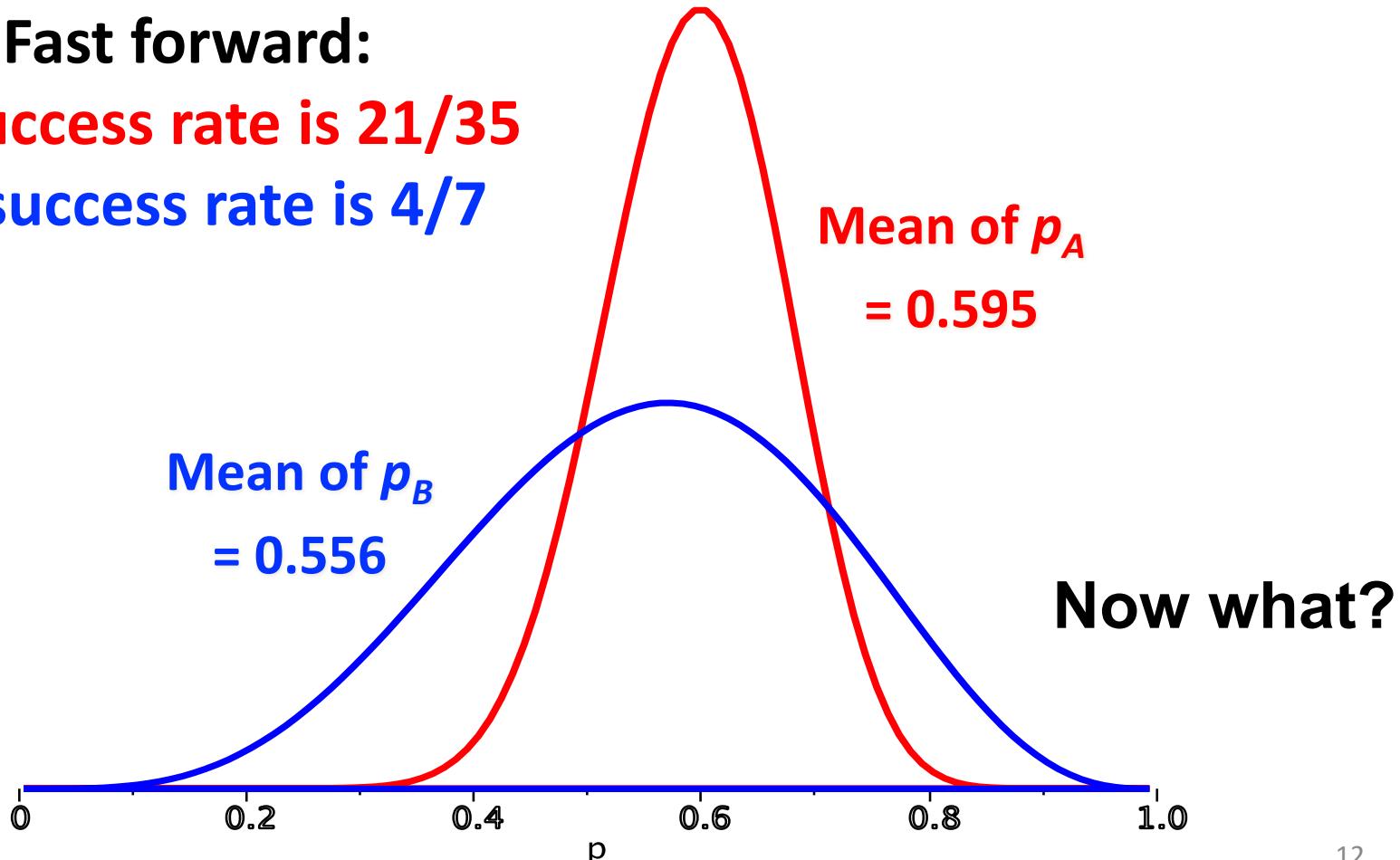
**Now what?**

**Suppose won on first pull and chose A again,  
losing on second pull of arm A ...**



**Now what?**

**Fast forward:**  
**A's success rate is 21/35**  
**B's success rate is 4/7**



# Optimal bandit strategies are complicated

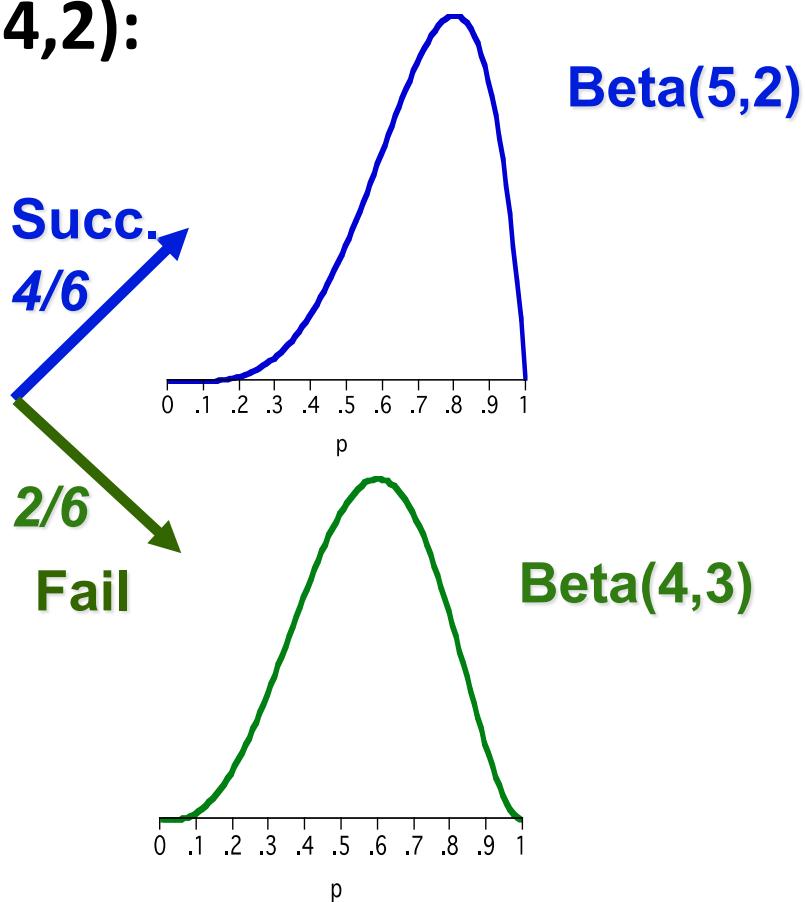
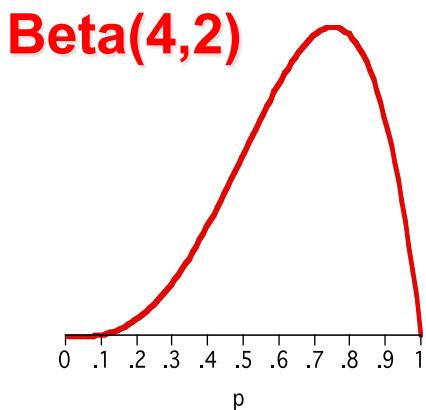
- “N-P hard”
- Conflicting desiderata (and ethical dilemmas!): make money today or get information to help make money later
- Wikipedia entry “Multi-armed bandit:”  
“Originally considered by Allied scientists in World War II, it proved so intractable that, according to Peter Whittle, the problem was proposed to be dropped over Germany so that German scientists could also waste their time on it.”

# Solution by dynamic programming

- Start at the “end,” laying out all possible outcomes at some “long time” hence, evaluate each outcome’s value, then evaluate the value of each penultimate outcome
- A fundamental requirement is the (Bayesian) predictive probability of each final outcome given each penultimate outcome
- Continue going backwards till you get to today
- Role of discounting the future?

Suppose after 4 pulls of an arm:

3 wins & 1 loss,  $p \sim \text{beta}(4,2)$ :



# **Characterizing optimal strategies: Stay with a winner**

## A BERNOULLI TWO-ARMED BANDIT<sup>1</sup>

BY DONALD A. BERRY

*University of Minnesota*

One of two independent Bernoulli processes (arms) with unknown expectations  $\rho$  and  $\lambda$  is selected and observed at each of  $n$  stages. The selection problem is sequential in that the process which is selected at a particular stage is a function of the results of previous selections as well as of prior information about  $\rho$  and  $\lambda$ . The variables  $\rho$  and  $\lambda$  are assumed to be independent under the (prior) probability distribution. The objective is to maximize the expected number of successes from the  $n$  selections. Sufficient conditions for the optimality of selecting one or the other of the arms are given and illustrated for example distributions. The stay-on-a-winner rule is proved.

**1. Introduction and statement of the problem.** Let  $\mathcal{R}$  and  $\mathcal{L}$  denote independent Bernoulli processes with parameters—probabilities of success— $\rho$  and  $\lambda$  respectively. Call  $\mathcal{R}$  the *right arm* and  $\mathcal{L}$  the *left arm*. An observation on either arm is called a *pull*. A right pull or a left pull is made at each of  $n$  stages and the result of the pull at each stage is known before a right or left pull is made at the next stage. The parameters  $\rho$  and  $\lambda$  associated with  $\mathcal{R}$  and  $\mathcal{L}$  are

**Approximating optimal strategies,  
while adding randomization ...**

## **W.R. Thompson (1933)**

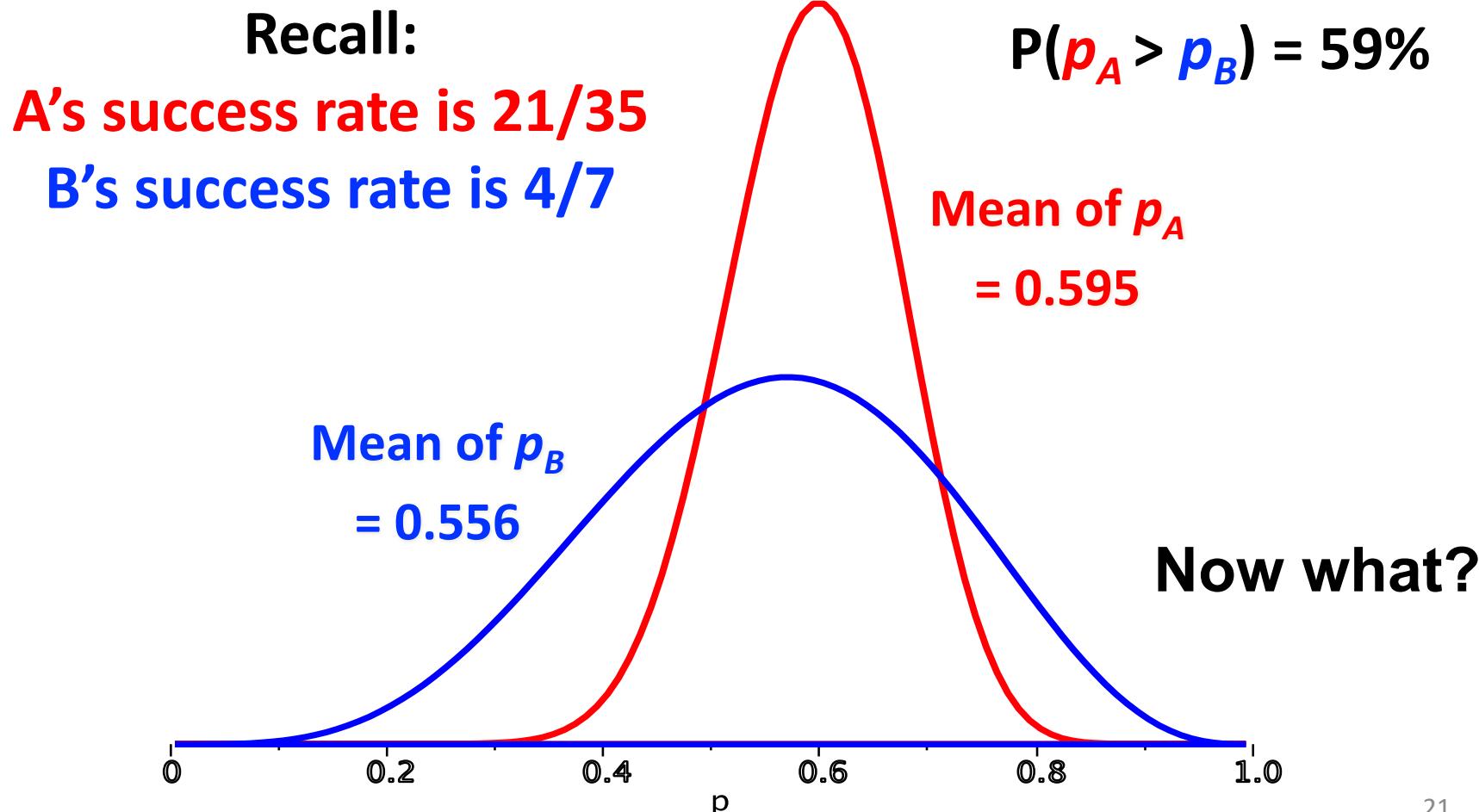
- Calculate Bayesian probability arm A better than arm B; assign next patient to arm A with that probability (or some function of it)
- Then: “even though [this strategy is] not the best possible, it seems apparent that a considerable saving of individuals otherwise sacrificed to the inferior treatment might be effected.”
- Predates RCT
- Adaptive randomization is a standard at MDACC & in many platform trials

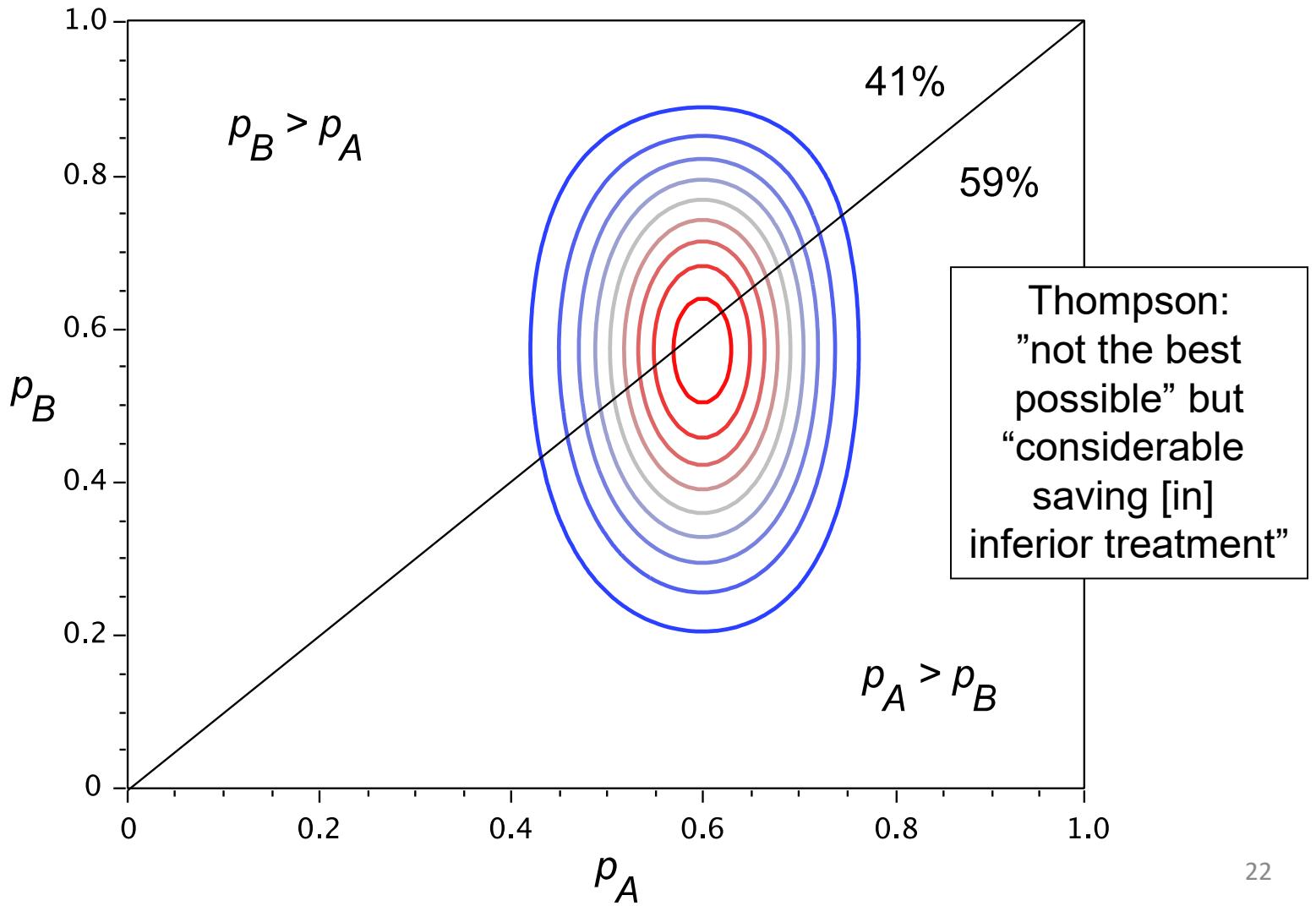
ON THE LIKELIHOOD THAT ONE UNKNOWN  
PROBABILITY EXCEEDS ANOTHER IN VIEW  
OF THE EVIDENCE OF TWO SAMPLES.

By WILLIAM R. THOMPSON. From the Department of Pathology,  
Yale University.

*Section 1.*

IN elaborating the relations of the present communication interest was not centred upon the interpretation of particular data, but grew out of a general interest in problems of research planning. From this point of view there can be no objection to the use of data, however meagre, as a guide to action required before more can be collected; although serious objection can otherwise be raised to argument based upon a small number of observations. Indeed, the fact that such objection can never be eliminated entirely—no matter how great the number of observations—suggested the possible value of seeking other modes of operation than that of taking a large number of observations before analysis or any attempt to direct our course. This problem is more general than that treated in *Section 2*, and is directly concerned with any case where probability criteria may be established by means of which we judge whether one mode of operation is *better* than another in some given sense or not.





# Multi-armed bandit experiments in the online service economy

Steven L. Scott 

First published: 16 January 2015 | <https://doi.org/10.1002/asmb.2104> | Citations: 30

[Read the full text >](#)



PDF



TOOLS



SHARE

## Abstract

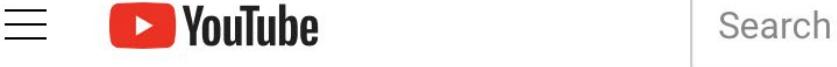
The modern service economy is substantively different from the agricultural and manufacturing economies that preceded it. In particular, the cost of experimenting is dominated by opportunity cost rather than the cost of obtaining experimental units. The different economics require a new class of experiments, in which stochastic models play an important role. This article briefly summarizes multi-armed bandit experiments, where the experimental design is modified as the experiment progresses to reduce the cost of experimenting. Special attention is paid to Thompson sampling, which is a simple and effective way to run a multi-armed bandit experiment. Copyright © 2015 John Wiley & Sons, Ltd.

**Hal Varian, Chief Economist, Google,  
in 2009:**

**“I keep saying that the sexy job in the  
next 10 years will be statisticians.”**

# Hal Varian at RSS 2012 Conference: Statistics at Google

<https://www.youtube.com/watch?v=p8R-UL6RPSg>



## Multi-armed bandits

Website Optimizer: allowed for A-B testing of web page design for users of Google Analytics

Optimize some objective, e.g., conversions

Experiments are expensive!

Could not easily model features  
(font, colors, images, layout)



## Google Analytics Content Experiments

Multiarmed bandit

Far more cost-effective testing

More natural interpretation

Can model features easily



# Modified Two-Armed Bandit Strategies for Certain Clinical Trials

DONALD A. BERRY\*

---

A procedure which maximizes the expected number of successes in a clinical trial involving two treatments can usually be found only by backward induction. Not only is it difficult to find an optimal procedure but, once found, it is difficult to describe and cumbersome to communicate. A procedure is suggested which depends on the information present concerning the treatments. This procedure is easy to calculate and approximates an optimal procedure quite well. It is applicable to trials for which the number of patients is unknown as well as those of known duration.

**KEY WORDS:** Clinical trials; Two-armed bandits; Sequential Bayesian decisions; Feldman's strategy.

## 1. INTRODUCTION

as a function of  $(p_1, p_2)$ . Another, the Bayesian approach, asks that current information concerning  $(p_1, p_2)$  be quantified in the form of a probability distribution. The Bayesian approach will be used in most of this article. A distinct advantage of this approach is that accumulating information can be handled in a unified way: Bayes' theorem is used to modify the probability distribution on  $(p_1, p_2)$ . The effectiveness of a procedure can then be averaged over  $(p_1, p_2)$  and, possibly, a procedure can be found that maximizes the expected number of successes.

If one of the  $p_i$  is known, say  $p_1$ , and  $p_2$  has probability

# Advice from biostatisticians

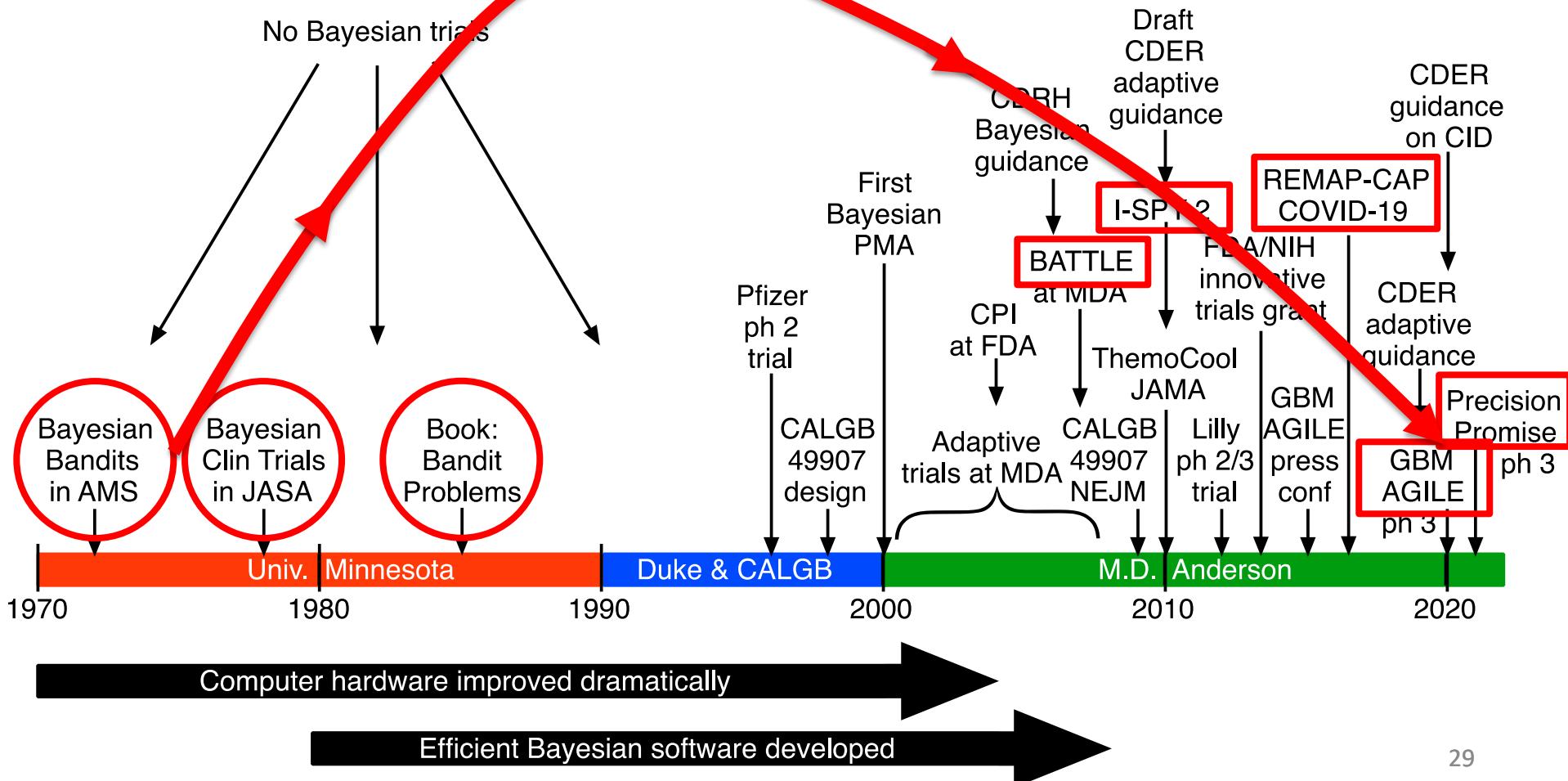
- Berry, you don't understand clinical trials.
- You're wet behind the ears. You should learn something about real clinical trials.
- Bayesian prior distributions are biased.
- Adaptive randomization is crazy; it will never fly.

**I took their advice**

**Coming attractions:**

**Some details in Parts 2 and 3**

# Bayesian Adaptive Clinical Trials over Time



# Platform Trial

AGILE Inc.



# **Platform Trials Past, Present, & Future**

- Many cancers
- Alzheimer's
- ALS
- Ebola
- Antibacterials
- Scleroderma
- Community acquired pneumonia
- COVID-19
- ...

# **Janet Woodcock, Director CDER FDA**

**2011, NEJM:** “In 2010, the [FNIH] Biomarkers Consortium [including] the NIH, the FDA, patient groups, and pharmaceutical and biotech initiated a groundbreaking trial in breast cancer to predict drug responsiveness based on the presence or absence of genetic and biological markers ... I-SPY 2.”

**2013:** FDA will need to "turn the clinical trial paradigm on its head" to allow personalized drug therapies to get on the market faster.

**2015, Press launch of GBM AGILE:** “This is the future.”

REVIEW ARTICLE

# We illustrate the concept using the Investigation of Serial Studies to Predict Your Therapeutic Response with Imaging and Molecular Analysis 2 (I-SPY 2) ...

Ph.D.,

Janet Woodcock, M.D., and Lisa M. LaVange, Ph.D.

From the Center for Drug Evaluation and Research, Food and Drug Administration, Silver Spring, MD. Address reprint requests to Dr. LaVange at the Office of Biostatistics, Office of Translational Sciences, Center for Drug Evaluation and Research, Food and Drug Administration, 10903 New Hampshire Blvd., Silver Spring, MD 20993, or at lisa.lavange@fda.hhs.gov.

**H**IGH-QUALITY EVIDENCE IS WHAT WE USE TO GUIDE MEDICAL PRACTICE. The standard approach to generating this evidence — a series of clinical trials, each investigating one or two interventions in a single disease — has become ever more expensive and challenging to execute. As a result, important clinical questions go unanswered. The conduct of “precision medicine” trials to evaluate targeted therapies creates challenges in recruiting patients with rare genetic subtypes of a disease. There is also increasing interest in performing mechanism-based trials in which eligibility is based on criteria other than traditional disease definitions. The common denominator is a need to answer more questions more ef-

# Ongoing Adaptive Bayesian Platform Trials in Oncology

- I-SPY 2: phase 2, neo-adjuvant breast cancer
- Seamless phase 2-3 trials (for registration)
  - GBM AGILE (glioblastoma)
  - Precision Promise (pancreatic cancer)

# Prototype Bayesian Adaptive Platform Trial: I-SPY 2

- <https://www.ispytrials.org/i-spy-platform/i-spy2>
- Everlasting Phase 2 trial in neoadjuvant breast cancer
- 24 experimental arms to date (17 pharma companies)
- Adaptively randomized (a la Thompson)
- Fixed randomization (20%) to control
- > 2000 patients randomized, 2010 – present
- 8 disease subtypes; 10 possible signatures (indications)
- 7 arms have “graduated,” in 7 different signatures

# The NEW ENGLAND JOURNAL of MEDICINE

ESTABLISHED IN 1812

JULY 7, 2016

VOL. 375 NO. 1

## Adaptive Randomization of Neratinib in Early Breast Cancer

J.W. Park, M.C. Liu, D. Yee, C. Yau, L.J. van 't Veer, W.F. Symmans, M. Paoloni, J. Perlmutter, N.M. Hylton, M. Hogarth, A. DeMichele, M.B. Buxton, A.J. Chien, A.M. Wallace, J.C. Boughey, T.C. Haddad, S.Y. Chui, K.A. Kemmer, H.G. Kaplan, C. Isaacs, R. Nanda, D. Tripathy, K.S. Albain, K.K. Edmiston, A.D. Elias, D.W. Northfelt, L. Pusztai, S.L. Moulder, J.E. Lang, R.K. Viscusi, D.M. Euhus, B.B. Haley, Q.J. Khan, W.C. Wood, M. Melisko, R. Schwab, T. Helsten, J. Lyndres, S.E. Davis, G.L. Hirst, A. Sanil, L.J. Esserman, and D.A. Berry, for the I-SPY 2 Investigators\*

### ABSTRACT

#### BACKGROUND

The heterogeneity of breast cancer makes identifying effective therapies challenging. The I-SPY 2 trial, a multicenter, adaptive phase 2 trial of neoadjuvant therapy for high-risk clinical stage II or III breast cancer, evaluated multiple new agents added to stan-

The authors' full names, academic degrees, and affiliations are listed in the Appendix. Address reprint requests to Dr. Esserman at the UCSF Carol Franc Buck

The NEW ENGLAND JOURNAL of MEDICINE

### EDITORIAL



## I-SPY 2 — Toward More Rapid Progress in Breast Cancer Treatment

Lisa A. Carey, M.D., and Eric P. Winer, M.D.

The NEW ENGLAND JOURNAL of MEDICINE

### ORIGINAL ARTICLE

## Adaptive Randomization of Veliparib–Carboplatin Treatment in Breast Cancer

H.S. Rugo, O.I. Olopade, A. DeMichele, C. Yau, L.J. van 't Veer, M.B. Buxton, M. Hogarth, N.M. Hylton, M. Paoloni, J. Perlmutter, W.F. Symmans, D. Yee, A.J. Chien, A.M. Wallace, H.G. Kaplan, J.C. Boughey, T.C. Haddad, K.S. Albain, M.C. Liu, C. Isaacs, Q.J. Khan, J.E. Lang, R.K. Viscusi, L. Pusztai, S.L. Moulder, S.Y. Chui, K.A. Kemmer, A.D. Elias, K.K. Edmiston, D.M. Euhus, B.B. Haley, R. Nanda, D.W. Northfelt, D. Tripathy, W.C. Wood, C. Ewing, R. Schwab, J. Lyndres, S.E. Davis, G.L. Hirst, A. Sanil, D.A. Berry, and L.J. Esserman, for the I-SPY 2 Investigators\*

### ABSTRACT

#### BACKGROUND

The genetic and clinical heterogeneity of breast cancer makes the identification of effective therapies challenging. We designed I-SPY 2, a phase 2, multicenter, adaptively randomized trial to screen multiple experimental regimens in combination with standard neoadjuvant chemotherapy for breast cancer. The goal is to match ex-

### PERSPECTIVE

#### STATISTICS IN MEDICINE

## I-SPY 2 — A Glimpse of the Future of Phase 2 Drug Development?

David Harrington, Ph.D., and Giovanni Parmigiani, Ph.D.

The articles by Rugo et al. (pages 23–34) and Park et al. (pages 11–22) in this issue of the *Journal* report results from the I-SPY (Investigation of Serial Studies to Predict Your Therapeutic Response with Imaging and Molecular Analysis) 2 platform, a

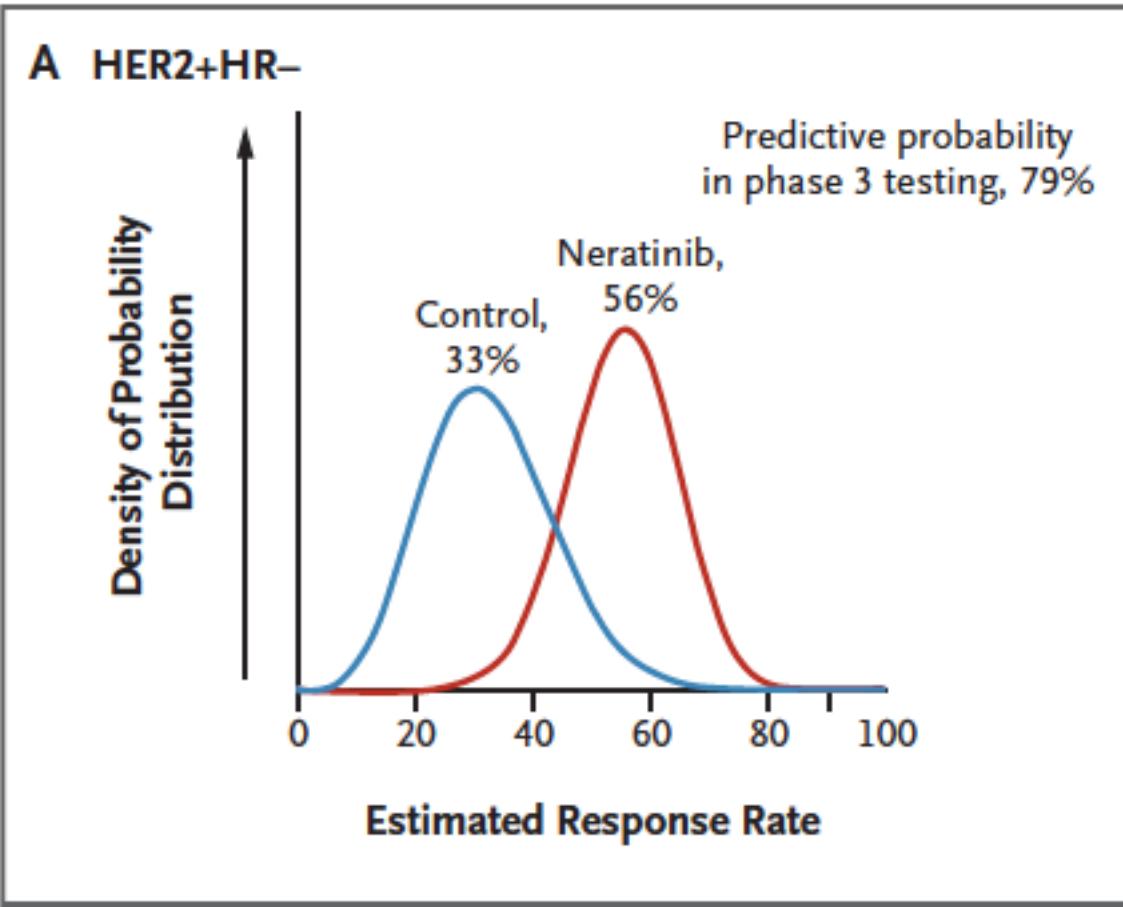
good responses and, equal portant, may be useful in ing patients to avoid treat when meaningful benefit likely. The challenges, how in identifying successful therapies in cancer are subti

I-SPY 2 — THE FUTURE OF PHASE 2 DRUG DEVELOP

**Benefits of having many arms –  
just one example ...**

# Neratinib's “graduation signature”

From  
NEJM:



# Randomization to neratinib partway thru its tour

		MP-		MP+	
		HR+	HR-	HR+	HR-
HER2+	+	++		+	++
HER2-	0	0		+	+

- Nightmare in a 2-armed trial
- Easy in a platform trial: no amendment needed, only DSMB knew it was happening

# **Seamless 2-3 Trials**

- **GBM AGILE (glioblastoma)**
- **Precision Promise (pancreatic cancer)**

# New & Transformative in a Phase 3 trial

1. Seamless shift, learn (phase 2) to confirm (phase 3)
2. Many arms, that enter and leave the trial
3. Common control (by patient subtype)
4. Controls include “contemporary controls” via time machine
5. Continuous learning and updating information
6. Adaptive randomization (in learn stage)
7. Identify and confirm arms’ biomarker indications, if any
8. Interpretation of Type I error
9. Decisions determined by predictive probability (PP)
10. Longitudinal model of disease burden
11. Hierarchical modeling of two control arms
12. Re-randomize patients to second-line therapy



March 11, 2018

Brian Alexander, M.D.  
Dana-Farber Cancer Institute  
Harvard Medical School  
25 Shattuck Street  
Boston, MA 02115

Dear Dr. Alexander,

FDA strongly supports the development of disease-specific platform trials. In particular, the agency is

We expect no increased regulatory risk to result from the innovative statistical design for GBM AGILE. Moreover, depending on the specific results, we anticipate that data from experimental arms that have “graduated” and been confirmed by GBM AGILE will be used as the foundation for new drug application (NDA) or biological drug application (BLA) submissions and registration.

the master protocol for GBM AGILE, and there have been extensive interactions between the trial design team and FDA biostatisticians. While you can be assured that each substudy will be reviewed prospectively and we will base all regulatory decision-making on our usual high standards, we agree with and support the design and objectives of the GBM AGILE trial.

We expect no increased regulatory risk to result from the innovative statistical design for GBM AGILE. Moreover, depending on the specific results, we anticipate that data from experimental arms that have “graduated” and been confirmed by GBM AGILE will be used as the foundation for new drug application (NDA) or biological drug application (BLA) submissions and registration.

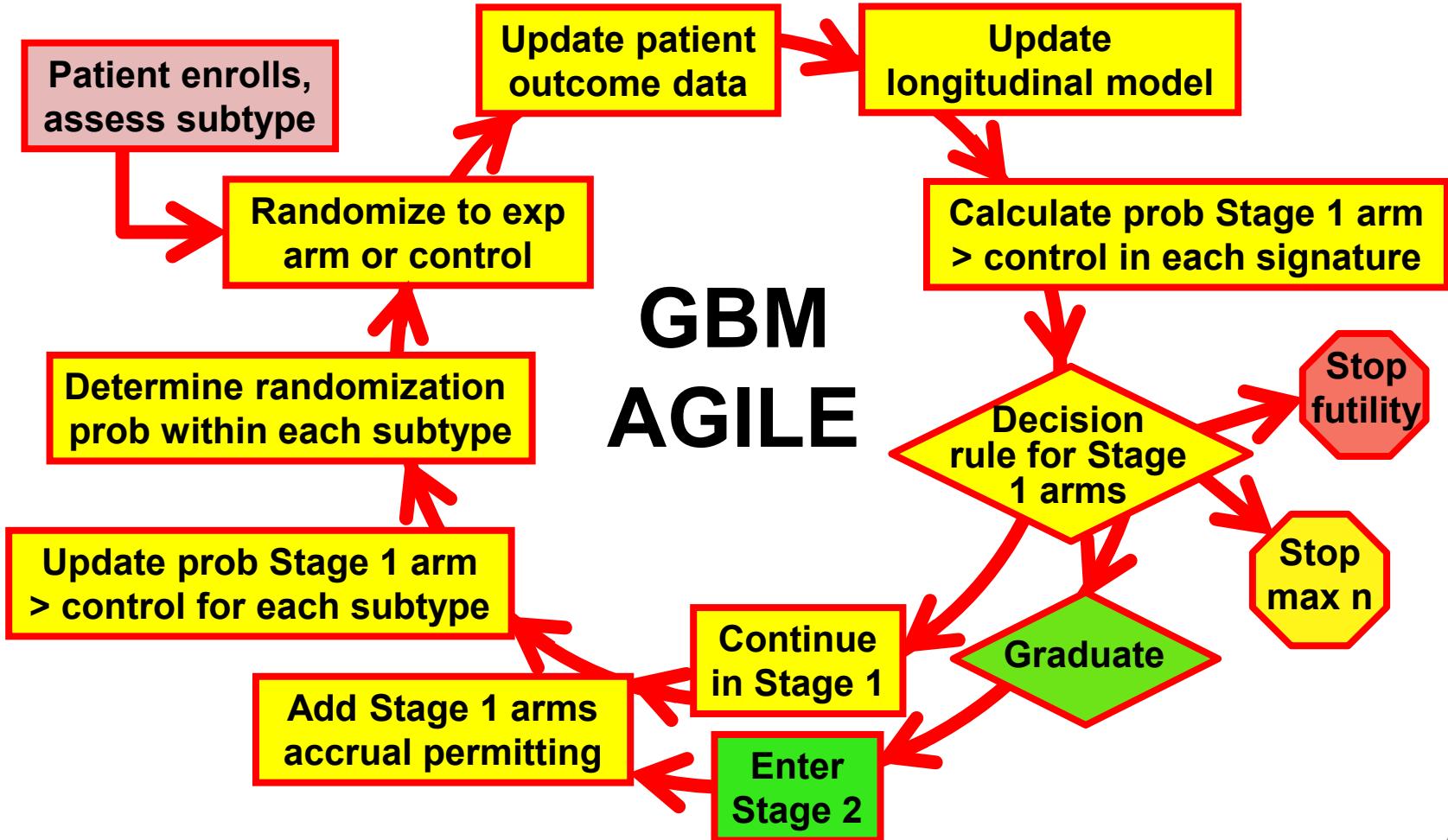
Thank you for your tireless efforts to address this debilitating, deadly disease.

