

## **Exploring estimation approaches for principal stratum estimands in CAR-T Phase III randomized trials**

Aiesha Zia

2020 Basel Biometrics Section webinar

September 7<sup>th</sup>, 2020

# Acknowledgement

Aby Buchbinder

Amy Racine

Antonella Maniero

Bjoern Bornkamp

Ekkehard Glimm

Emmanuel Zuber

Eric Bleickardt

Evgeny Degtyarev

Frank Bretz

**Jessie Gu**

Kapildeb Sen

Oezlem Anak

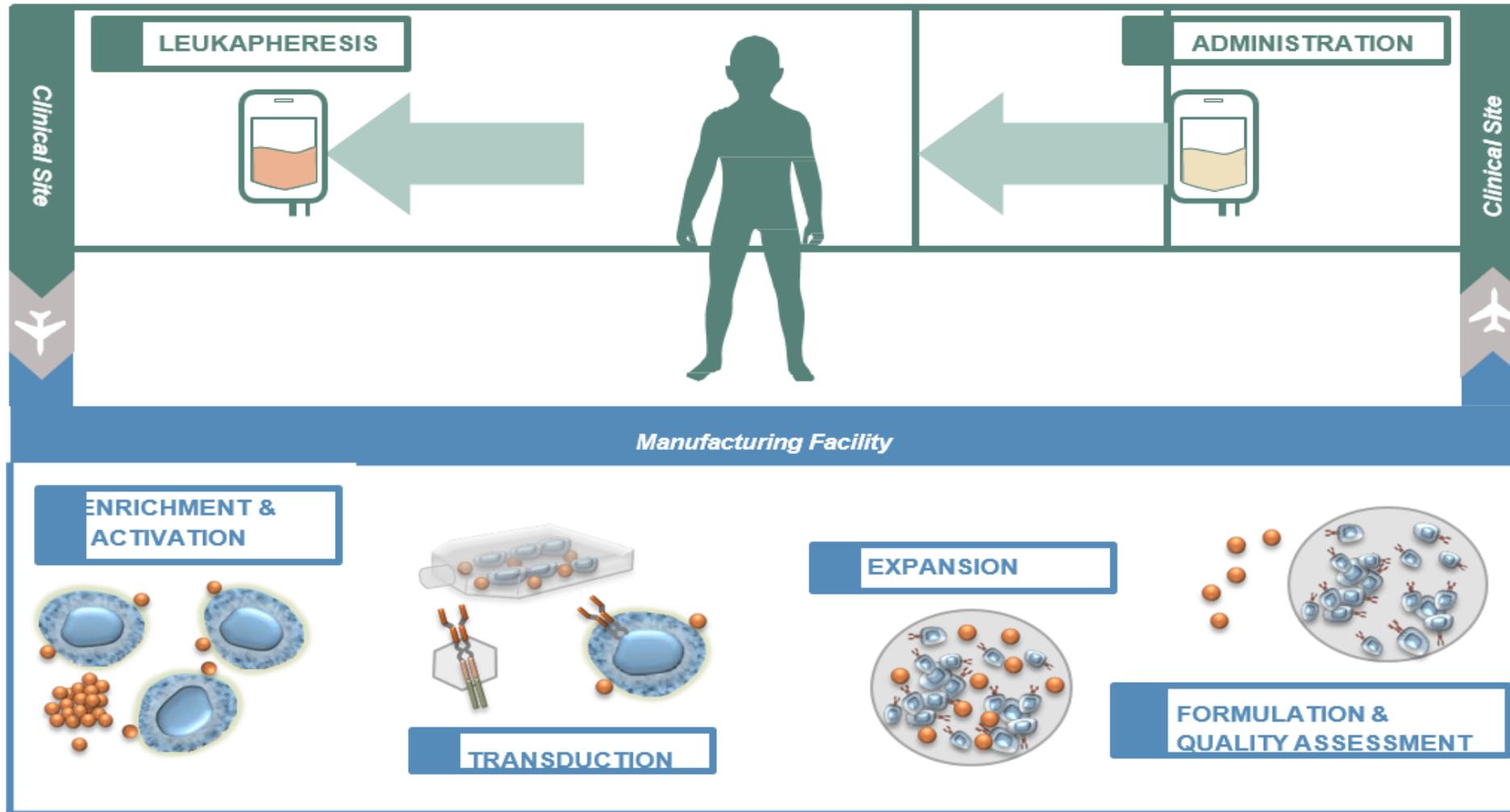
**Yiyun (Michael) Zhang**

Yuanbo Song

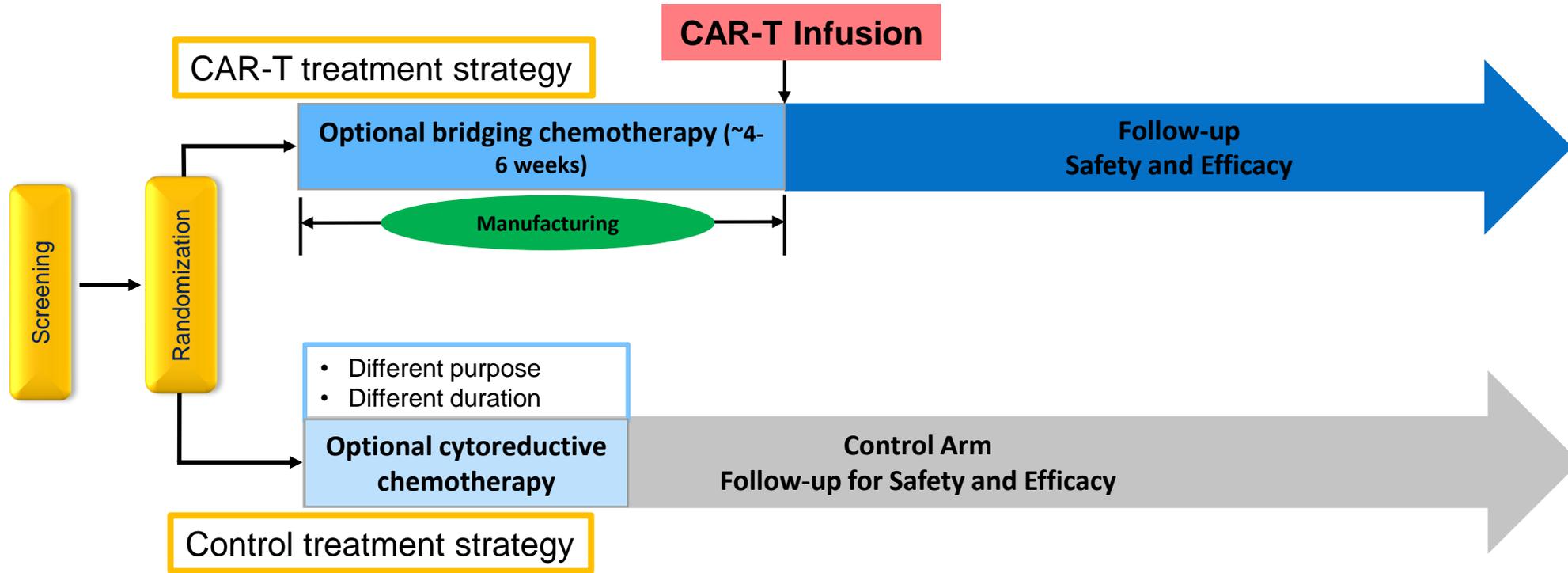
# Agenda

- The CAR-T manufacturing process
- CAR-T randomized phase III trial:
  - The scientific question of interest, the primary estimand and study design
- FDA request
- Principal stratum estimand
  - How to implement the principal strategy estimand
  - Different approaches
- Simulation results
- Summary & conclusion

# CAR-T manufacturing process



# Study design: randomized phase III trial



## Scientific question of interest (Primary objective):

Efficacy (Overall survival) of the entire sequence of interventions is the most relevant question of interest from a patient's perspective.

# FDA request for information

## FDA Comment

Subjects in the CAR-T arm may receive extensive bridging chemotherapy while awaiting CAR-T manufacture, and some, especially those experiencing extended delays in product manufacture, could achieve a CR/CRi [...] status in response to aggressive bridging chemotherapy even before initiation of CAR-T treatment. Since **these responses cannot be directly attributed to CAR-T treatment, the statistical assessment plan should prospectively create rules for appropriately censoring CR [...] subjects from secondary endpoints [...]**.

## *FDA' suggestion*

Censor patients who are responding to bridging chemotherapy in CAR-T arm and comparison with the complete control arm.

## Concerns:

- Censoring in CAR-T arm only: response status unknown in control arm;
- Targeting hypothetical scenario in which no patient would respond to bridging chemotherapy in CAR-T arm, which is unlikely;
- Timing of censoring: relatively close to the randomization, before CAR-T infusion. Similar as a “naive” comparison based on grouping patients on response status after bridging chemotherapy and comparing to complete control arm.

