FRIENDS OF CANCER RESEARCH

Annual Meeting

Panel 1

Augmenting Randomized Confirmatory Trials for Breakthrough Therapies with Historical Clinical Trials Data

#FriendsAM18

Wifi: RitzCarlton_CONFERENCE

Password: fcr18

Panel 1 Participants

Moderator: Elizabeth Stuart, Johns Hopkins University

- Andrea Ferris, LUNGevity Foundation
- Antoine Yver, Daiichi Sankyo
- Joohee Sul, U.S. FDA
- Pallavi Mishra-Kalyani, U.S. FDA
- Rajeshwari Sridhara, U.S. FDA
- Ruthie Davi, Medidata Solutions

Augmenting Randomized Confirmatory Trials for Breakthrough Therapies with Historical Clinical Trials Data:
Landscape Review and Topic Introduction

Elizabeth Stuart, PhD

November 13, 2018

JOHNS HOPKINS
BLOOMBERG SCHOOL
of PUBLIC HEALTH

Big picture

- Need for non-experimental study designs
- But how do we know when we can “trust” the results from non-randomized studies?
- Growing literature to try to answer this question
- Today will discuss two case studies examining this, within a broader conversation about the possibilities (and issues)

©2016, Johns Hopkins University. All rights reserved.

The need for non-experimental studies

- Some questions cannot be answered using randomized experiments:
  - Effects of approved treatments in real-world use
  - Effects of treatments it would be unethical to deny from some individuals
  - Effects of potentially harmful exposures (e.g., carcinogens)

- So then left with non-experimental studies where we did not control who received which exposure/intervention
  - Many designs exist
  - Today will focus on comparison group designs

©2016, Johns Hopkins University. All rights reserved.

The challenge of non-experimental studies

- But in non-experimental studies we have to worry about confounding
  - The people that get one treatment may be quite different from those that get another
- We can use statistical methods to deal with observed differences
- But to interpret difference in outcomes as due to the treatment need to assume no unobserved confounders that we didn't adjust for
- So how do we know when to trust the results?

©2016, Johns Hopkins University. All rights reserved.

Topic of this session

- Two case studies seeing whether using historical comparison data (instead of randomized controls) can yield accurate effect estimates
- In this case, does the average outcome under the control condition match between the historical data and the randomized arm?
  - (After adjusting for observed characteristics)
- A start at developing a body of literature on this topic

©2016, Johns Hopkins University. All rights reserved.

Related literature

- This idea of “design replication” or “within study comparison” studies common in other fields, esp. social sciences (esp. Tom Cook)
- Growing body of research learning lessons from that field
- Factors that lead to accurate results:
  - Consistent measurement of covariates and outcomes
  - Match groups based on the key confounders, esp. strong predictors of outcome (e.g., baseline measures)
  - Match “locally” if possible

The question: Which of these findings carry over to cancer research?

©2016, Johns Hopkins University. All rights reserved.

Plan for the session

- Brief introductions of panel
- Two case studies (10 min. each)
- Panel discussion (45 min.)
- Audience Q+A

©2016, Johns Hopkins University. All rights reserved.

medidata

Exploration of whether a Synthetic Control Arm derived from Historical Clinical Trials by Matching of Baseline Characteristics can Replicate the Overall Survival of a Randomized Control Arm

Case Study in Non-Small Cell Lung Cancer

Working Group

Ruthie Davi
Medidata

Andrea Ferris
LUNGevity Foundation

Andrew Howland
Medidata

Dominic Labriola
Bristol-Myers Squibb

Michael LeBlanc
Fred Hutchinson Cancer Research Center, University of Washington

David Lee
Medidata

Antara Majumdar
Medidata

Pallavi Mishra-Kalyani
U.S. FDA

Zhenming Shun
Daiichi Sankyo

Rajeshwari Sridhara
U.S. FDA

Larry Strianese
Medidata

Elizabeth Stuart
Johns Hopkins University

Joohee Sul
U.S. FDA

Xiang Yin
Medidata

Antoine Yver
Daiichi Sankyo

© 2018 Medidata Solutions, Inc. – Proprietary and Confidential

The Challenge: Recruitment, Retention, and Compliance in Randomized Control

- Maintaining a concurrent control arm can be difficult due to rarity of the disease or availability of the investigational agent outside the study
- BRAVO study in BRCA+ breast cancer[1]
  - Unusually high rate of censoring in the control arm likely associated with increased availability of PARP inhibitors
  - Unlikely to produce data that is interpretable
- Sunitinib Malate in gastrointestinal stromal tumor[2]
  - Large effect on progression free survival (HR 0.33, 95% CI (0.24, 0.47))
  - After 84% of placebo patients elected to receive sunitinib malate, effect on overall survival was not observed (HR 0.88, 95% CI (0.68, 1.1))

1. Tesaro press release (2017)
2. Sunitinib malate capsule prescribing information (2006)

© 2018 Medidata Solutions, Inc. – Proprietary and Confidential

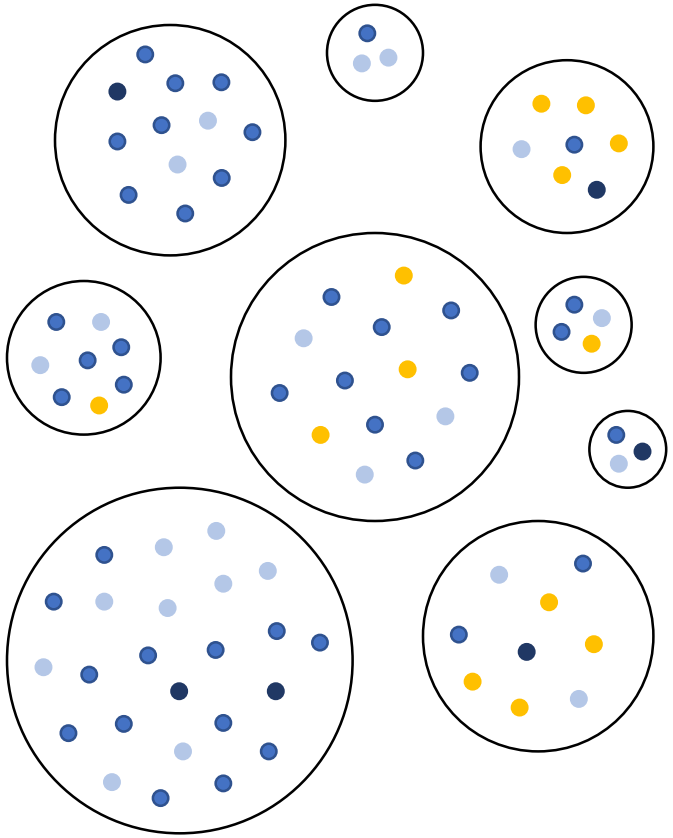
A Possible Solution – Synthetic Control Arm

A Synthetic Control Arm is

- *Patient level data from multiple historical clinical trials in the same indication*
- *Carefully selected historical patients*
  - Who meet eligibility criteria and were assigned to receive the appropriate standard of care
  - With baseline characteristics that statistically match those in the current-day experimental arm
- *For a setting where a randomized control is problematic*