Sincerely,



Janet Woodcock, M.D.  
Director,  
Center for Drug Evaluation and Research  
U.S. Food and Drug Administration

# GBM AGILE



# **Design characteristics of GBM AGILE**

- **Subtypes of GBM (& 5 possible signatures):**
  - Newly diagnosed MGMT methylated
  - Newly diagnosed unmethylated
  - Recurrent disease
- **Exp arm randomization updated monthly within subtypes:**
  - Adaptive (Thompson) in arm's Stage 1 (n between 100 & 150)
  - Fixed in Stage 2 (n = 50)
- **Control arms 20% randomization throughout**
- **Exp arm graduates with a signature if any PP  $\geq$  80%**
- **Exp arm drops for futility,  $50 \leq n \leq 100$ , if PP  $< 25\%$  for all sigs.**
- **All patients followed for 12 months after last patient accrued**

# Other design characteristics

- Final efficacy analysis of OS: Probability distribution of HR, including whether  $\Pr(\text{HR} < 1.0)$ , for each signature
- Based on both Stages 1 and 2 (plus 12 months follow-up)
- If Stage 2 then claim “significance” if  $\Pr(\text{HR} < 1.0) \geq 0.98$ , which controls 1-sided type I error rate to  $< 0.025$ )
- Each arm has own appendix, that trumps master protocol
- E.g., an arm may have added subtypes and signatures defined by an prospectively defined enrichment biomarker

# **Operating Characteristics by Simulation**

- **To find operating characteristics:**
  - Control Type I error rate—complicated
  - Find power—complicated
  - Sample size distribution
  - Arm's duration in trial
  - Amount of drug required
  - Other?
- **Prospective design essential**
- **Many scenarios**
- **Accrual rate matters**
- **Numbers of drugs coming and going matters**
- **Efficacies of non-index drugs matter**

# Simulated trial ... spoiler alert!

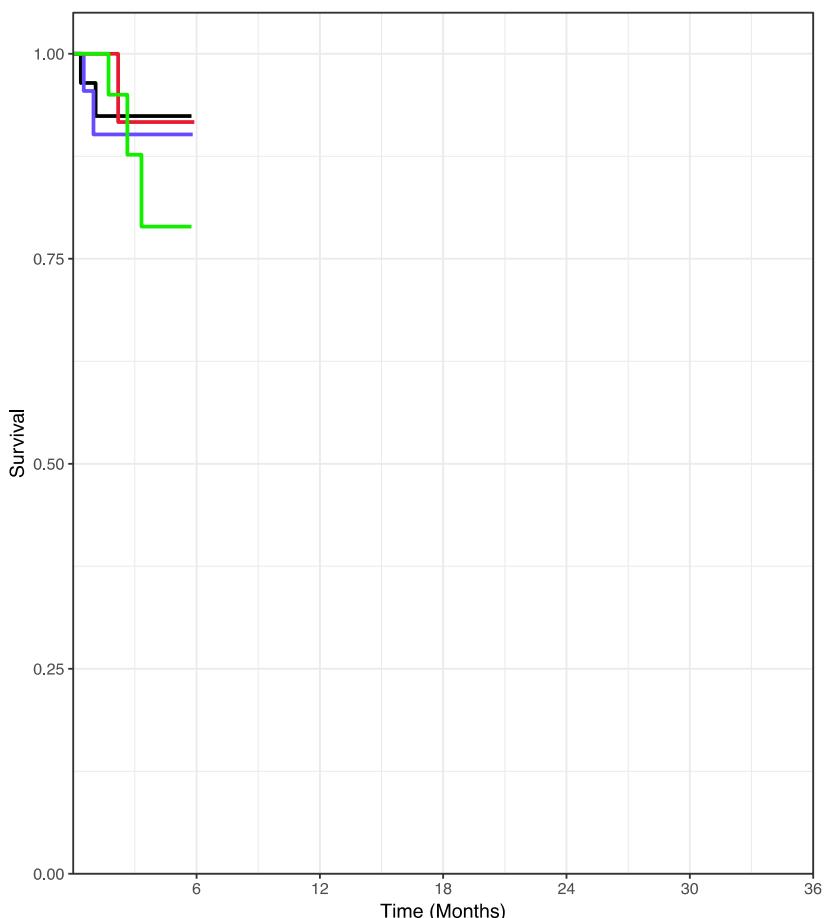
Arm 0 is Control.

Arm	Month enters	Month accrual ends	Months accrued	Sample size	HR (Probability interval)	
					at accrual end	12 mos later
1	0	16	16	61	1.05 (0.59, 1.84)	1.33 (0.90, 1.96)
2	0	19	19	85	0.98 (0.60, 1.61)	1.12 (0.78, 1.60)
3	0	22	22	150	1.00 (0.67, 1.49)	1.02 (0.75, 1.39)
4	12	26	14	70	1.29 (0.76, 2.18)	1.10 (0.76, 1.58)
5	24	41*	17	157	0.54 (0.37, 0.80)	0.44 (0.33, 0.59)
6	24	35	11	54	1.24 (0.74, 2.08)	1.00 (0.69, 1.47)
7	36	47	11	96	0.96 (0.60, 1.55)	0.91 (0.69, 1.20)
8	36	54	18	150	1.04 (0.74, 1.45)	0.98 (0.74, 1.45)

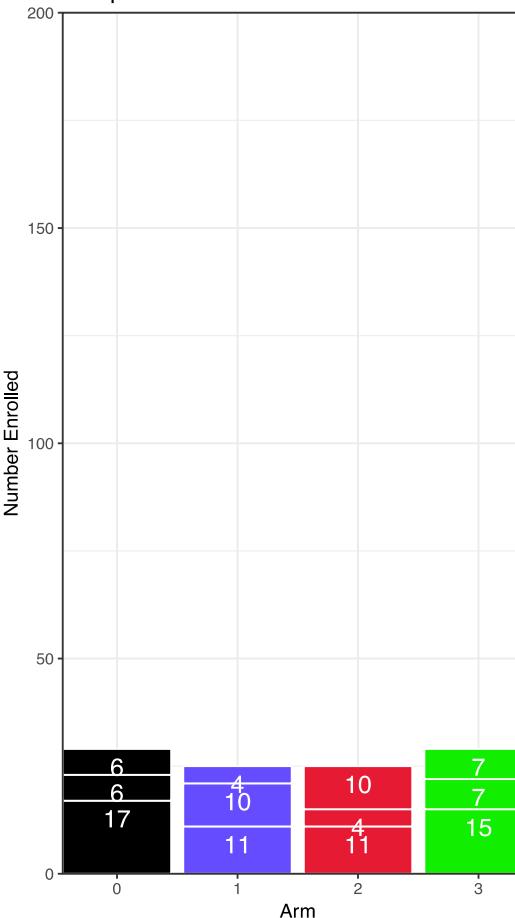
\*Arm 5 graduated in “All” signature after 10 mos (34 mos after trial’s start) & after enrolling 107 patients in its Stage 1.

## Trial begins with control and three investigational arms

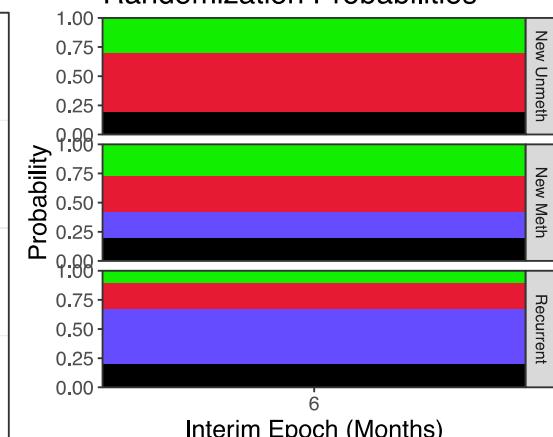
Overall Survival: All Patients. 6 Months After Trial Start



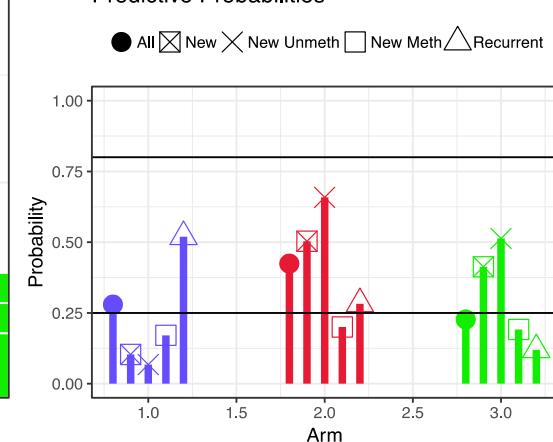
Sample Size



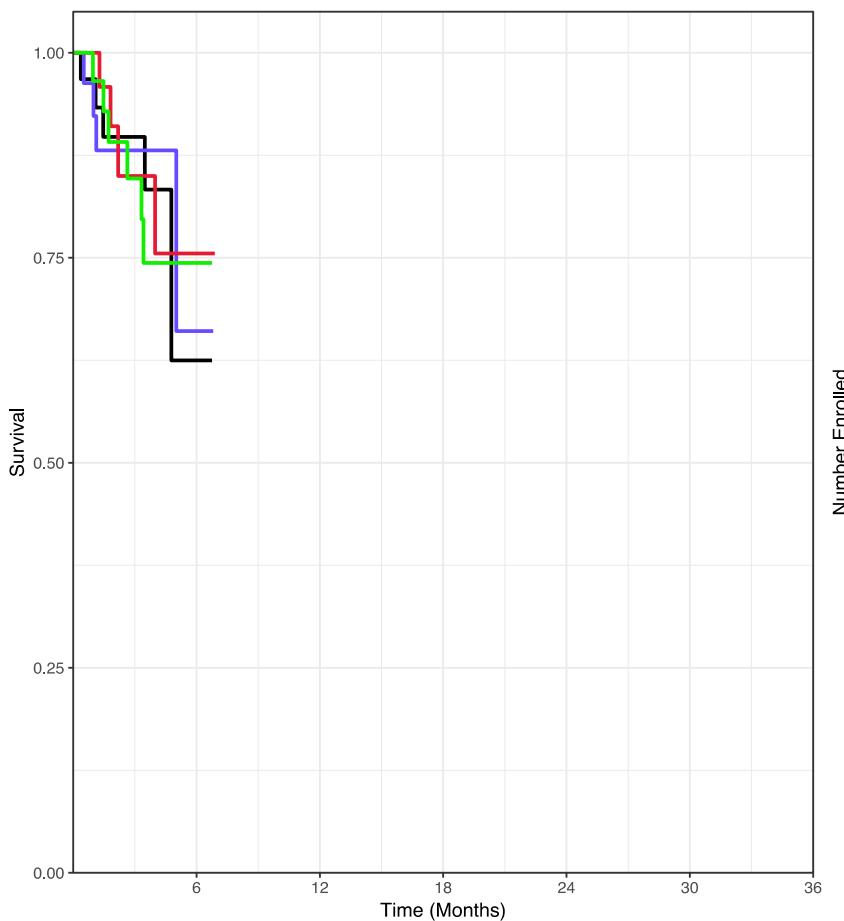
Randomization Probabilities



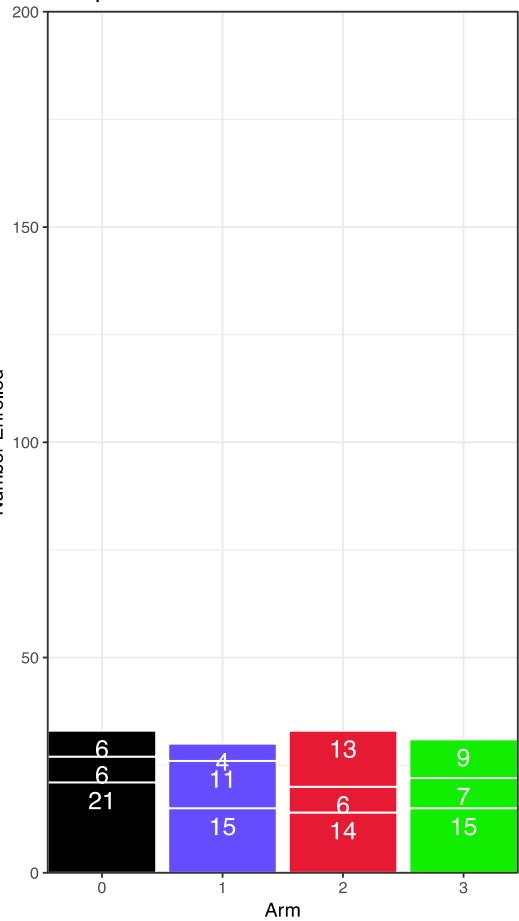
Predictive Probabilities



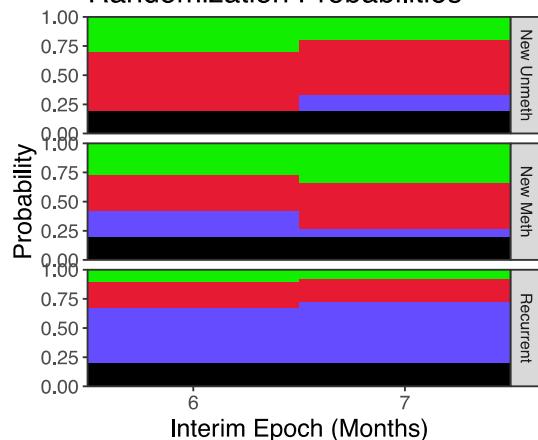
Overall Survival: All Patients. 7 Months After Trial Start



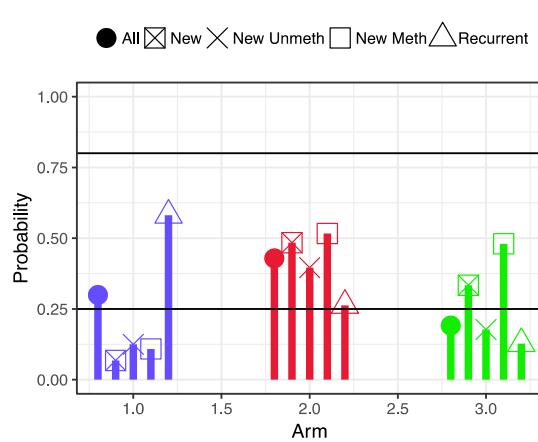
Sample Size



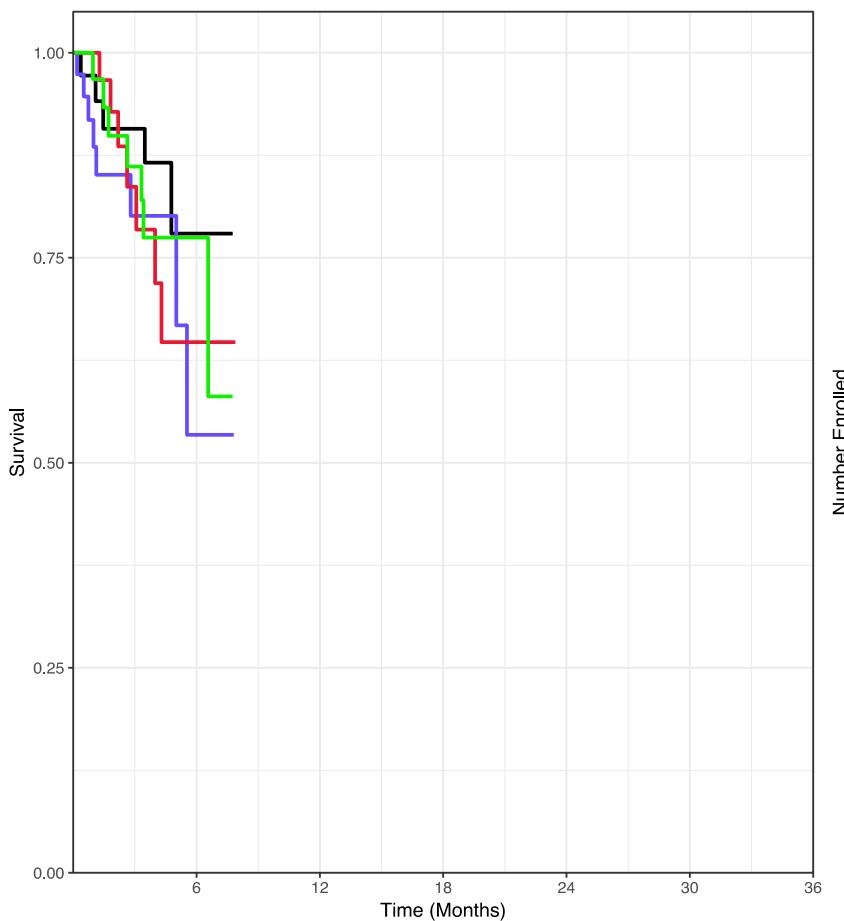
Randomization Probabilities



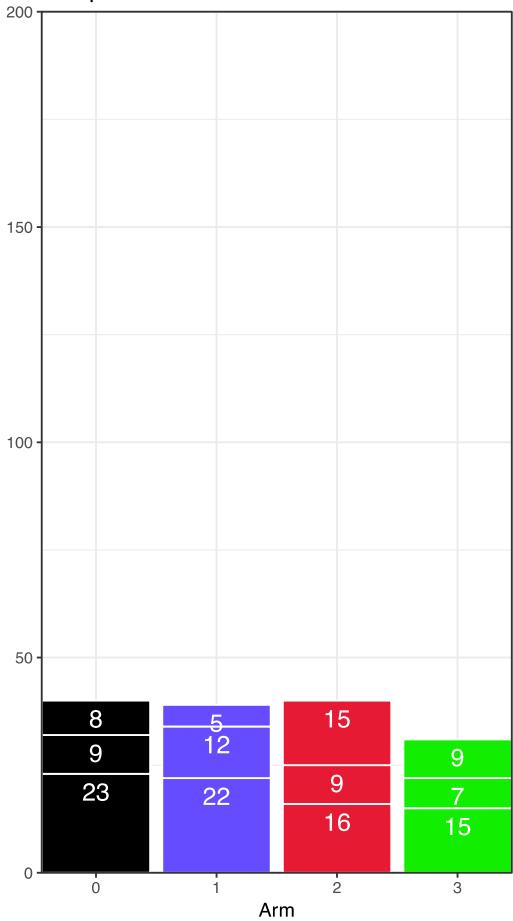
Predictive Probabilities



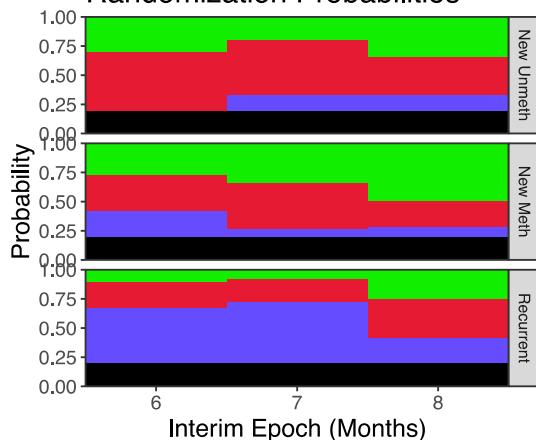
Overall Survival: All Patients. 8 Months After Trial Start



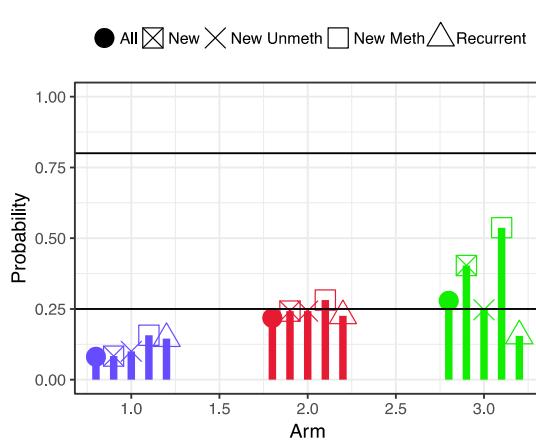
Sample Size



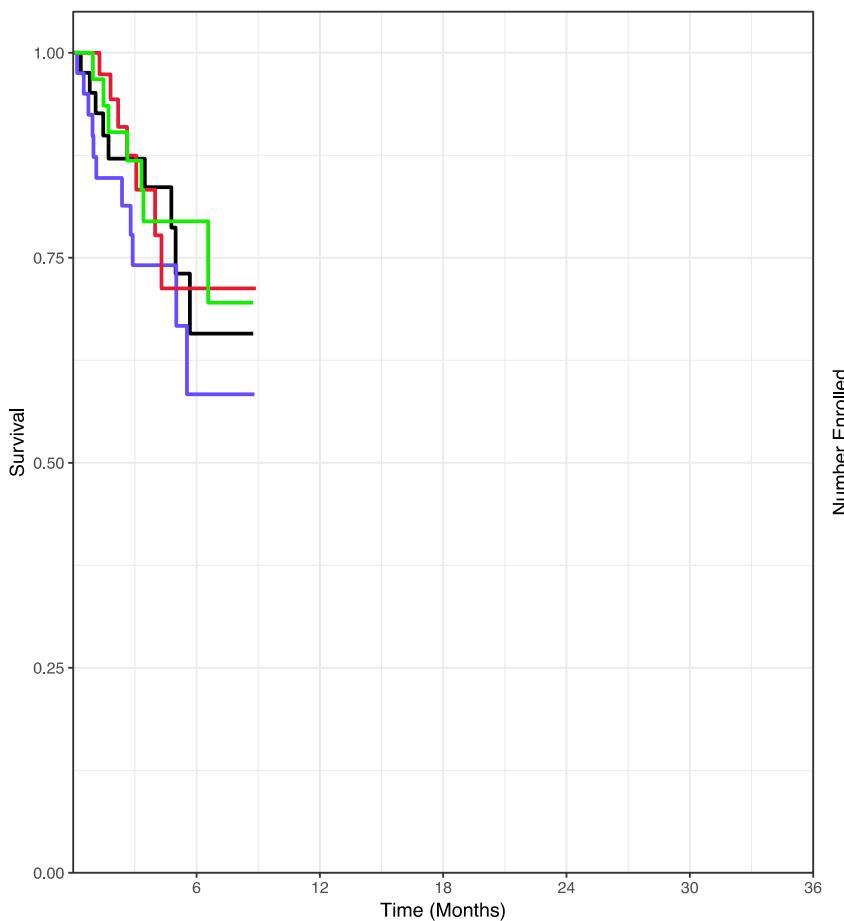
Randomization Probabilities



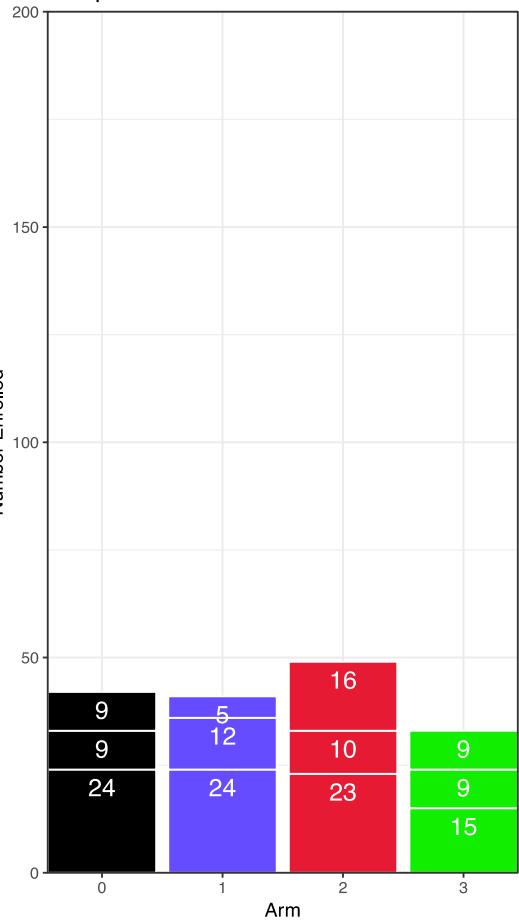
Predictive Probabilities



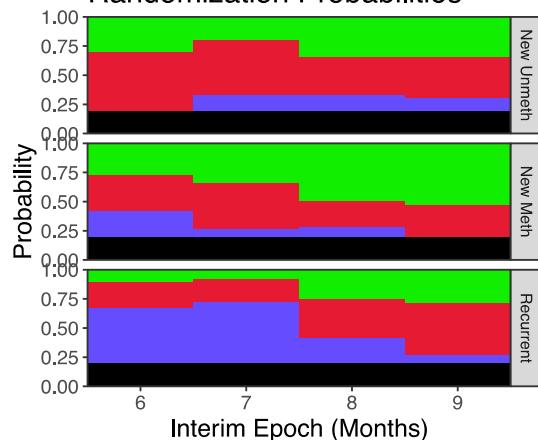
Overall Survival: All Patients. 9 Months After Trial Start



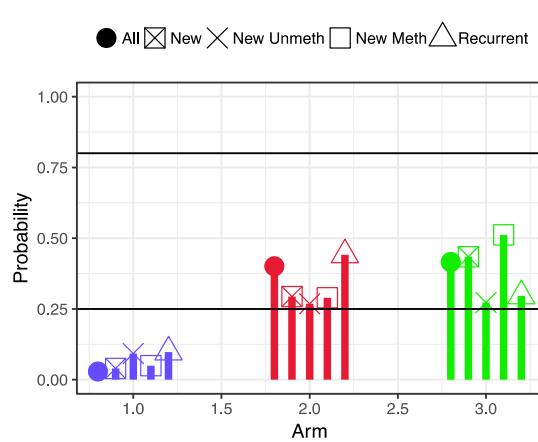
Sample Size



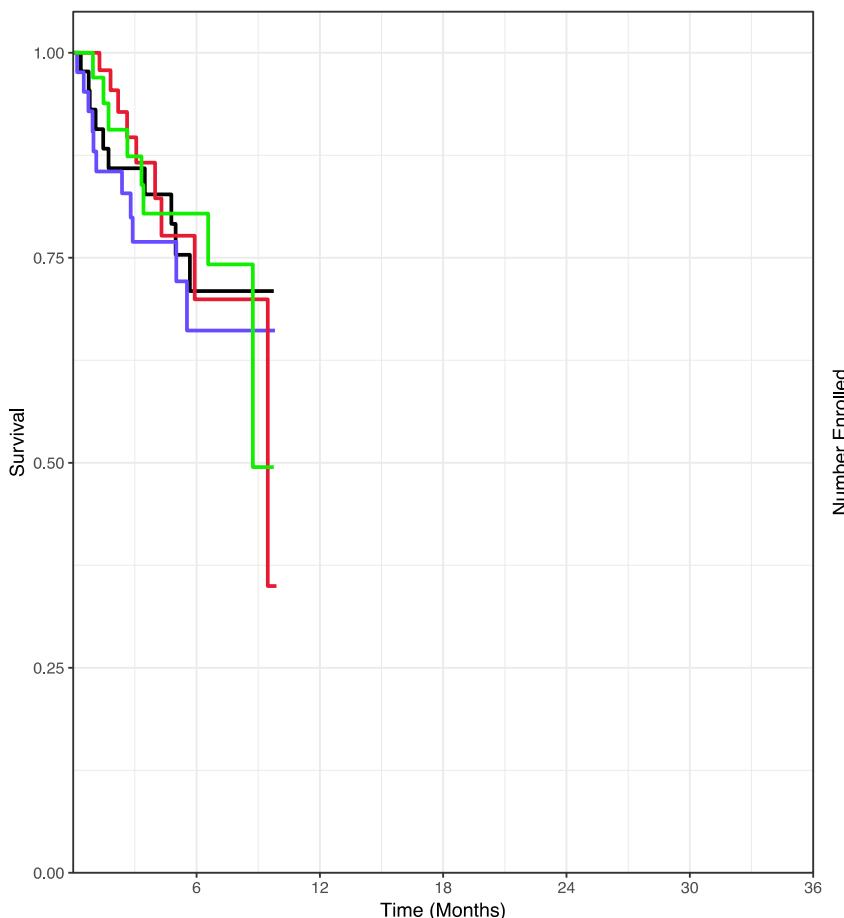
Randomization Probabilities



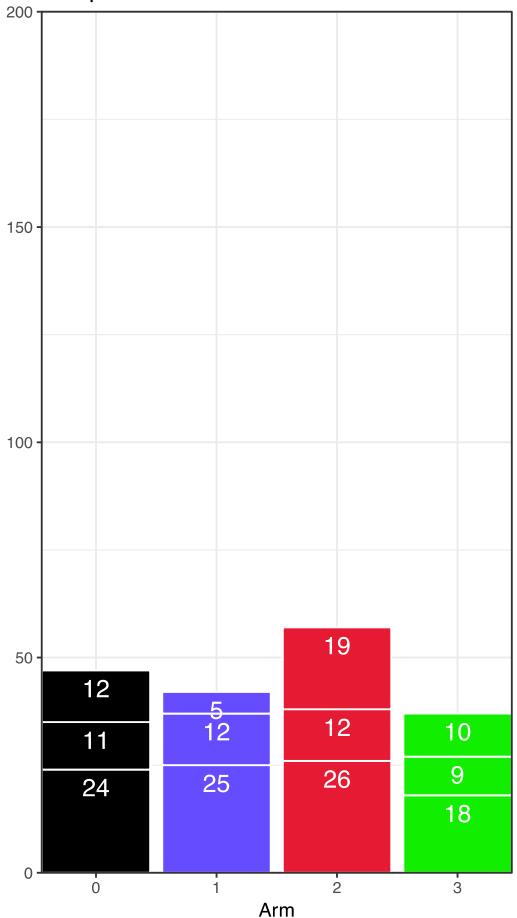
Predictive Probabilities



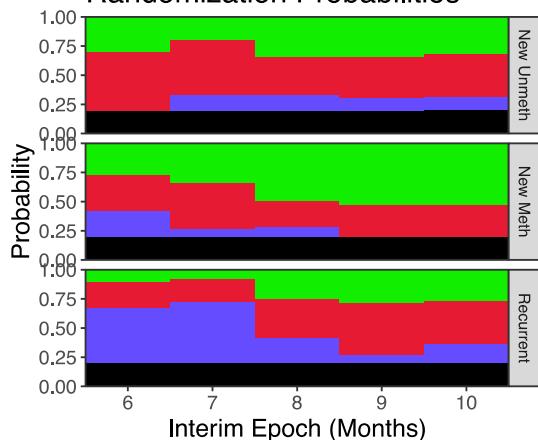
Overall Survival: All Patients. 10 Months After Trial Start



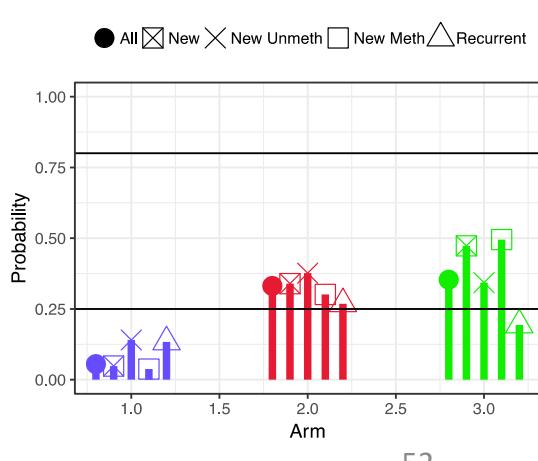
Sample Size



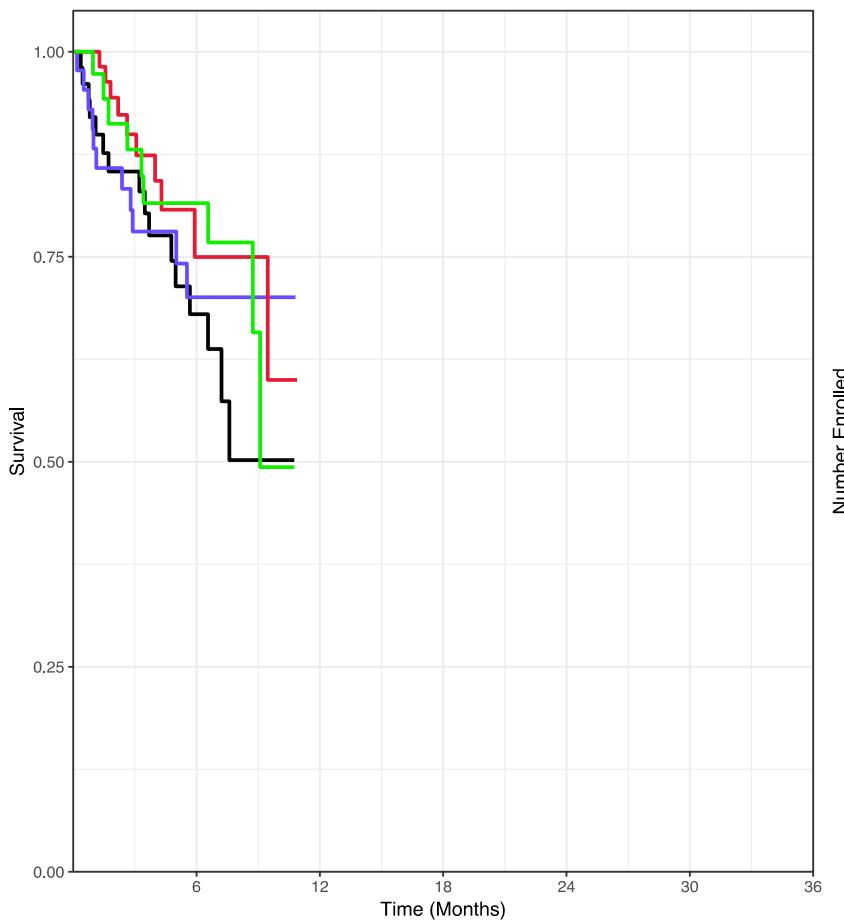
Randomization Probabilities



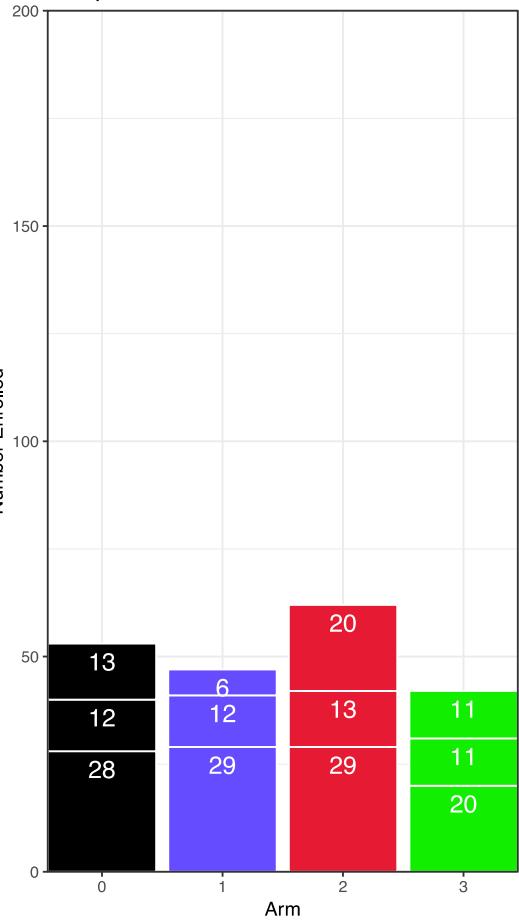
Predictive Probabilities



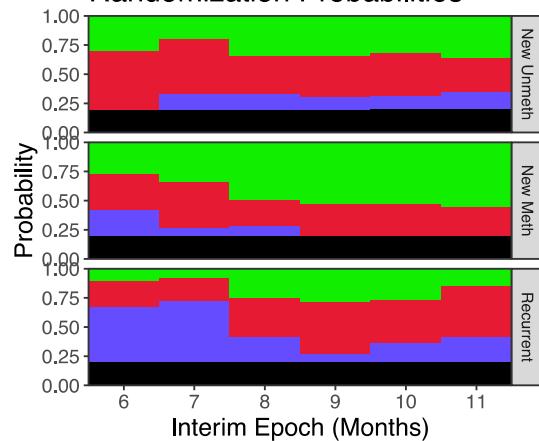
Overall Survival: All Patients. 11 Months After Trial Start



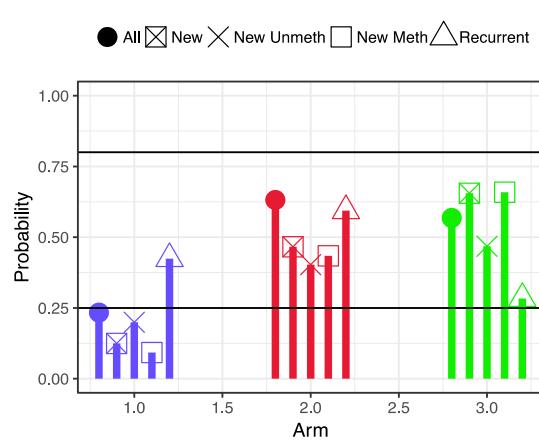
Sample Size



Randomization Probabilities

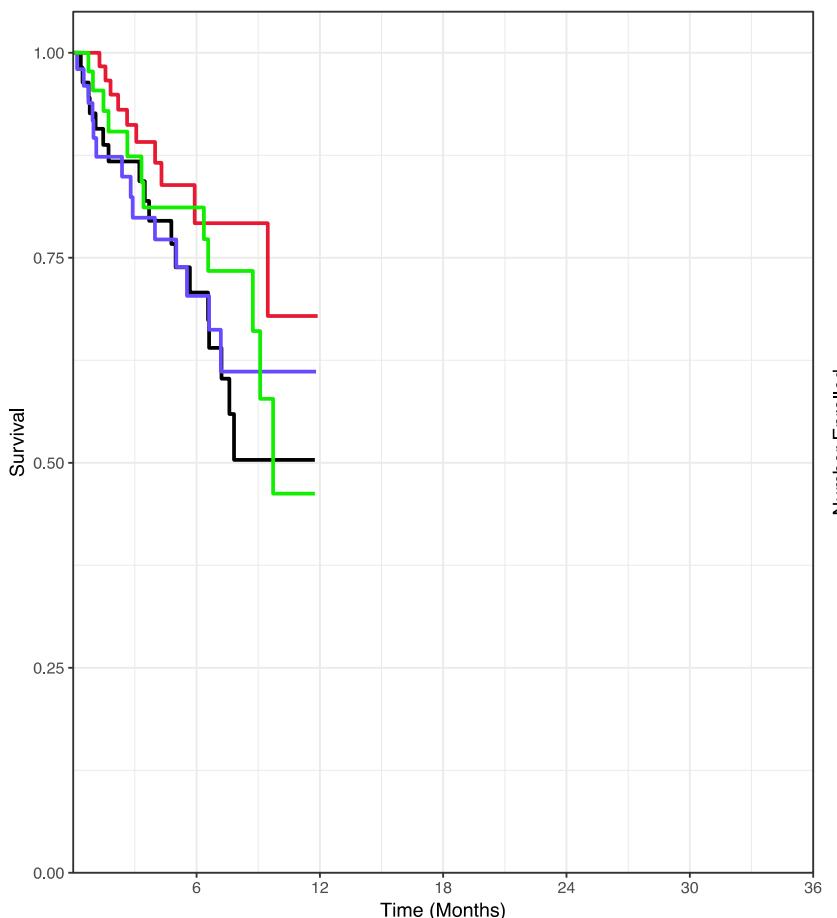


Predictive Probabilities

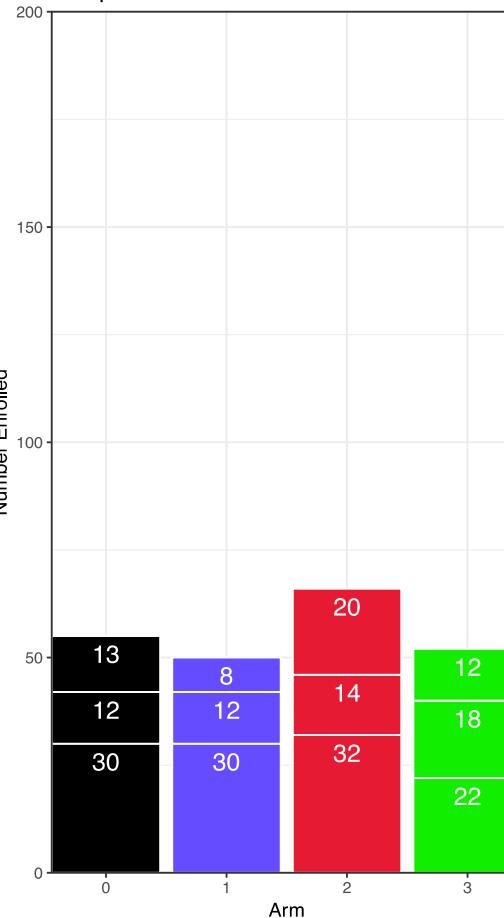


## Investigational Arm 4 enters the trial, 12 months after the start

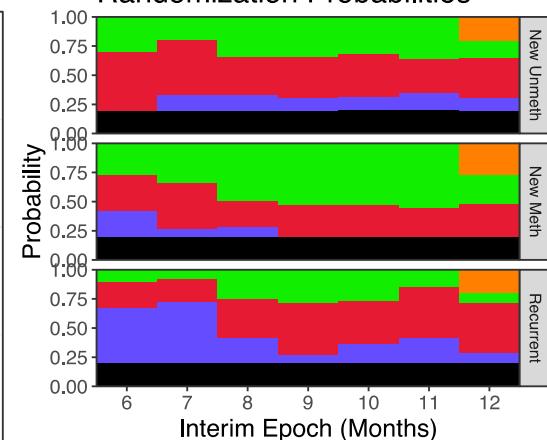
Overall Survival: All Patients. 12 Months After Trial Start