# Principal stratum estimand

## *Principal stratum strategies*

This relates to the population of interest (see A.3.3.). The target population might be taken to be the “principal stratum” (see Glossary) in which an intercurrent event would occur. Alternatively, the target population might be taken to be the principal stratum in which an intercurrent event would not occur. The clinical question of interest relates to the treatment effect only within the principal stratum. For example, it might be desired to know a treatment effect on severity of infections in the principal stratum of patients becoming infected after vaccination. Alternatively, a toxicity might prevent some patients from continuing the test treatment, but it would be desired to know the treatment effect among patients who are able to tolerate the test treatment.

It is important to distinguish “principal stratification” (see Glossary), which is based on potential intercurrent events (for example, subjects who would discontinue therapy if assigned to the test product), from subsetting based on actual intercurrent events (subjects who discontinue therapy on their assigned treatment). The subset of subjects who experience an intercurrent event on the test treatment will often be a different subset from those who experience the same intercurrent event on control. Treatment effects defined by comparing outcomes in these subsets confound the effects of the different treatments with the differences in outcomes possibly due to the differing characteristics of the subjects.

# How to implement the principal stratum strategy

**The Scientific question:** What is the **long term efficacy (EFS)** of the **CAR-T treatment strategy** relative to **control treatment strategy** in patients who would not respond to bridging chemotherapy if they were given the bridging chemotherapy?

Denote:

- $S(T)$  as potential EFS outcome for treatment  $T$  ( $T = 0$ : control;  $T = 1$ : CAR-T)
- $R(1)$  as potential outcome for patients who would not respond to bridging chemotherapy in CAR-T arm ( $R = 0$ : non-responder;  $R = 1$ : responder).

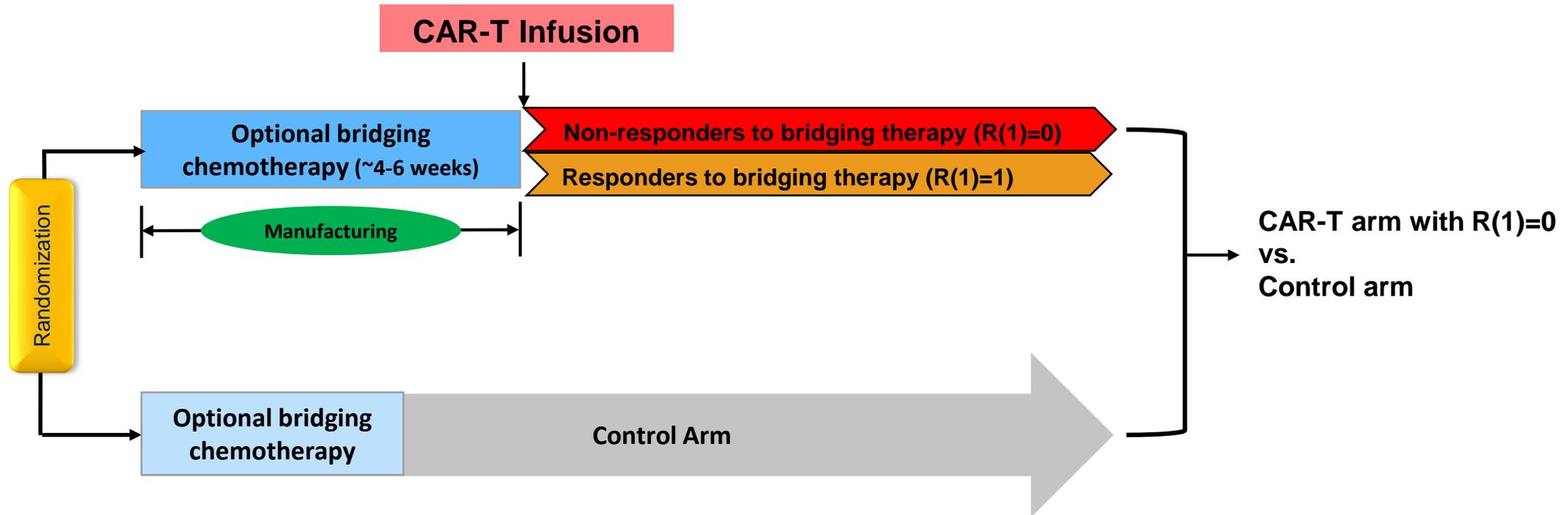
Interest in contrasting the distribution of:

- EFS for stratum  $R(1)=0$ : patients who still had measurable disease prior to infusion (did not respond to bridging chemotherapy):  $\{S(T=1)|R(1)=0\}$  vs.  $\{S(T=0)|R(1)=0\}$ ;

with hazard ratio as the effect measure

# 1 - A naive comparison...

## Only valid under strong assumptions

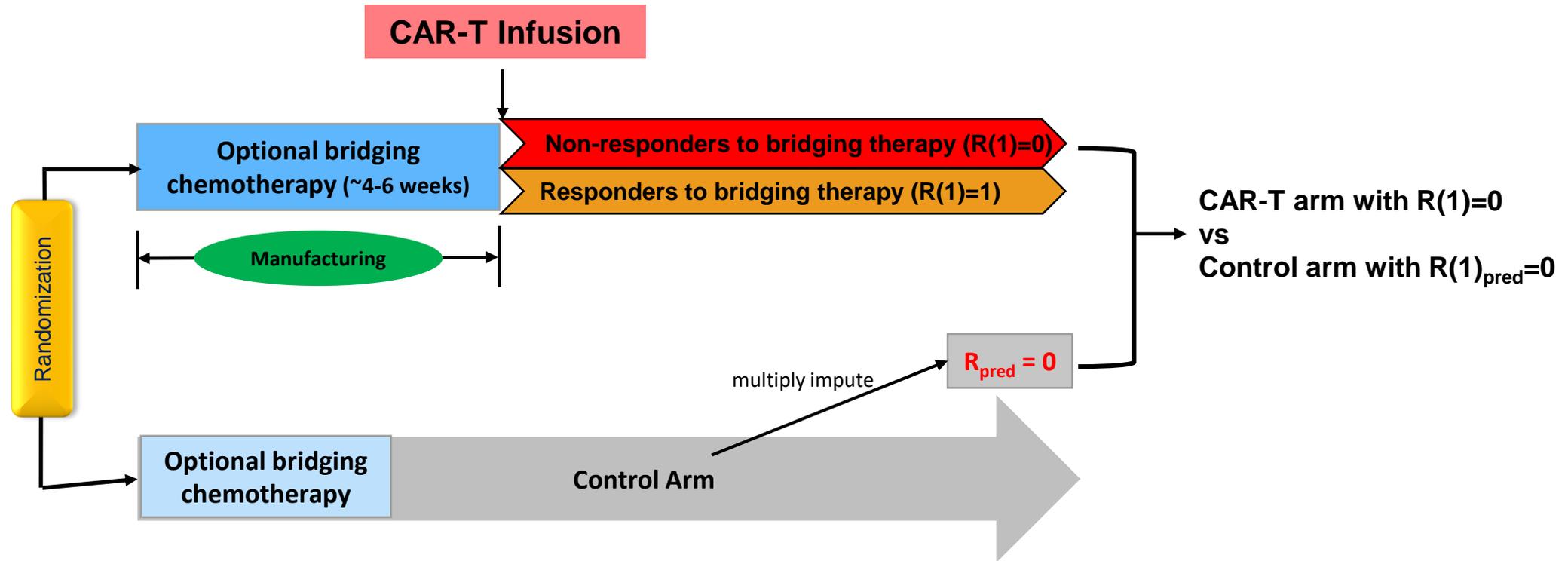


- R(1) is a post-baseline event → Patient population with R(1)=0 on CAR-T arm might systematically differ from the control population. Non-randomized comparison (not comparing „like with like“).
- The naive comparisons are only valid (i.e. give unbiased estimate of treatment effect) if the following assumption is true: **S(T=0) and R(1) are independent**: all patients in control arm share the same EFS distribution regardless of their response to bridging chemotherapy if they were given bridging chemotherapy.

# Use baseline characteristics to identify a matching control group

- Basic idea
  - For patients who did not respond to bridging therapy on CAR-T arm, i.e.  $R(1)=0$  try to find the matching comparator group on control (i.e. subgroup of control patients) to compare „like with like“.
- How? Identify confounder variables
  - Affect EFS (on control) and remission status (on CAR-T arm).  
(formal assumption: conditional independence given confounders)
  - Choose patients on the control arm so that baseline characteristics are comparable to group of patients on CAR-T with  $R(1)=0$ .
- No unmeasured confounders assumption
  - Often also used in observational studies (where we are interested in confounders predicting both outcome and treatment assignment)
  - Note: Unverifiable assumption; more plausible than assuming  $R(1)$  does not affect the potential EFS outcome on control at all.
- Once confounders are decided upon, different statistical analyses are possible to achieve balance between groups.