© 2018 Medidata Solutions, Inc. – Proprietary and Confidential

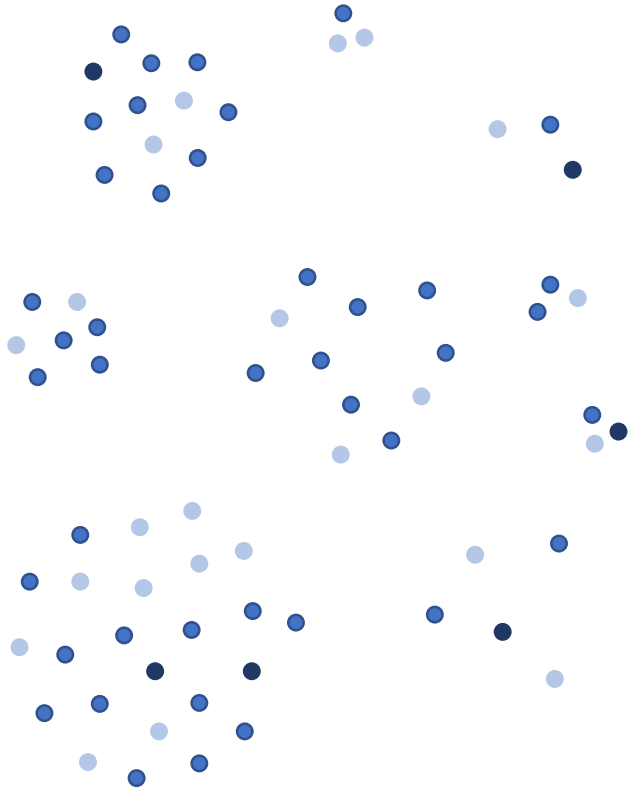
Historical Clinical Trials Data



Treatment Arm of Target Trial

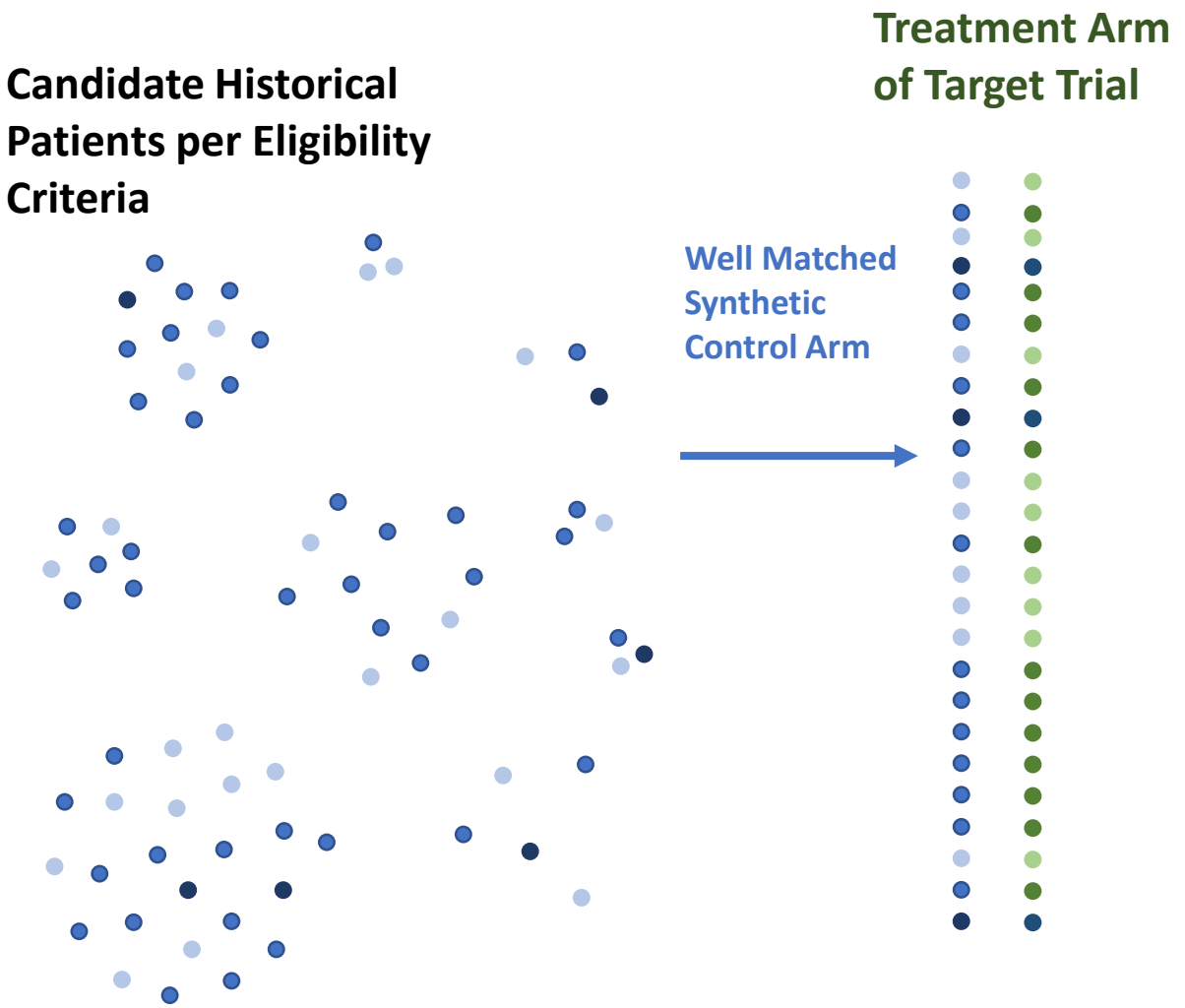
© 2018 Medidata Solutions, Inc. – Proprietary and Confidential

Candidate Historical Patients per Eligibility Criteria

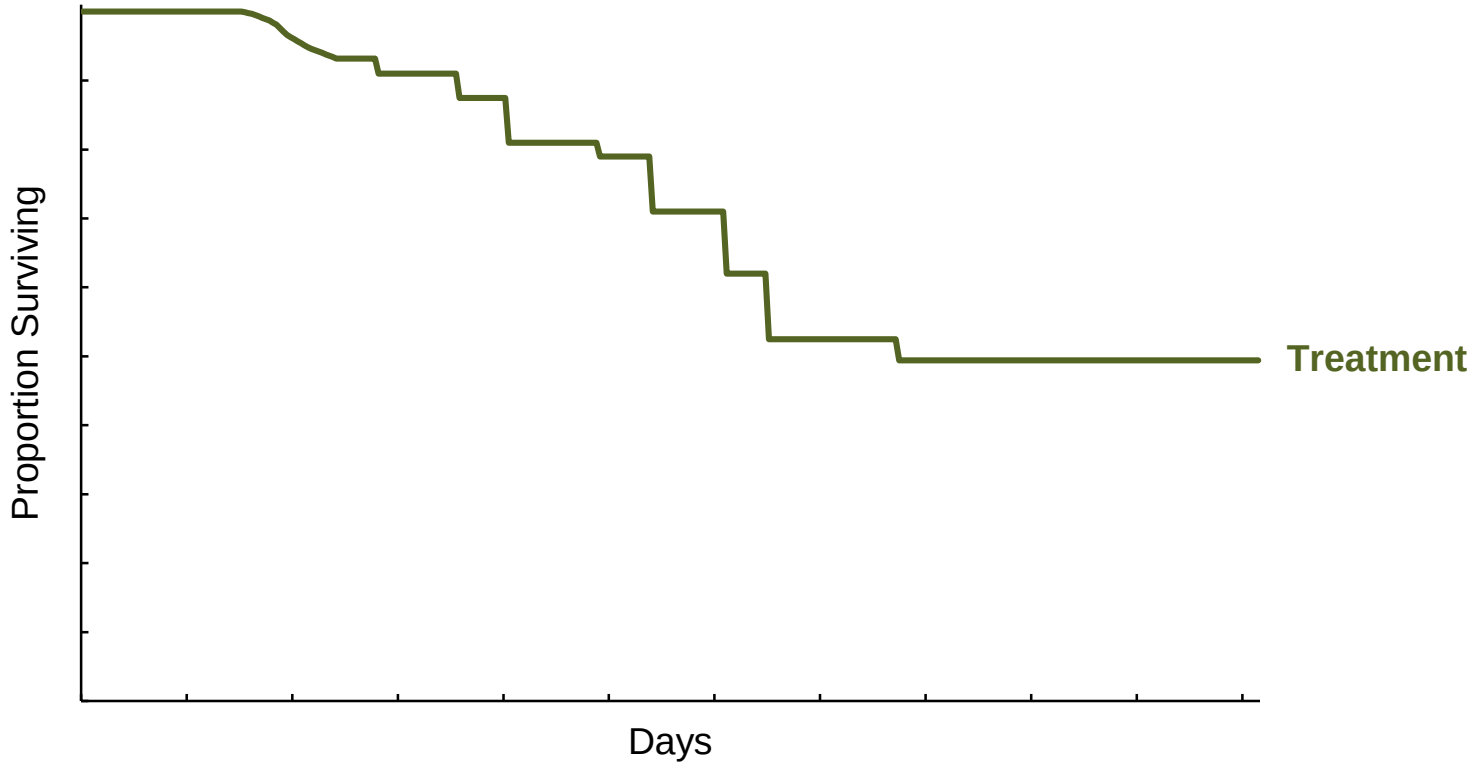


Treatment Arm of Target Trial

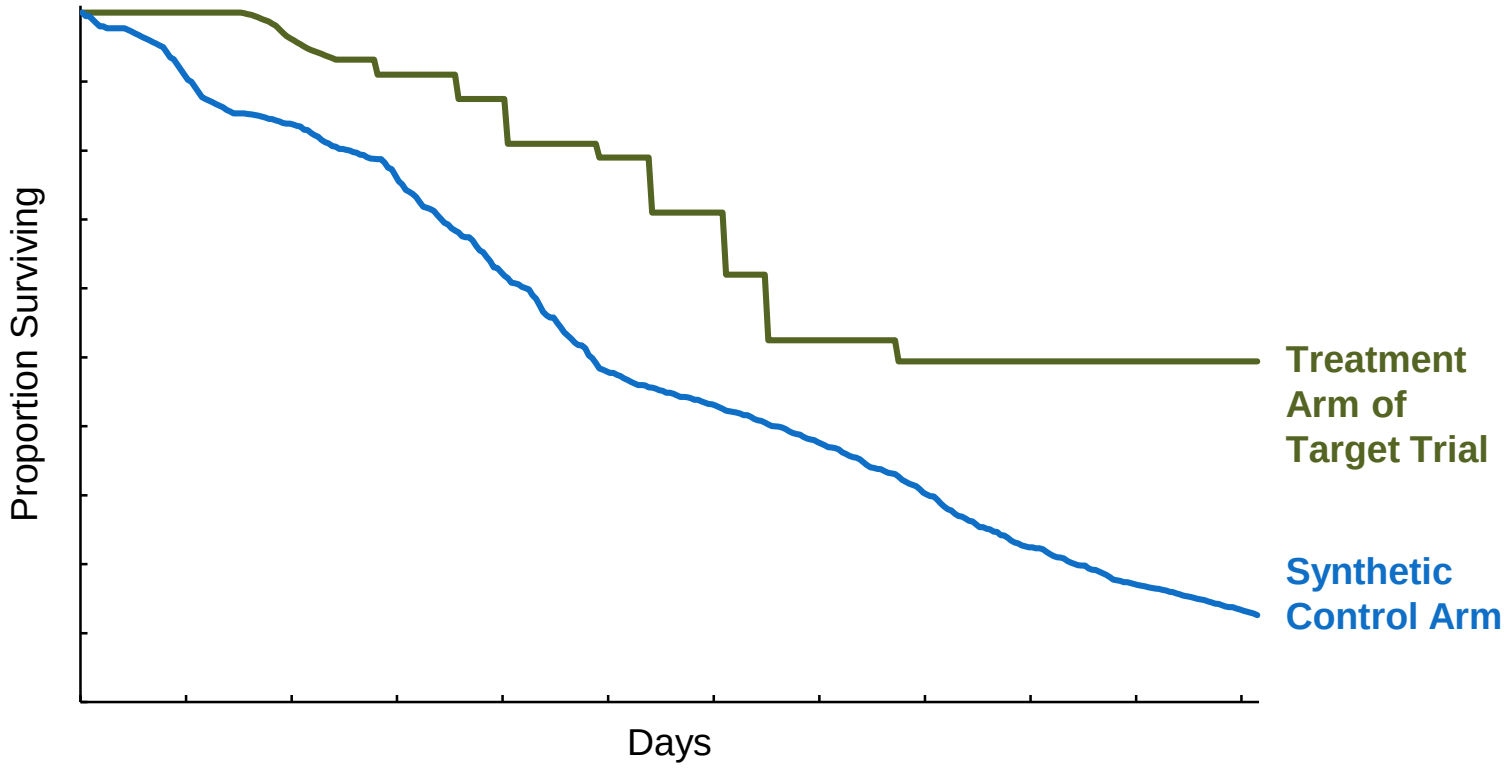
© 2018 Medidata Solutions, Inc. – Proprietary and Confidential