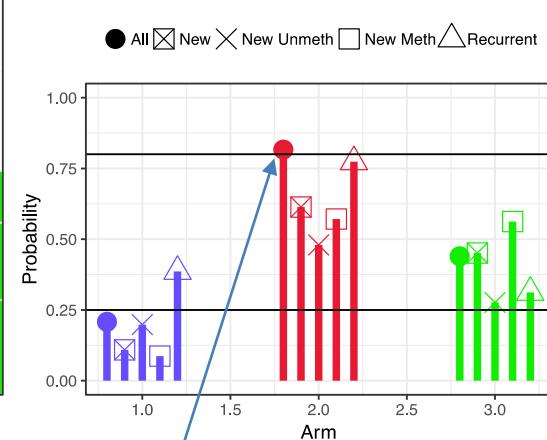
Sample Size



Randomization Probabilities

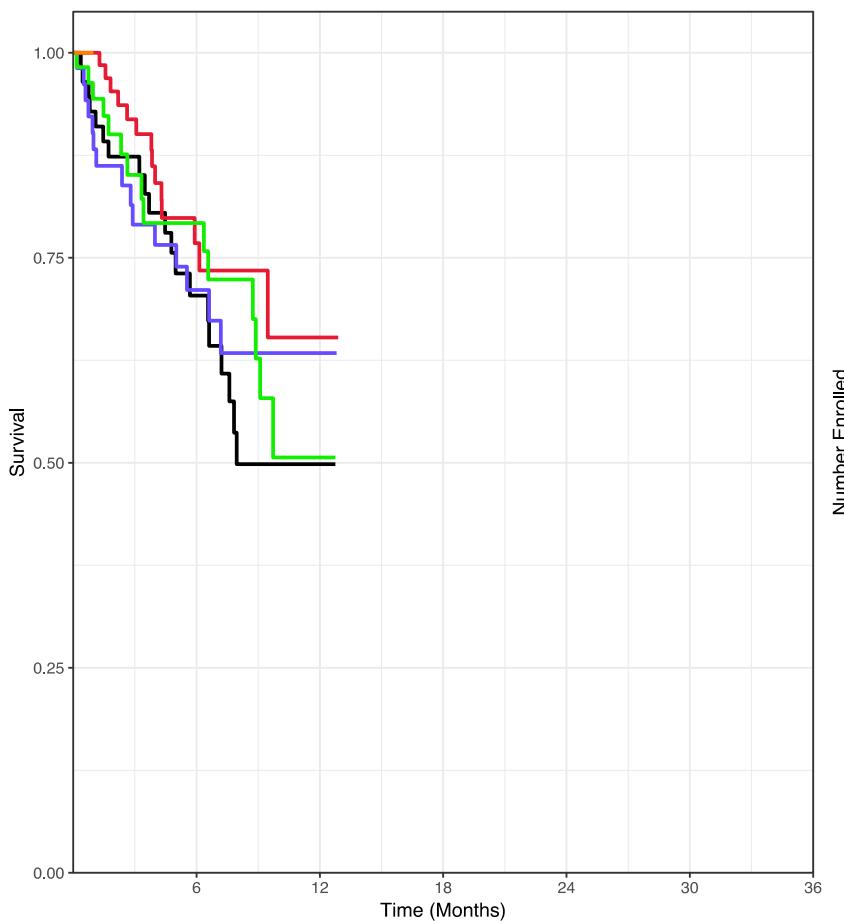


Predictive Probabilities

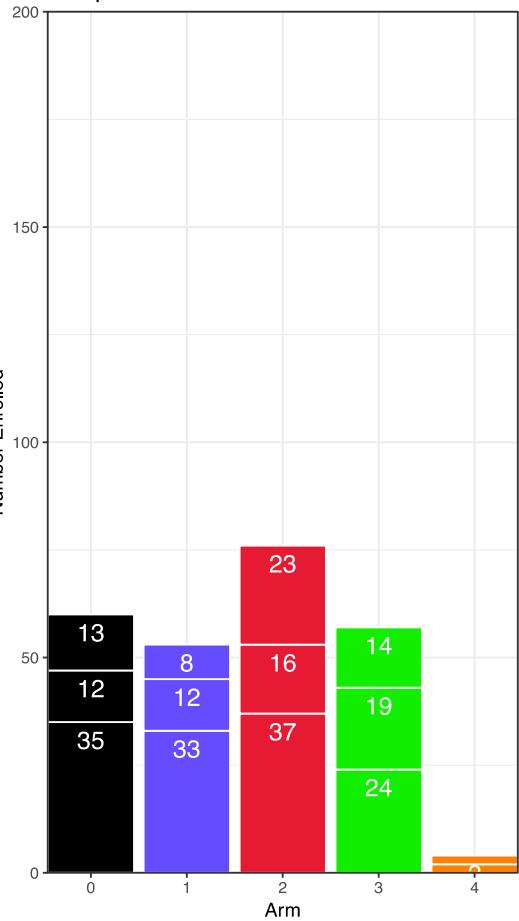


Arm 2 has a predictive probability above the 0.80 threshold, but cannot graduate with fewer than 100 patients

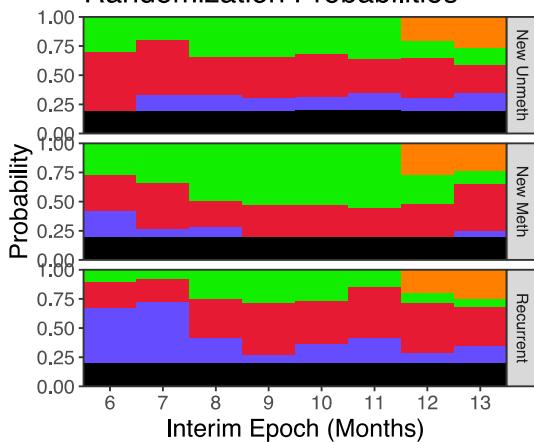
Overall Survival: All Patients. 13 Months After Trial Start



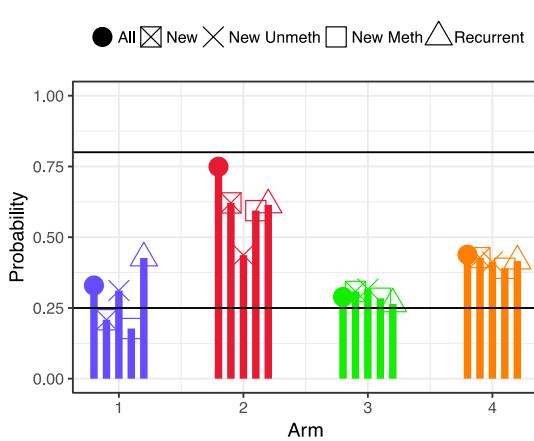
Sample Size



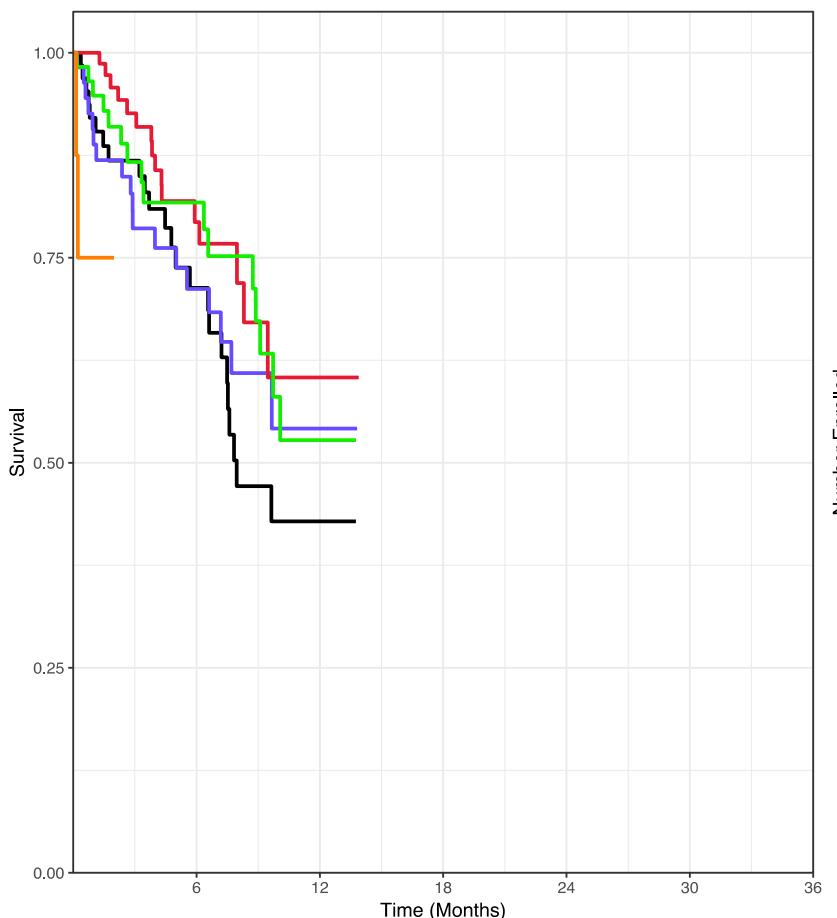
Randomization Probabilities



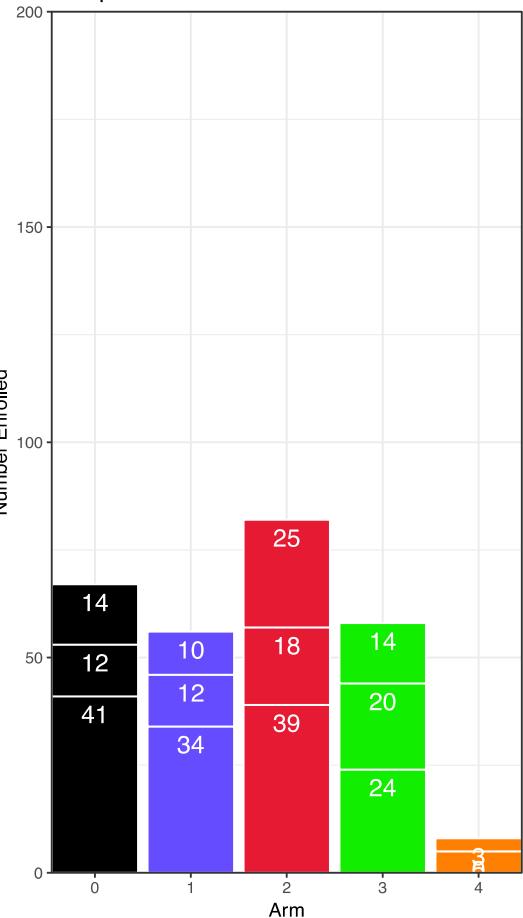
Predictive Probabilities



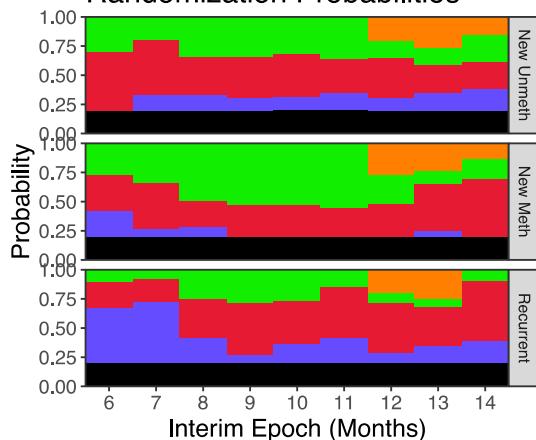
Overall Survival: All Patients. 14 Months After Trial Start



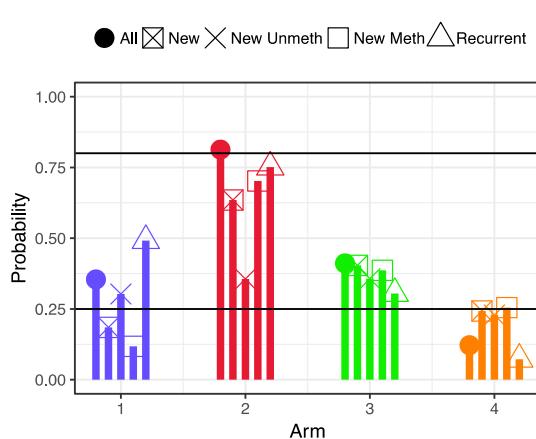
Sample Size



Randomization Probabilities

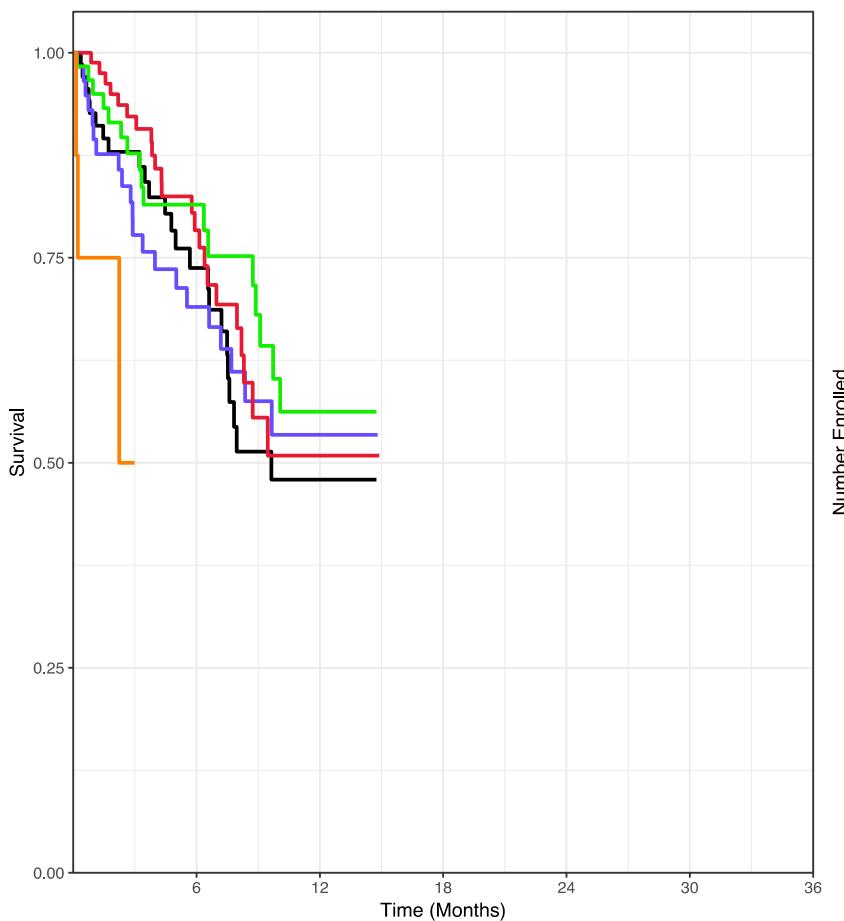


Predictive Probabilities

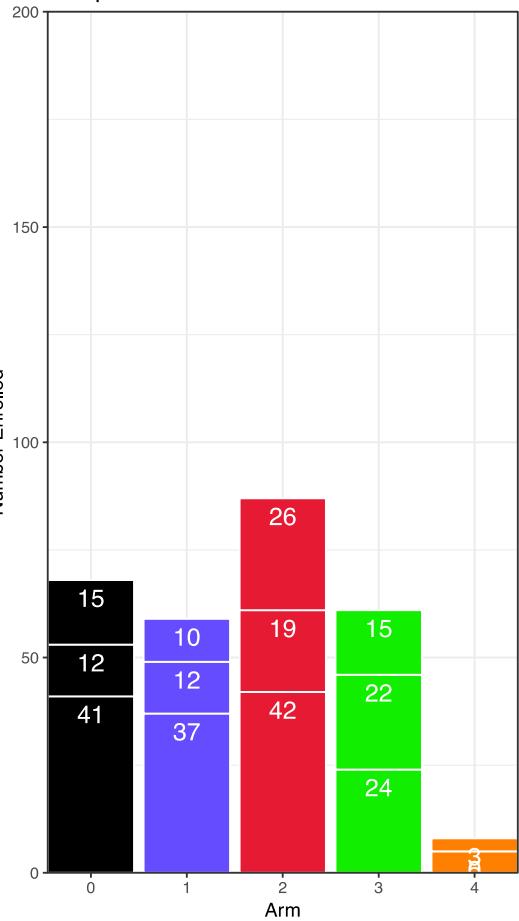


Arm 2 has a predictive probability above the 0.80 threshold, but cannot graduate with fewer than 100 patients

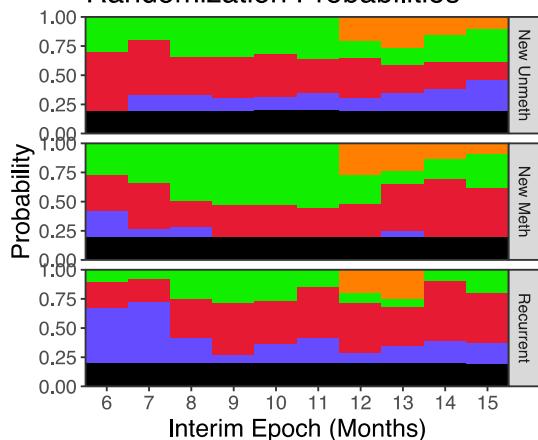
Overall Survival: All Patients. 15 Months After Trial Start



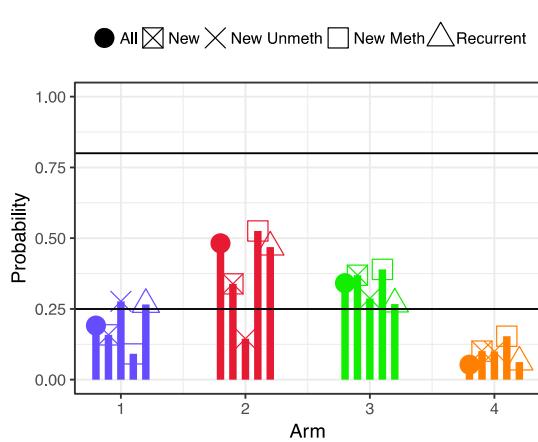
Sample Size



Randomization Probabilities

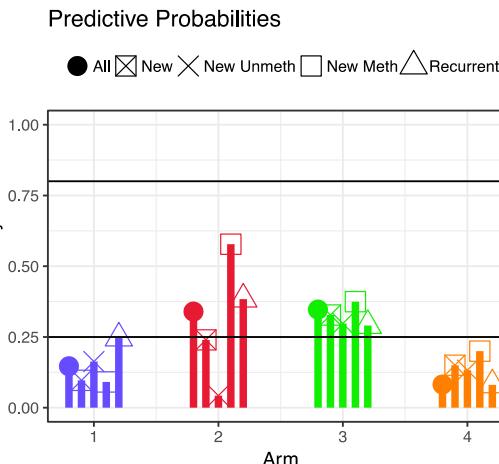
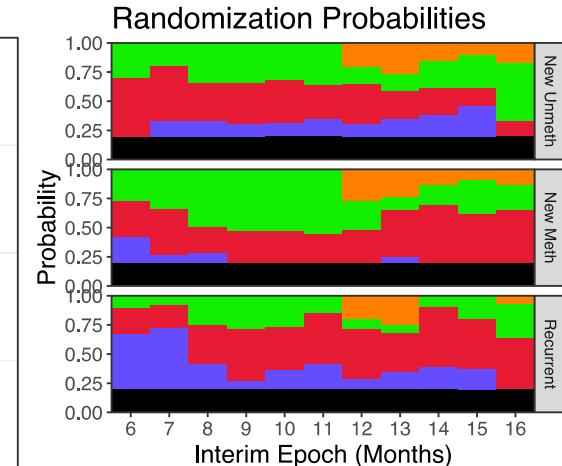
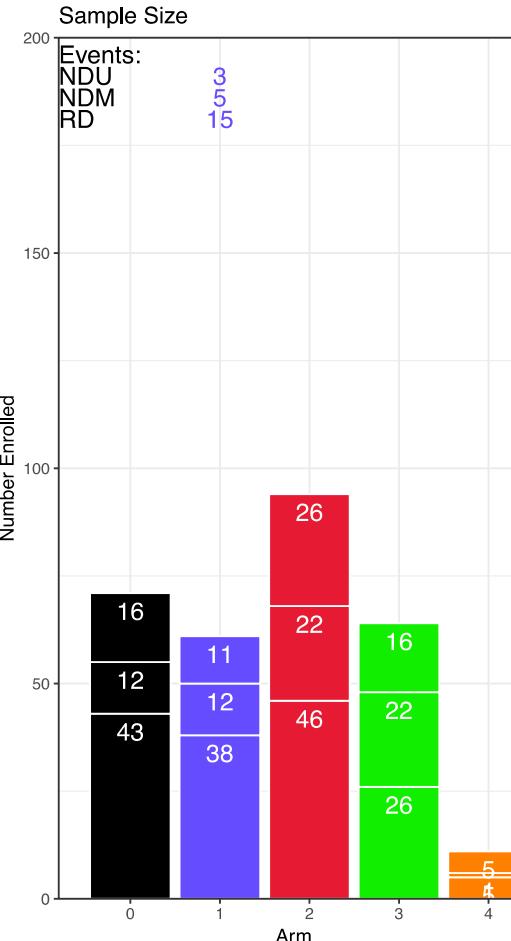
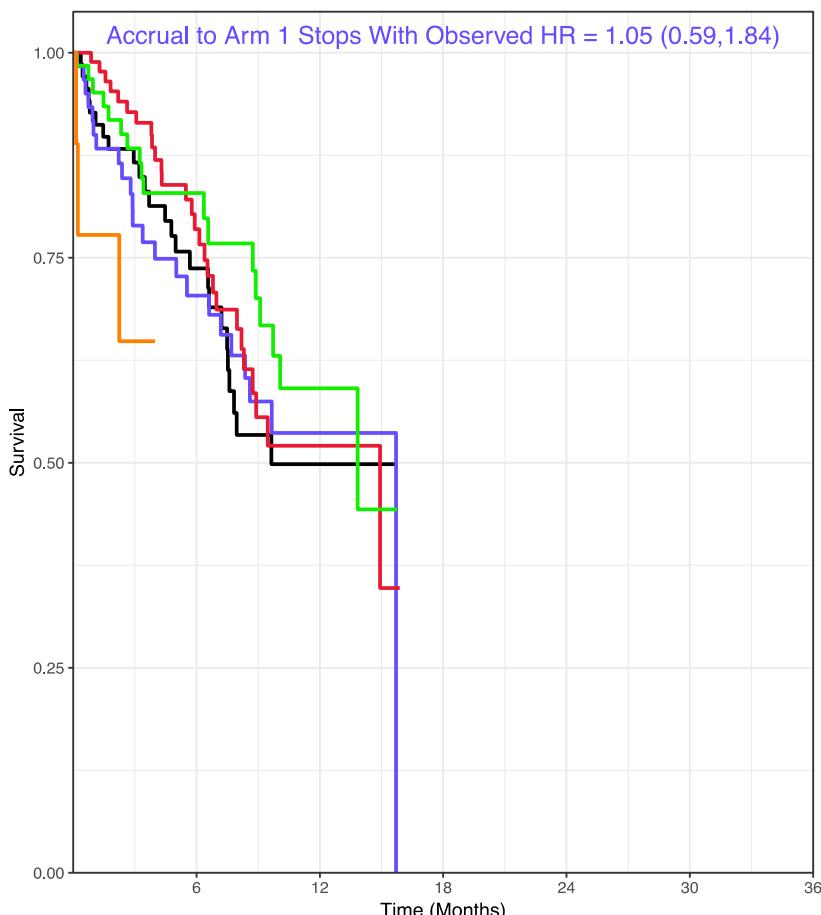


Predictive Probabilities



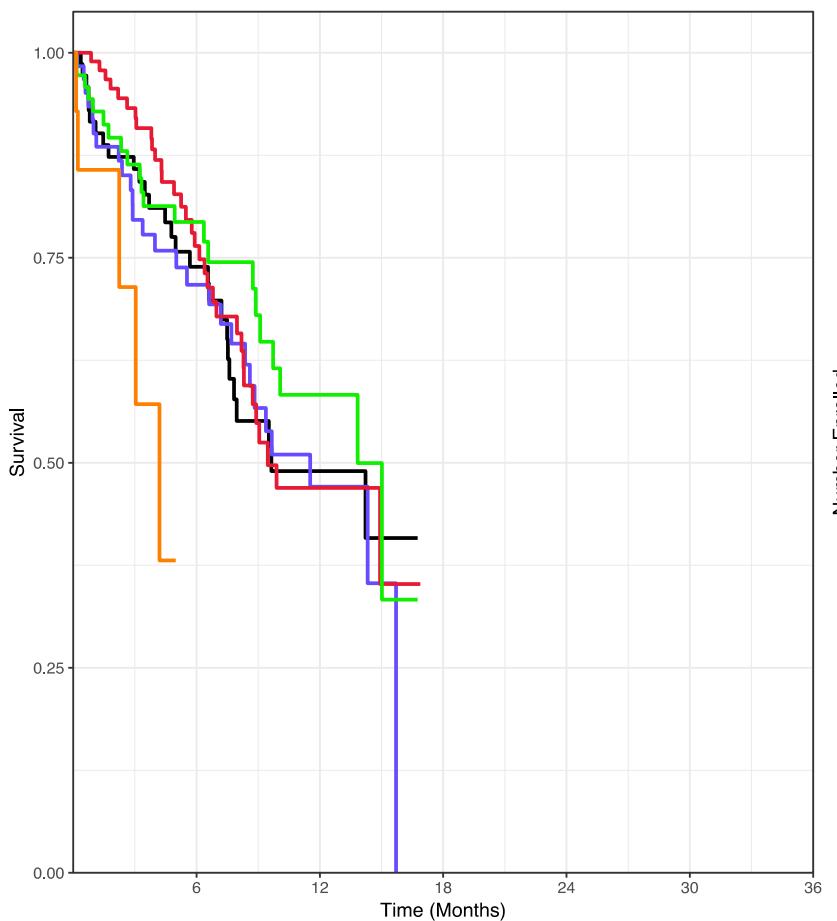
# Month 16: accrual to Arm 1 stops for futility with 61 patients

Overall Survival: All Patients. 16 Months After Trial Start

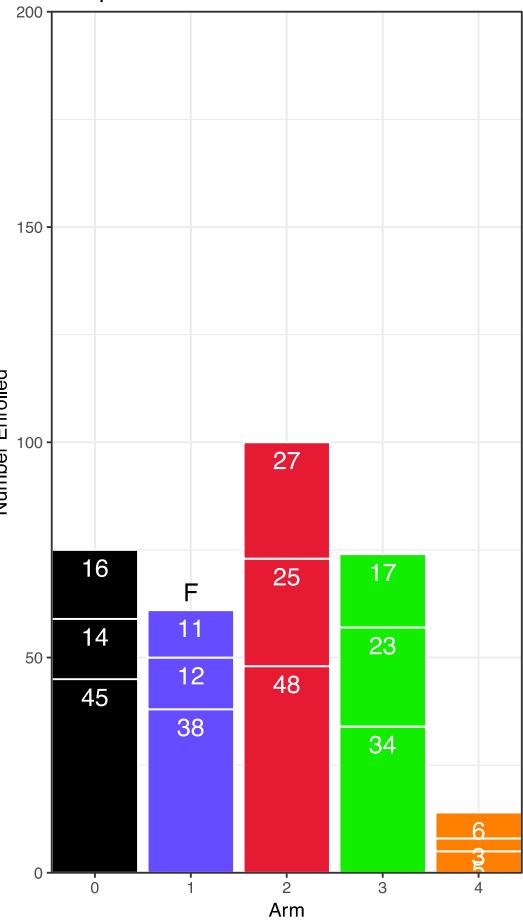


NDU: Newly Diagnosed Unmethylated; NDM: Newly Diagnosed Methylated; RD: Recurrent Disease

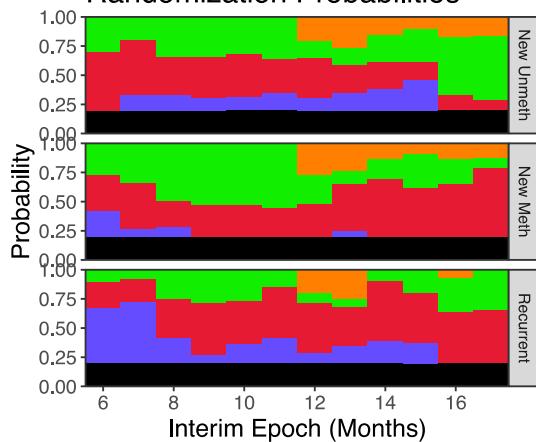
Overall Survival: All Patients. 17 Months After Trial Start



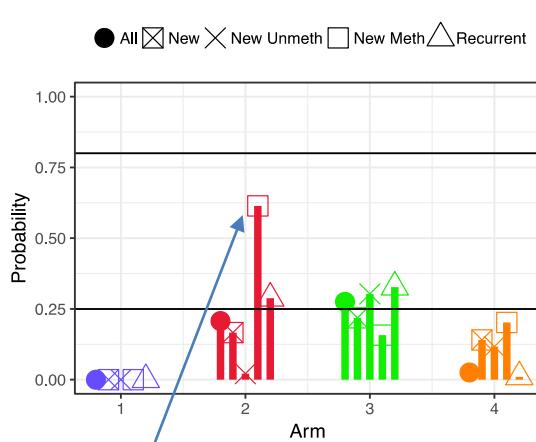
Sample Size



Randomization Probabilities

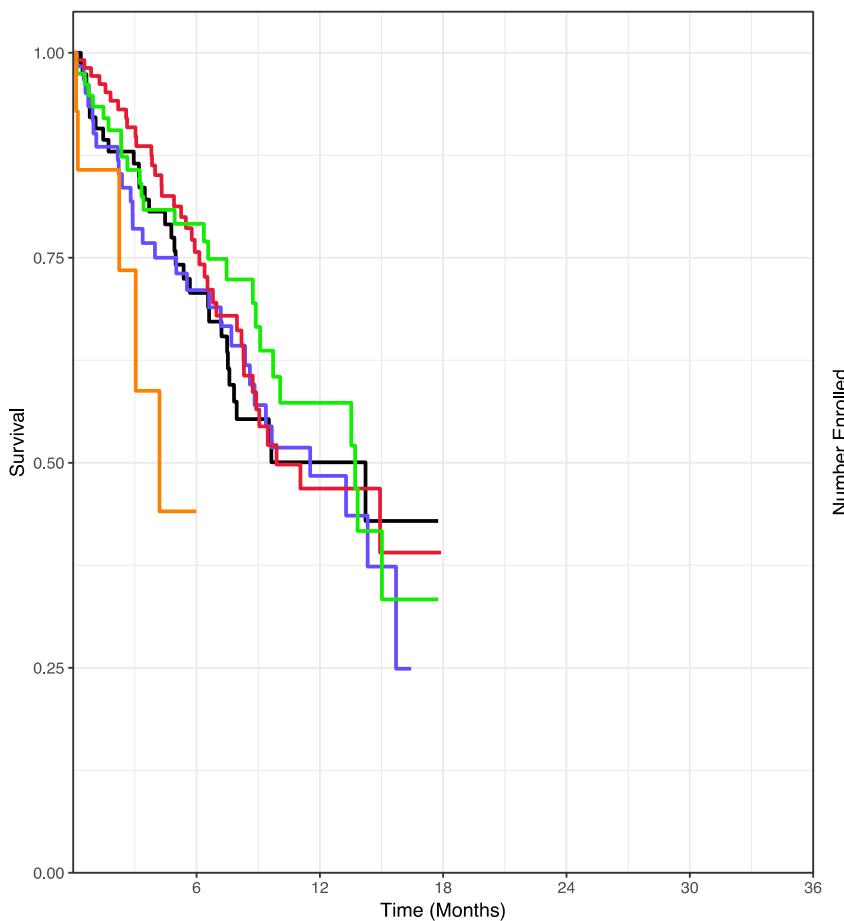


Predictive Probabilities

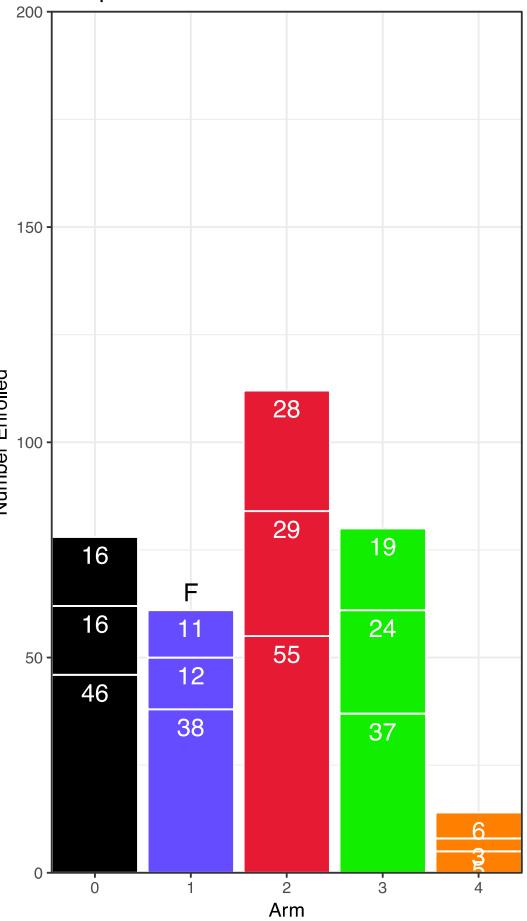


Arm 2 reaches 100 patients, but its predictive probabilities are below the 80% threshold

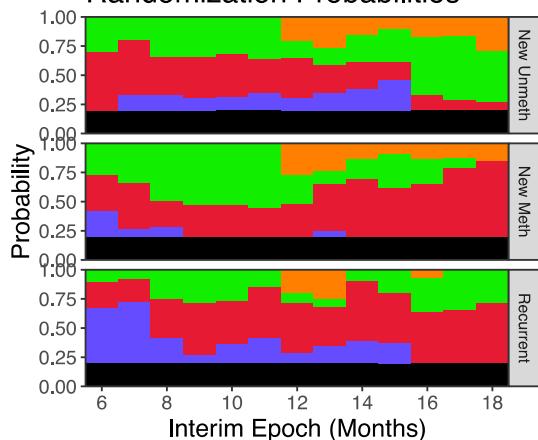
Overall Survival: All Patients. 18 Months After Trial Start



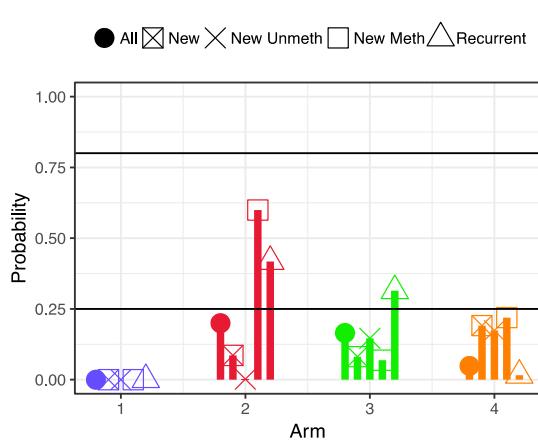
Sample Size



Randomization Probabilities

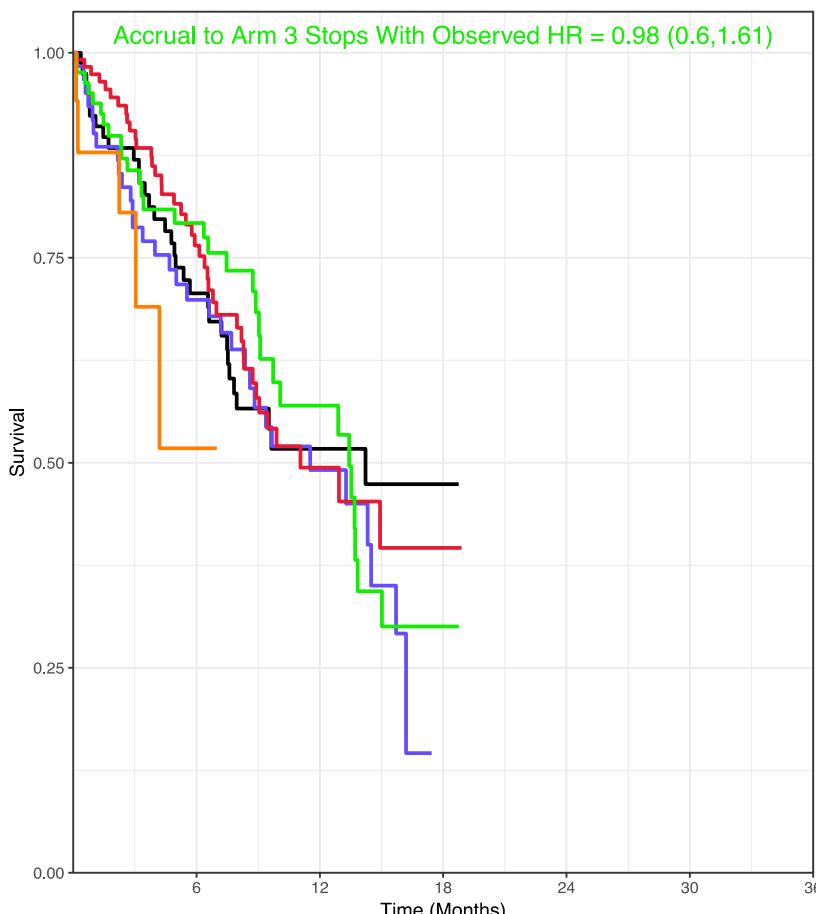


Predictive Probabilities

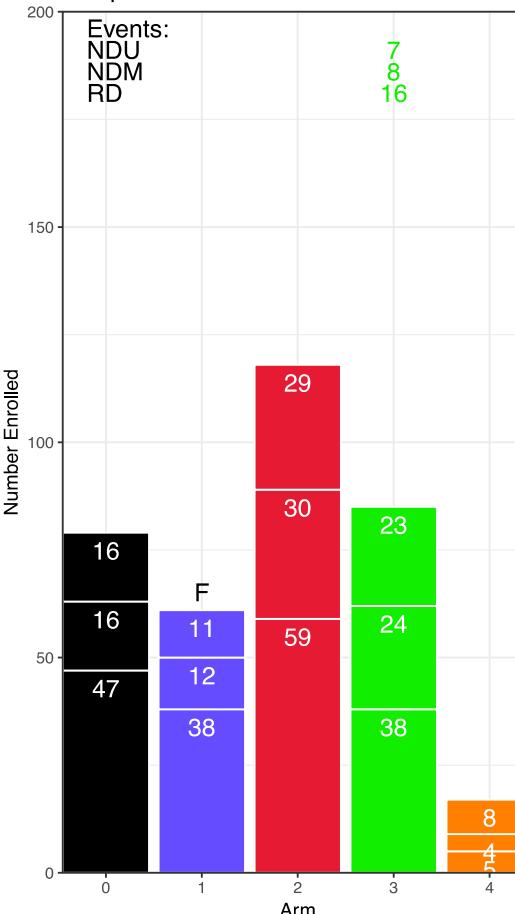


# Month 19: accrual to Arm 3 stops for futility with 85 patients

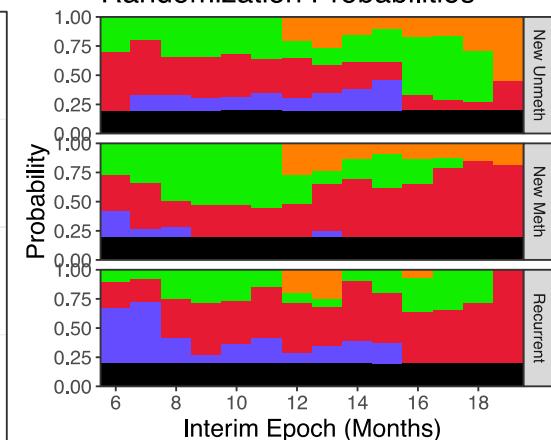
Overall Survival: All Patients. 19 Months After Trial Start