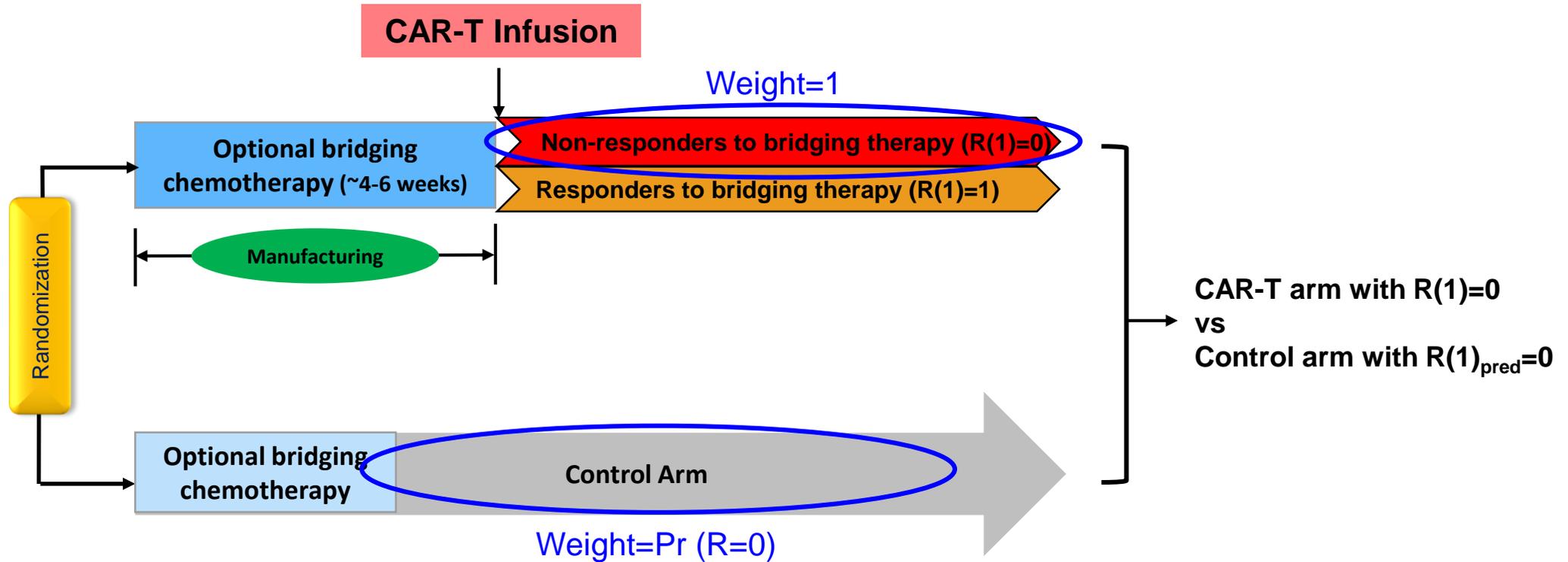
## 2 - Multiple imputation approach



### Predicting a patient's response to bridging chemotherapy for those in the control arm:

- 1) Fit logistic regression model for  $R(1)$  on CAR-T arm using confounders as predictors:  $R(1) \sim \text{covariates}$
- 2) Multiply impute  $R(1)_{\text{pred}}$  for every patient on the control arm using the fitted model in step 1)
- 3) For each imputed complete data-set perform main analysis: **Cox regression model** to estimate HR of CAR-T arm ( $R(1)=0$ ) + control arm ( $R(1)_{\text{pred}}=0$ )
- 4) At end combine results across imputations (e.g. Rubin's rules)

# 3 – Weighting approach



Probability of being a non-responder to bridging chemotherapy will be used as weight in the weighted cox regression model.