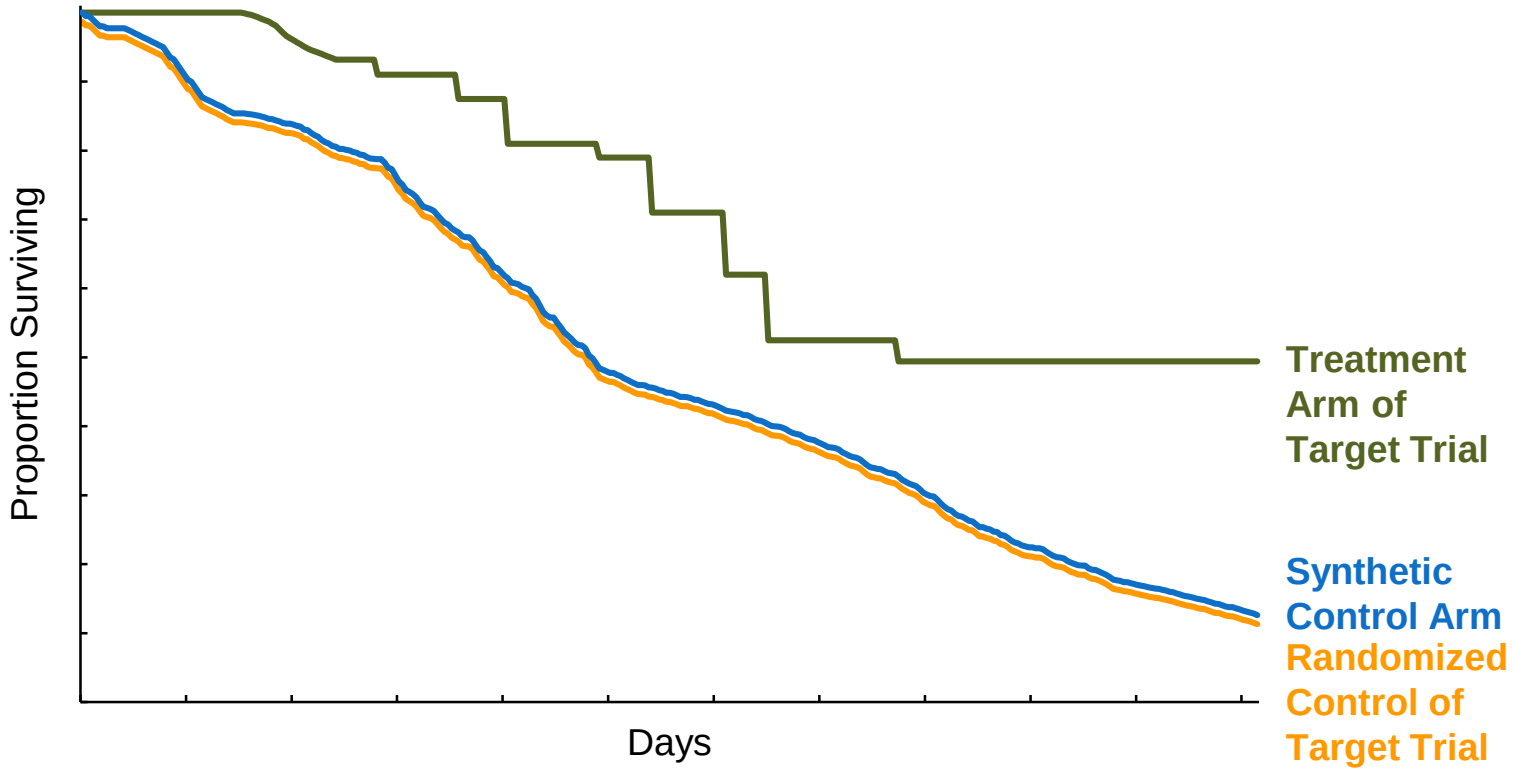
© 2018 Medidata Solutions, Inc. – Proprietary and Confidential



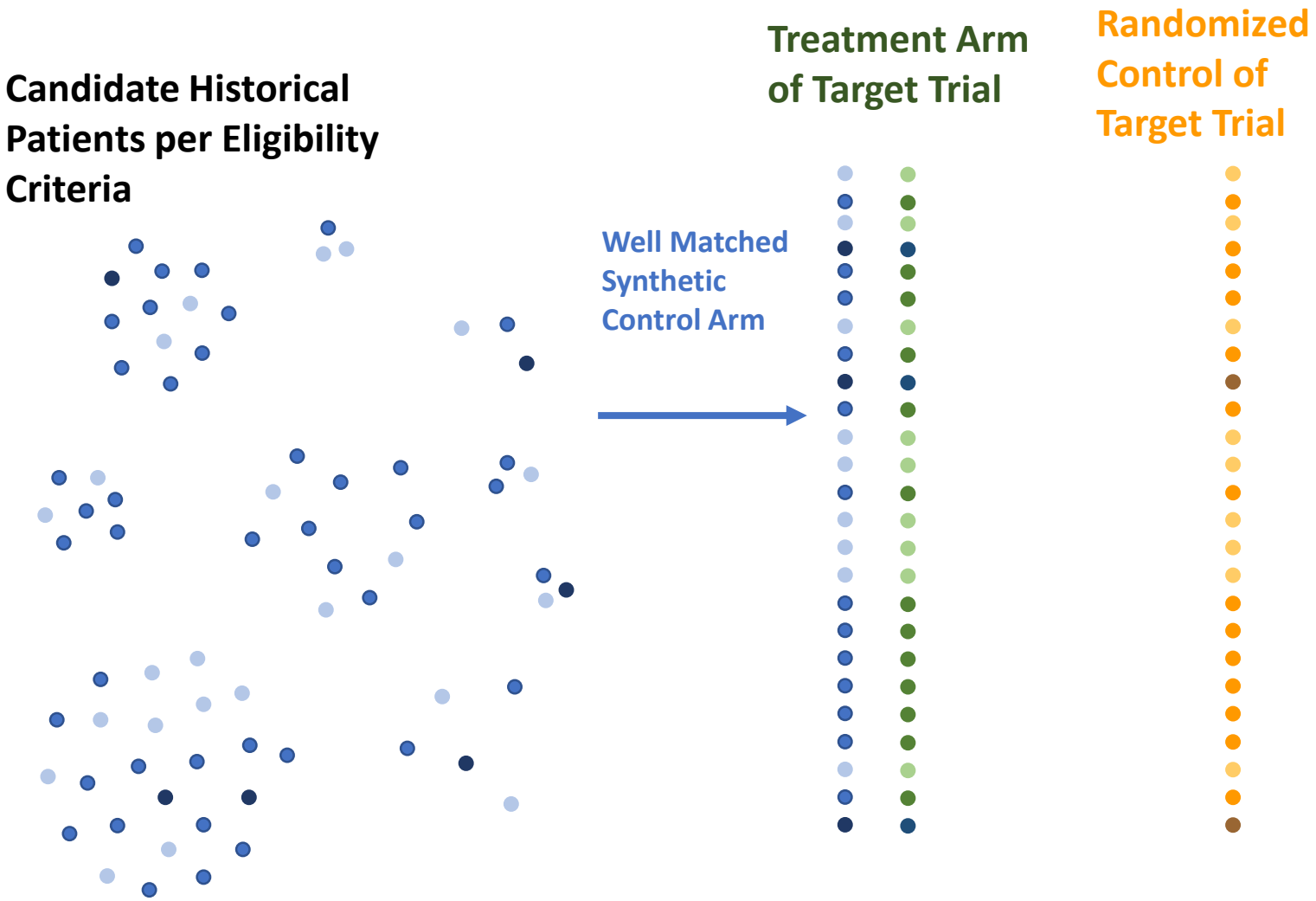
© 2018 Medidata Solutions, Inc. – Proprietary and Confidential



© 2018 Medidata Solutions, Inc. – Proprietary and Confidential

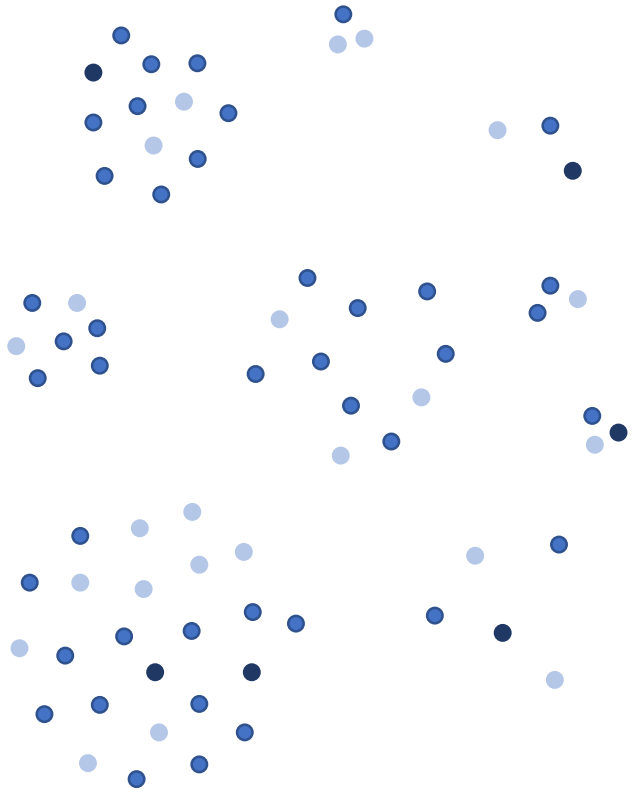


© 2018 Medidata Solutions, Inc. – Proprietary and Confidential



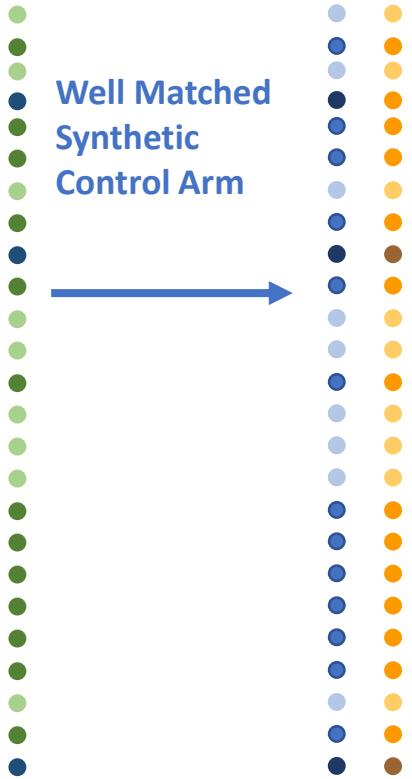
© 2018 Medidata Solutions, Inc. – Proprietary and Confidential