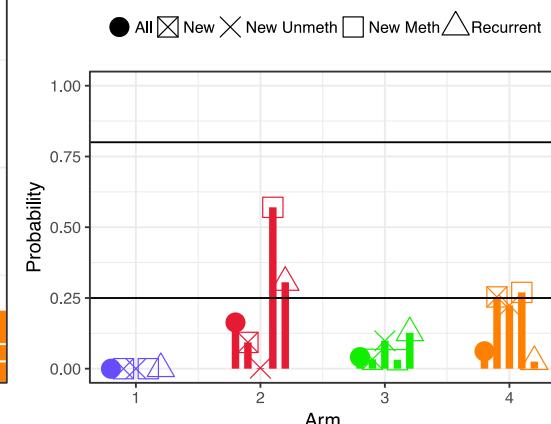
Sample Size



Randomization Probabilities

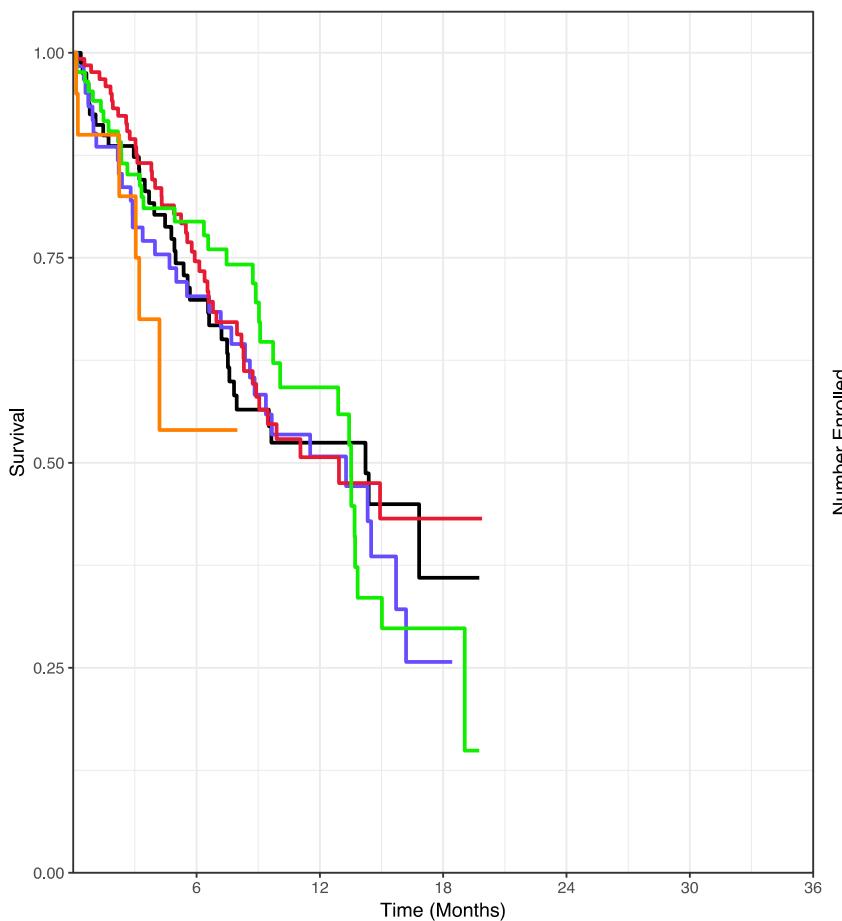


Predictive Probabilities

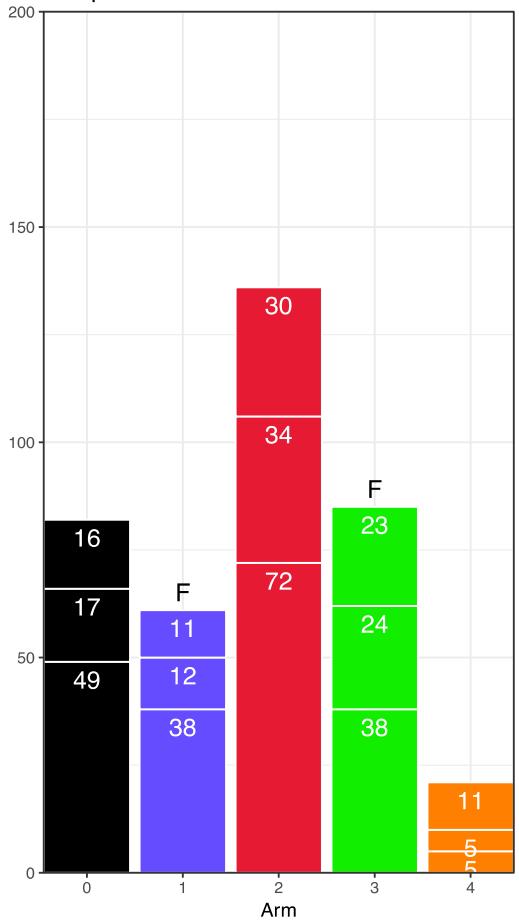


NDU: Newly Diagnosed Unmethylated; NDM: Newly Diagnosed Methylated; RD: Recurrent Disease

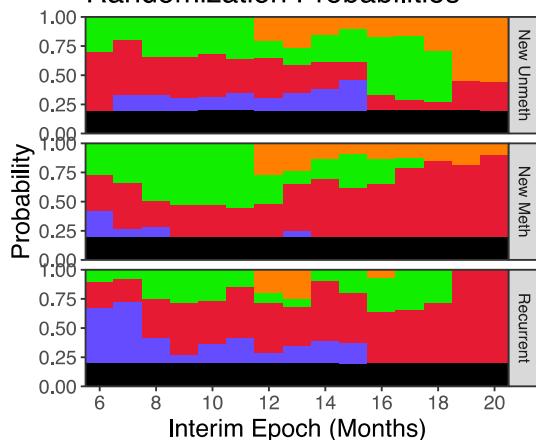
Overall Survival: All Patients. 20 Months After Trial Start



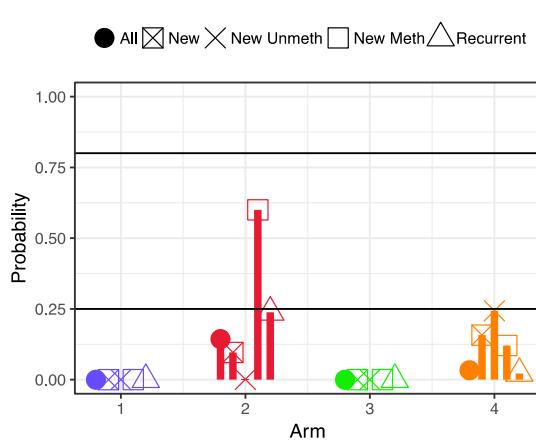
Sample Size



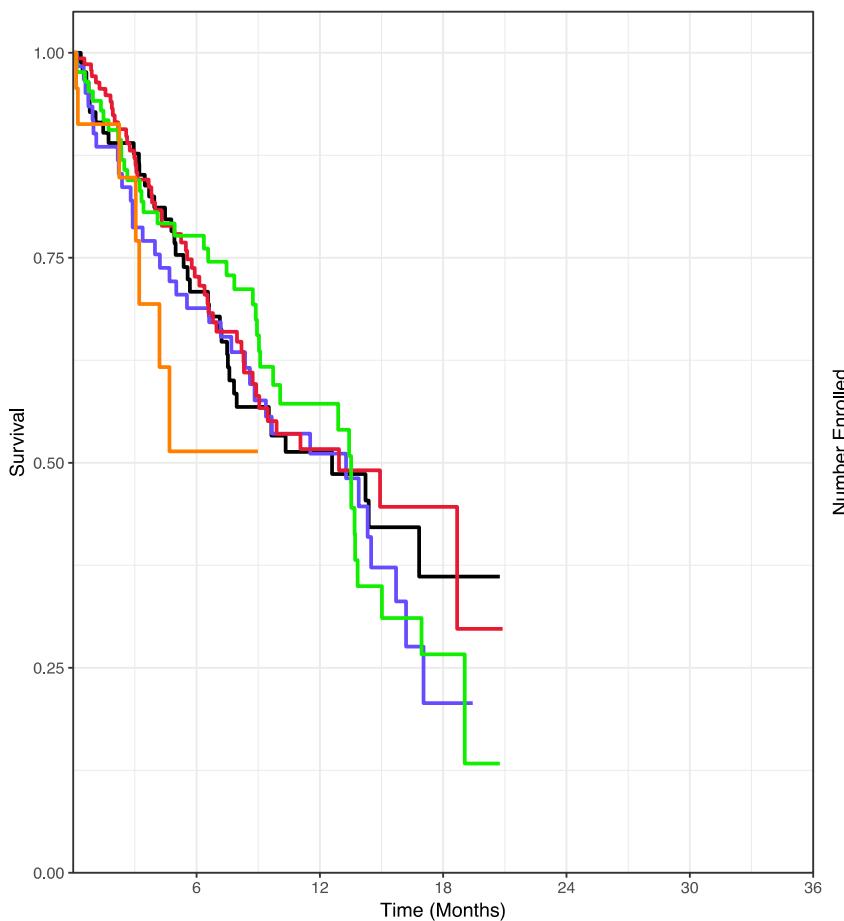
Randomization Probabilities



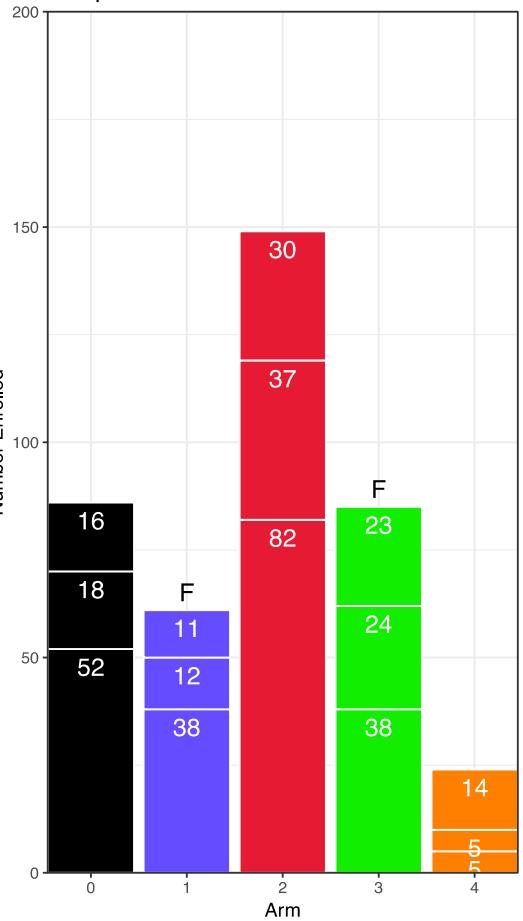
Predictive Probabilities



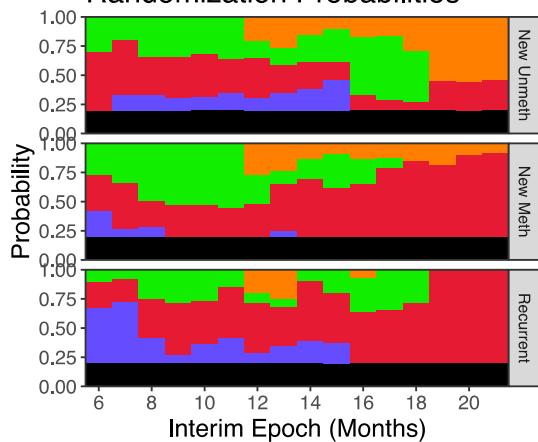
Overall Survival: All Patients. 21 Months After Trial Start



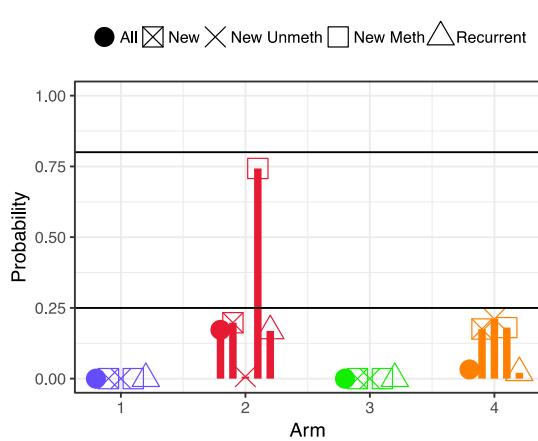
Sample Size



Randomization Probabilities

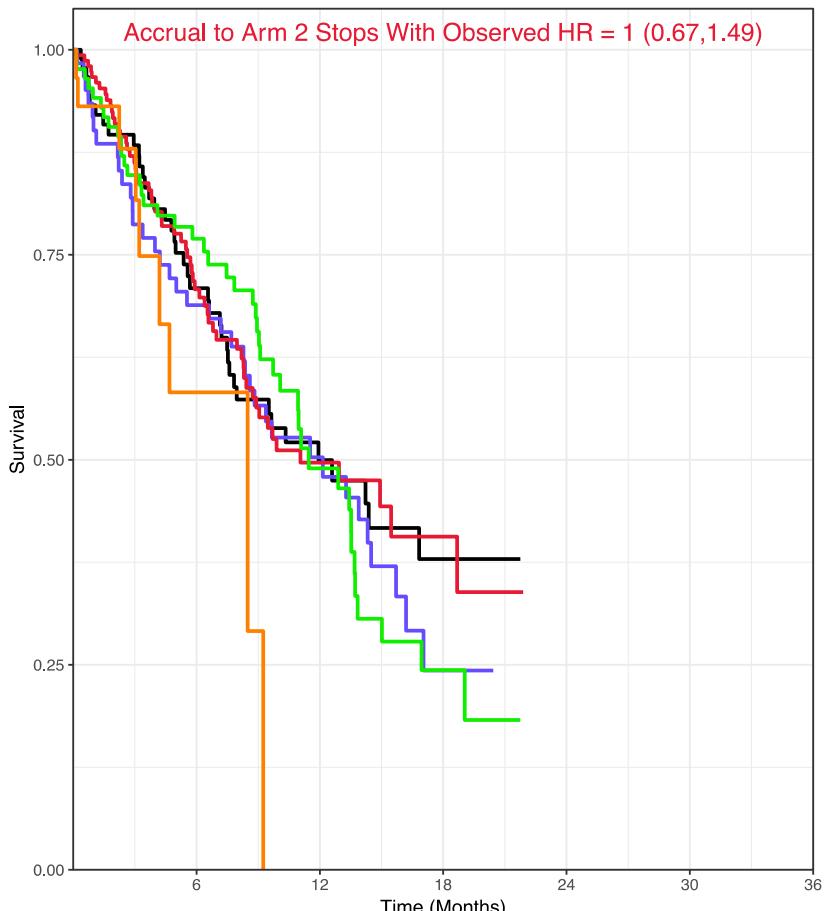


Predictive Probabilities

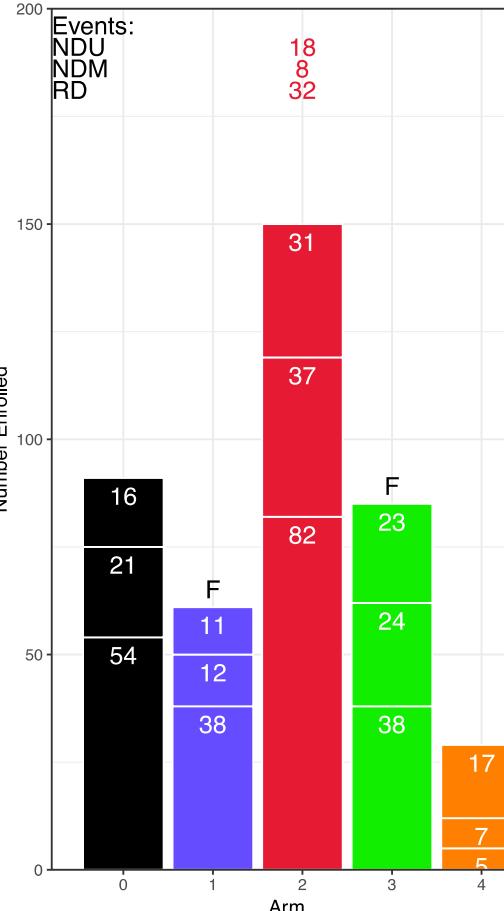


## Month 22: accrual to Arm 2 stops after it reaches 150 patients

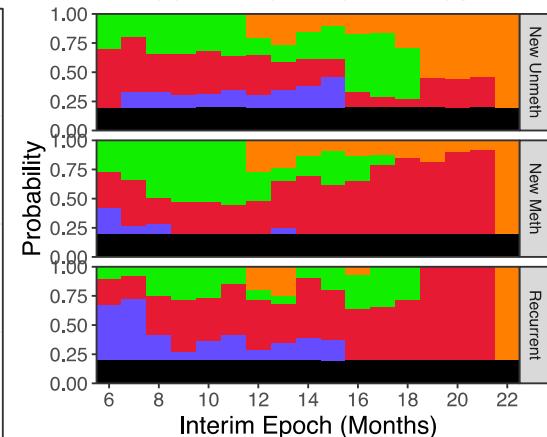
Overall Survival: All Patients. 22 Months After Trial Start



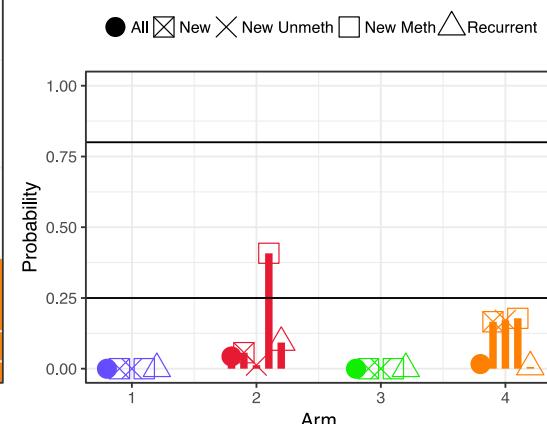
Sample Size



Randomization Probabilities

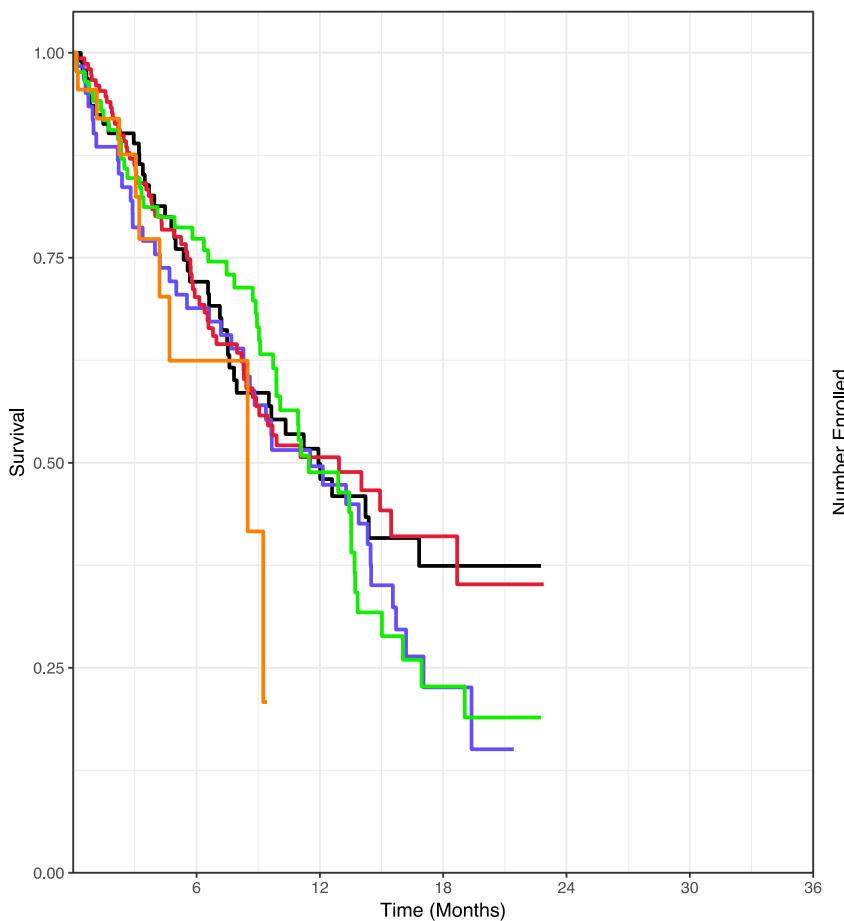


Predictive Probabilities

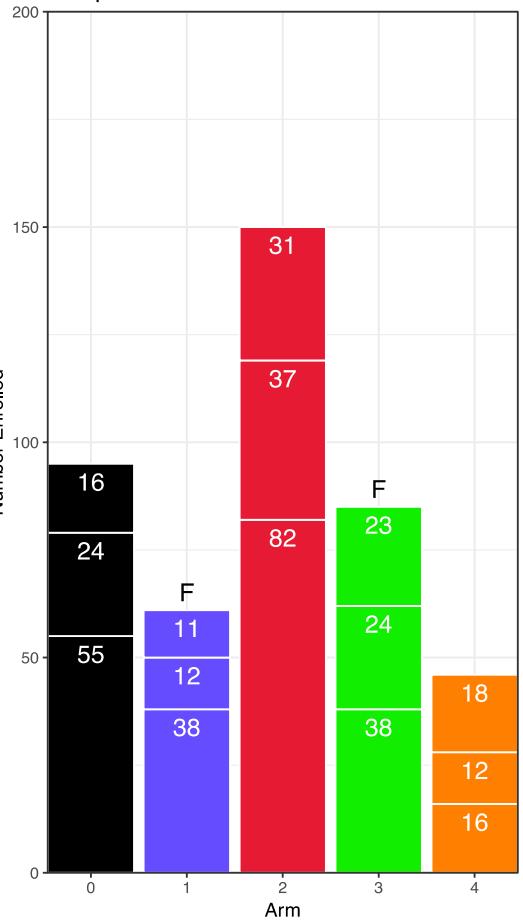


NDU: Newly Diagnosed Unmethylated; NDM: Newly Diagnosed Methylated; RD: Recurrent Disease

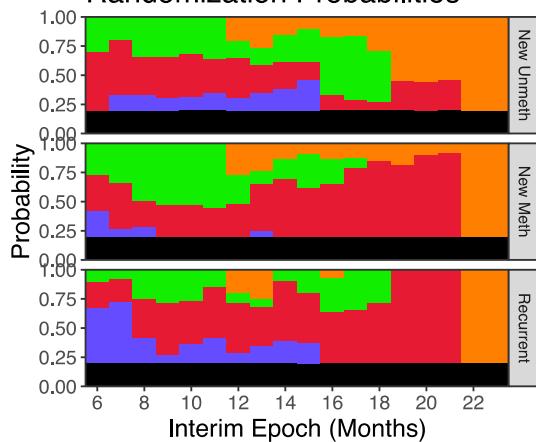
Overall Survival: All Patients. 23 Months After Trial Start



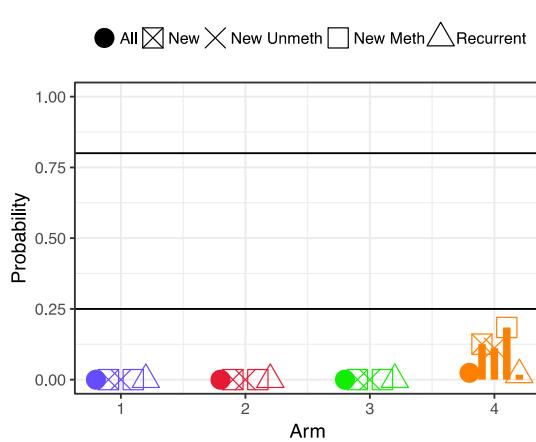
Sample Size



Randomization Probabilities

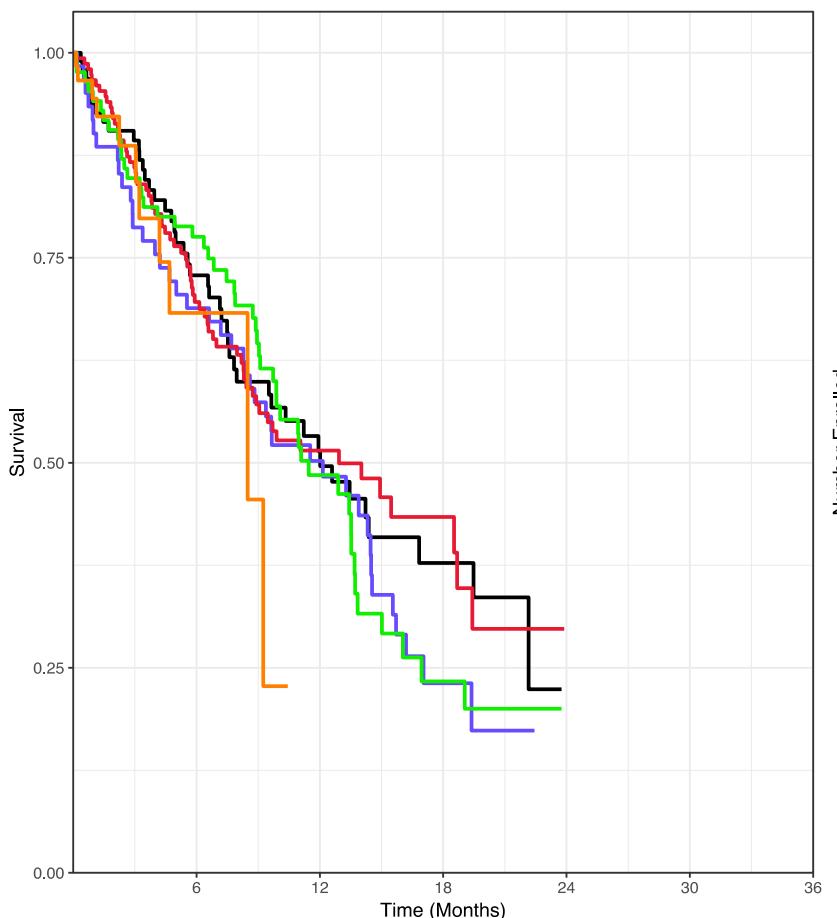


Predictive Probabilities

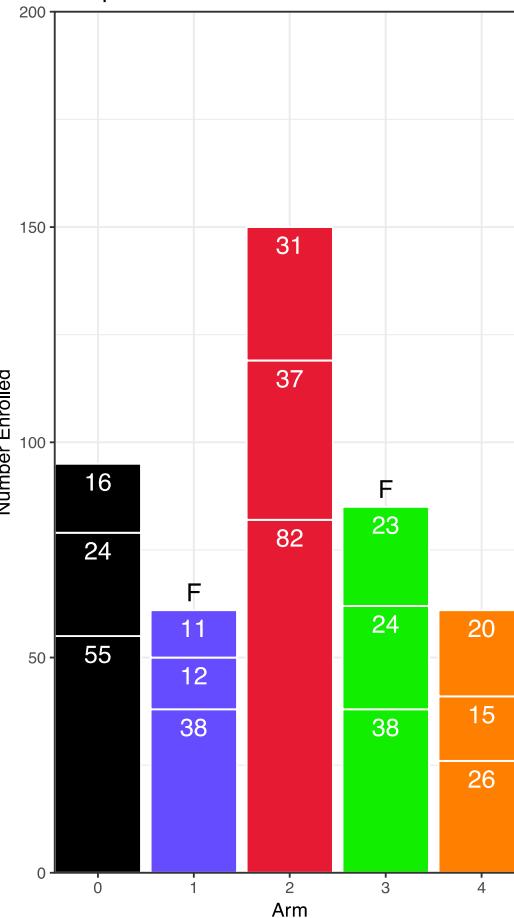


## Month 24: Investigational arms 5 and 6 enter the trial

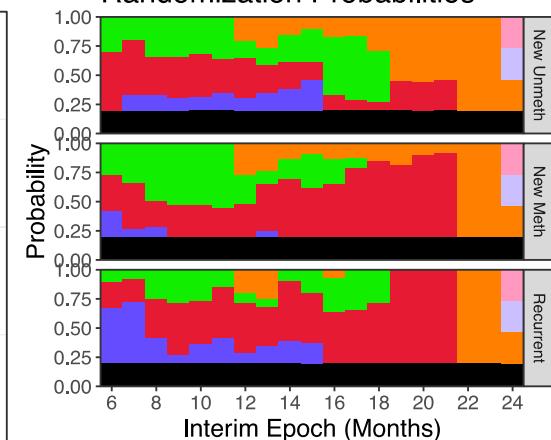
Overall Survival: All Patients. 24 Months After Trial Start



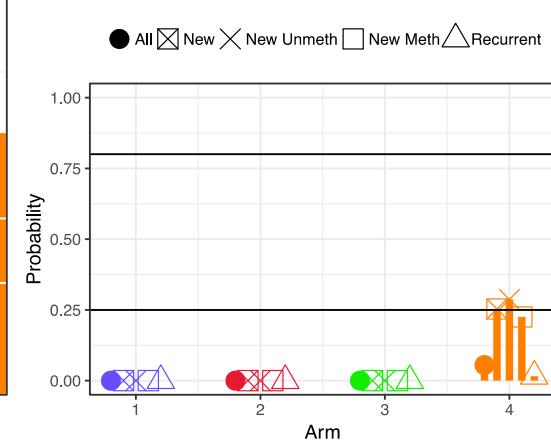
Sample Size



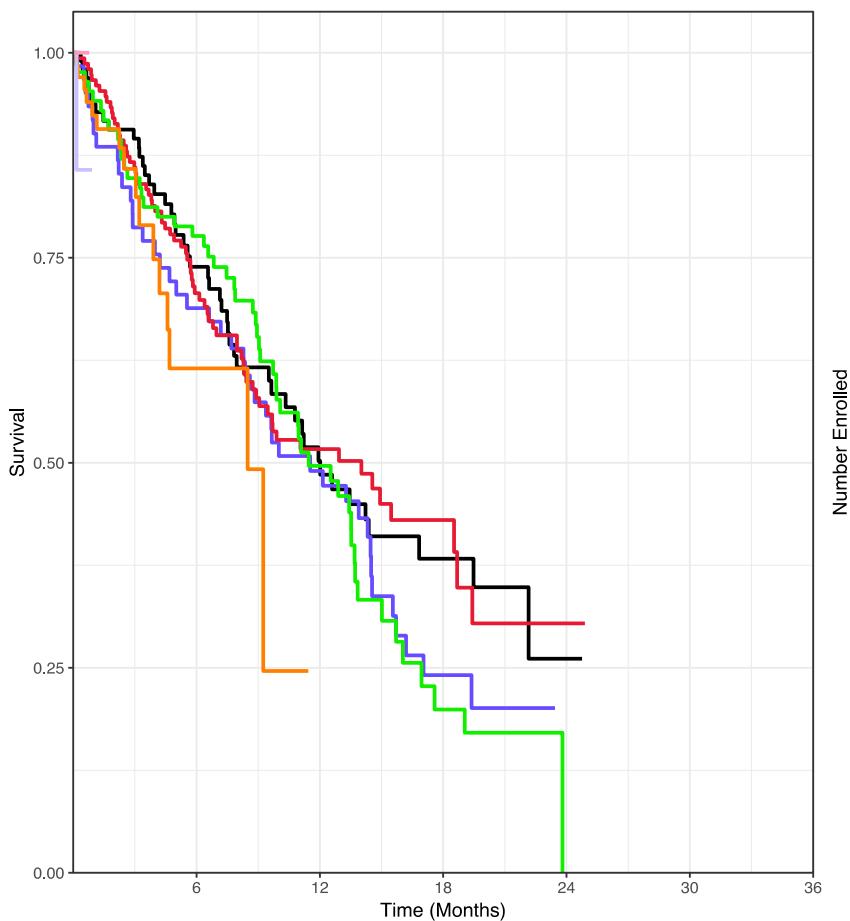
Randomization Probabilities



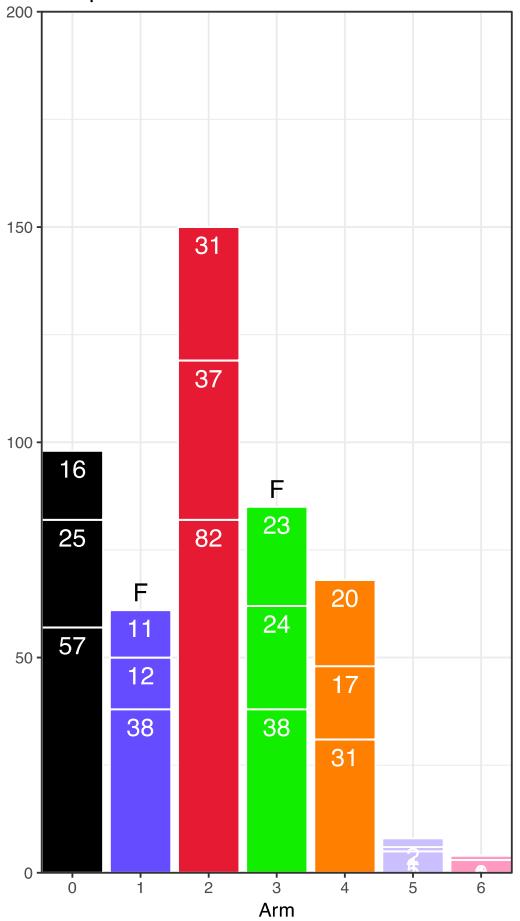
Predictive Probabilities



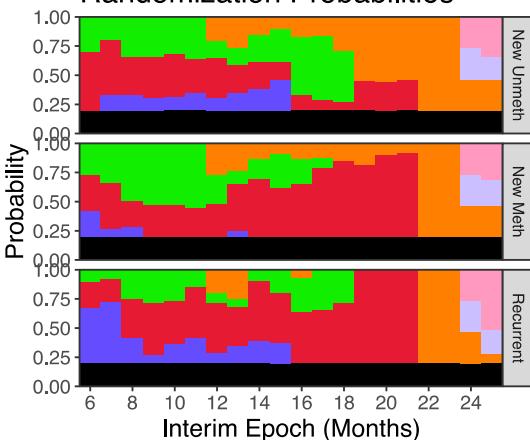
Overall Survival: All Patients. 25 Months After Trial Start



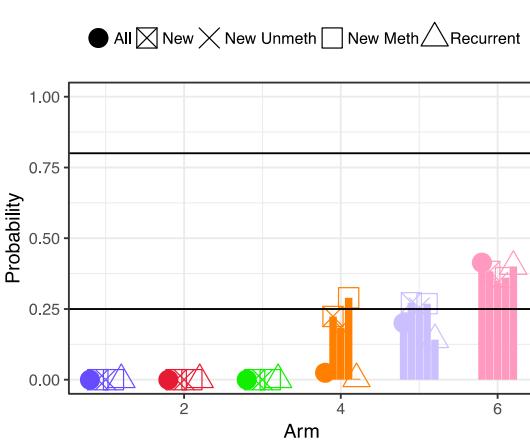
Sample Size



Randomization Probabilities

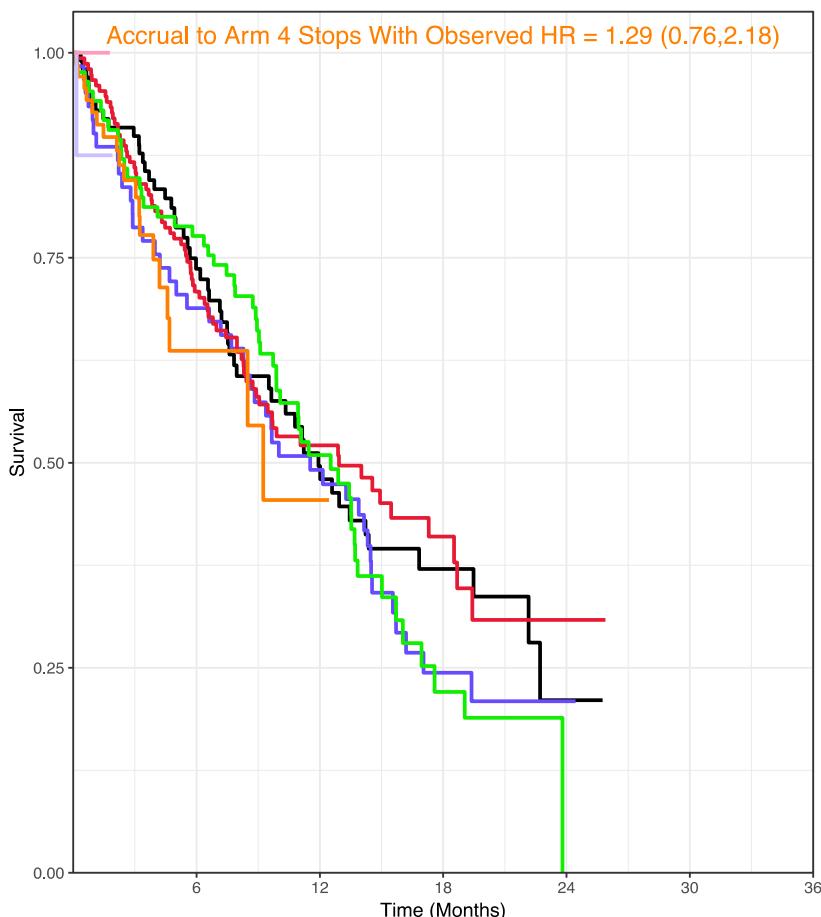


Predictive Probabilities

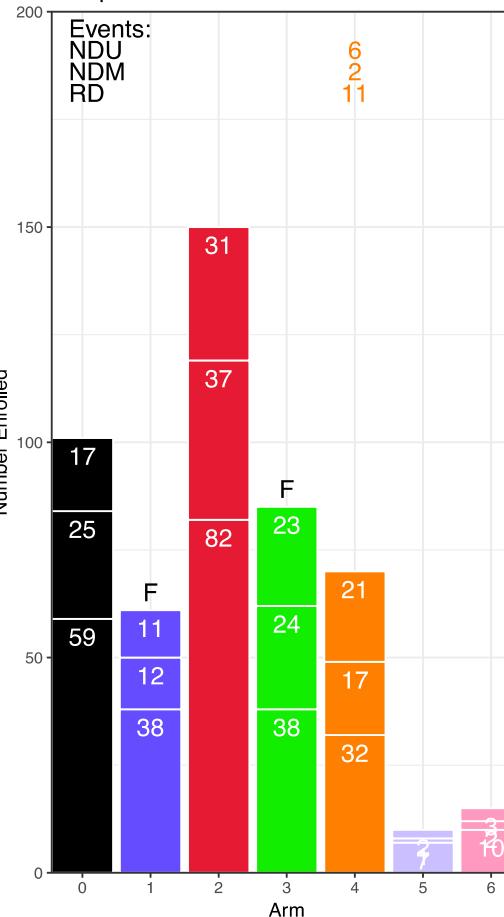


# Accrual to Arm 4 stops for futility with 70 patients, after 14 months in the trial

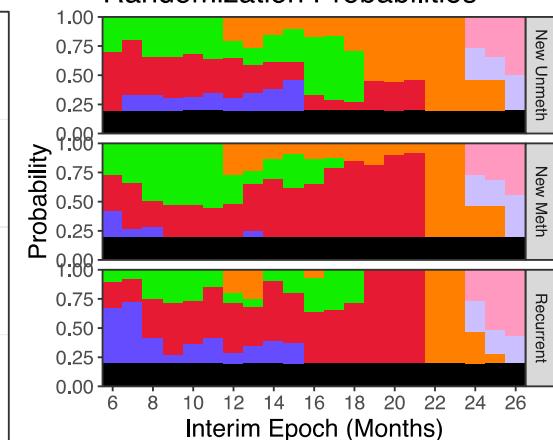
Overall Survival: All Patients. 26 Months After Trial Start



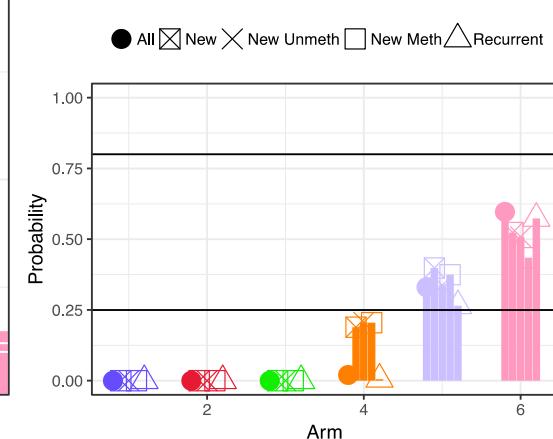
Sample Size



Randomization Probabilities

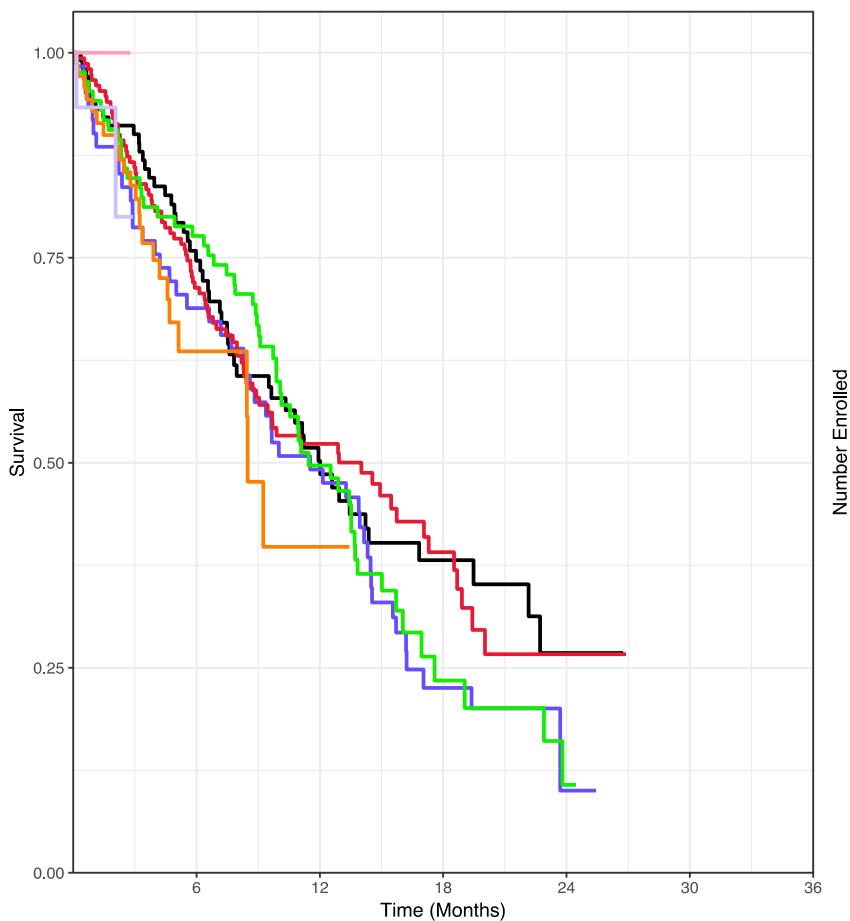


Predictive Probabilities

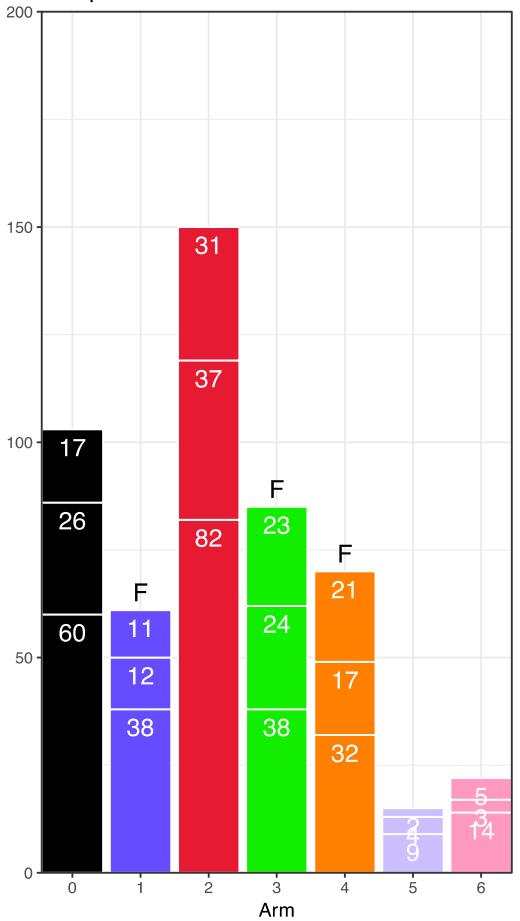


NDU: Newly Diagnosed Unmethylated; NDM: Newly Diagnosed Methylated; RD: Recurrent Disease