- 1) Fit logistic regression model for  $R(1)$  on CAR-T arm using confounders as predictors:  $R(1) \sim \text{covariates}$
- 2) Apply the fitted model to control arm and predict for all patients  $\text{Pr}(R(1)=0)$
- 3) Perform the analysis: weighted Cox regression models to estimate HR of CAR-T arm ( $R(1)=0$ ,  $\text{weight}=1$ ) + all control arm ( $\text{weight}=\text{Pr}(R(1)=0)$ )

# Simulation setting

- Simulated dataset:

- Assumption on hazard ratio of CAR-T arm vs. Control arm in patients who did not respond to bridging chemotherapy of 0.5 based on previous trial
- Simulate CAR-T patients' R status using baseline tumor burden as predictor: Larger tumor burden → Less likely to be responder; shorter EFS time
- Simulate EFS outcome based on treatment group and tumor burden
- **R(1) and S(T=0) are both generated as a function of X, they are conditionally independent.**

- Simulations:

- Different scenarios varying the predictiveness of the logistic regression model: how well tumor burden predicts R(1)

# Simulation results

	Model's predictability (AUC under ROC curve*)	Bias in the estimated HR for patients not responding to bridging chemotherapy		
		Naive	MI	Weighted
<i>Sim 1</i>	<i>0.60</i>	0.029 (SE=0.170)	<0.001 (SE=0.205)	0.001 (SE=0.184)
<i>Sim 2</i>	<i>0.72</i>	0.065 (SE=0.169)	<0.001 (SE=0.208)	0.001 (SE=0.184)
<i>Sim 3</i>	<i>0.90</i>	0.138 (SE=0.170)	-0.001 (SE=0.213)	<0.001 (SE=0.189)

Results averaged across 500 simulations

\* Higher the AUC, better the model is at predicting

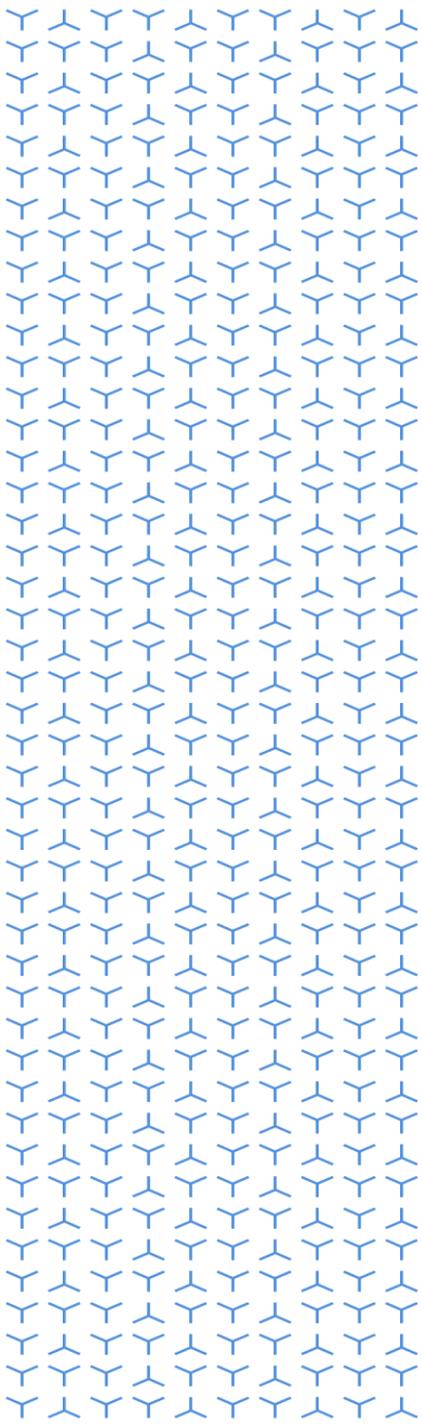
- Naive approach:
  - Underestimates the treatment effect
  - Bias decreases as model becomes less predictive → Less dependent between R(1) and S(T=0)
- MI/Weighted approach: Small bias regardless of model's predictability
  - If conditional independence holds -> good prediction of R(1) not required

# Summary & conclusion

- CAR-T treatment not readily available at randomization
  - Primary estimand based on treatment strategies: CAR-T arm versus Control arm, regardless of a patient's response to bridging chemotherapy
  - FDA request:
    - Principal stratum estimand !?
    - Hypothetical strategy (hypothetical scenario, where no one would be in remission), discouraged in the ICH E9(R1)
- Naive comparison to address principal stratum estimand provides biased estimates
- Other approaches (e.g. MI, weighted) exist to obtain unbiased estimates under assumptions ( $S(T=0)$  and  $R(1)$  are independent, conditional on  $X$ )

# Summary & conclusion

- Use of principal stratum strategy
  - Probably not very often for primary analyses (due to assumptions)
    - e.g. could run a different study design if scientific question is of main interest
  - Very valuable for important secondary or exploratory questions
- In this case FDA agreed to use the principal stratum strategy as supportive analysis
  - But asked for details and justification on the assumptions *before start of the trial*



**Thank you**