Candidate Historical Patients per Eligibility Criteria



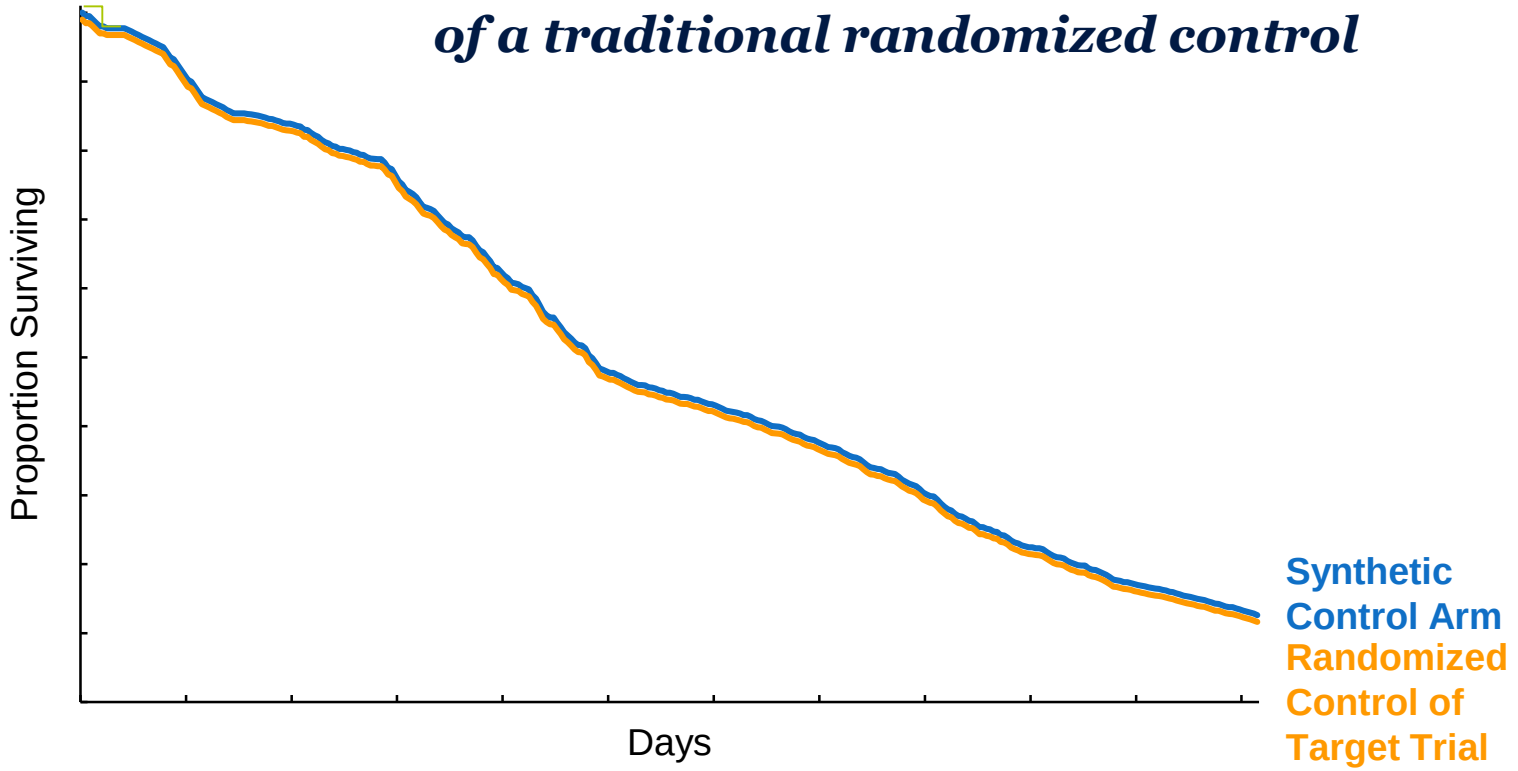
Treatment Arm of Target Trial

Randomized Control of Target Trial



© 2018 Medidata Solutions, Inc. – Proprietary and Confidential

Case Study Objective: Assess whether a Synthetic Control Arm can be created to replicate the outcomes of a traditional randomized control



© 2018 Medidata Solutions, Inc. – Proprietary and Confidential

Non-Small Cell Lung Cancer Case Study

© 2018 Medidata Solutions, Inc. – Proprietary and Confidential

Building the SCA

Propensity Score Matching

© 2018 Medidata Solutions, Inc. – Proprietary and Confidential

Patient level data from multiple previous NSCLC trials

Data Sources

- Project Data Sphere[1]
- Medidata Enterprise Data Store (MEDS)[2]

Trial Characteristics

- Open label and blinded phase 2 & 3 trials
- Multinational
- Timespan of starts of trials (2004 to 2013)
- Overall survival measured

SCA Patient Eligibility Criteria

- Inclusion in a historical clinical trial accessible within this project
- NSCLC at stage III or IV
- Received prior platinum-based chemotherapy
- Men and women ≥ 18 years of age
- ECOG performance status of ≤ 2
- Measurable disease
- Received treatment with docetaxel

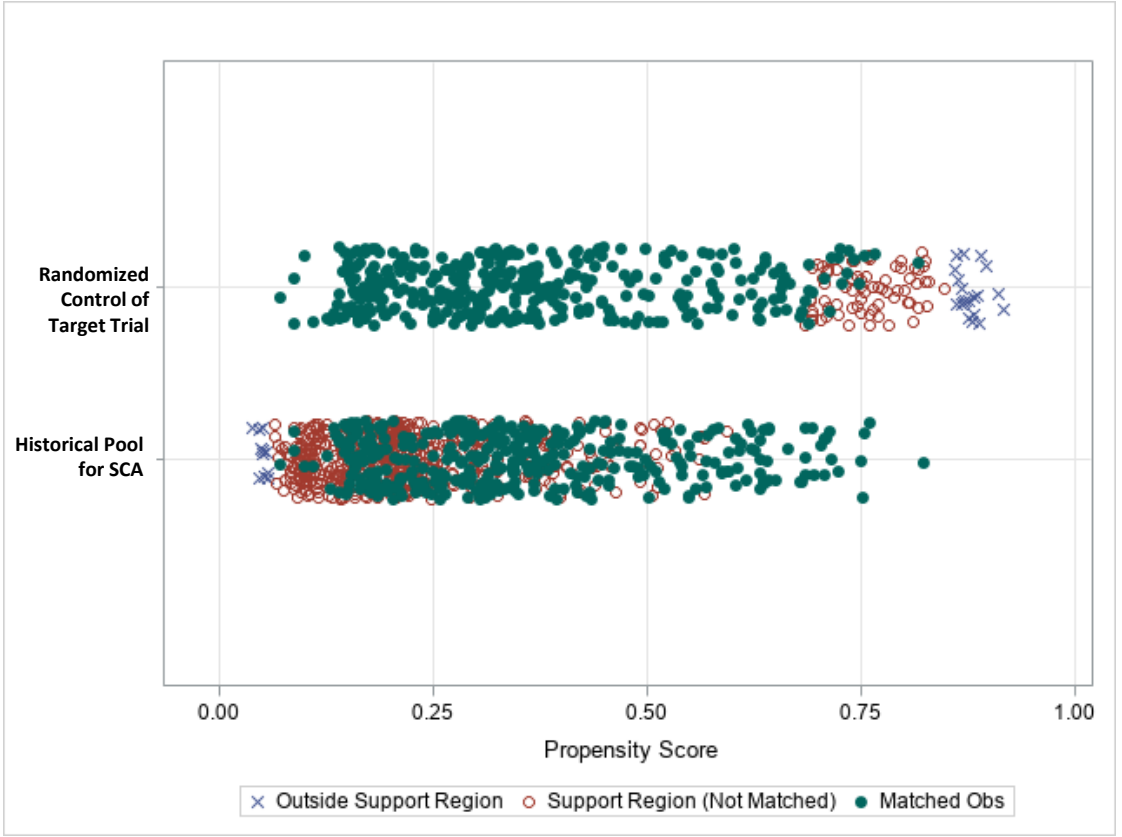
1. These analyses are based on research using information obtained from www.projectdatasphere.org, which is maintained by Project Data Sphere, LLC. Neither Project Data Sphere, LLC nor the owner(s) of any information from the web site have contributed to, approved or are in any way responsible for the contents of this work.
2. Includes thousands of previous clinical trials conducted by the pharmaceutical industry for drug or medical product development with patient level data recorded through the Medidata electronic data capture system. Legal agreements permit use in deidentified (i.e., patients and original sponsor of the trial cannot be identified) and aggregated (i.e., every analysis must include data from two or more sponsors) form.