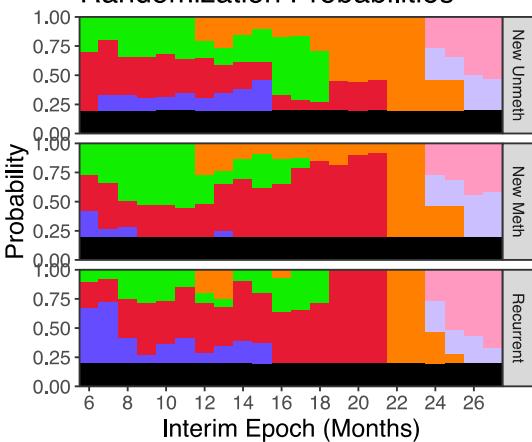
Overall Survival: All Patients. 27 Months After Trial Start



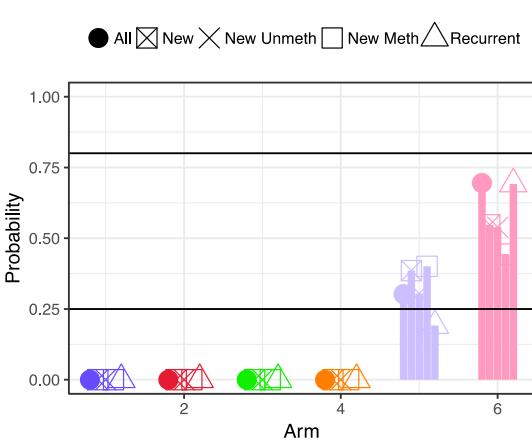
Sample Size



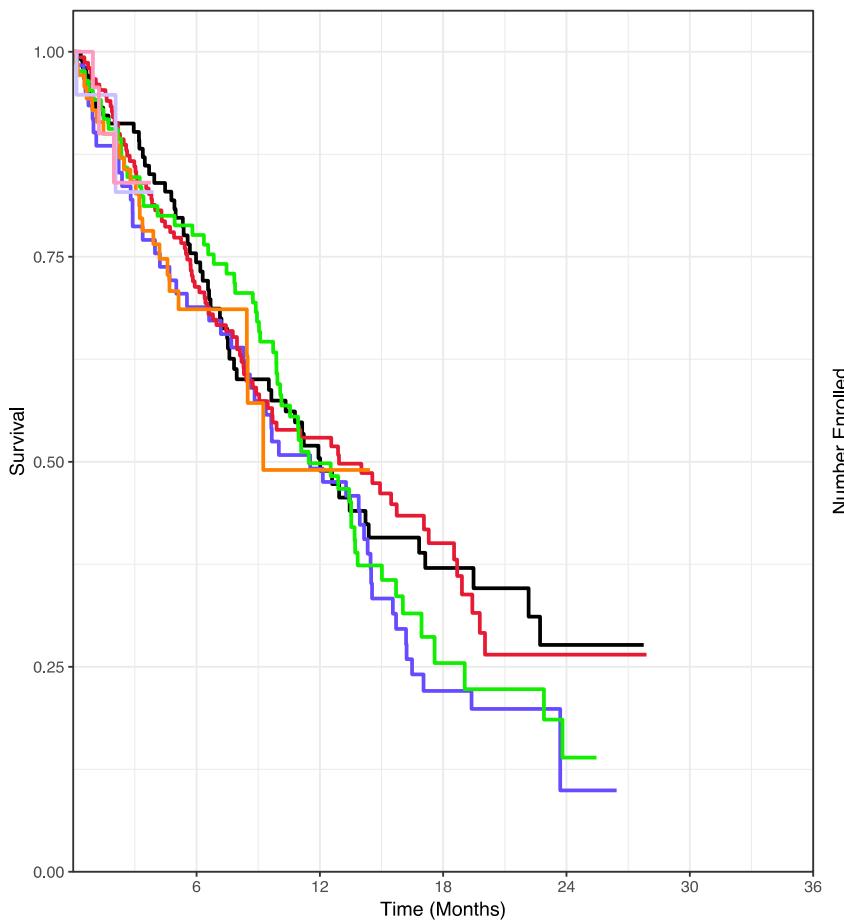
Randomization Probabilities



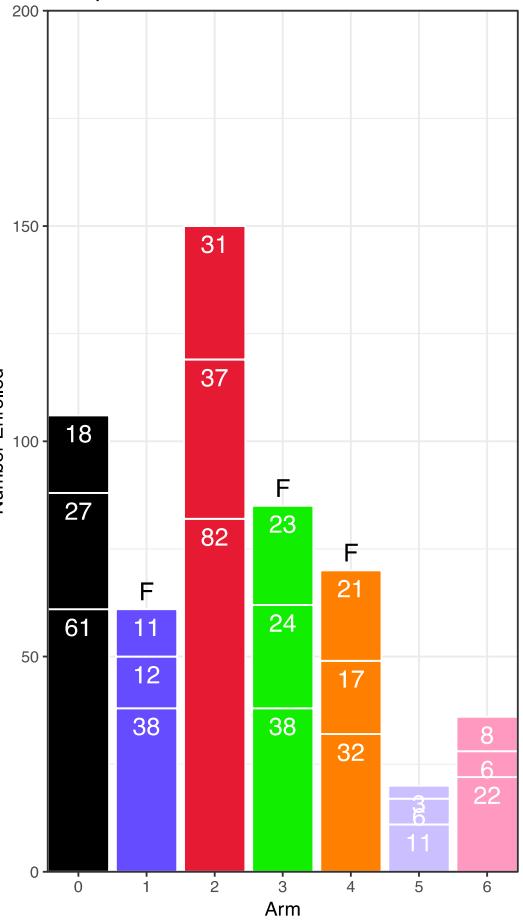
Predictive Probabilities



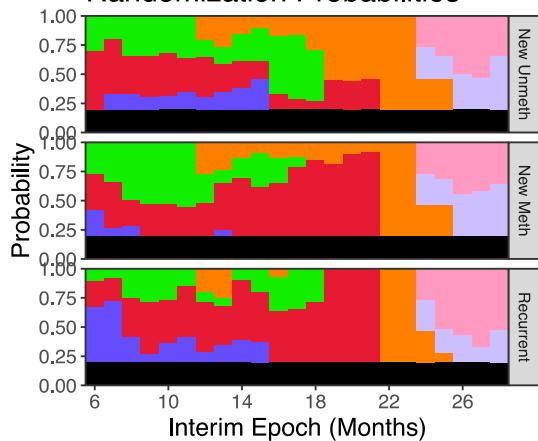
Overall Survival: All Patients. 28 Months After Trial Start



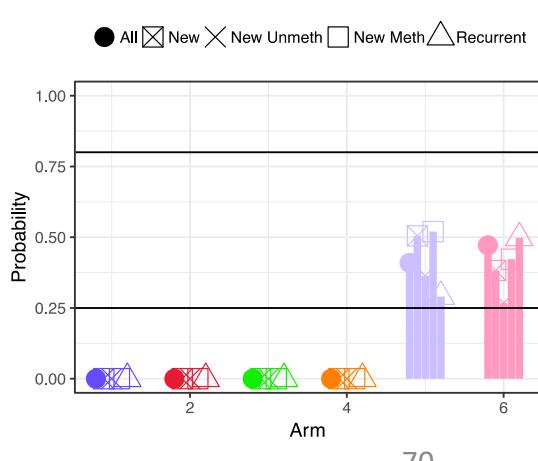
Sample Size



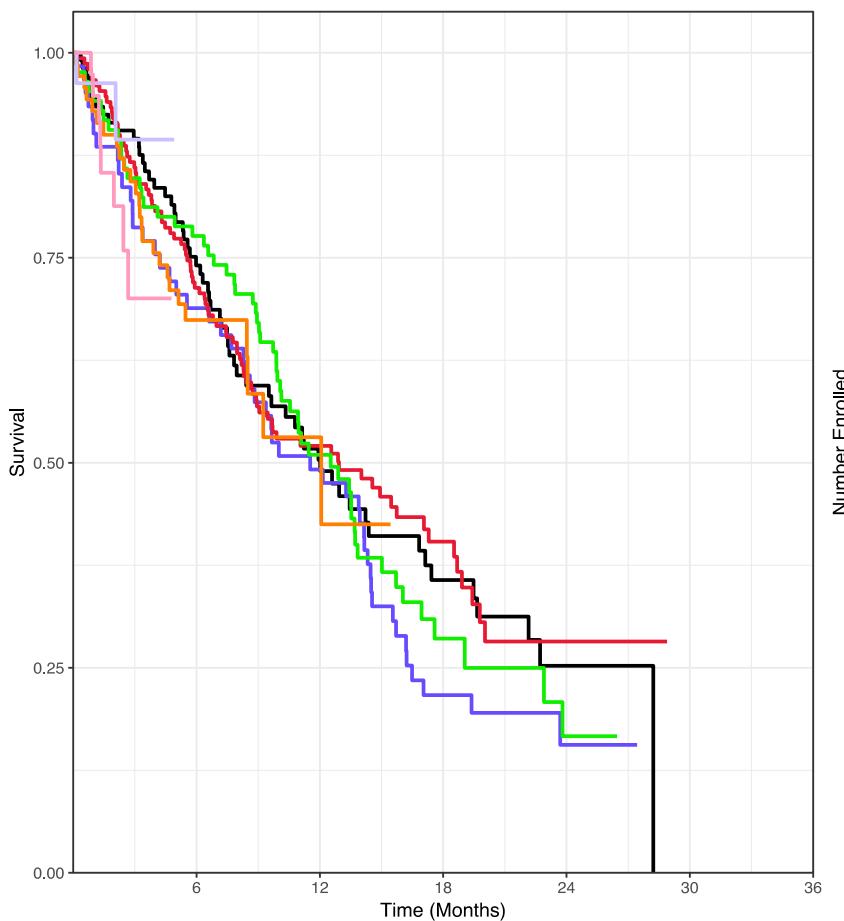
Randomization Probabilities



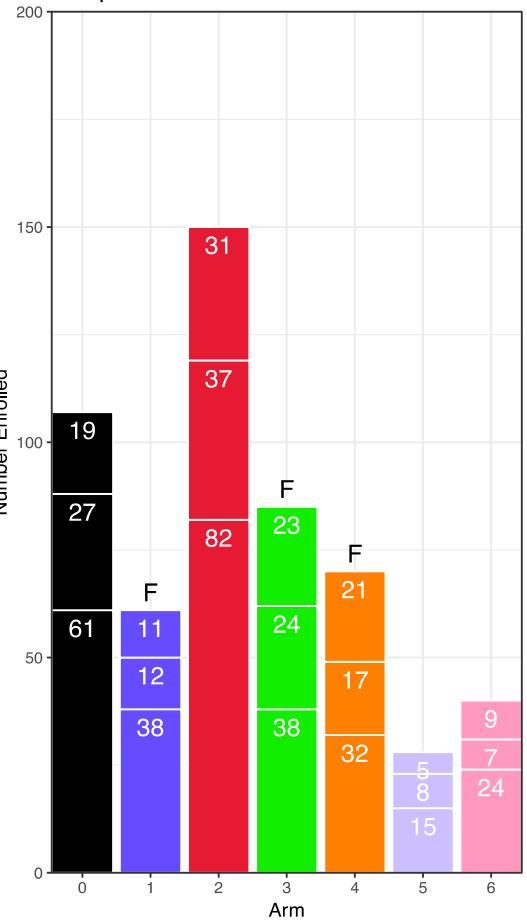
Predictive Probabilities



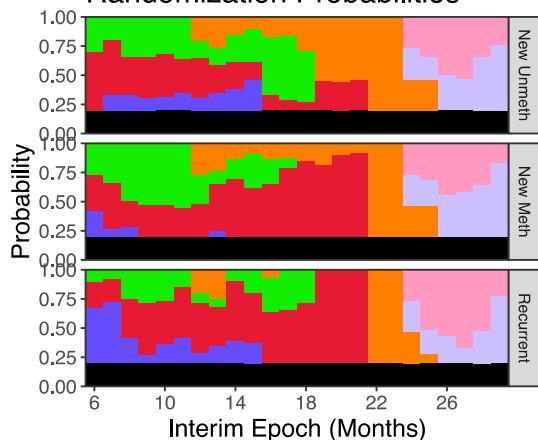
Overall Survival: All Patients. 29 Months After Trial Start



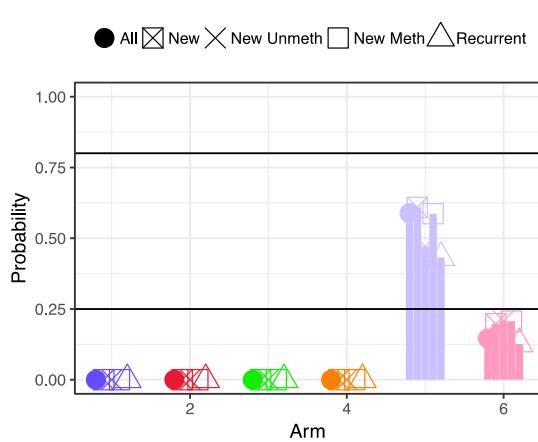
Sample Size



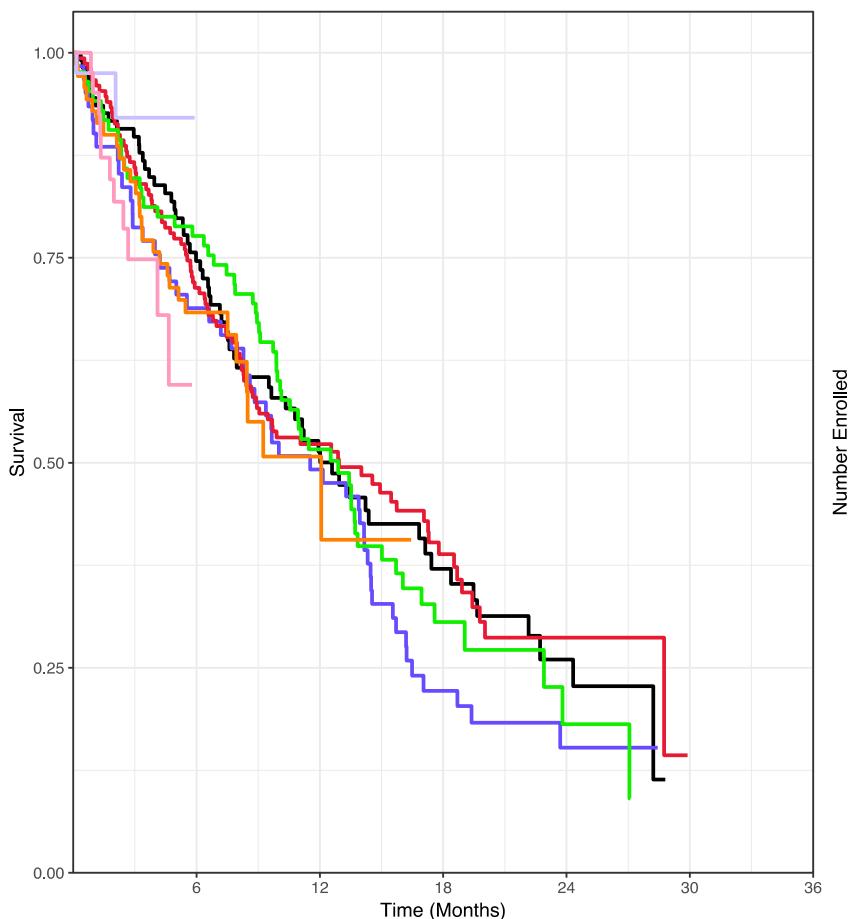
Randomization Probabilities



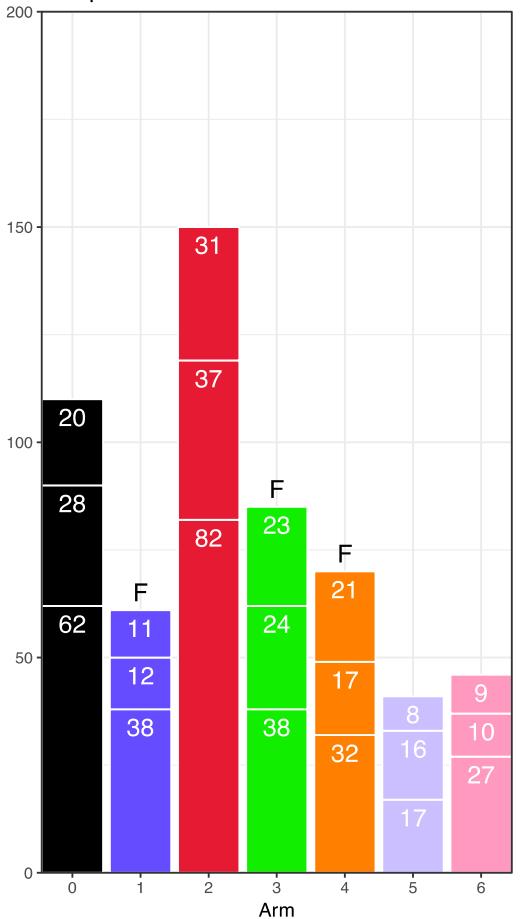
Predictive Probabilities



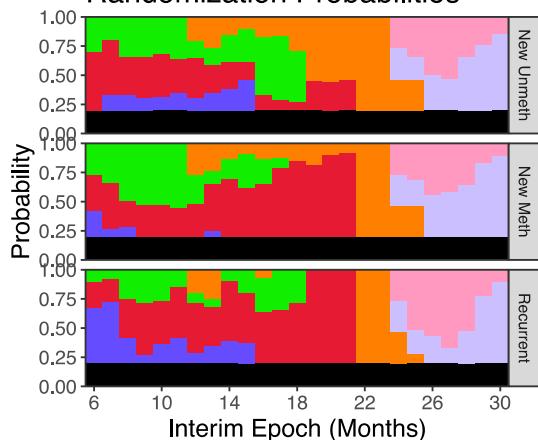
Overall Survival: All Patients. 30 Months After Trial Start



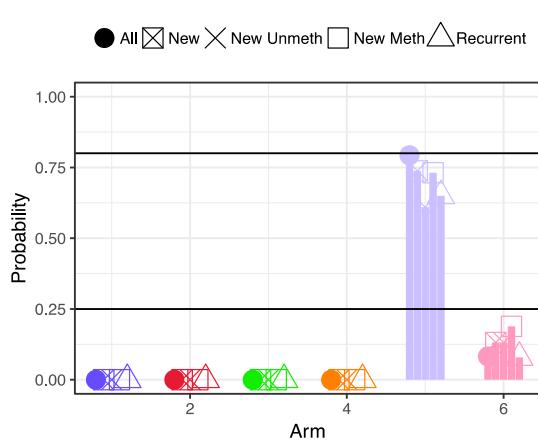
Sample Size



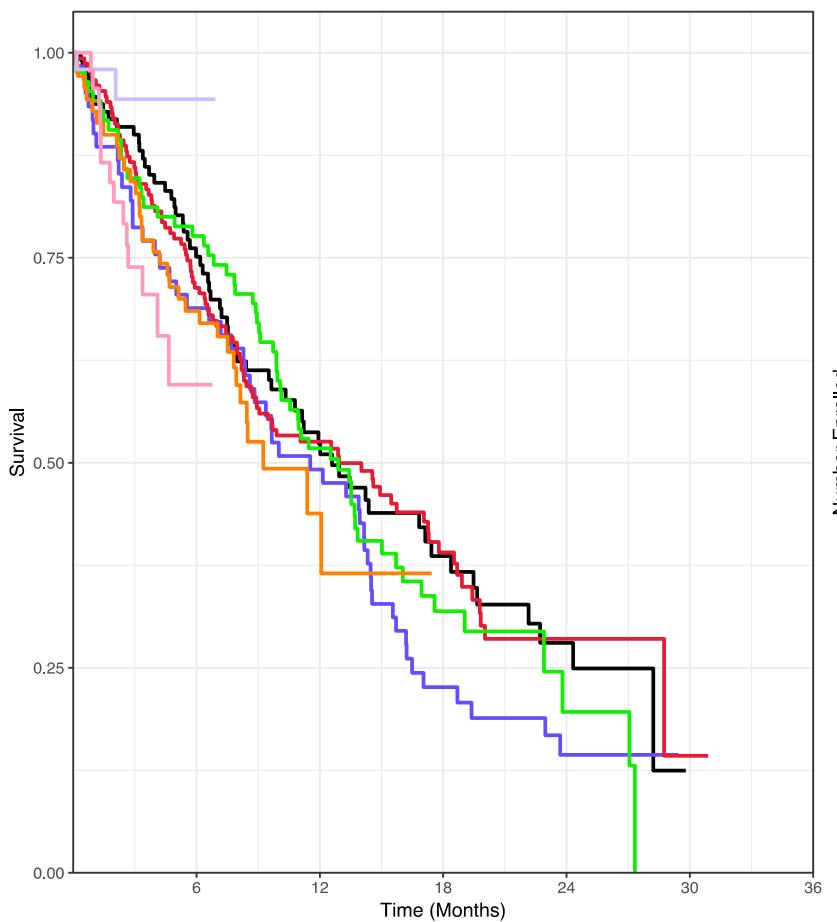
Randomization Probabilities



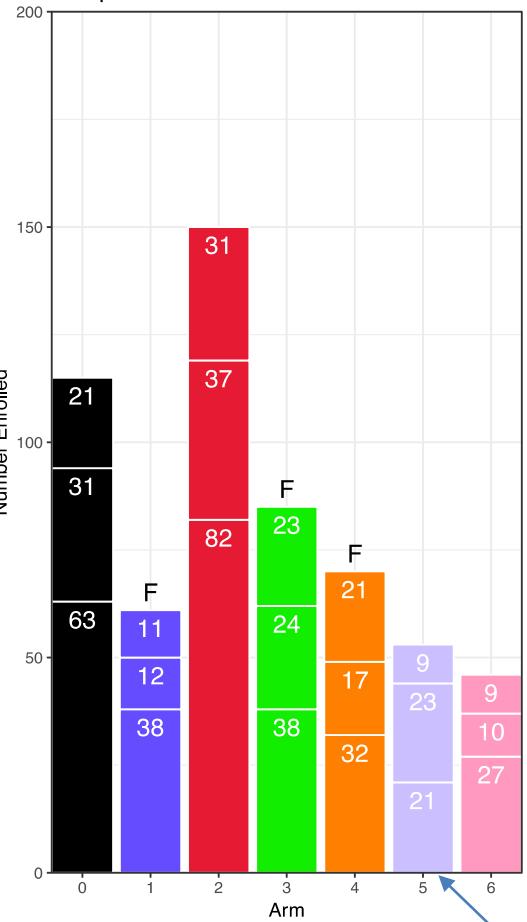
Predictive Probabilities



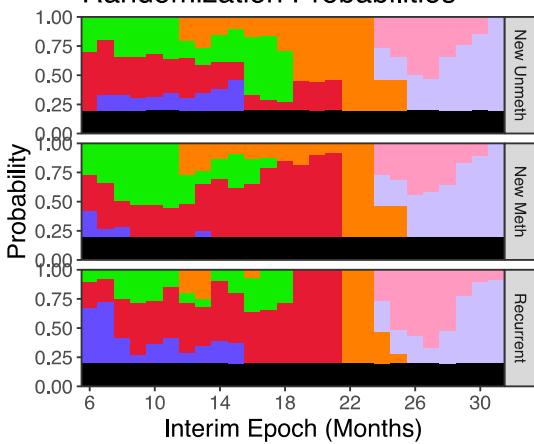
Overall Survival: All Patients. 31 Months After Trial Start



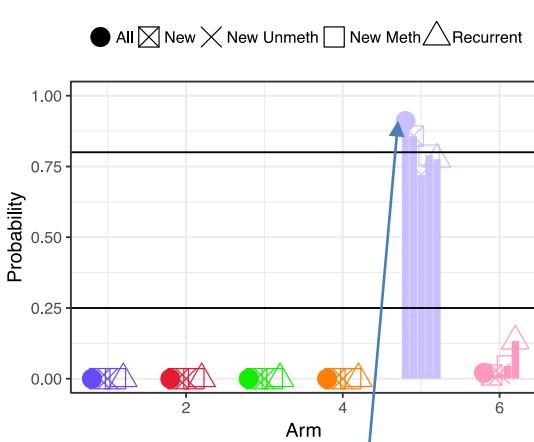
Sample Size



Randomization Probabilities

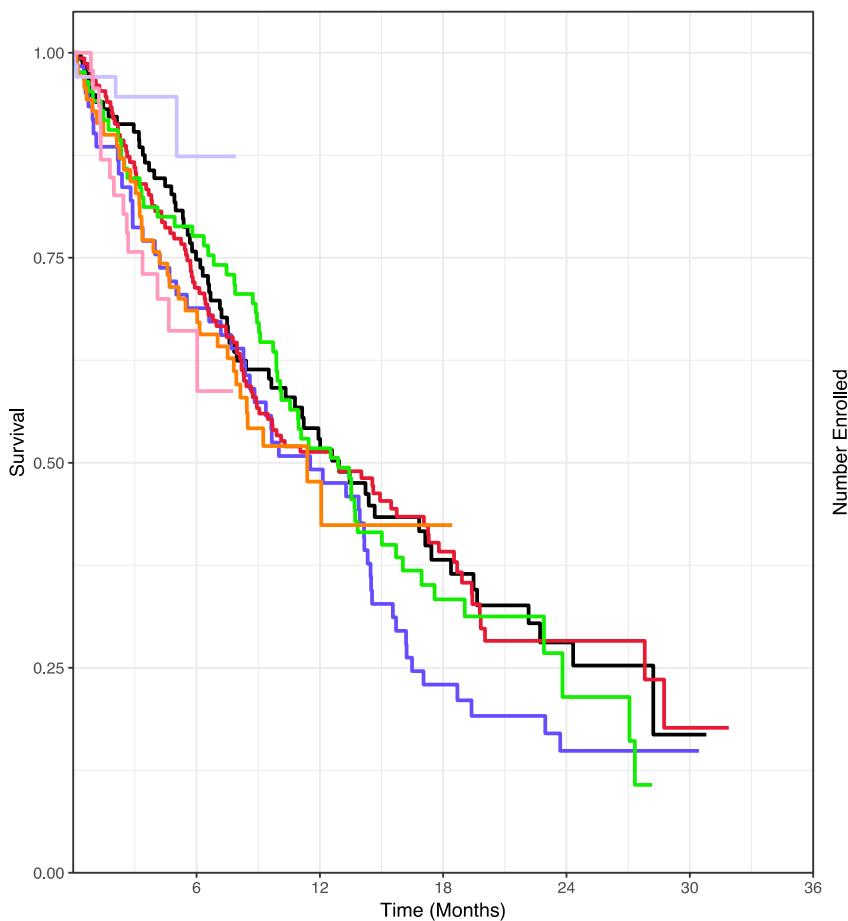


Predictive Probabilities

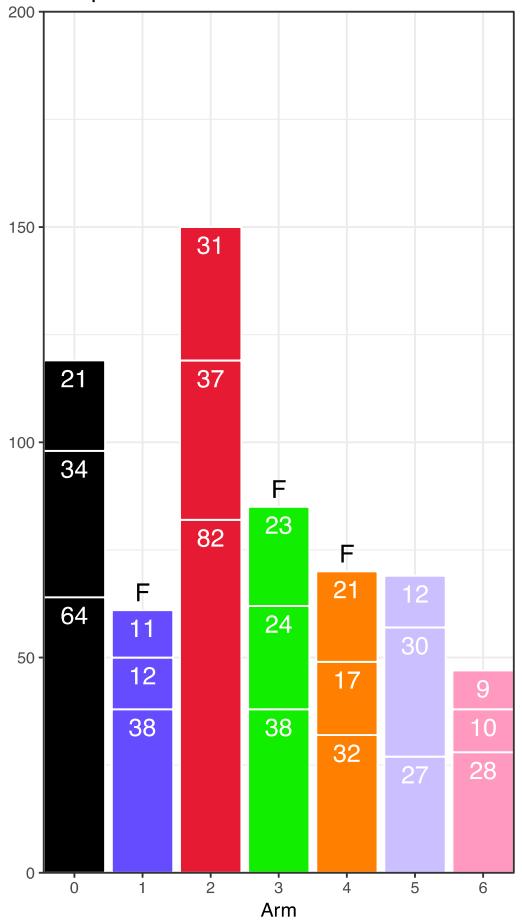


Arm 5 has a predictive probability above the 0.80 threshold, but cannot graduate with fewer than 100 patients

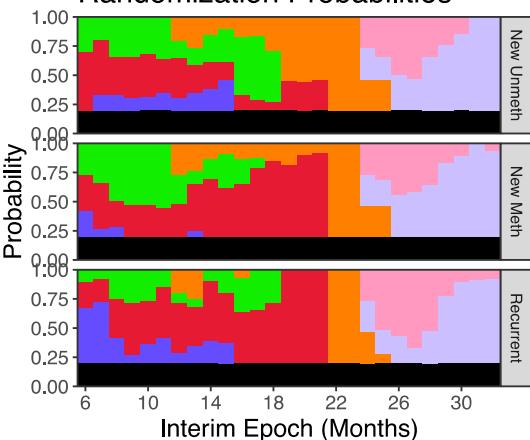
Overall Survival: All Patients. 32 Months After Trial Start



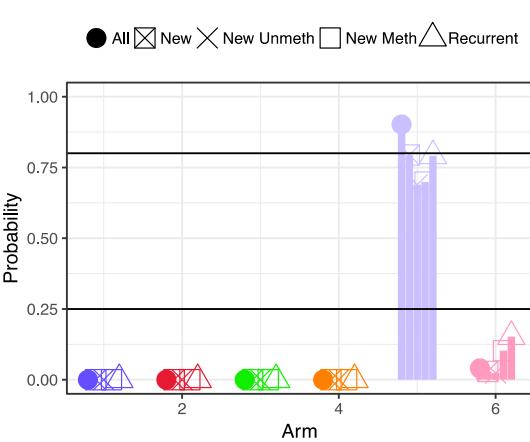
Sample Size



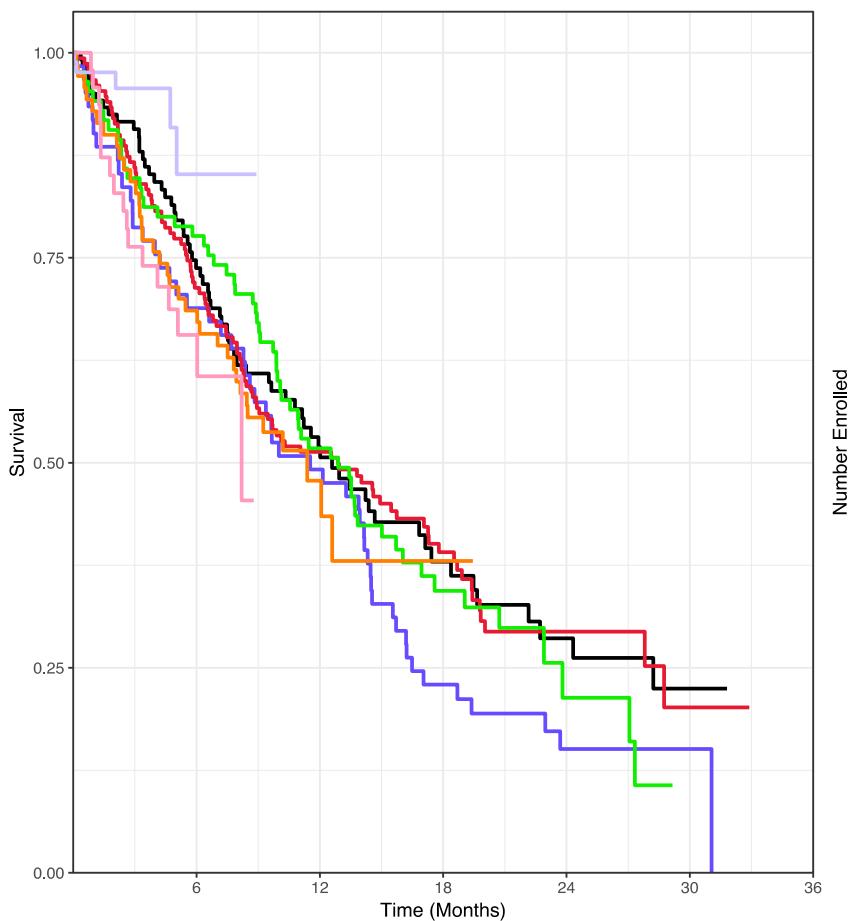
Randomization Probabilities



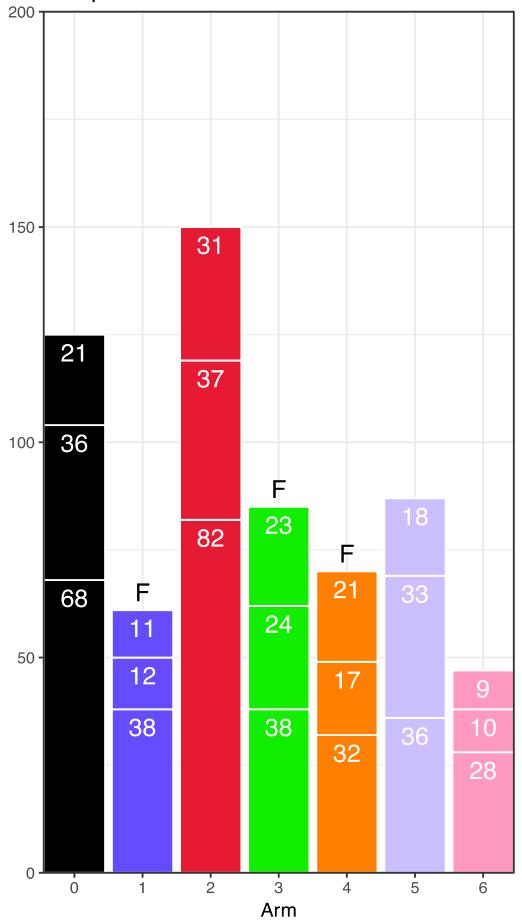
Predictive Probabilities



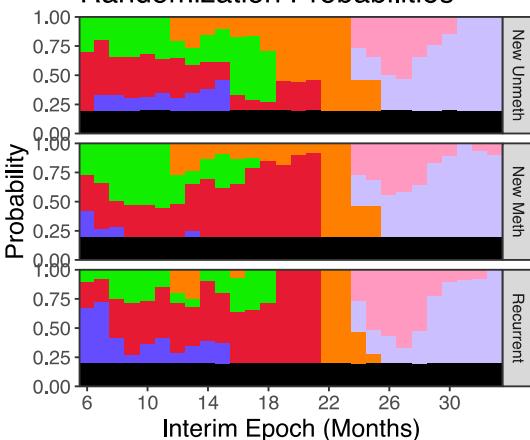
Overall Survival: All Patients. 33 Months After Trial Start



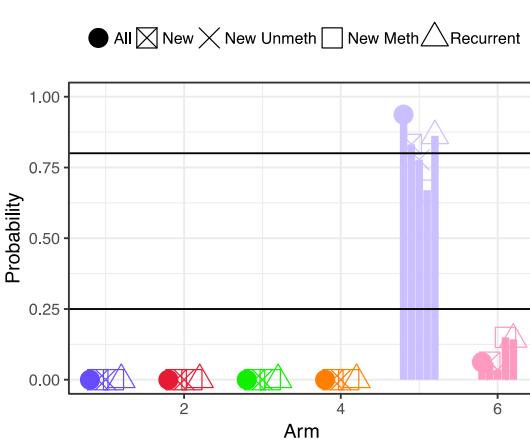
Sample Size



Randomization Probabilities

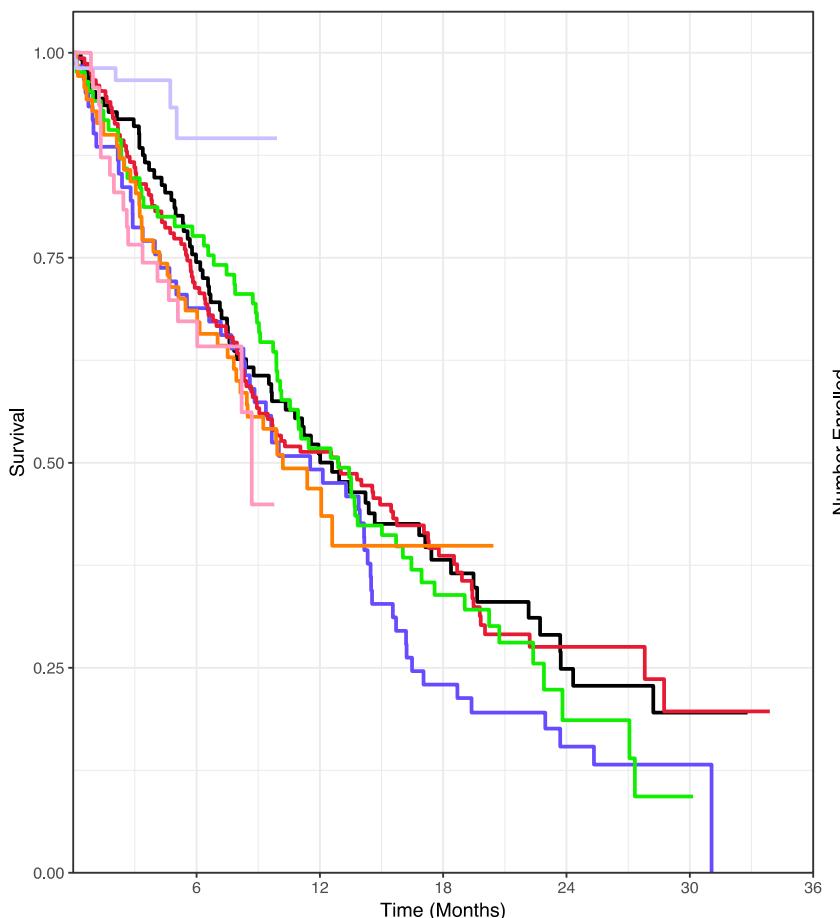


Predictive Probabilities

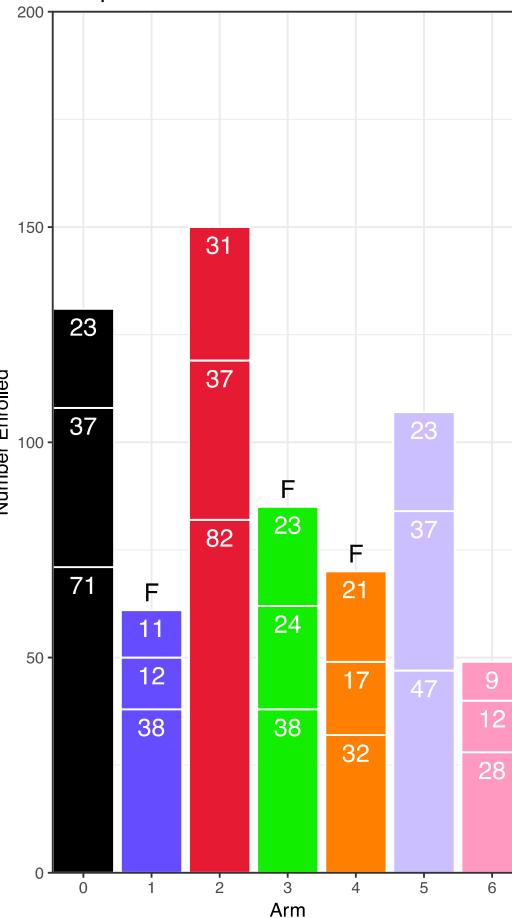


# After 10 months in the trial, Arm 5 graduates in the All Patients signature with 107 patients with HR 0.26 (0.10,

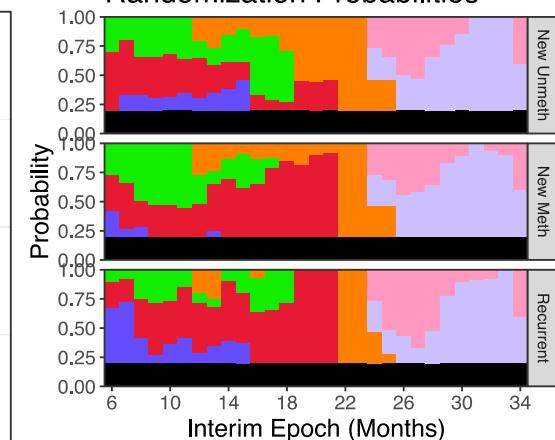
Overall Survival: All Patients. 34 Months After Trial Start