© 2018 Medidata Solutions, Inc. – Proprietary and Confidential

Prespecified Propensity Score Matching

Baseline Characteristics for Matching

- Age at baseline (continuous)
- Years from cancer diagnosis (continuous)
- Race (White vs Others)
- Sex (Female vs Male)
- Smoking (Current vs Former vs Never)
- Histology (Squamous vs Non-squamous)
- Stage (III vs IV)
- ECOG (0 vs 1 vs 2)
- Prior surgery (Yes/Maybe vs No)
- EGFR/KRAS mutation (Positive vs No/Unknown)



© 2018 Medidata Solutions, Inc. – Proprietary and Confidential

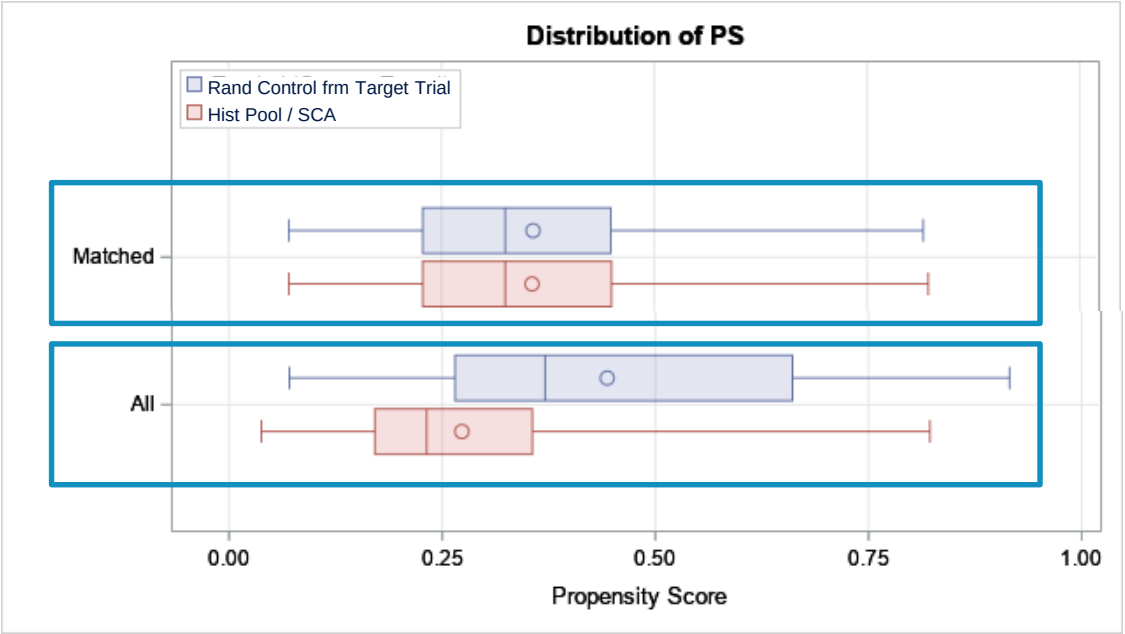
Matching Performance

Baseline Comparability

© 2018 Medidata Solutions, Inc. – Proprietary and Confidential

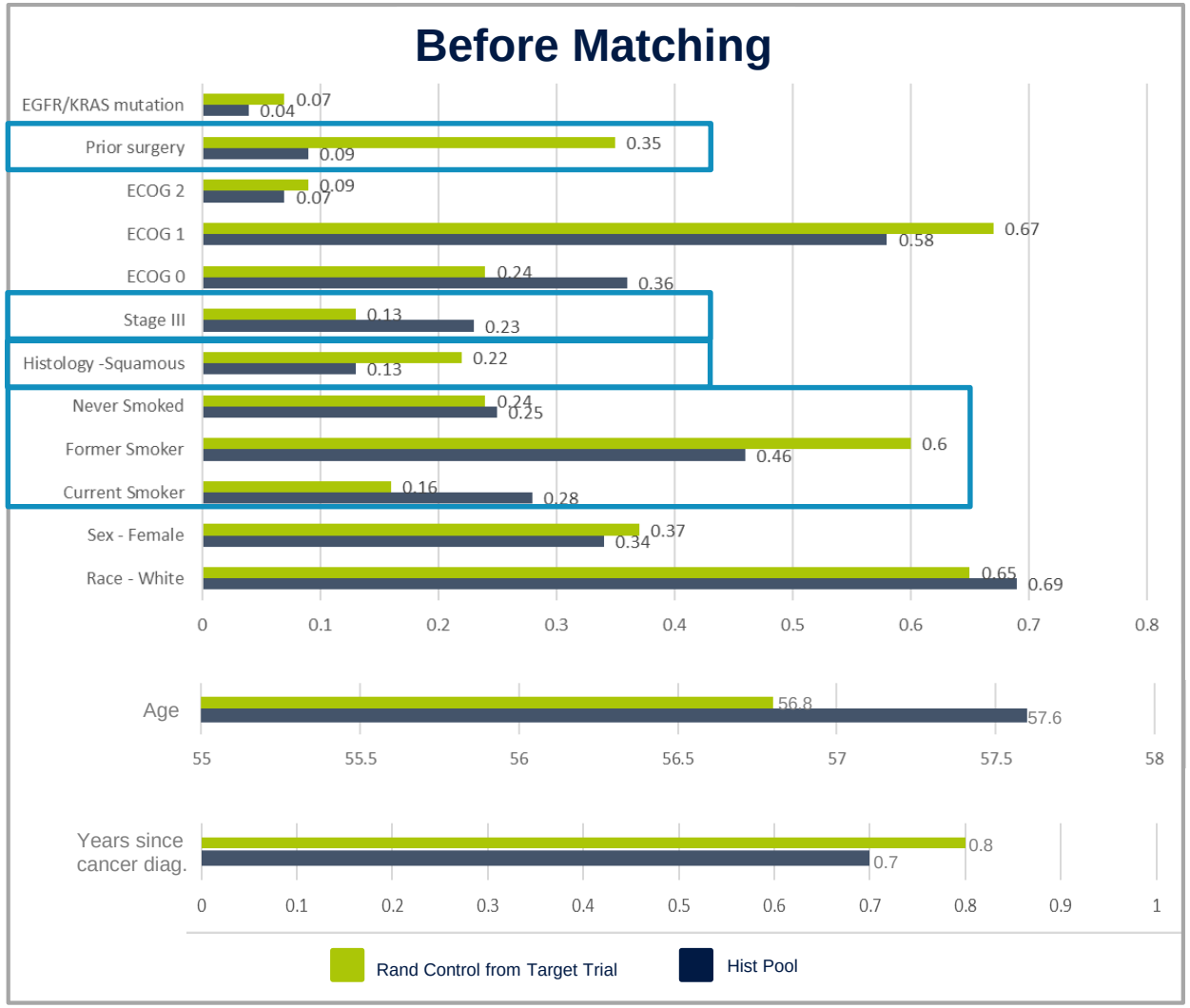
Comparability of SCA & Target Control

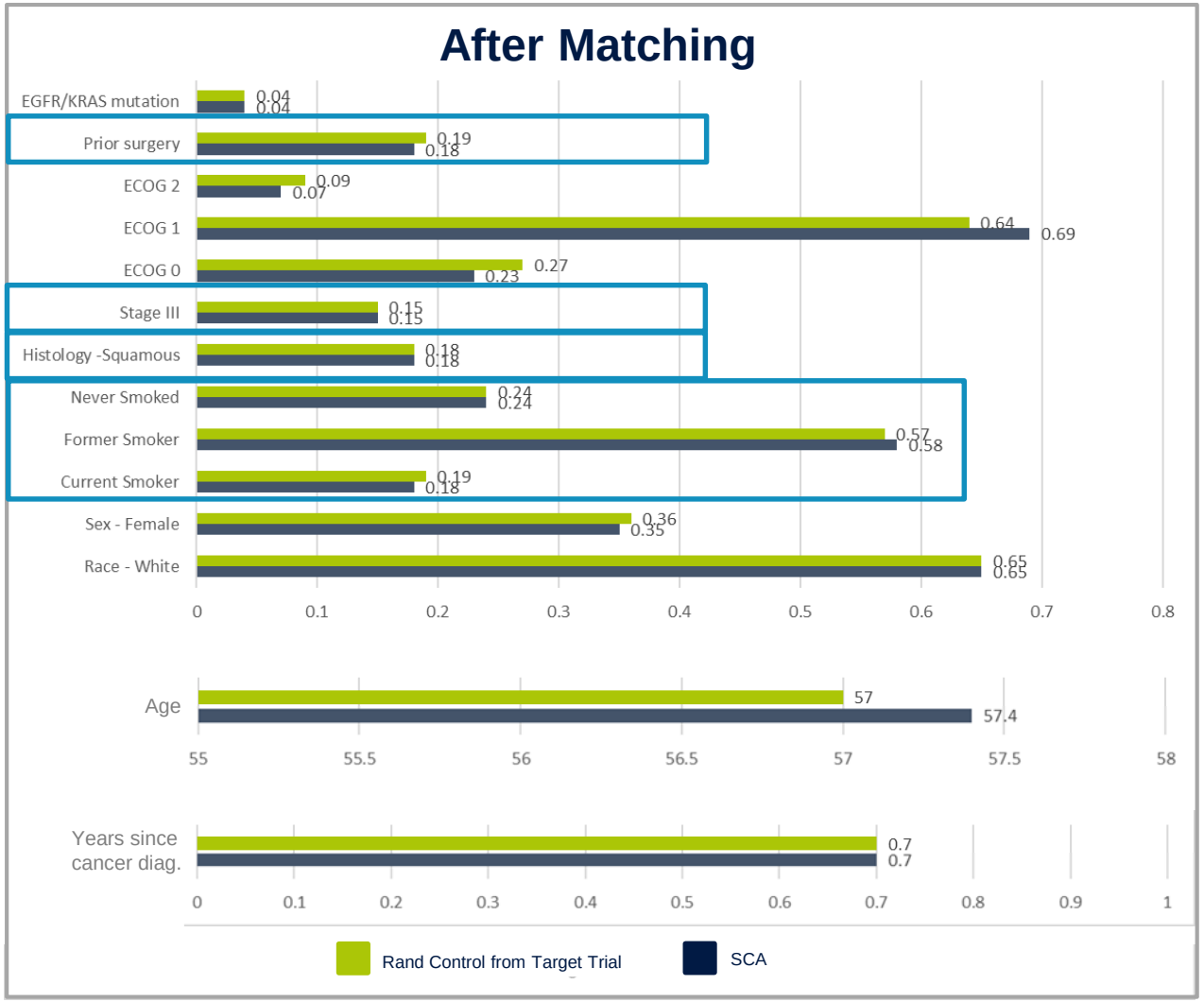
- *Considerable mismatch of propensity scores before matching*
- *After matching, SCA and Randomized Control from Target Trial have similar propensity score distributions*



© 2018 Medidata Solutions, Inc. – Proprietary and Confidential

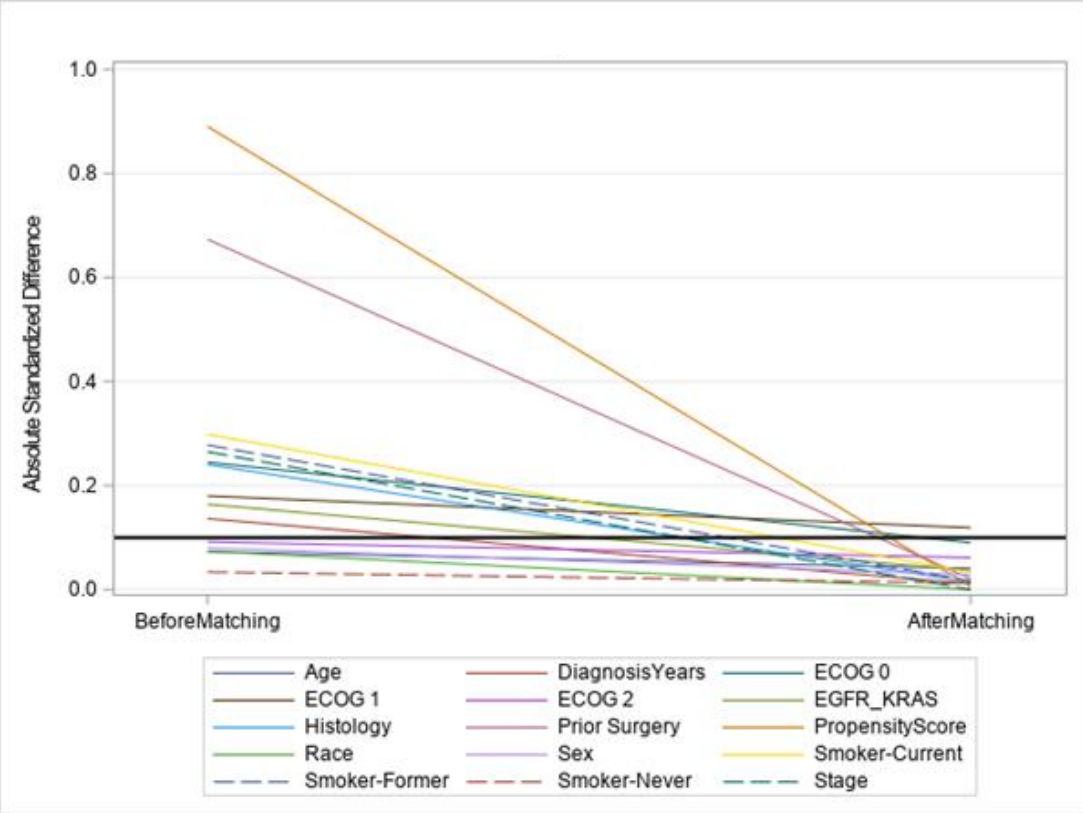
Baseline Characteristics Well Balanced After Matching





© 2018 Medidata Solutions, Inc. – Proprietary and Confidential

Absolute Standardized Differences in Baseline Characteristics are Small/Negligible After Matching



Propensity score matching has achieved good balance of baseline characteristics between the SCA and the randomized control from target trial.

© 2018 Medidata Solutions, Inc. – Proprietary and Confidential

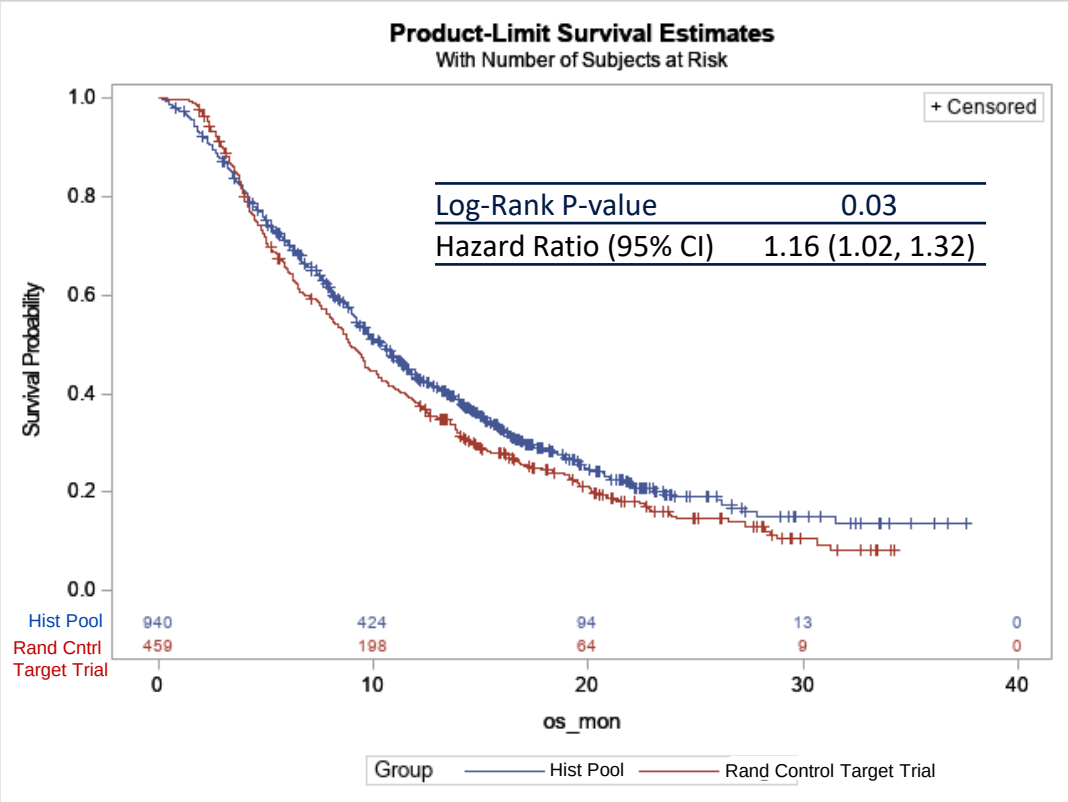
SCA Validation

Overall Survival

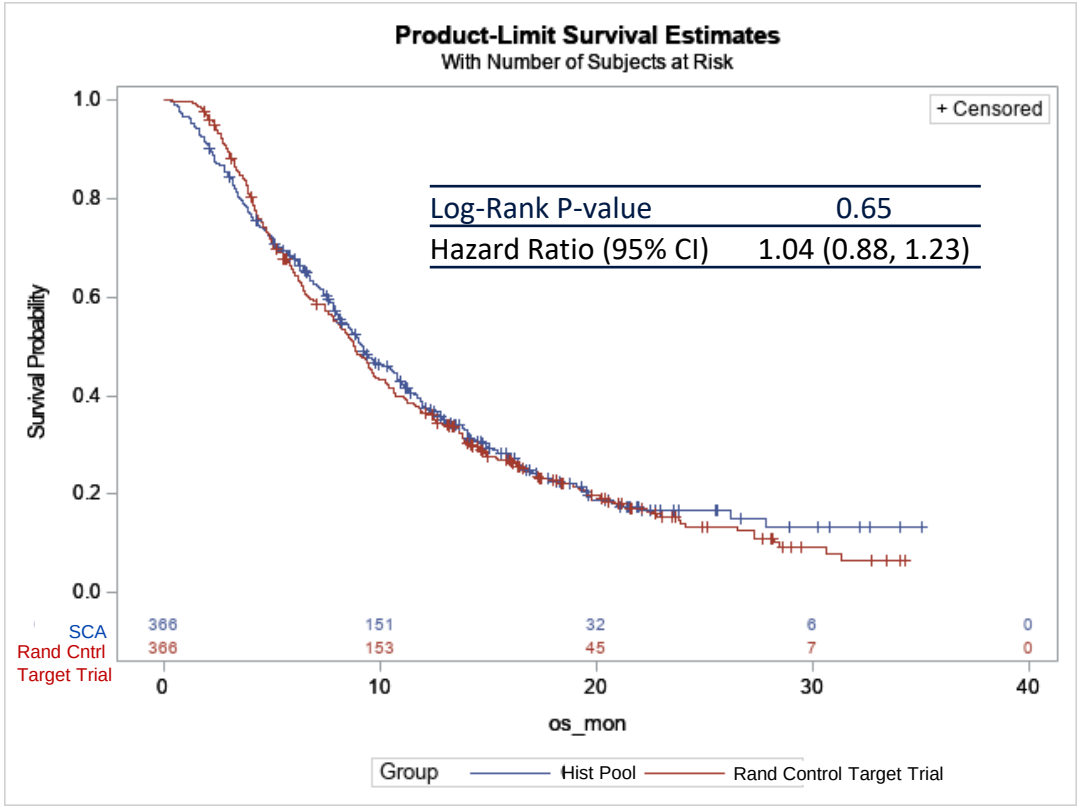
© 2018 Medidata Solutions, Inc. – Proprietary and Confidential