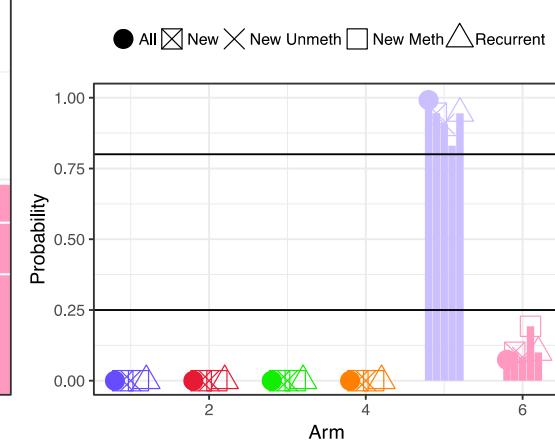
Sample Size



Randomization Probabilities

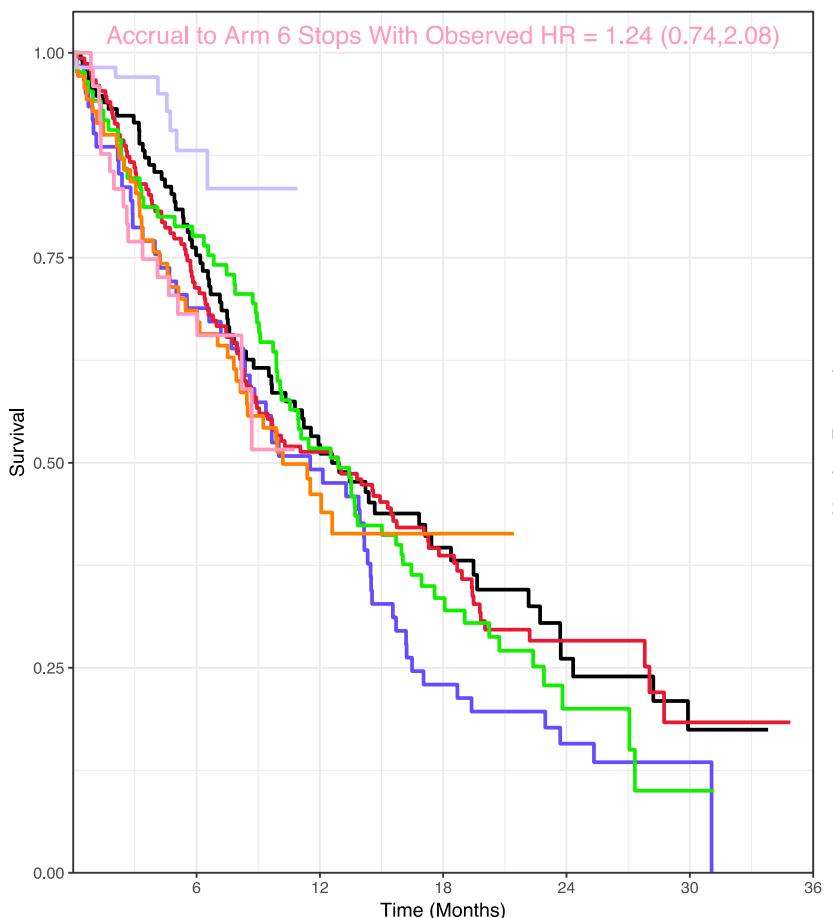


Predictive Probabilities

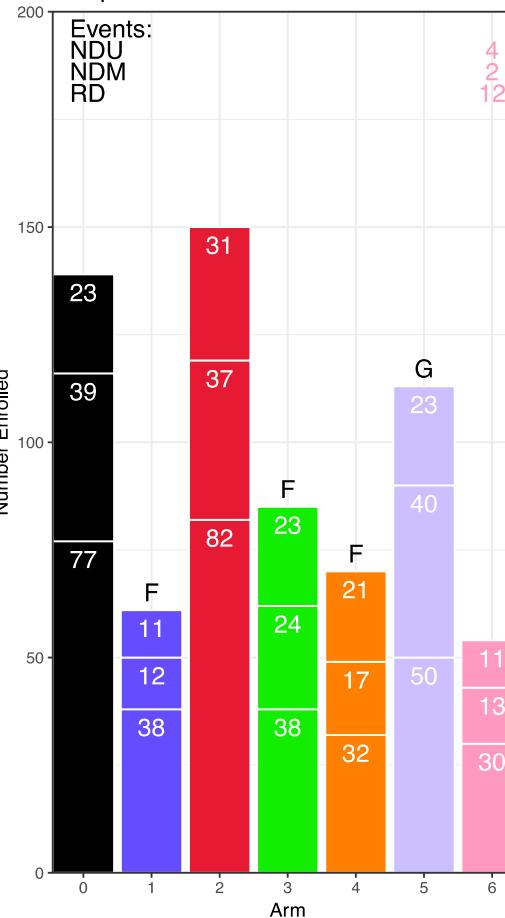


# Accrual to Arm 6 stops for futility after 54 patients and 11 months

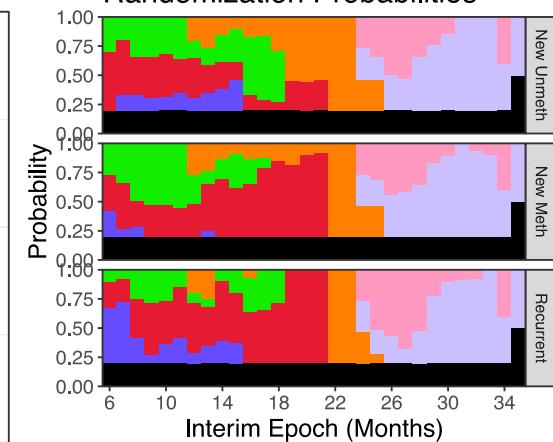
Overall Survival: All Patients. 35 Months After Trial Start



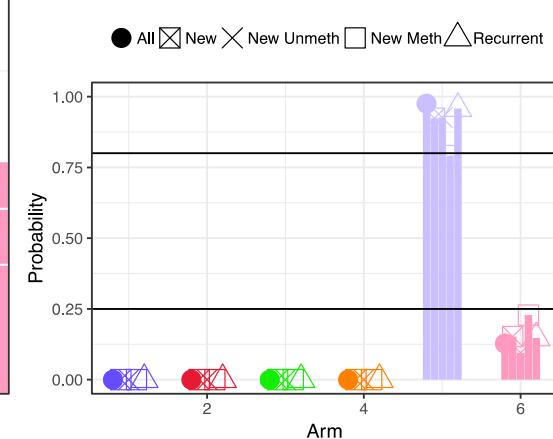
Sample Size



Randomization Probabilities

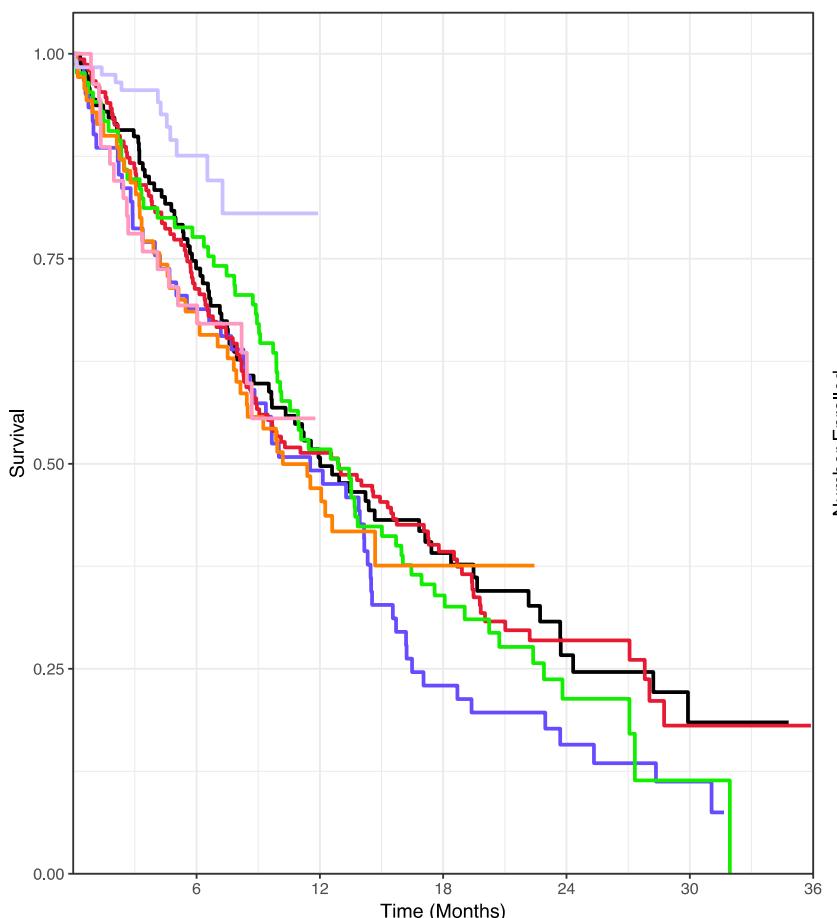


Predictive Probabilities

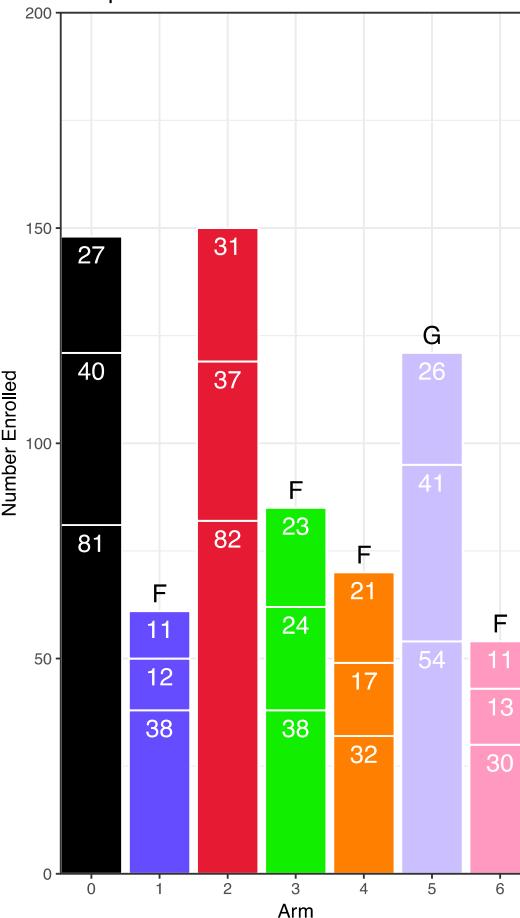


## Month 36: Investigational arms 7 and 8 enter the trial

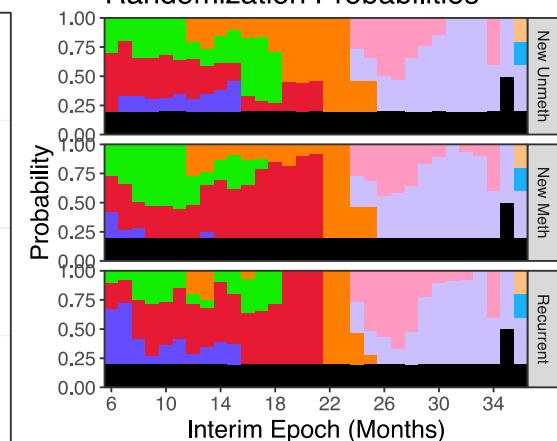
Overall Survival: All Patients. 36 Months After Trial Start



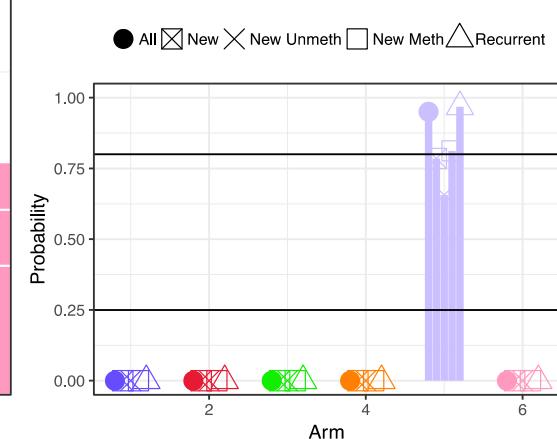
Sample Size



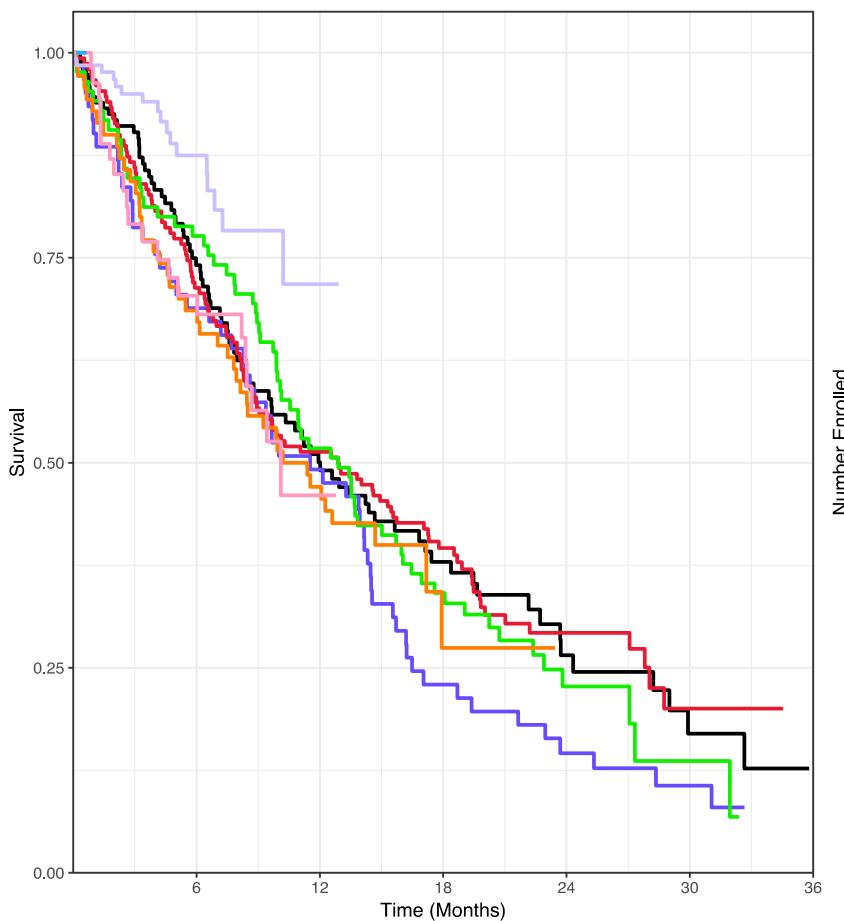
Randomization Probabilities



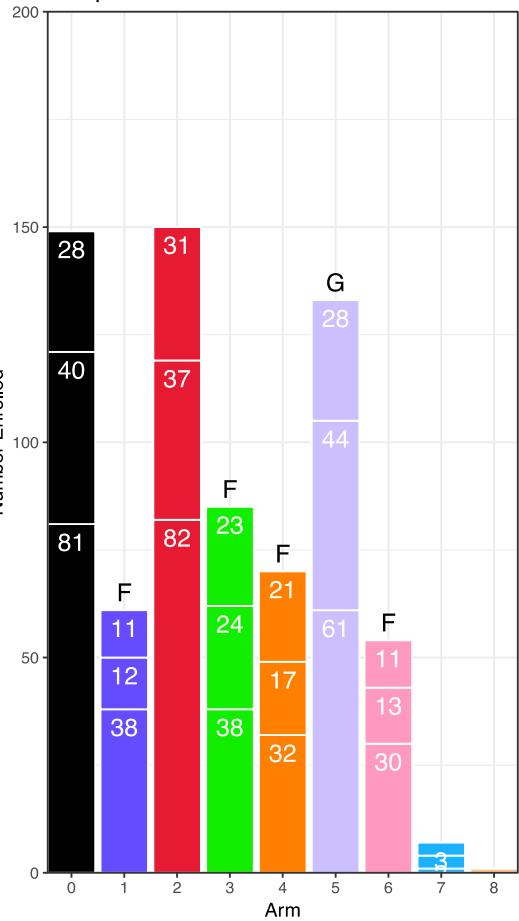
Predictive Probabilities



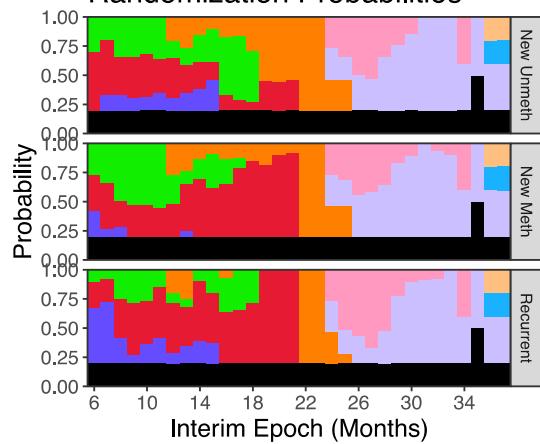
Overall Survival: All Patients. 37 Months After Trial Start



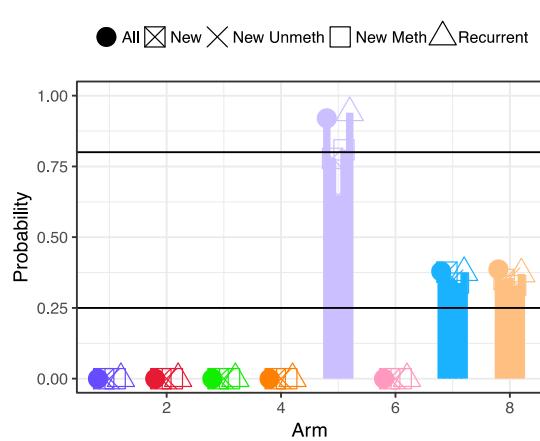
Sample Size



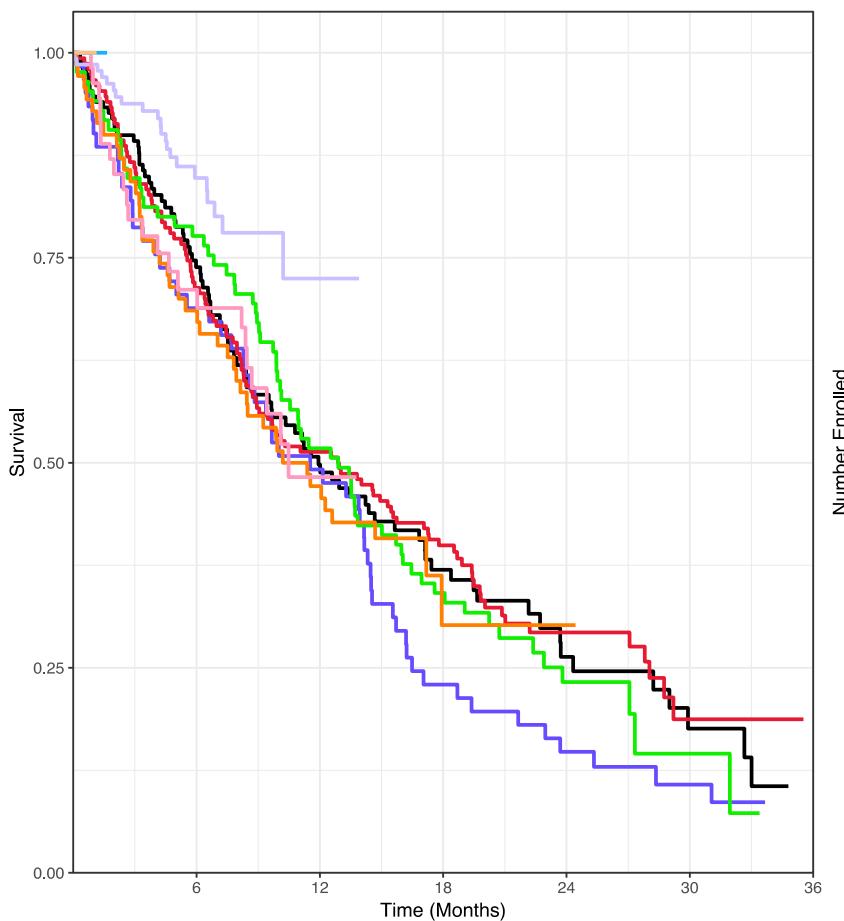
Randomization Probabilities



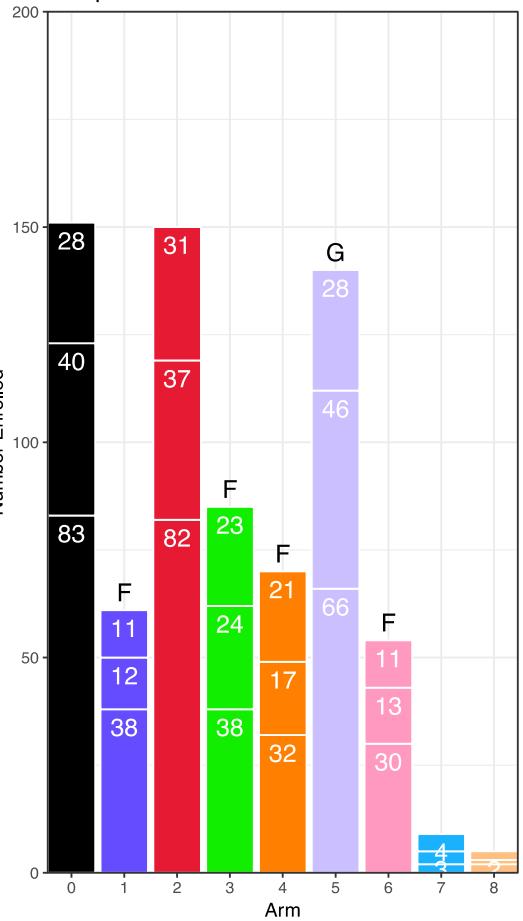
Predictive Probabilities



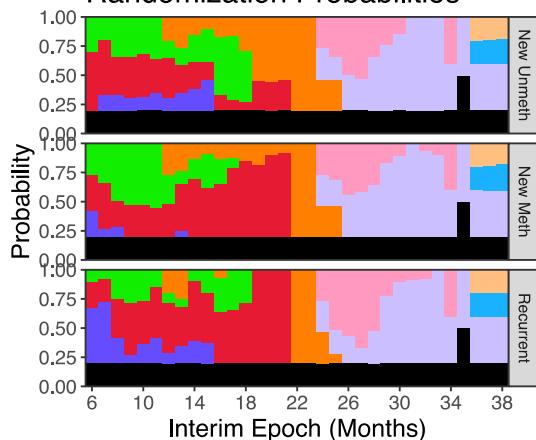
Overall Survival: All Patients. 38 Months After Trial Start



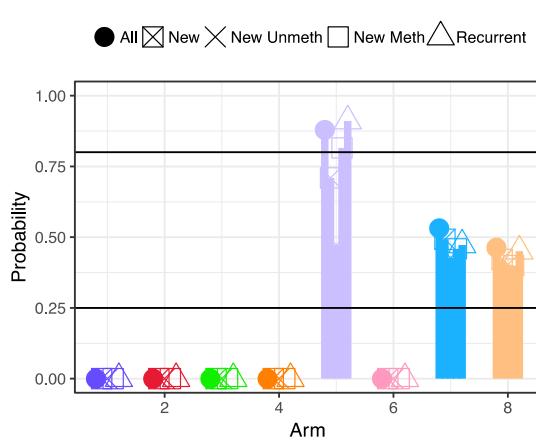
Sample Size



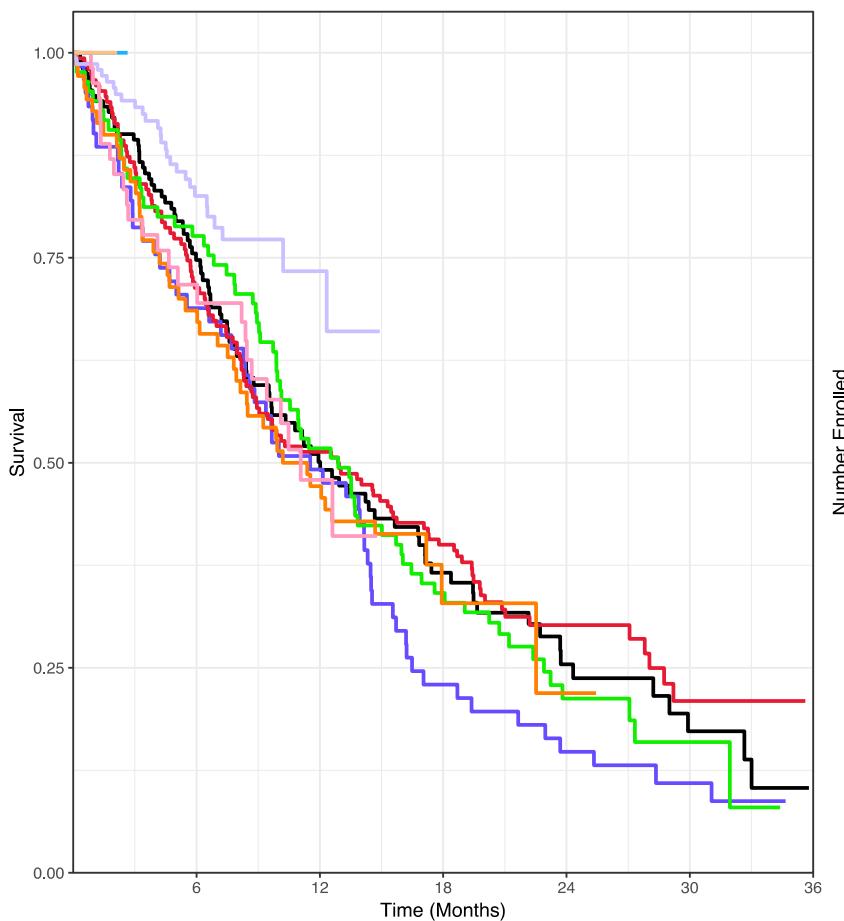
Randomization Probabilities



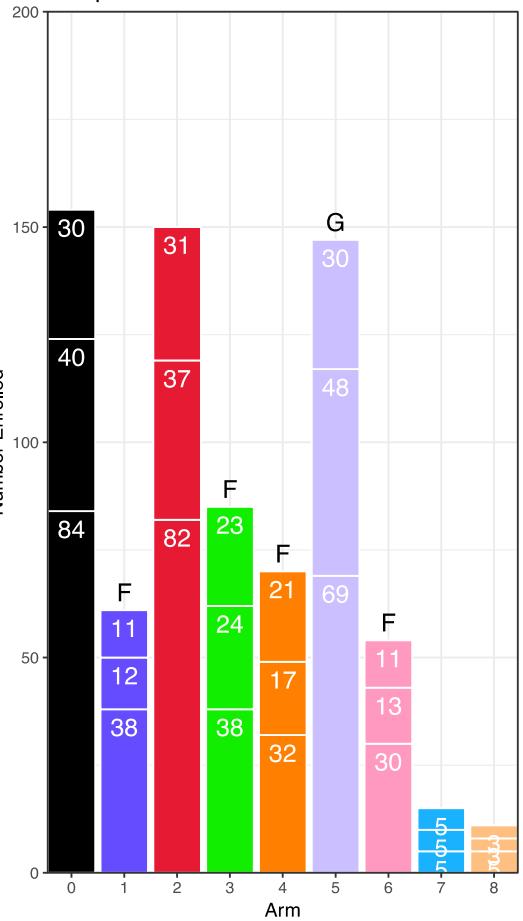
Predictive Probabilities



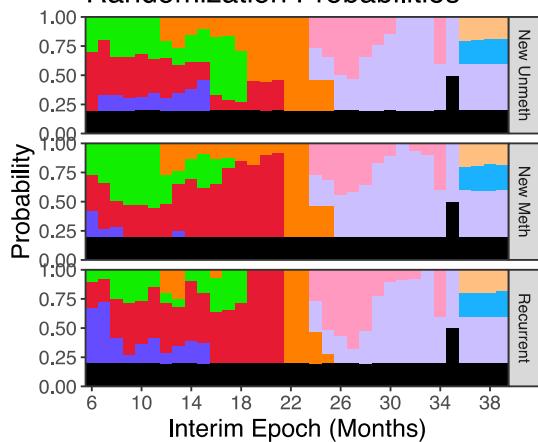
Overall Survival: All Patients. 39 Months After Trial Start



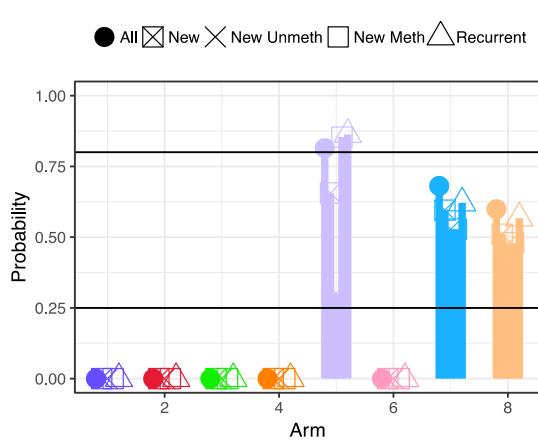
Sample Size



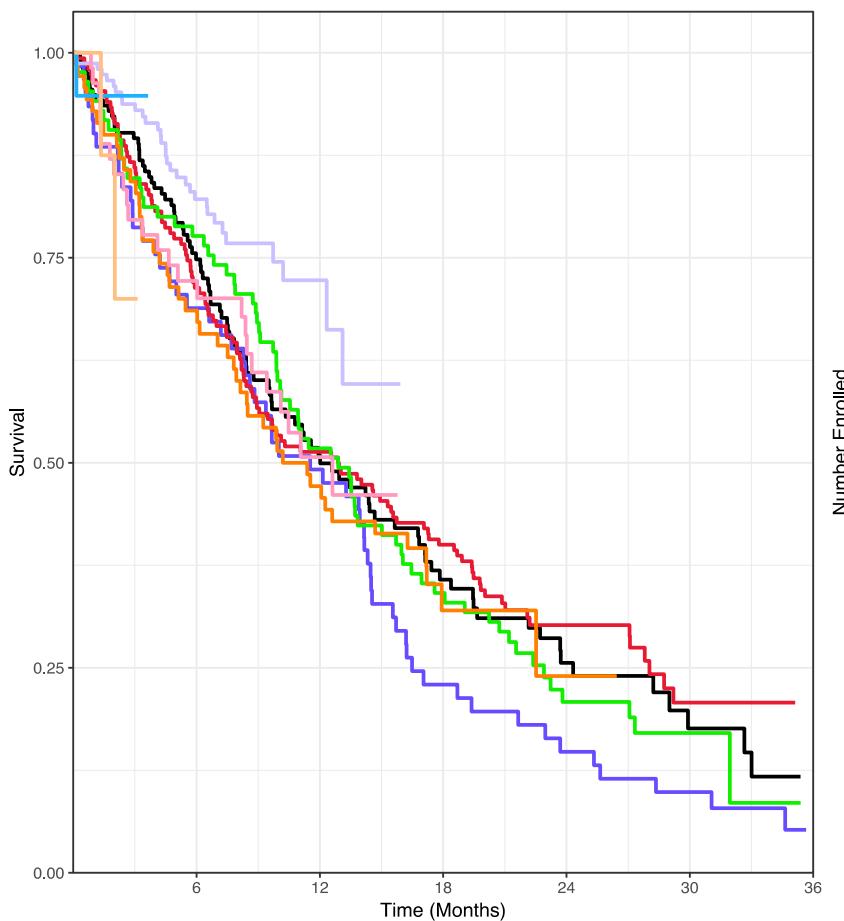
Randomization Probabilities



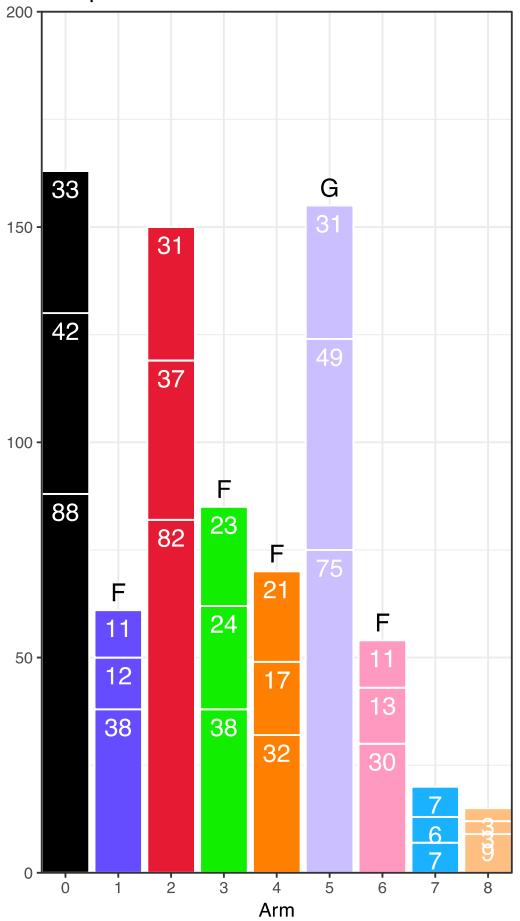
Predictive Probabilities



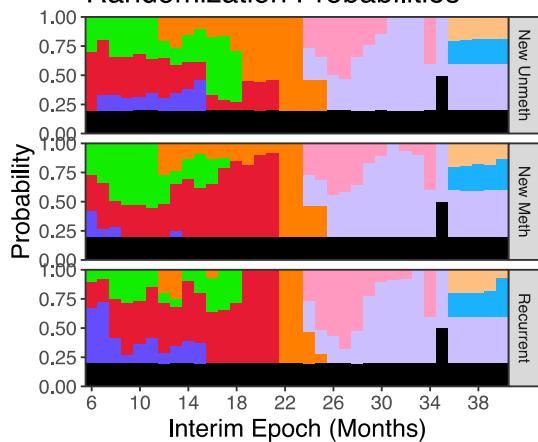
Overall Survival: All Patients. 40 Months After Trial Start



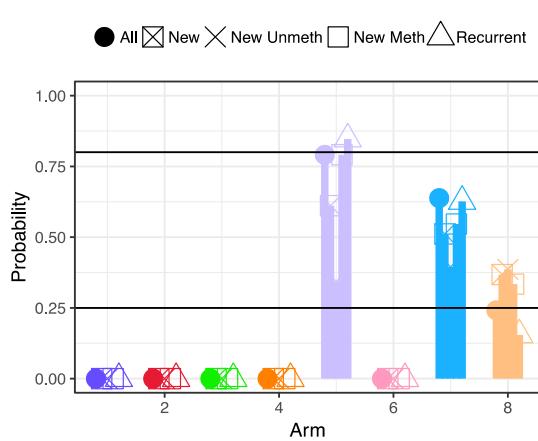
Sample Size



Randomization Probabilities

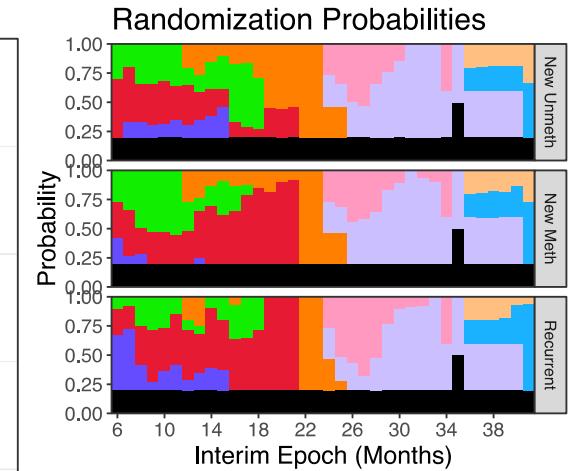
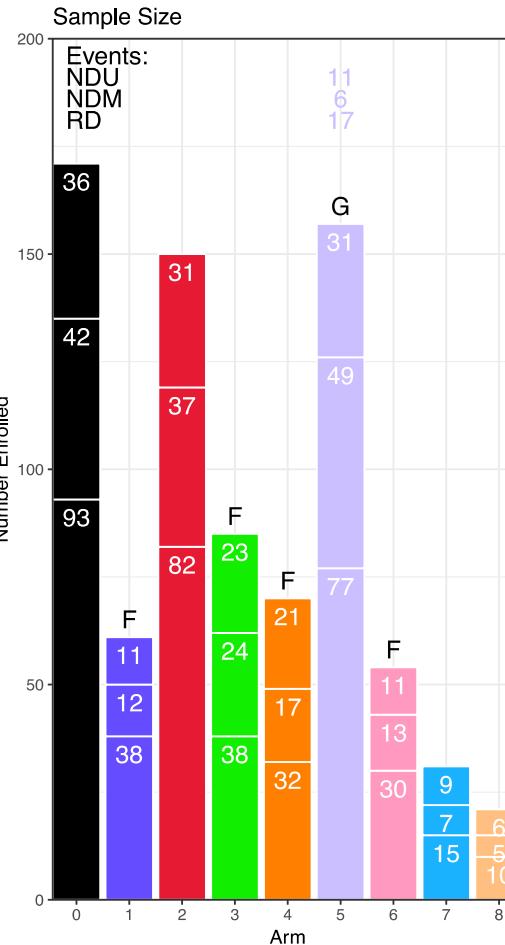
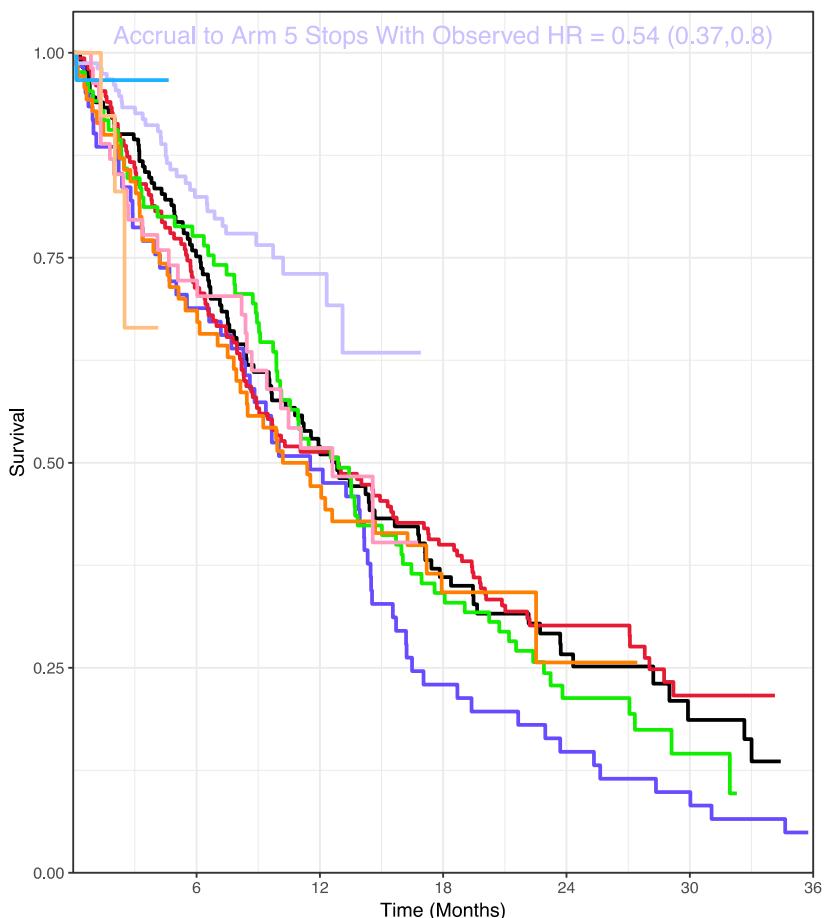