Overall Survival

Before Matching



After Matching



© 2018 Medidata Solutions, Inc. – Proprietary and Confidential

Summary

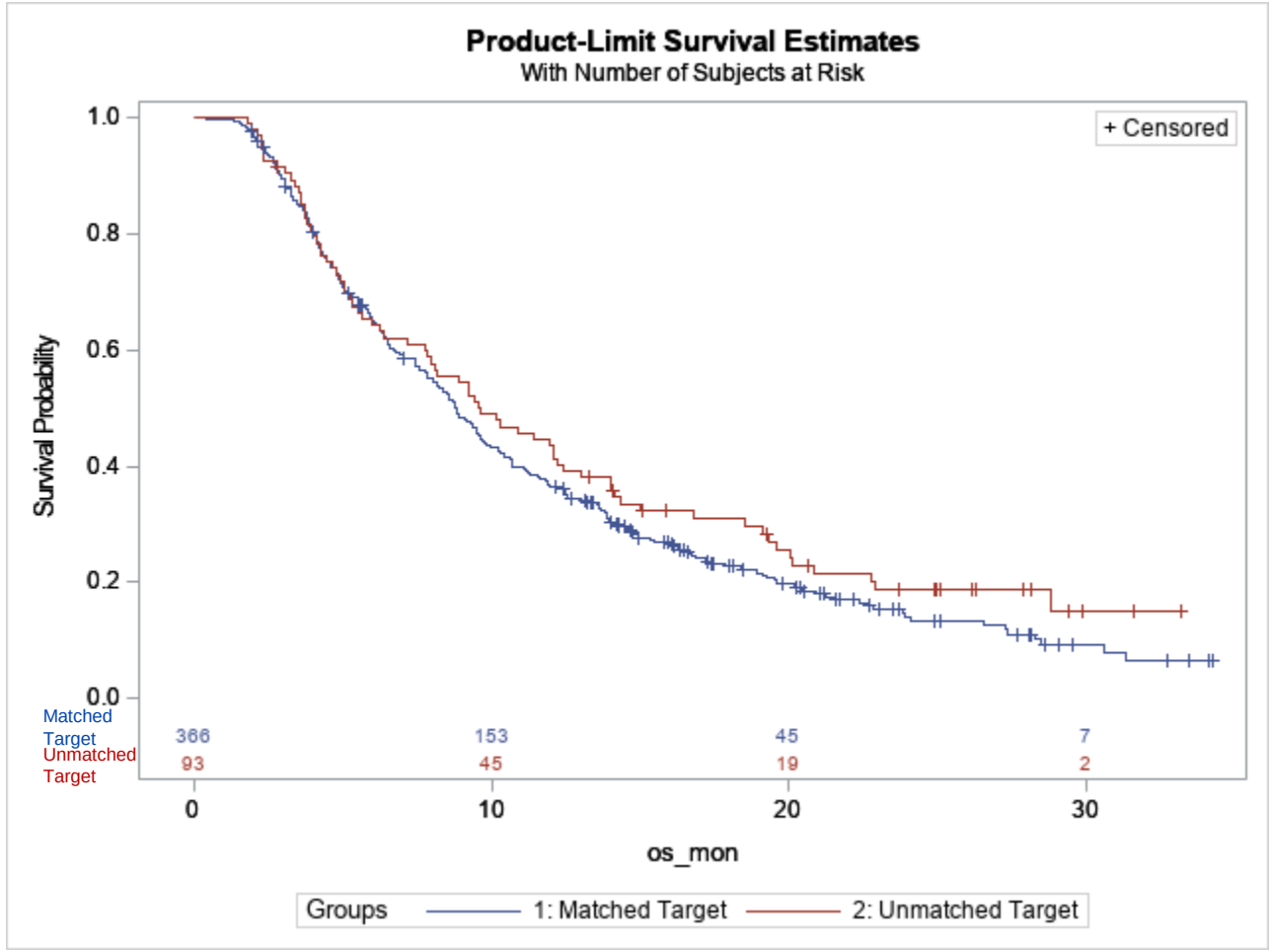
- *In this NSCLC case study, **SCA successfully replicated the control arm** of the target trial*
- *Important step in understanding how **SCA may mitigate challenges faced with a concurrent control arm** in difficult-to-study indications*
- *Future work*
  - Assessment of whether the treatment effect can be replicated with the use of SCA
  - Exploration of tweaks to the matching methods to reduce the proportion of patients who are not matched
  - Expand to additional indications

© 2018 Medidata Solutions, Inc. – Proprietary and Confidential

Thank you.

© 2018 Medidata Solutions, Inc. – Proprietary and Confidential

Overall Survival – Matched versus Unmatched Patients from Randomized Control



© 2018 Medidata Solutions, Inc. – Proprietary and Confidential

Use of External Controls: A Case Study in Melanoma

November 13, 2018

Pallavi S. Mishra-Kalyani
Division of Biometrics V, Office of Biostatistics, CDER, FDA

Disclosures

- I have no financial relationships to disclose
- This presentation reflects the views of the author and should not be construed to represent FDA's views or policies

Introduction

- Randomized clinical trials (RCT) remain the gold standard for the identification of treatment effect
- However, disease or population characteristics may require a non-randomized study design
- In the case that a randomized control arm is not feasible, an external control arm may be an option for estimating comparative treatment effect

Use of External Controls

- External controls could be considered to demonstrate:
  - Natural history of disease
  - Contribution of components
  - Established efficacy from prior trials (e.g. comparison to a single arm trial)
- Source of data for controls would determine potential use
  - Registry
  - Electronic health records (EHR)
  - Prior clinical trials/synthetic control arm

A Case Study in Melanoma

- Trial 1:
  - Design: Vemurafenib (Vem) vs. Dacarbazine
  - Primary endpoints: PFS (IRC assessed) and OS
  - First patient enrolled in 2010
- Trial 2:
  - Design: Cobimetinib (Cobi) + Vem vs. Vem
  - Primary endpoint: PFS (investigator assessed)
  - First patient enrolled in 2013

Could the data from the Vem arm of Trial 1 be used as an external/synthetic control arm for the Cobi+Vem experimental arm of Trial 2 instead of concurrent controls (CC)?

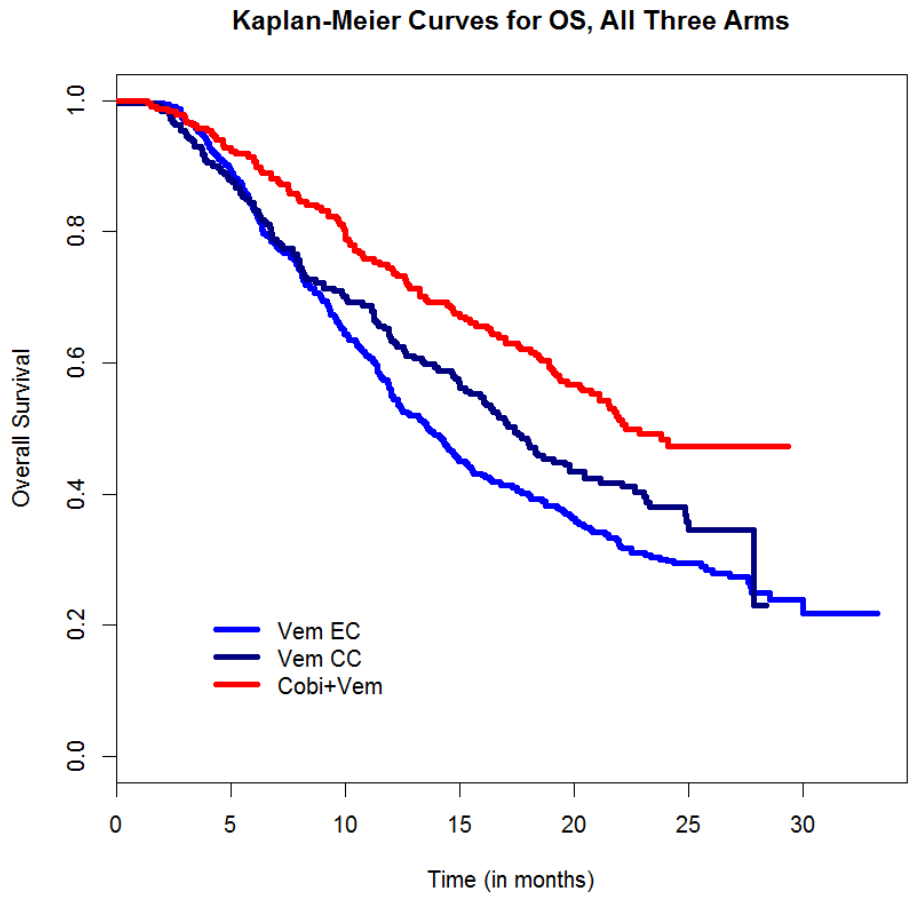
A Case Study in Melanoma

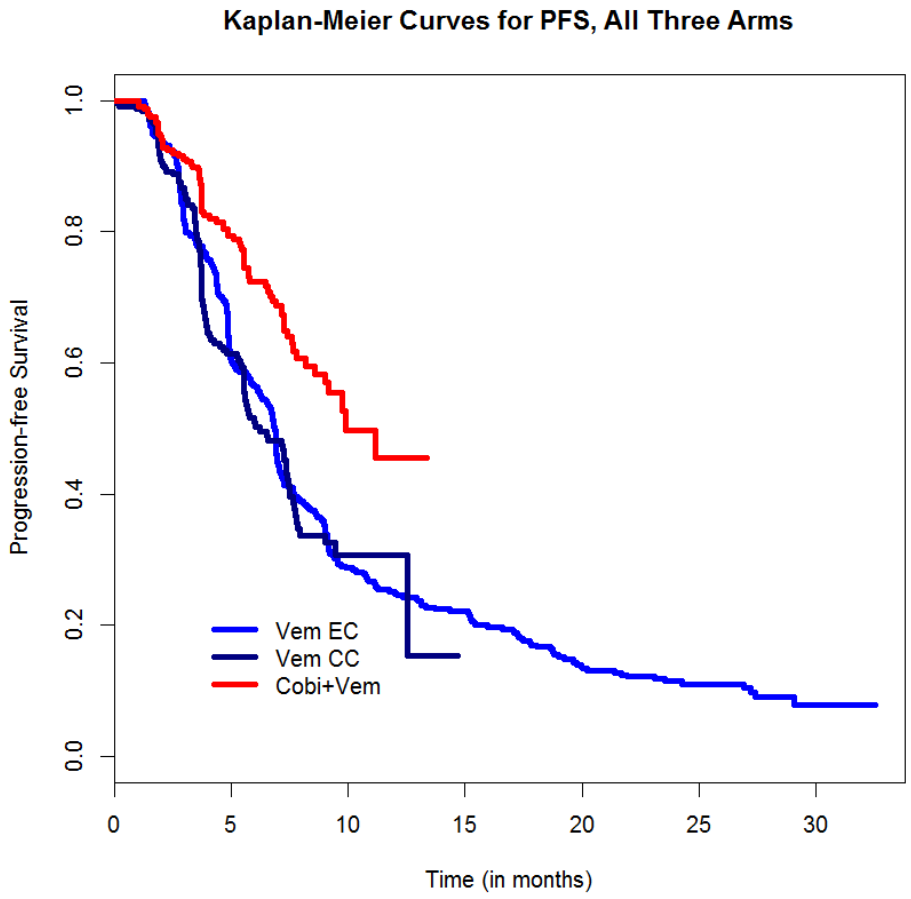
- In order to compare the two trials:
  - Inclusion/exclusion criteria must match
  - Consistency of endpoint measurement (INV vs. IRC, follow-up, etc.)
  - Methods needed to make sure comparative efficacy is not subject to bias
- Propensity Score Matching
  - PS calculated using a logistic regression model for treatment
  - Includes baseline covariates measured in both trials:
    - Age, sex, region, BMI, ECOG, prior adjuvant therapy, disease stage, LDH, BRAF mutation genotype