Predictive Probabilities

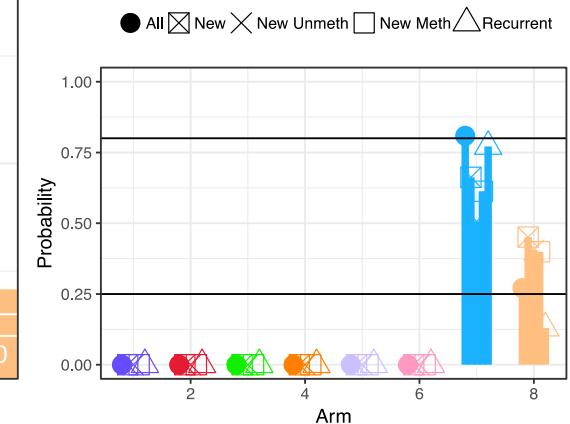


# Arm 5 completes Stage 2 of its accrual. It has enrolled 157 patients over 17 months

Overall Survival: All Patients. 41 Months After Trial Start

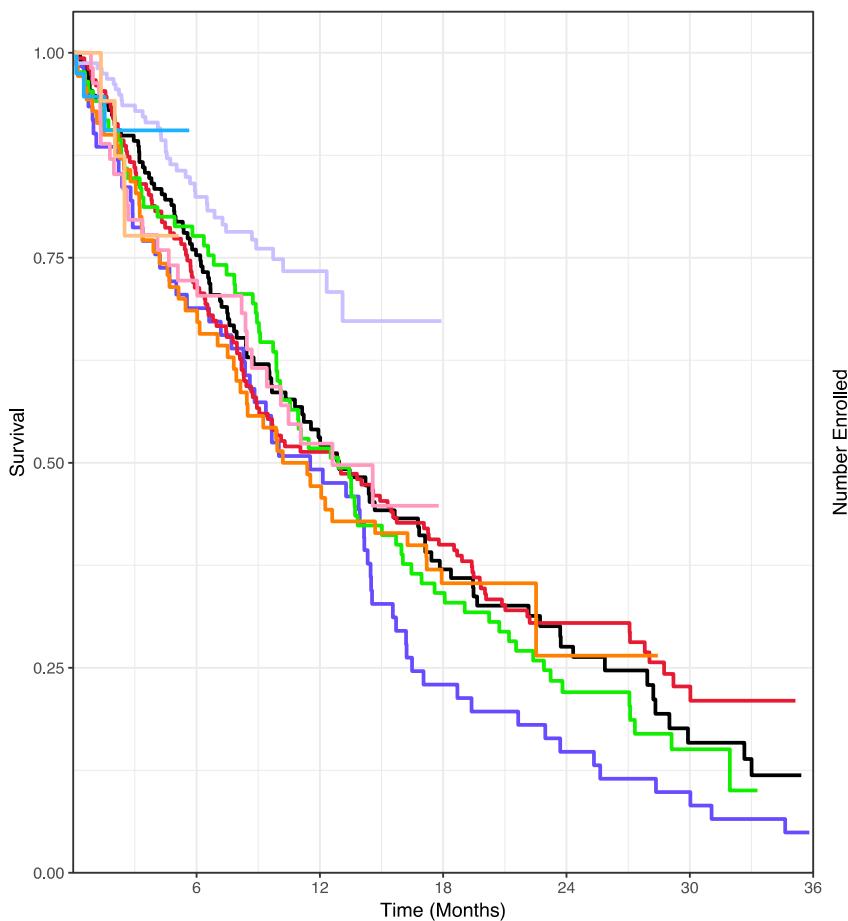


## Predictive Probabilities

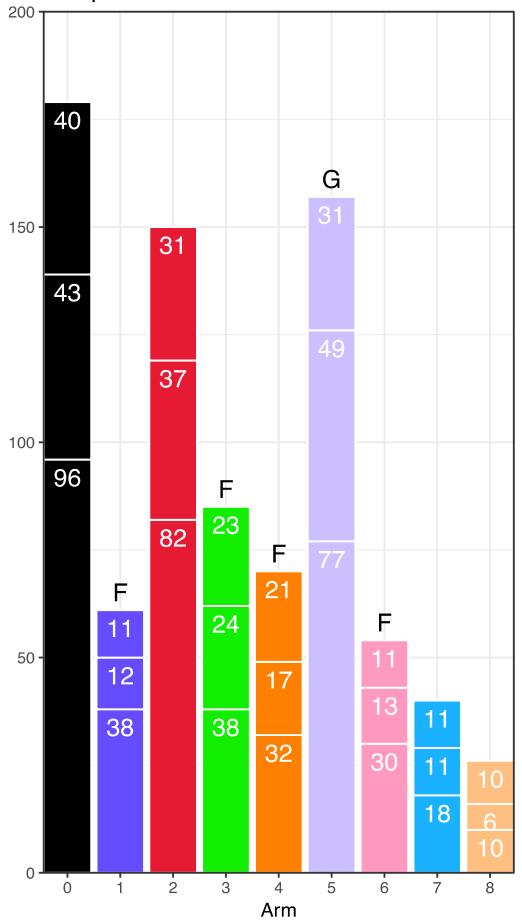


NDU: Newly Diagnosed Unmethylated; NDM: Newly Diagnosed Methylated; RD: Recurrent Disease

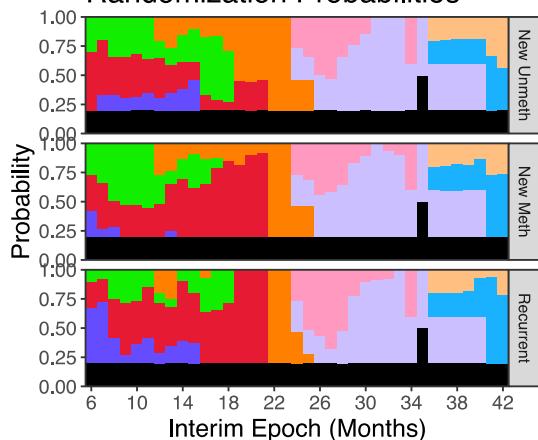
### Overall Survival: All Patients. 42 Months After Trial Start



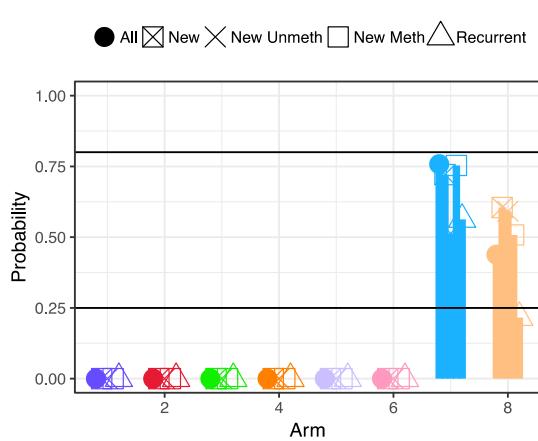
### Sample Size



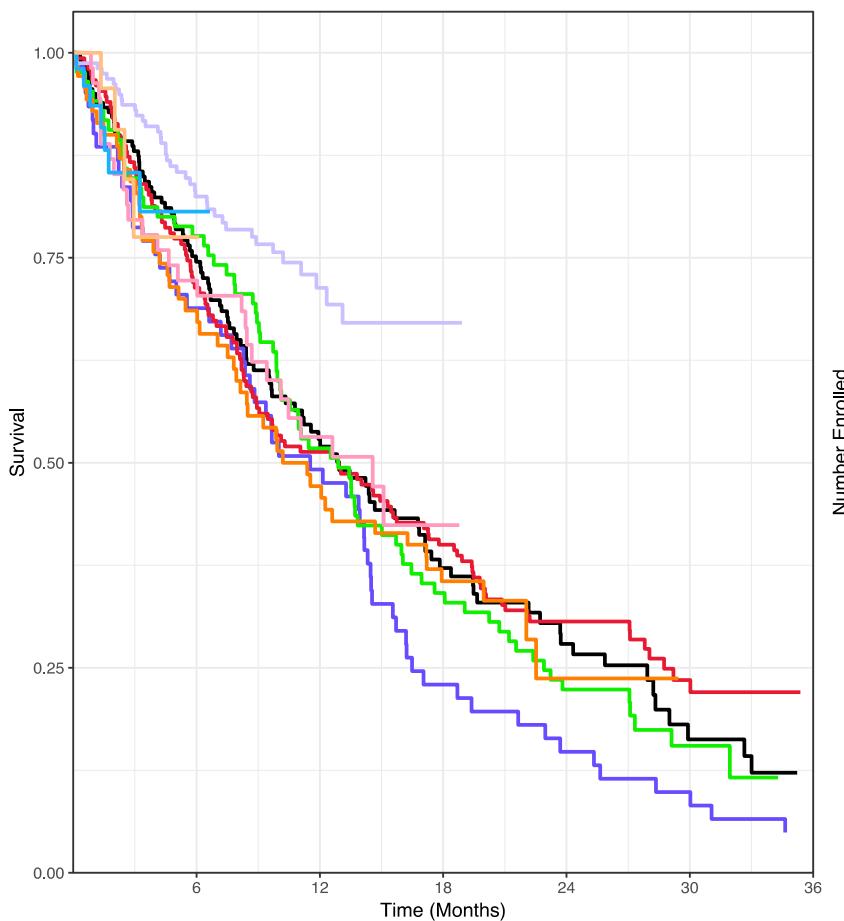
### Randomization Probabilities



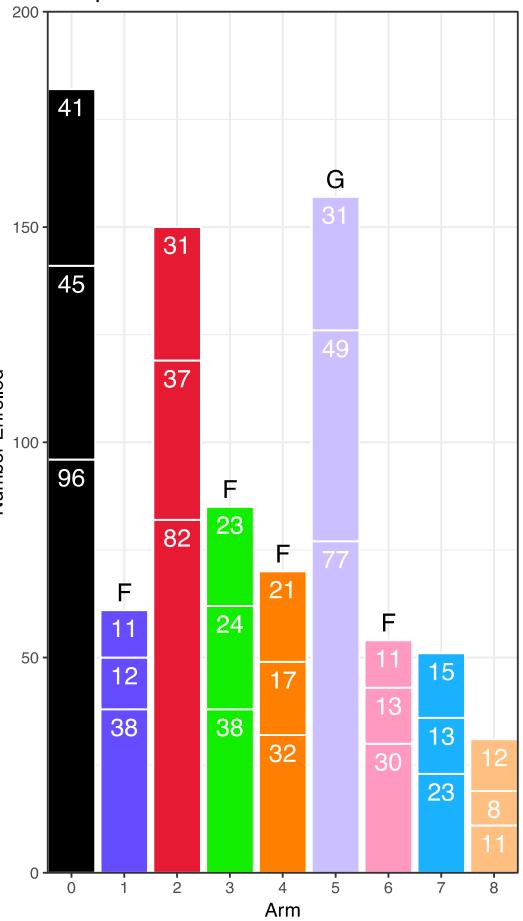
### Predictive Probabilities



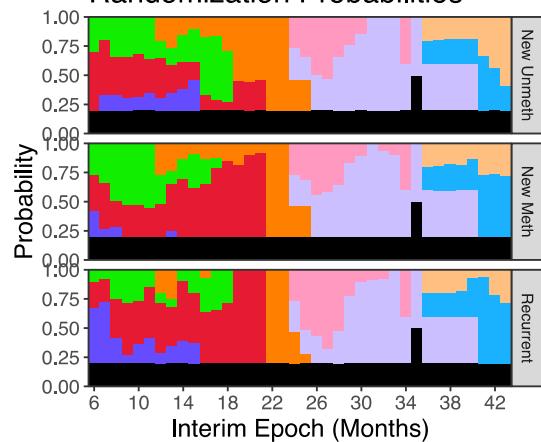
Overall Survival: All Patients. 43 Months After Trial Start



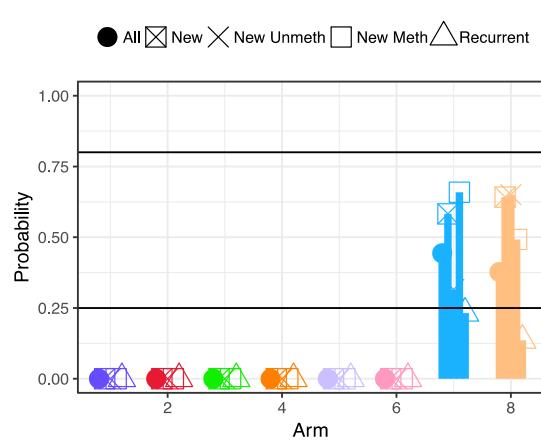
Sample Size



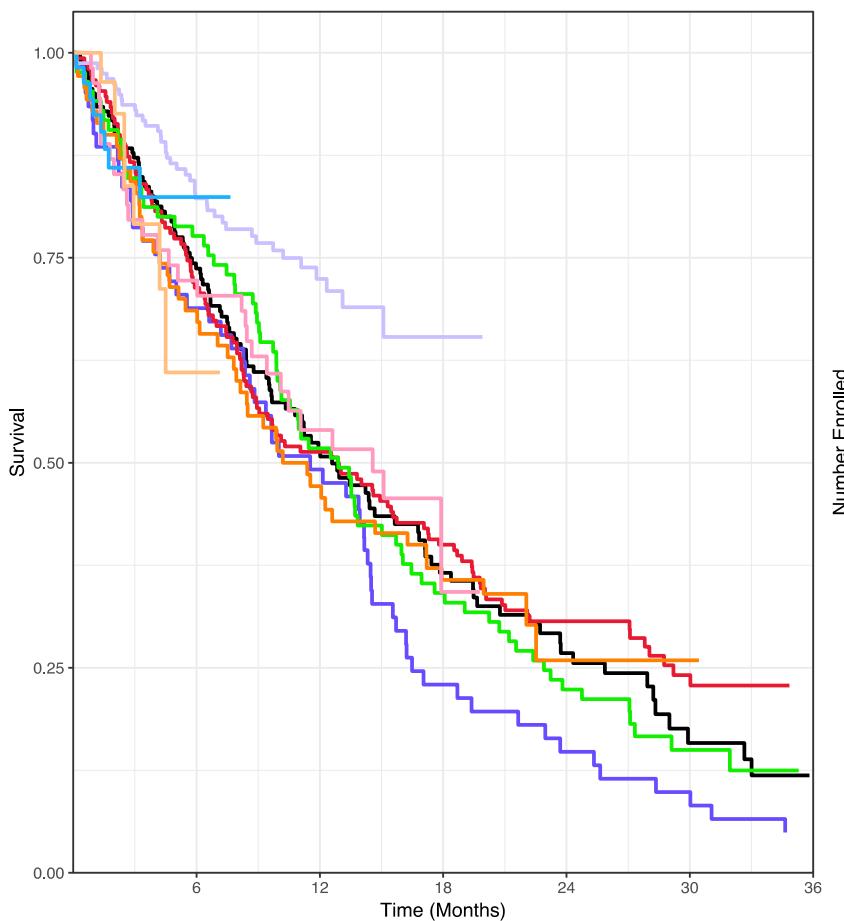
Randomization Probabilities



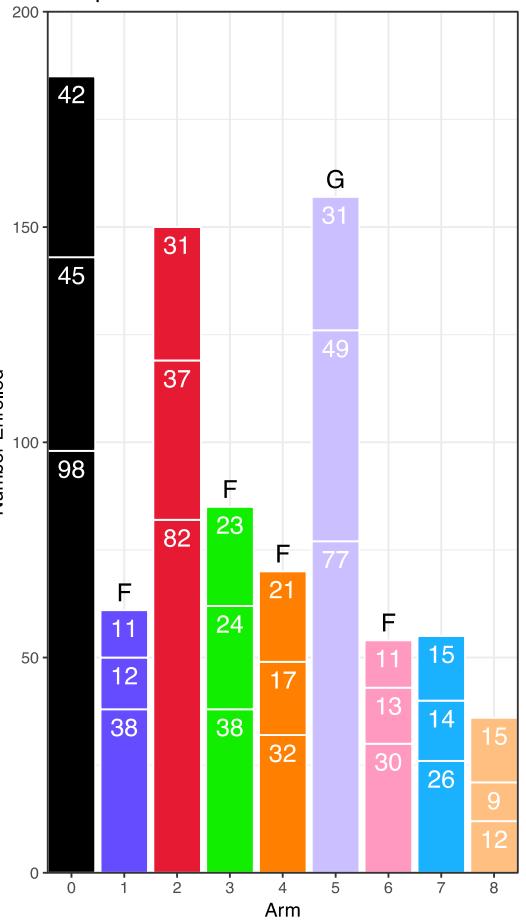
Predictive Probabilities



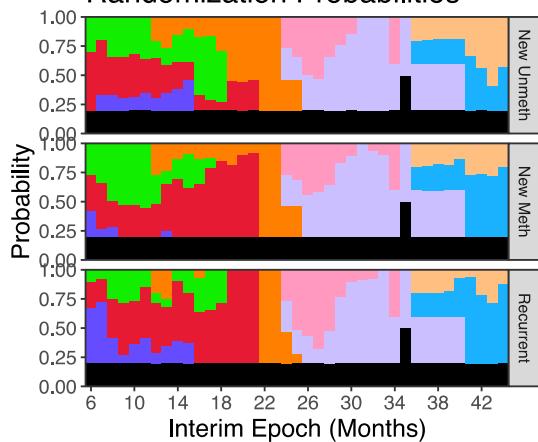
Overall Survival: All Patients. 44 Months After Trial Start



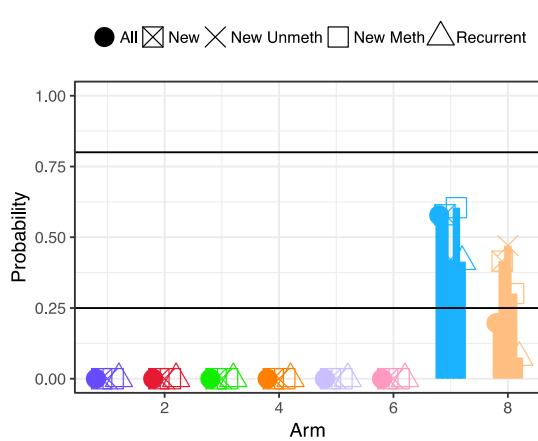
Sample Size



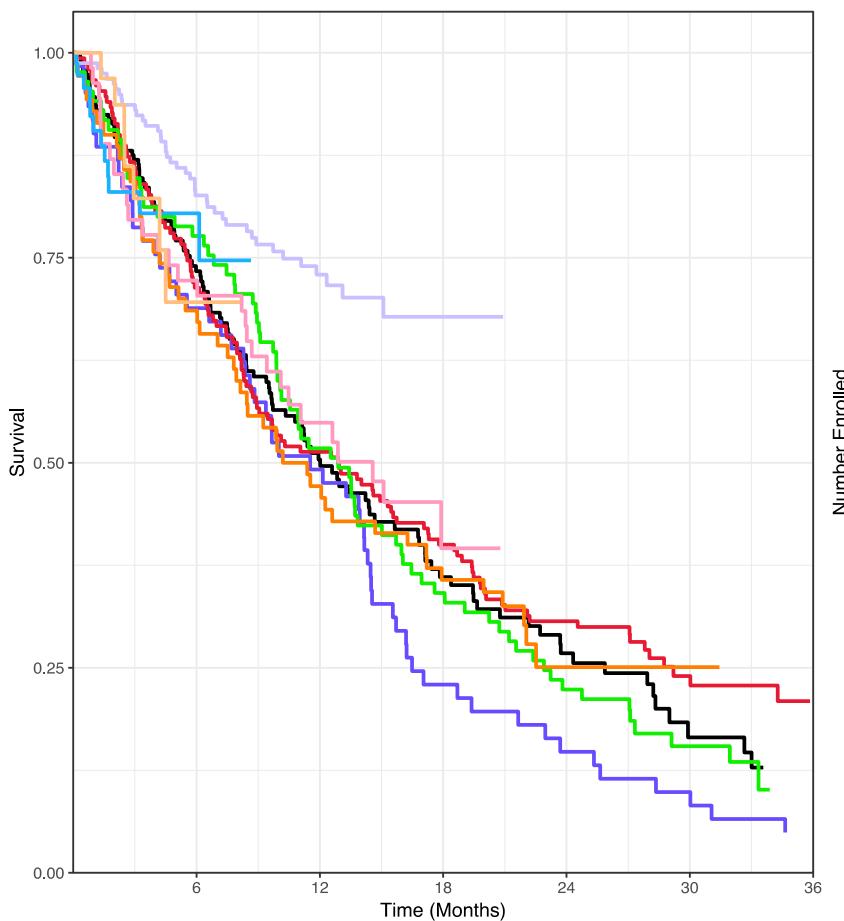
Randomization Probabilities



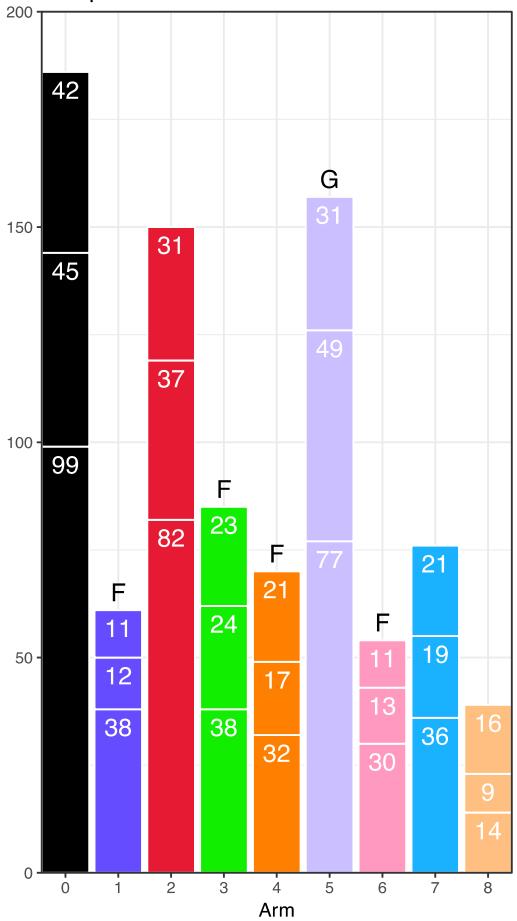
Predictive Probabilities



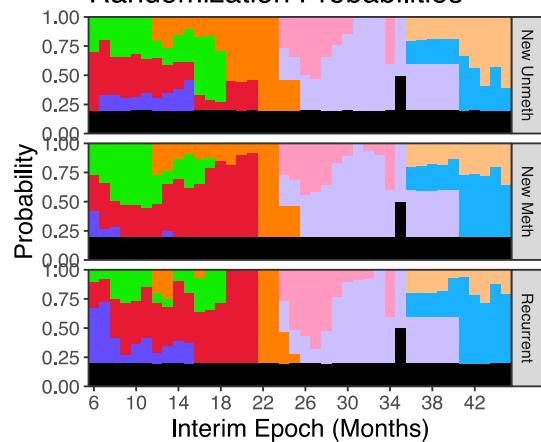
Overall Survival: All Patients. 45 Months After Trial Start



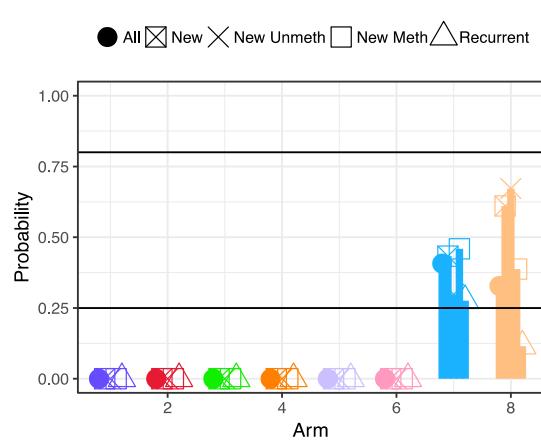
Sample Size



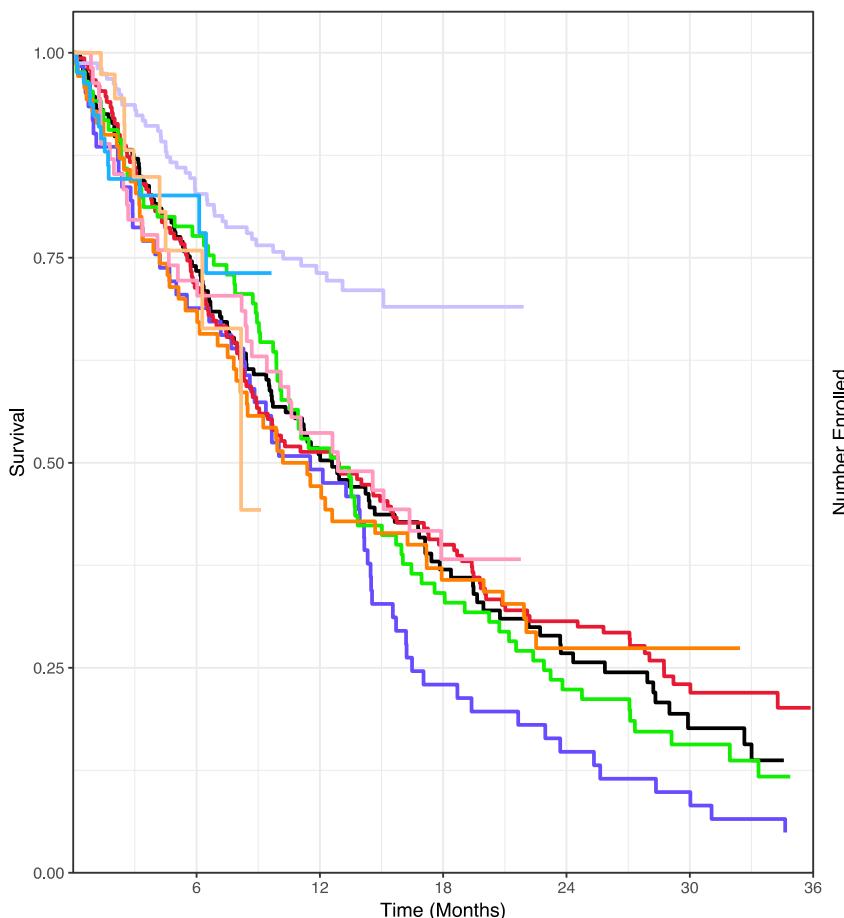
Randomization Probabilities



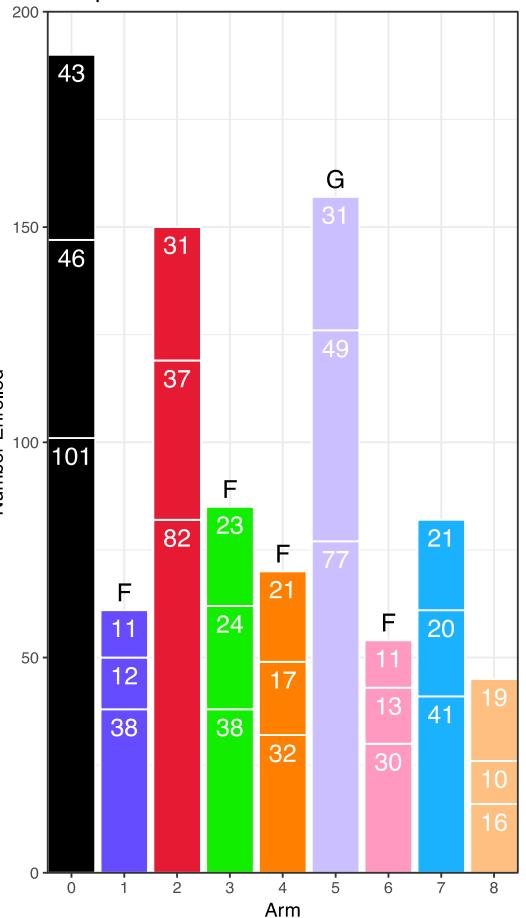
Predictive Probabilities



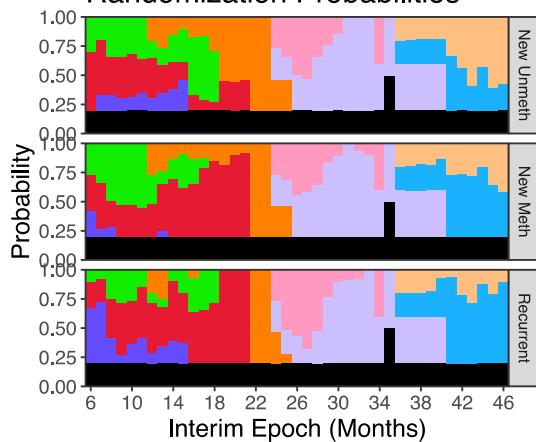
Overall Survival: All Patients. 46 Months After Trial Start



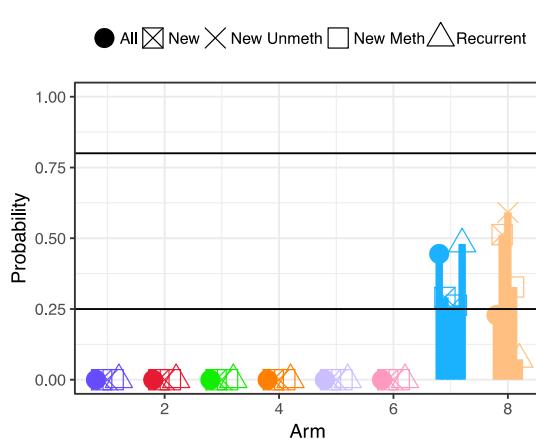
Sample Size



Randomization Probabilities

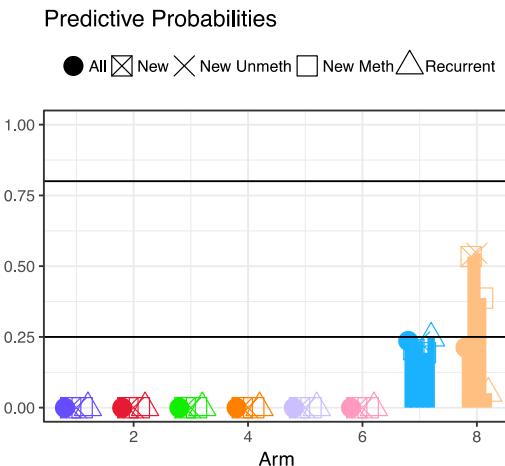
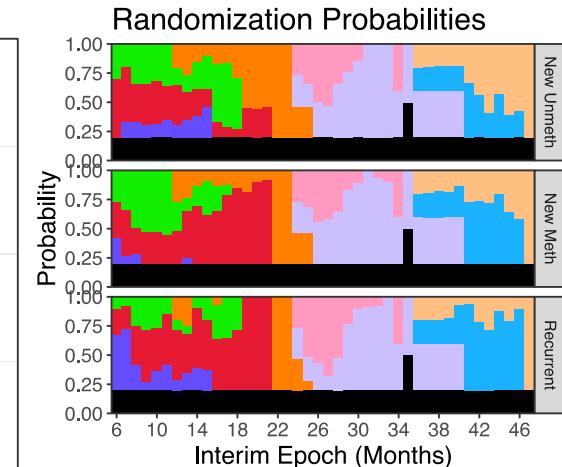
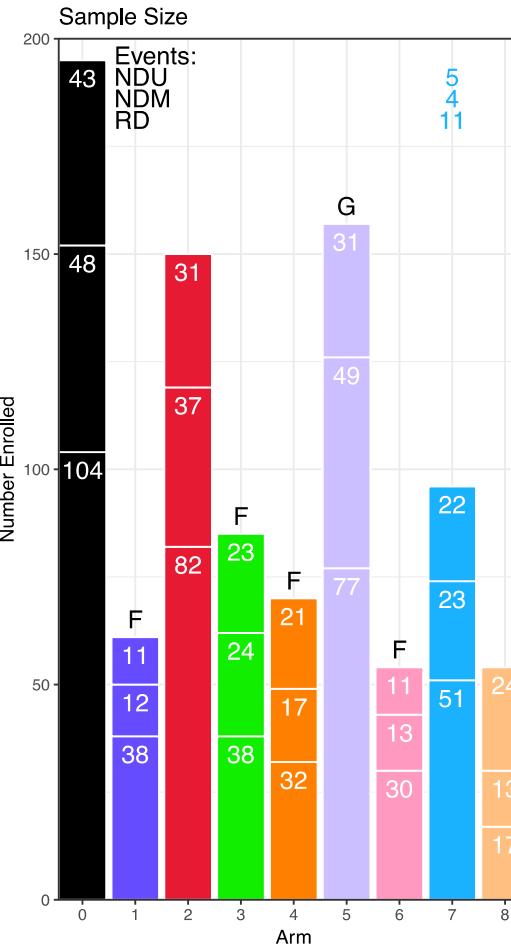
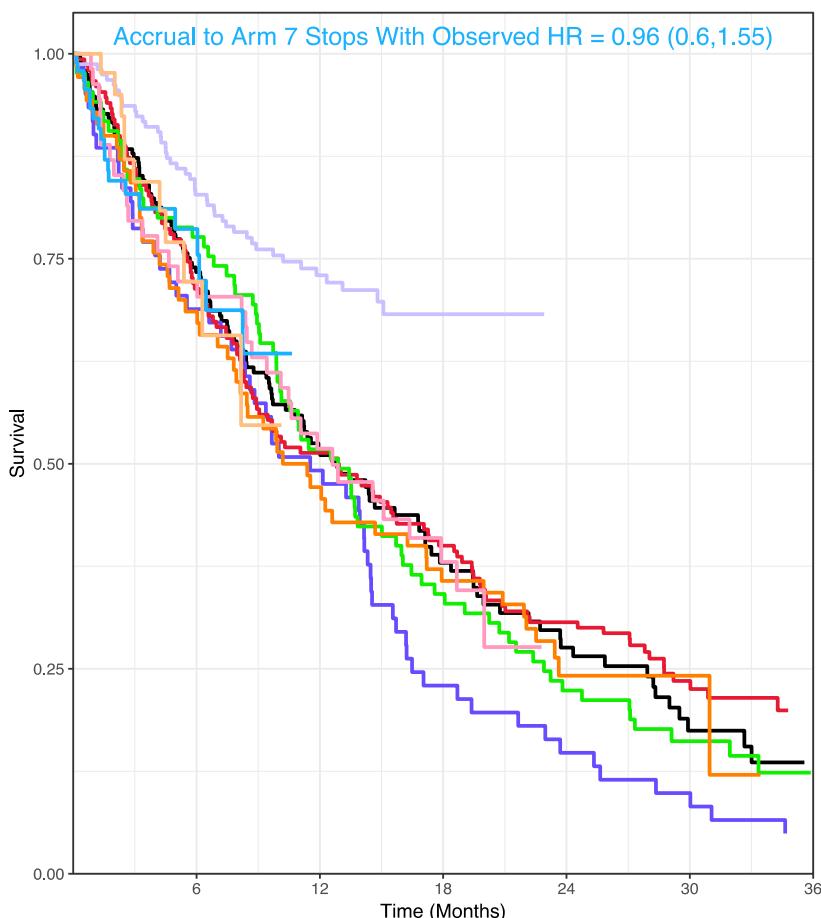


Predictive Probabilities



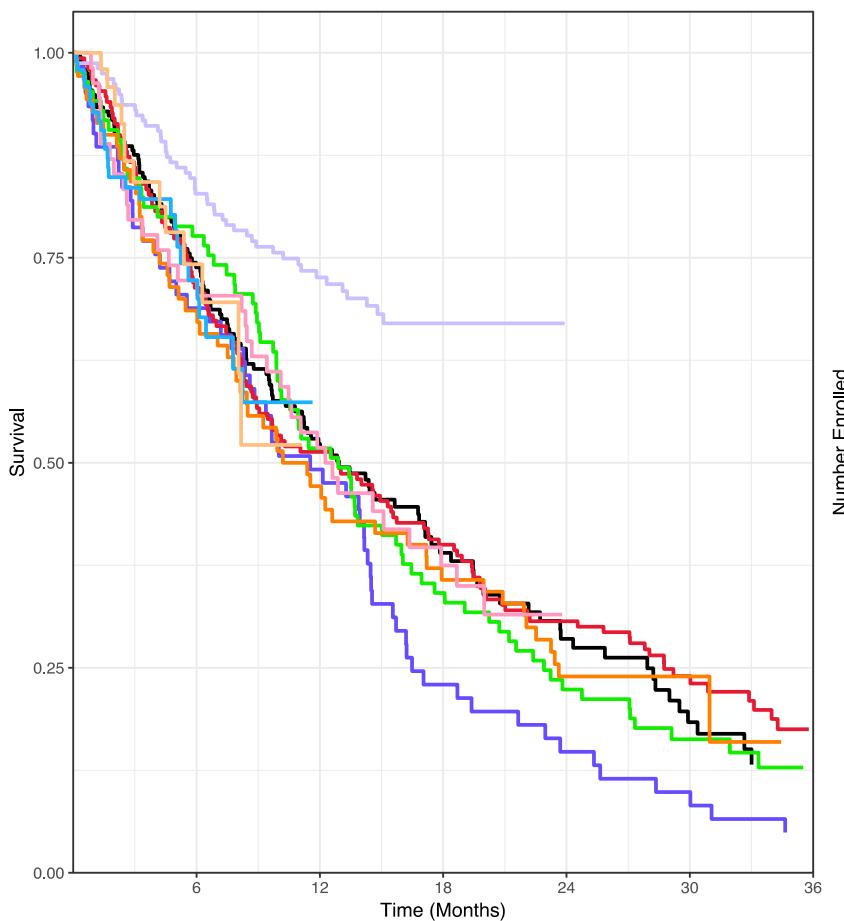
# Accrual to Arm 7 stops for futility with 96 patients, 11 months after it entered the trial

Overall Survival: All Patients. 47 Months After Trial Start

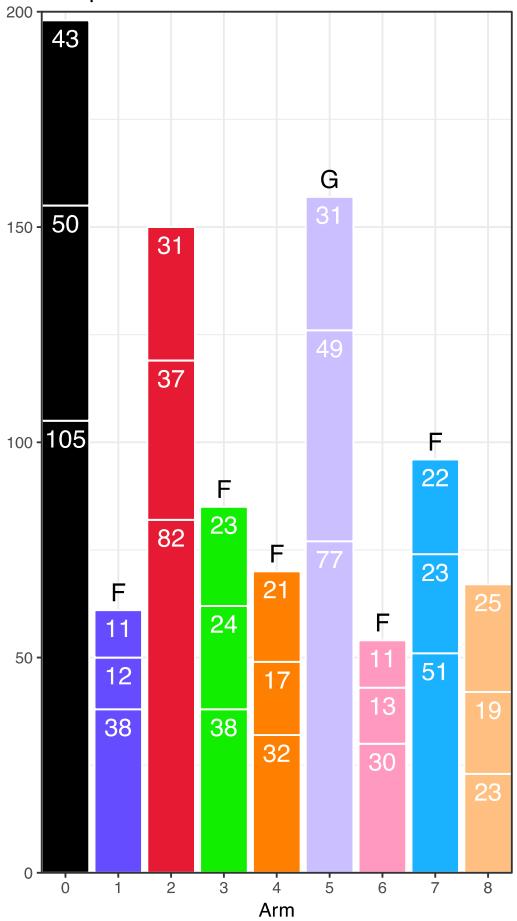


NDU: Newly Diagnosed Unmethylated; NDM: Newly Diagnosed Methylated; RD: Recurrent Disease

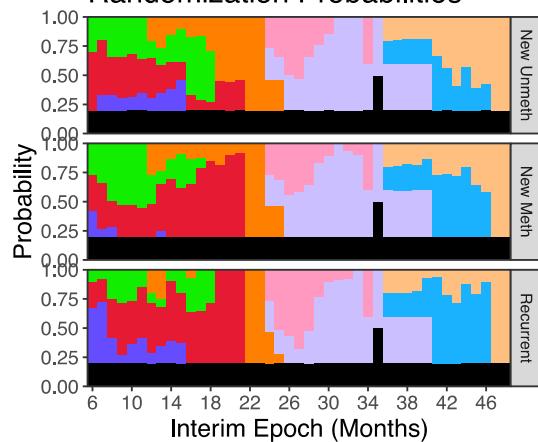
Overall Survival: All Patients. 48 Months After Trial Start