Baseline Population Characteristics

	Trial 1: Vem (EC) N = 337	Trial 2: Vem (CC) N = 248	Trial 2: Cobi + Vem N = 247
Age	55.2 (13.8)	55.3 (13.8)	54.9 (14.0)
BMI	26.8 (5.4)	27.1 (5.1)	27.7 (5.9)
Sex: Female	200 (59)	140 (56)	146 (59)
Race: White	86 (26)	26 (10)	25 (10)
Region: N. America	86 (26)	26 (10)	25 (10)
Europe	205 (61)	184 (74)	182 (74)
Other	46 (14)	38 (15)	40 (16)
ECOG: 1	108 (32)	80 (32)	58 (23)
Stage: Unresectable IIIC	13 (4)	13 (5)	21 (16)
M1a	40 (12)	20 (16)	40 (16)
M1b	65 (19)	42 (17)	40 (59)
M1c	220 (65)	153 (62)	146 (59)
Prior Adjuvant Therapy	68 (20)	24 (10)	24 (10)
Prior Brain Metastasis	0 (0)	2 (1)	1 (0)
Elevated LDH	142 (42)	104 (42)	112 (45)

OS and PFS Across Trial Arms

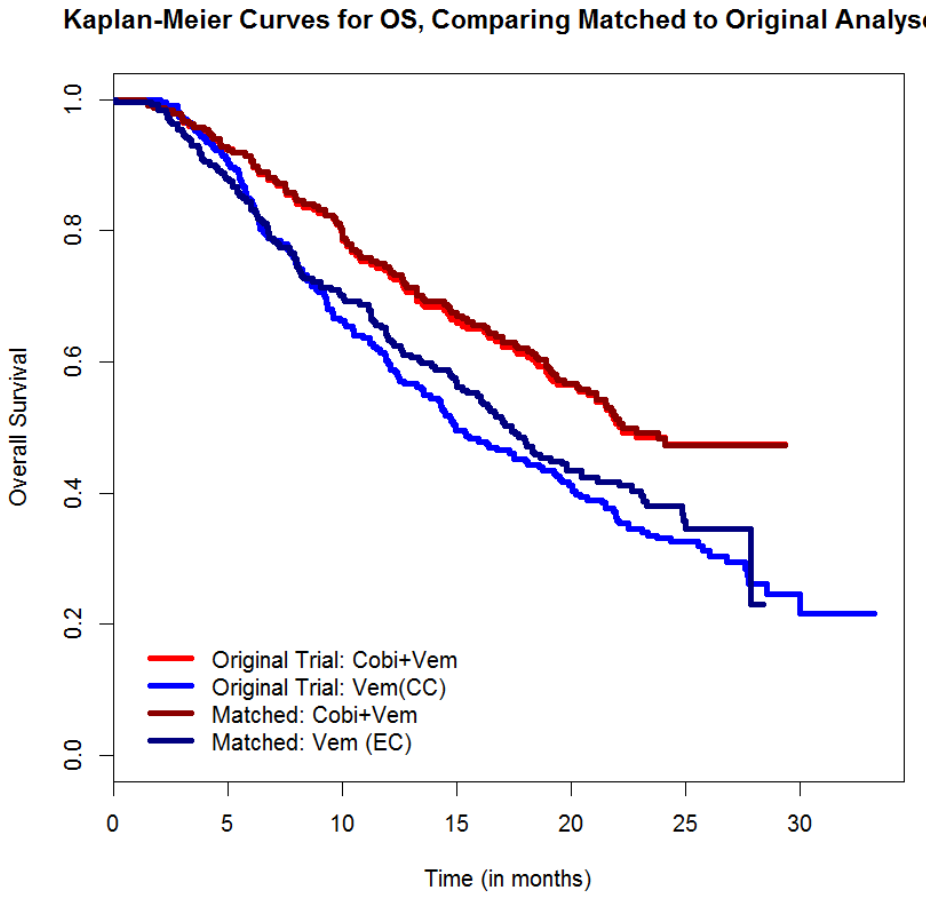


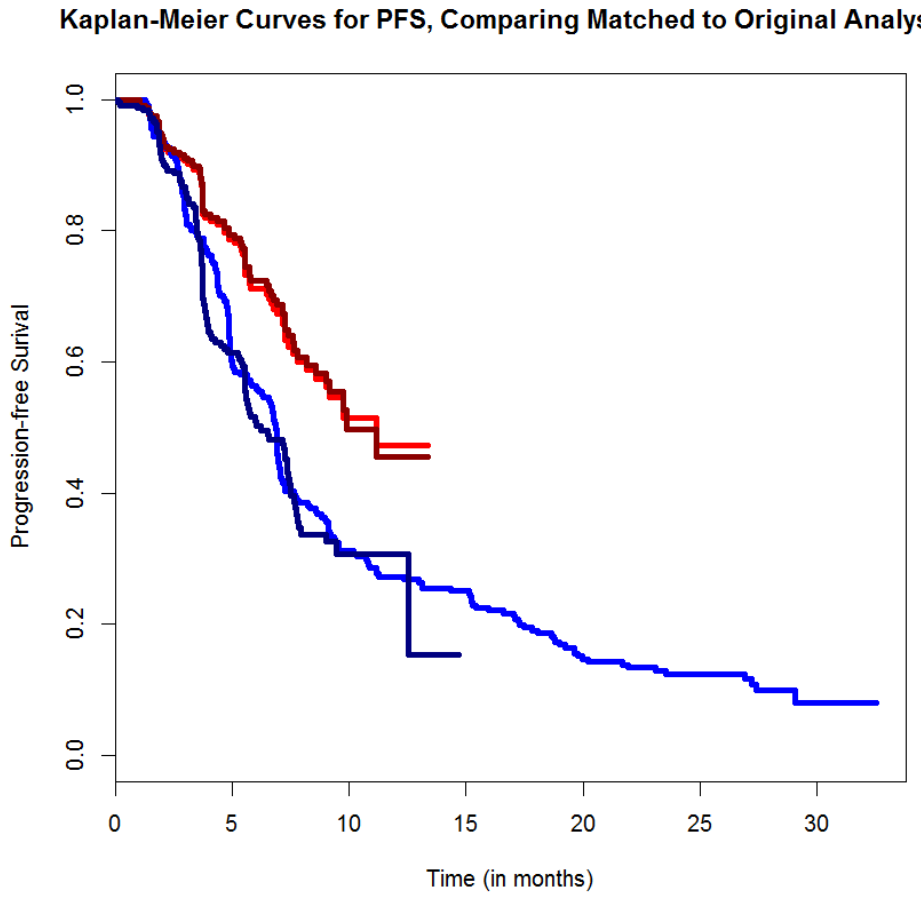


Results

		Original Trial		All External Controls		Matched EC	
		Vem N = 248	Cobi + Vem N = 247	Vem N = 337	Cobi + Vem N = 247	Vem N = 230	Cobi + Vem N = 230
PFS	Medians (95% CI)	6.2 (5.6, 7.5)	9.9 (9.0, NR)	6.9 (6.3, 7.0)	9.9 (9.0, NR)	6.9 (6.6, 7.2)	11.1 (8.6, NR)
	HR (95% CI)	0.51 (0.39, 0.68)		0.53 (0.41, 0.68)		0.57 (0.43, 0.75)	
OS	Medians (95% CI)	17.4 (15.3, 20.5)	22.3 (20.4,NR)	13.6 (12.6, 15.4)	22.3 (20.4,NR)	15.4 (13.6, 19.6)	22.3 (20.3,NR)
	HR (95% CI)	0.70 (0.54, 0.89)		0.57 (0.45, 0.71)		0.65 (0.51, 0.83)	

Comparing Original Trial to Matched Analysis





Discussion and Future Considerations

- When may EC be most appropriate?
- Data considerations when using EC
  - Compatibility of endpoints
  - Temporality of data – even small lags may make a big difference (such as for prior adjuvant therapy)
  - Treatment/disease/patient population idiosyncrasies
- Future work to include
  - Other disease areas
  - Combination of several trial arms for SCA

Acknowledgements

- Susan Jin, FDA
- Raji Sridhara, FDA
- Gideon Blumenthal, FDA
- Marc Theoret, FDA

APPENDIX

Results: Sensitivity to Caliper

		Caliper = 1		Caliper = 0.025		Caliper = 0.01	
		Vem N = 230	Cobi + Vem N = 230	Vem N = 184	Cobi + Vem N = 184	Vem N = 153	Cobi + Vem N = 153
PFS	Medians (95% CI)	6.9 (6.6, 7.2)	11.1 (8.6, NR)	6.9 (6.6, 7.2)	NR (8.6, NR)	6.9 (6.6, 7.9)	NR (11, NR)
	HR (95% CI)	0.56 (0.43, 0.74)		0.53 (0.39, 0.72)		0.51 (0.36, 0.72)	
OS	Medians (95% CI)	15.4 (13.6, 19.6)	22.3 (20.3,NR)	15.4 (13.5, 19.5)	22.1 (19.2, NR)	16.0 (13.4, 19.9)	22.3 (20.3,NR)
	HR (95% CI)	0.65 (0.51, 0.83)		0.64 (0.49, 0.85)		0.64 (0.47, 0.86)	

Panel 1 Participants

Moderator: Elizabeth Stuart, Johns Hopkins University

- Andrea Ferris, LUNGevity Foundation
- Antoine Yver, Daiichi Sankyo
- Joohee Sul, U.S. FDA
- Pallavi Mishra-Kalyani, U.S. FDA
- Rajeshwari Sridhara, U.S. FDA
- Ruthie Davi, Medidata Solutions