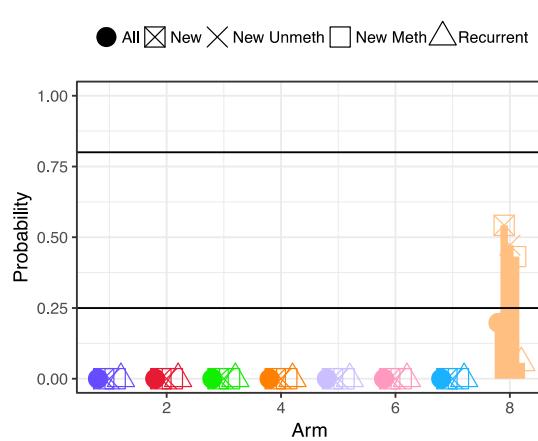
Sample Size



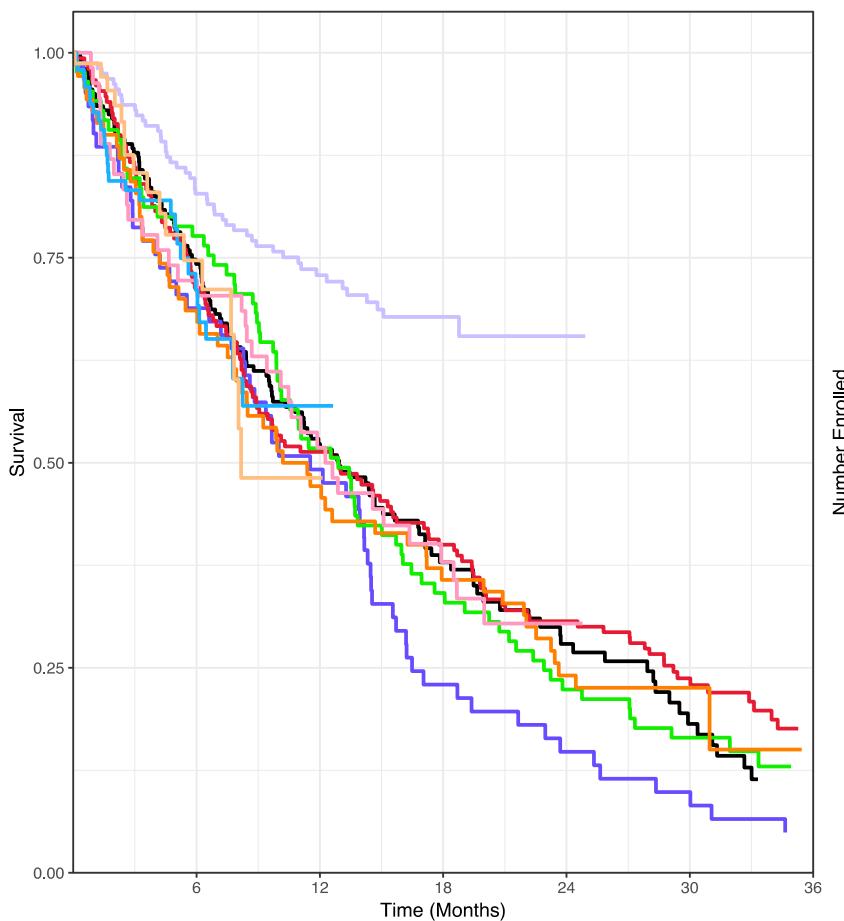
Randomization Probabilities



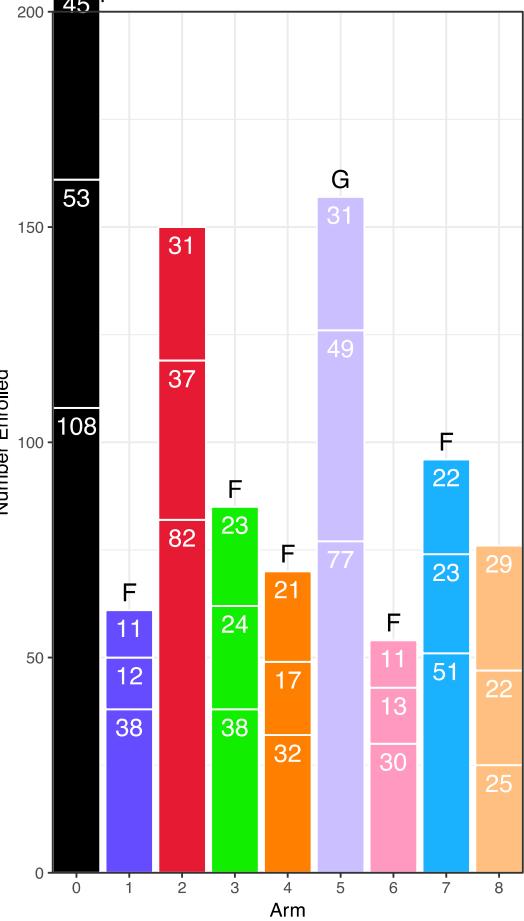
Predictive Probabilities



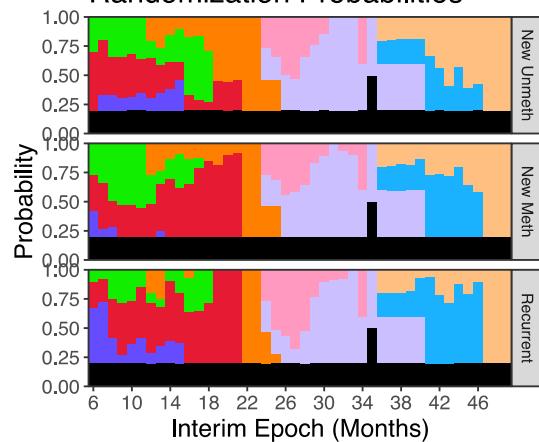
Overall Survival: All Patients. 49 Months After Trial Start



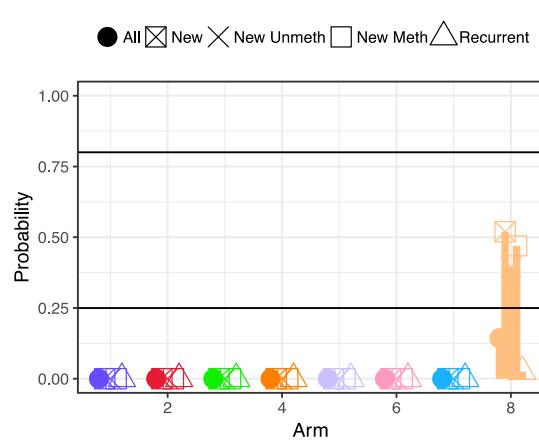
Sample Size



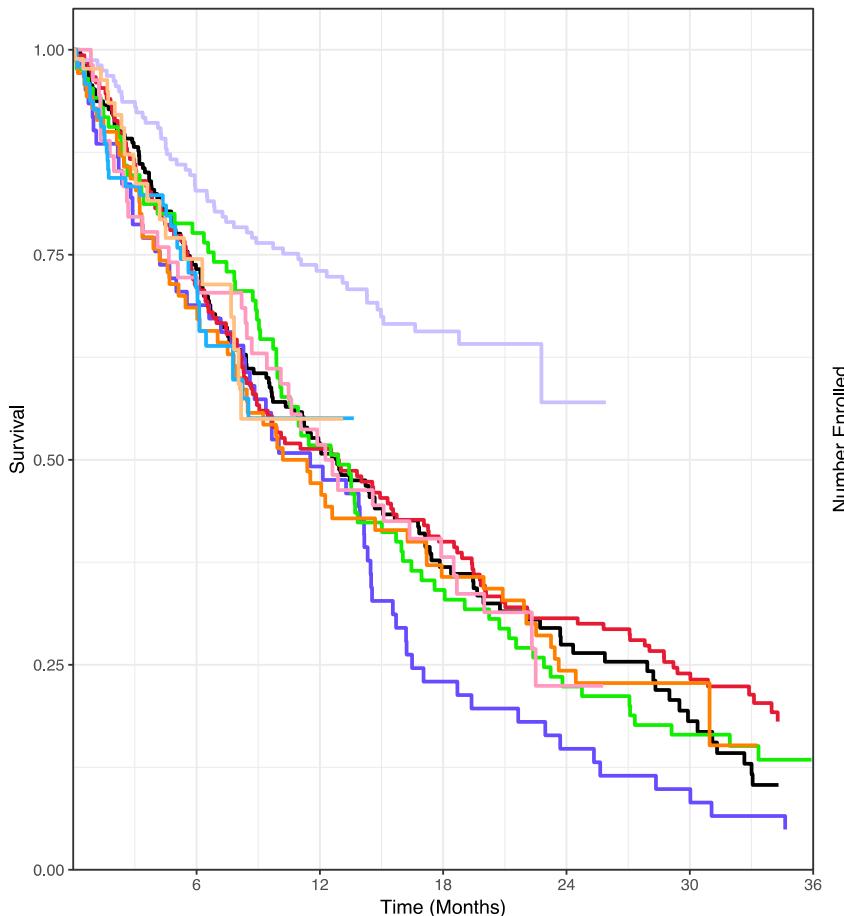
Randomization Probabilities



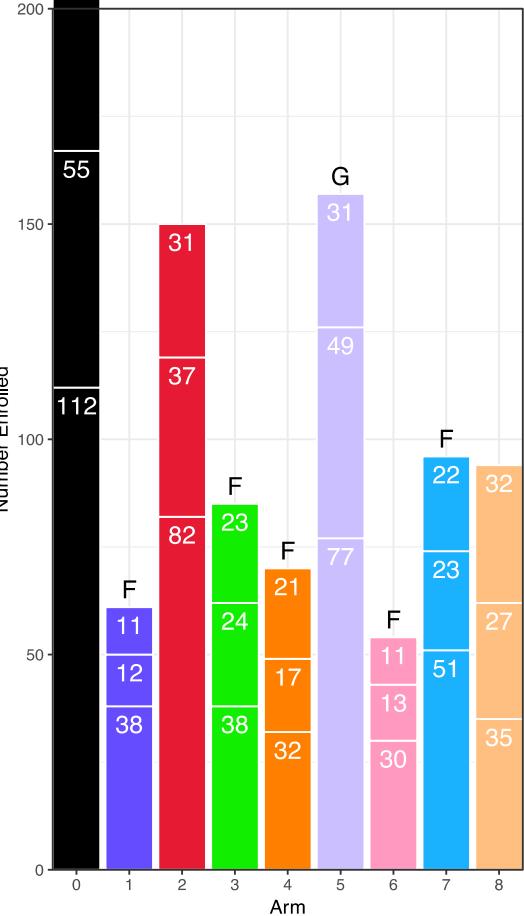
Predictive Probabilities



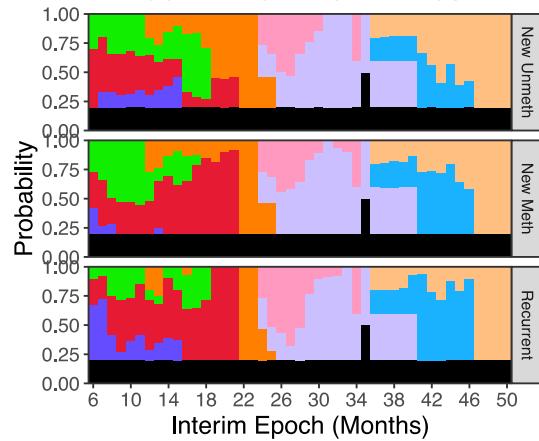
Overall Survival: All Patients. 50 Months After Trial Start



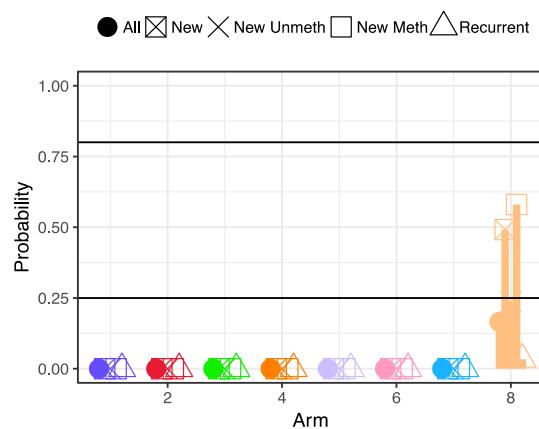
Sample Size



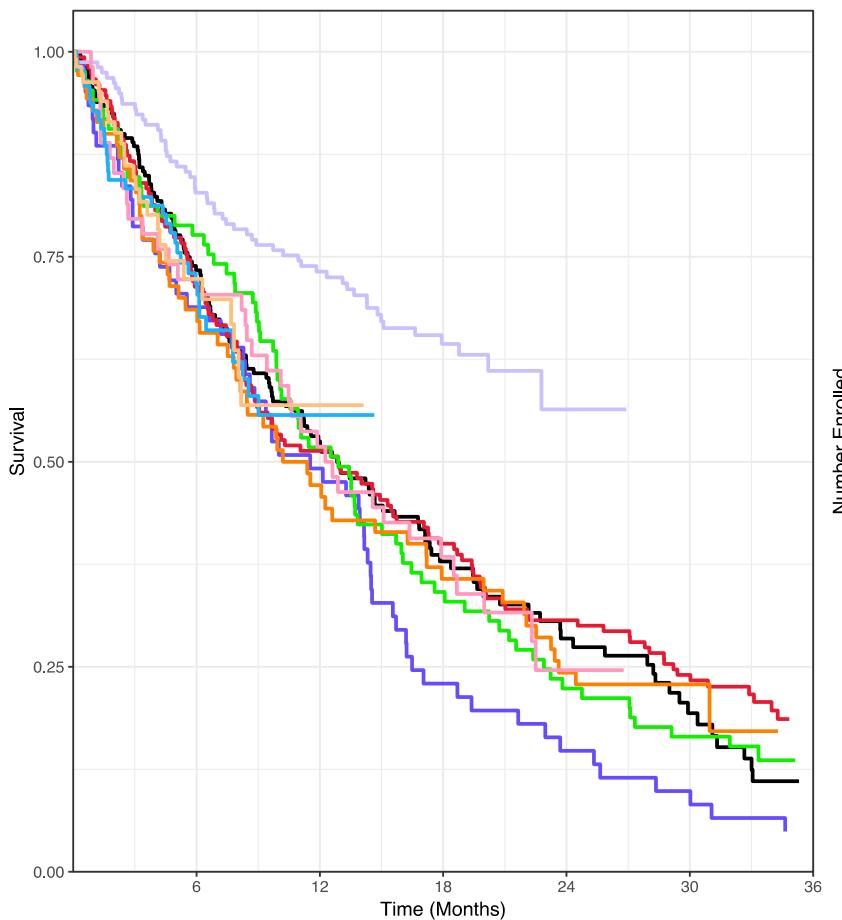
Randomization Probabilities



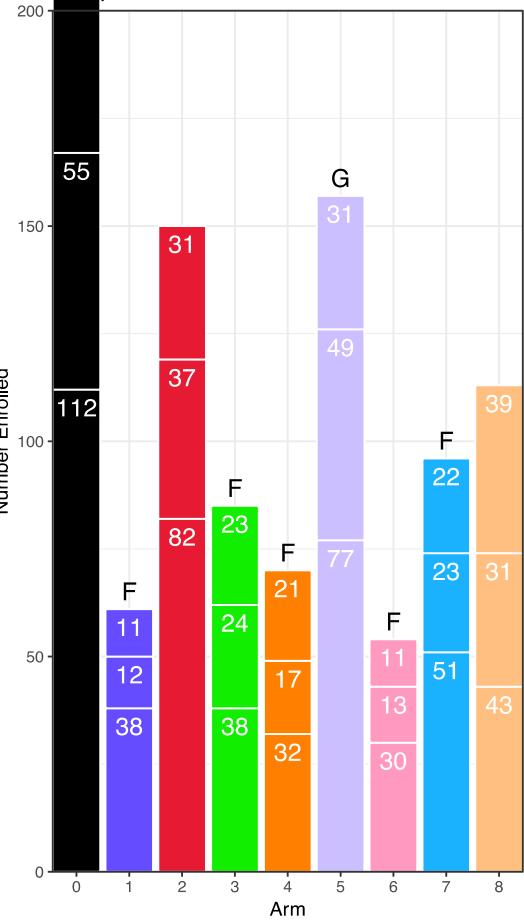
Predictive Probabilities



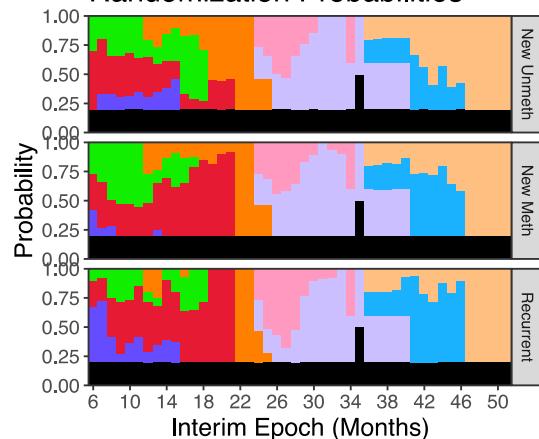
Overall Survival: All Patients. 51 Months After Trial Start



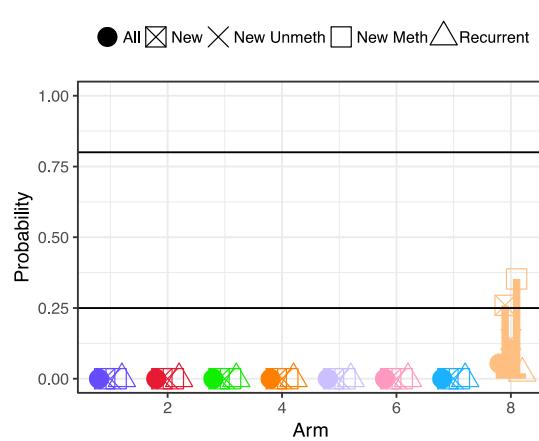
Sample Size



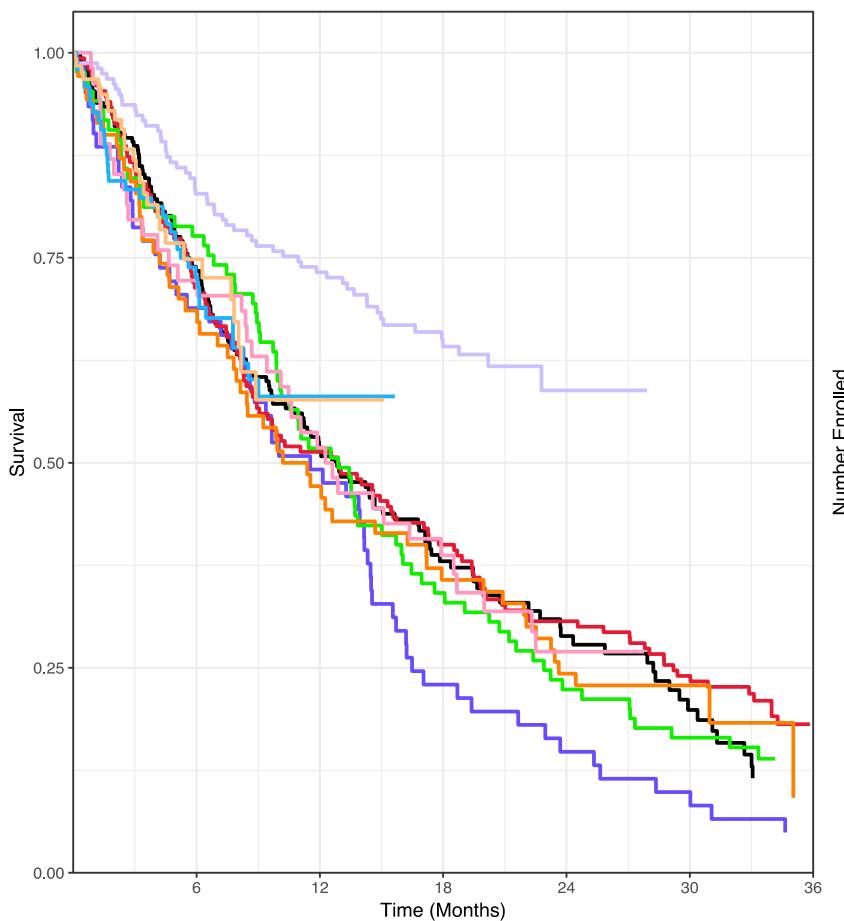
Randomization Probabilities



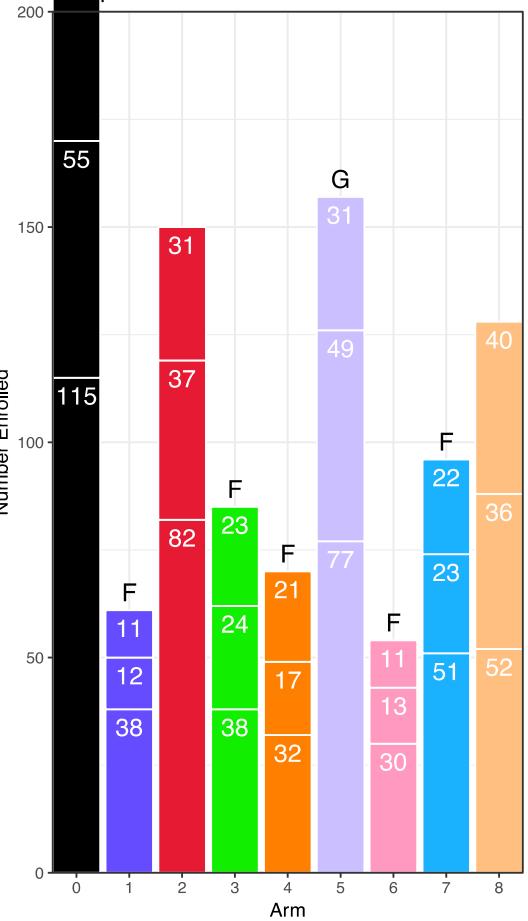
Predictive Probabilities



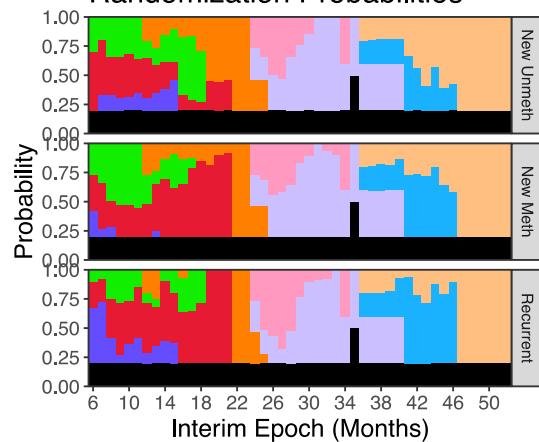
### Overall Survival: All Patients. 52 Months After Trial Start



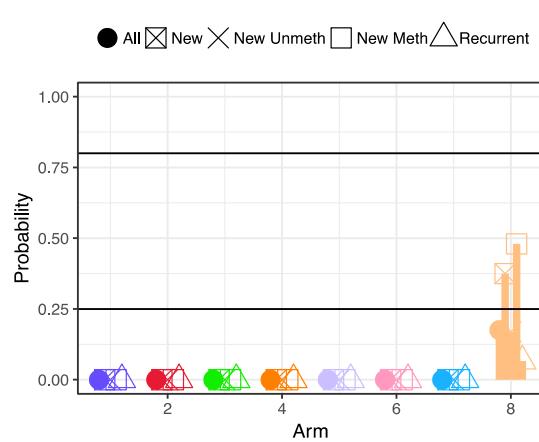
### Sample Size



### Randomization Probabilities

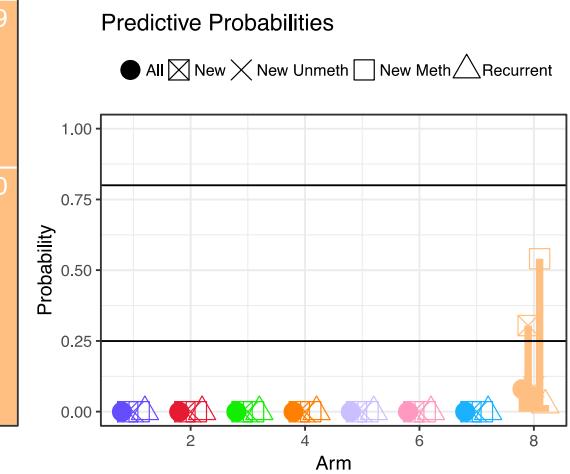
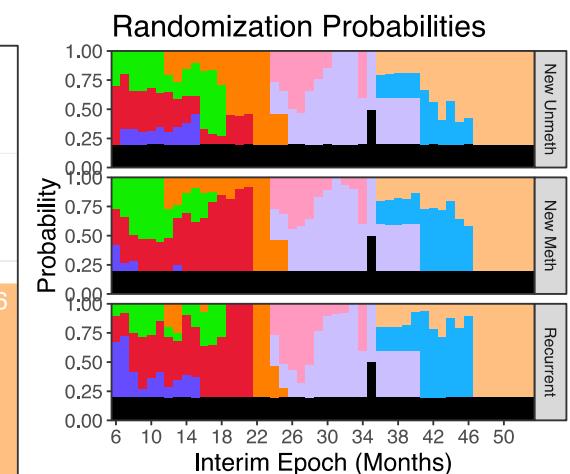
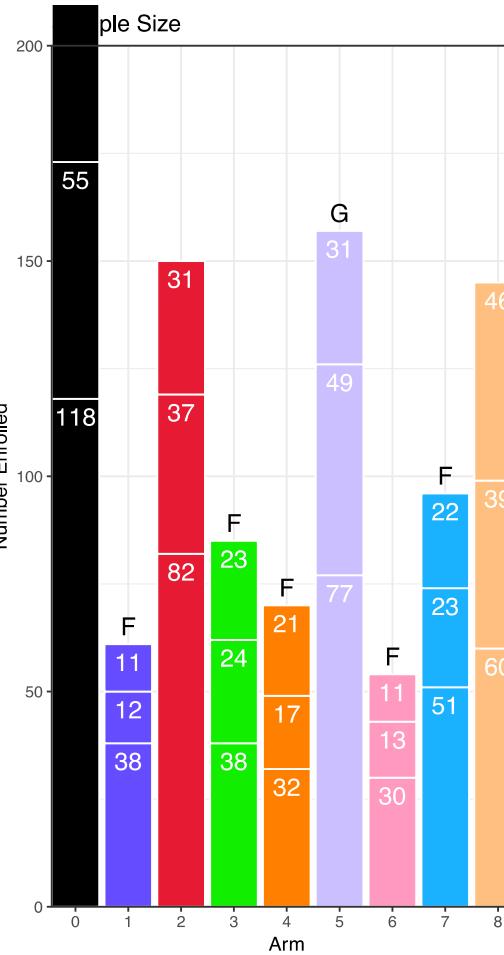
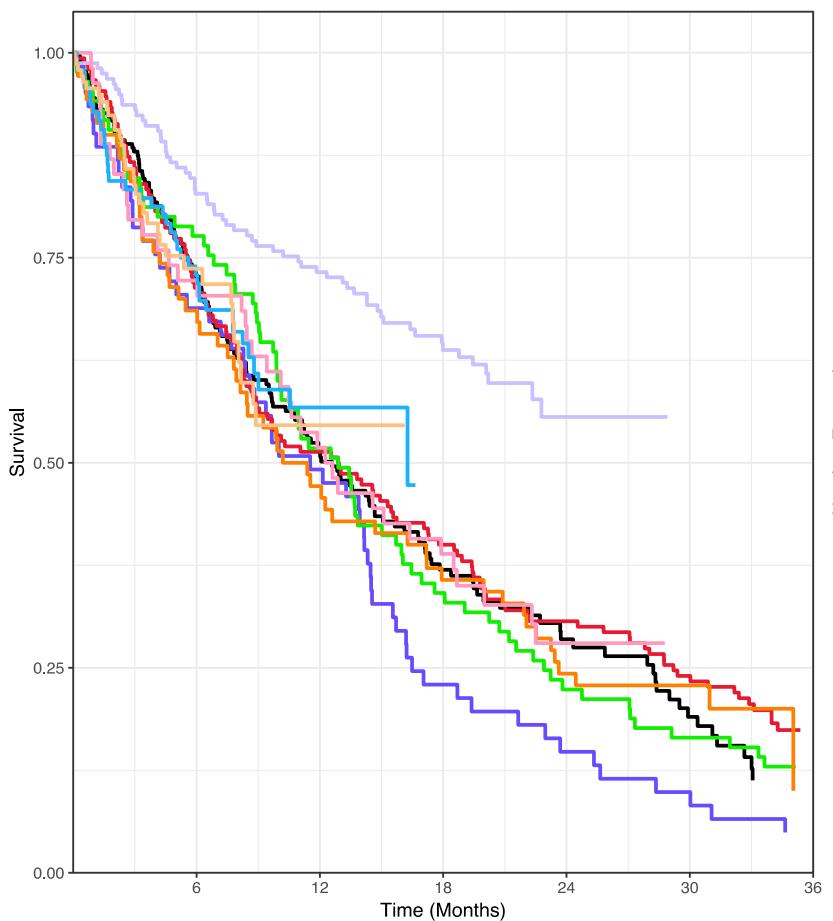


### Predictive Probabilities



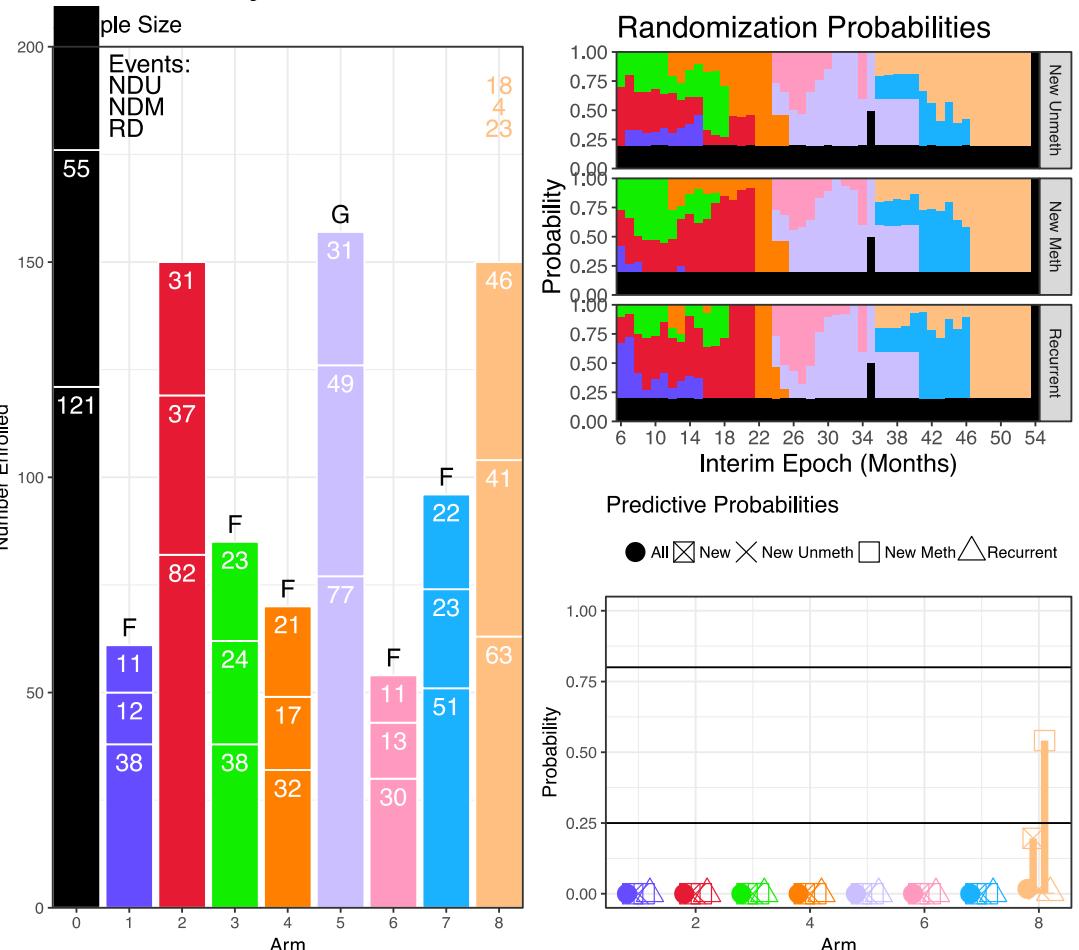
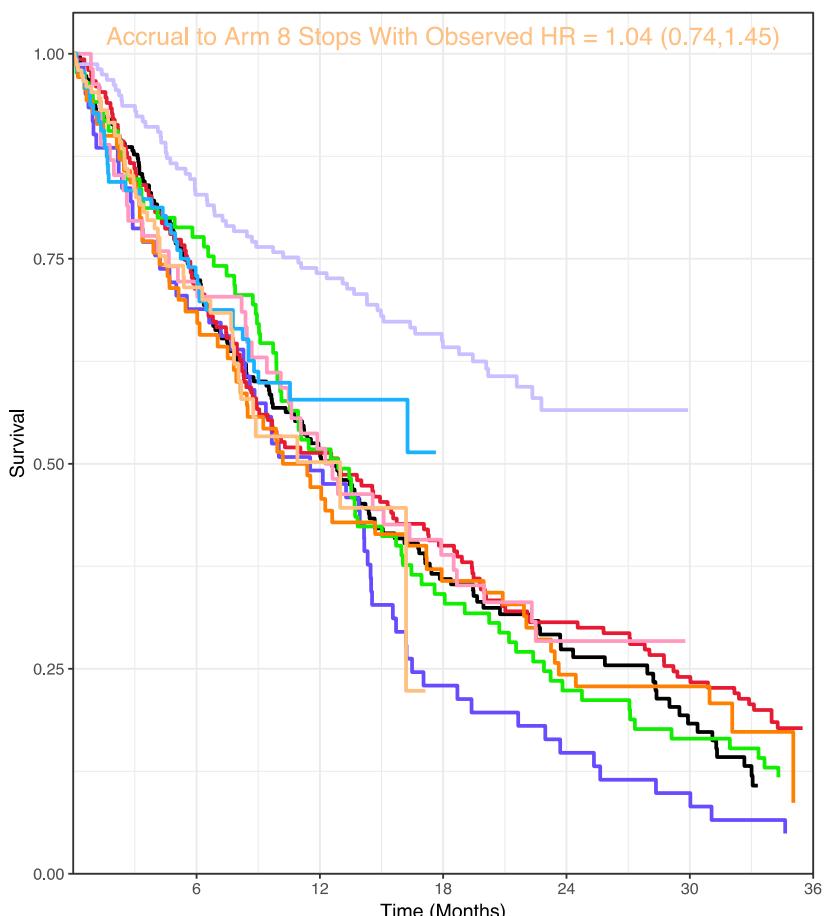
# Primary analysis for Arm 5. Probability of Benefit > 0.9999. HR = 0.44 (0.33, 0.59), 29 months after it entered the trial

Overall Survival: All Patients. 53 Months After Trial Start



# Accrual to Arm 8 stops after it enrolls its 150<sup>th</sup> patient, 18 months after it entered the trial

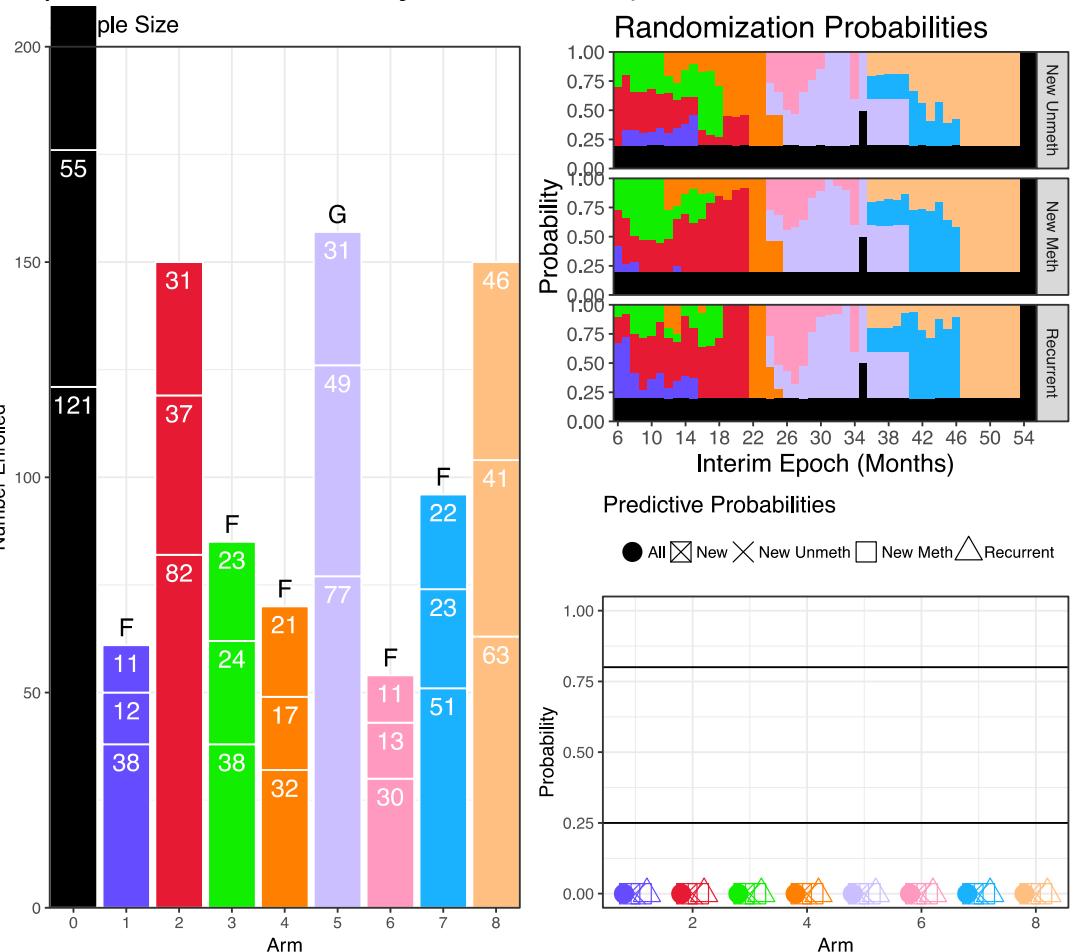
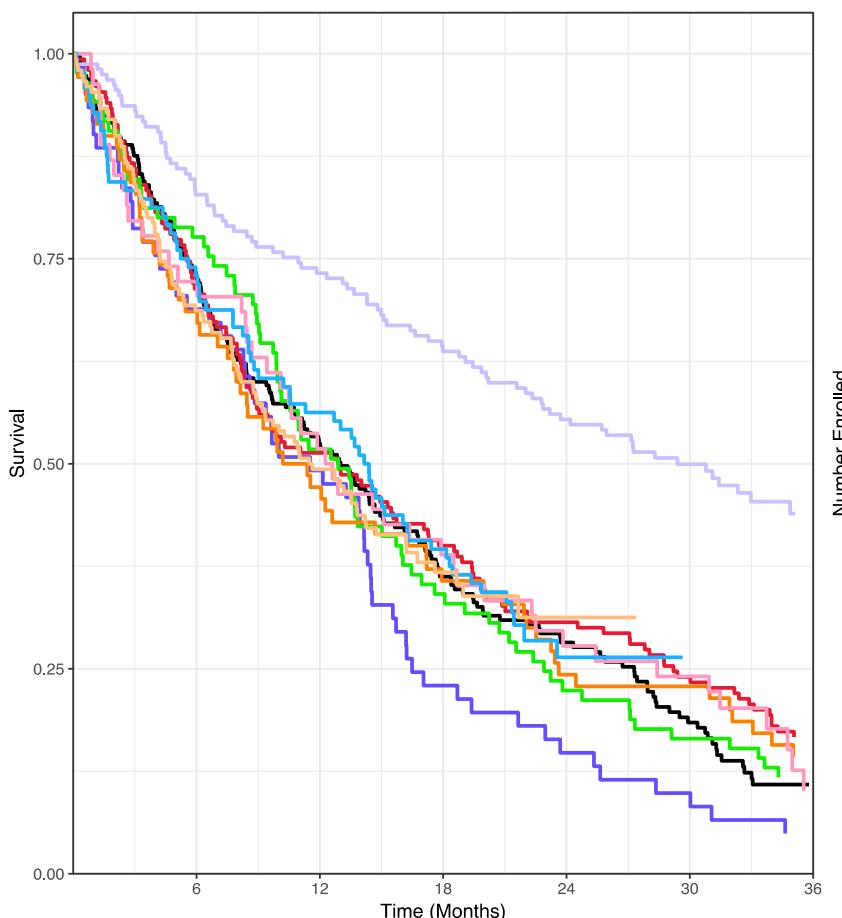
Overall Survival: All Patients. 54 Months After Trial Start



NDU: Newly Diagnosed Unmethylated; NDM: Newly Diagnosed Methylated; RD: Recurrent Disease

# Month 66: (12 additional months of follow-up for the patients enrolled in the first 54 months)

Overall Survival: All Patients. 66 Months After Trial Start



# Outline

- Why Bayes?
- What's a bandit strategy?
- Adaptive randomization
- Adaptive Bayesian platform trials, incl FDA views
- I-SPY 2 (neoadjuvant breast cancer)
- GBM AGILE (and Precision Promise)
- Need for simulations
- GBM AGILE ... the movie!

# **Adventures in Statistics**

- **Clinical trials**
  - Part 1: Bayesian bandits to platforms (today)
  - Part 2: Embracing adaptive Bayesian clinical trials
  - Part 3: Regulators and “complex innovative designs”

# Adventures in Statistics

- Clinical trials
  - Part 1: Bayesian bandits to platforms (today)
  - Part 2: Embracing adaptive Bayesian clinical trials
  - Part 3: Regulators and “complex innovative designs”
- Other adventures?
  - Multiplicities, *p*-values, & observational studies
  - Attributing decreases in breast cancer mortality to treatment vs screening: A Bayesian CISNET model (early ABC)
  - Science of doping ... or lack thereof
  - Corny but billion-dollar story of genetics of maize ... & Bayes
  - BRCAPRO: Risk model of carrying mutation of BRCA1 or BRCA2 based on family history ... a marriage of Mendel and Bayes