MULTI-REGIONAL CLINICAL TRIALS

THE MRCT CENTER OF BRIGHAM AND WOMEN'S HOSPITAL and HARVARD

AI Digital Twins and Synthetic Data: Practical Use Cases for Clinical Research

DATE
November 18, 2025

TIME
12–1pm ET

Disclaimer

The views expressed today do not imply endorsement or reflect the views or policies of Mass General Brigham, Harvard Medical School, or any affiliated organization or entity.

The MRCT Center is supported by voluntary contributions from foundations, corporations, international organizations, academic institutions, and government entities (see www.MRCTCenter.org), as well as by grants.

We are committed to autonomy in our research and to transparency in our relationships. The MRCT Center and its directors retain responsibility and final control of the content of any products, results, and deliverables.

Welcome~

Thank you for joining this webinar today!

Tips for today's session:

- Please use the Q&A for questions. We will do our best to answer live.
- Feel free to use the Closed Captioning available on the Zoom toolbar.
- Most of the links in our presentations will be shared in the Chat.

This meeting is being recorded.

The recording, slides, and any additional materials will be available next week. If you registered, you will receive an email about their availability and a notification about future webinars in this series.

The Multi-Regional Clinical Trials Center (MRCT Center)

The MRCT Center is a research and policy center focused on addressing the conduct, oversight, ethics and regulatory environment for clinical trials.

Our Vision

Improve the integrity, safety, and rigor of global clinical trials.

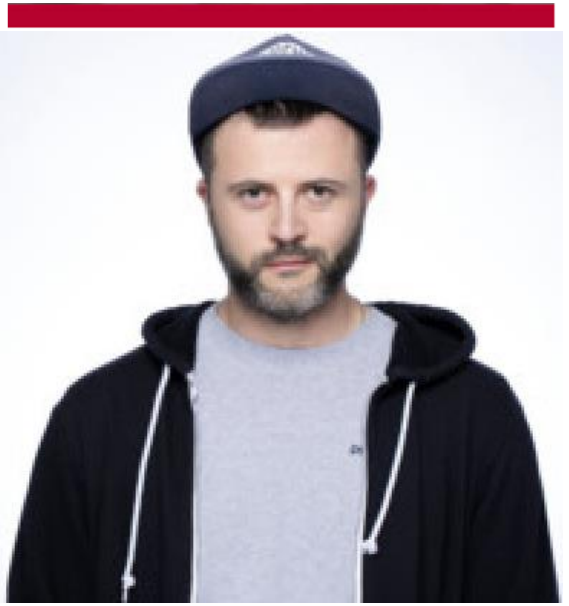
Our Mission

Engage diverse stakeholders to define emerging issues in global clinical trials and to create and implement ethical, actionable, and practical solutions.

AI Digital Twins and Synthetic Data: Practical Use Cases for Clinical Research

MULTI-REGIONAL CLINICAL TRIALS
THE MRCT CENTER OF BRIGHAM AND WOMEN'S HOSPITAL and HARVARD

Barbara E. Bierer, MD
Professor of Medicine, HMS
Faculty Director,
MRCT Center of BWH & Harvard
bbierer@bwh.harvard.edu



Daniele Bertolini
Machine Learning Scientist

UNLEARN

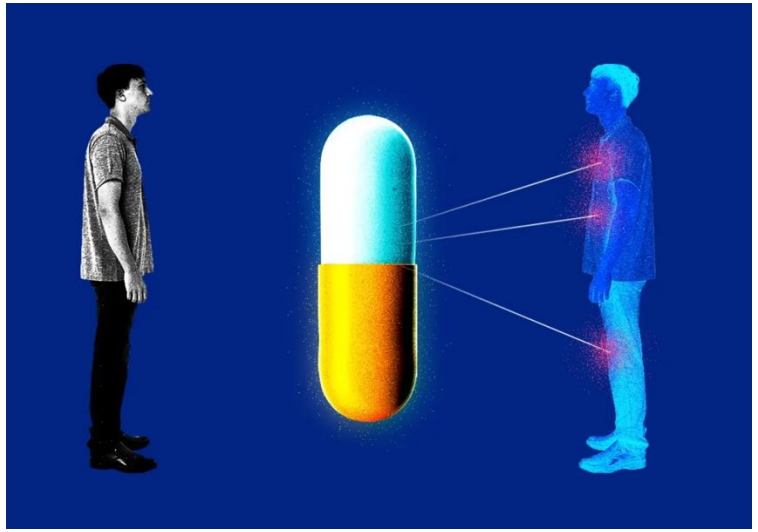
Digital Twins and Synthetic Data

Digital twins refer to simulated models of individual patients designed to predict disease trajectories and/or treatment responses, offering the potential to enhance statistical power, optimize trial design, and limit the number of participants assigned to a control arm.

Synthetic data are artificially generated datasets that mirror the statistical properties of verified clinical data.

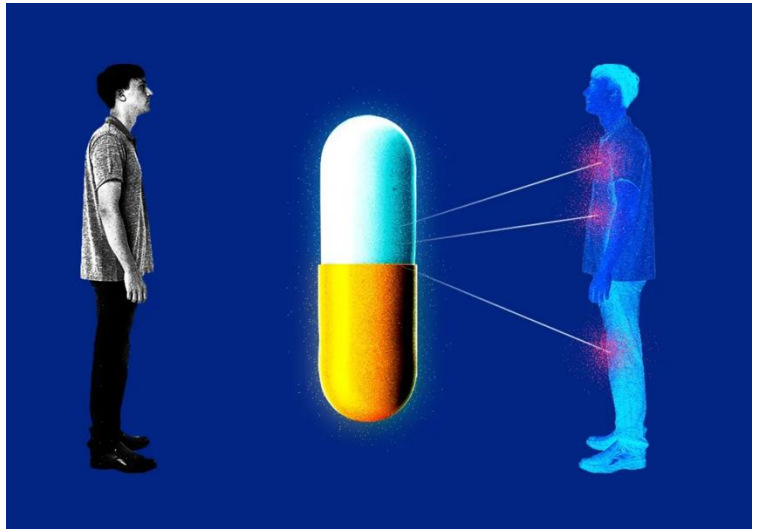
Both technologies are being developed for use in the design, conduct, and analysis of clinical trials, and with the intention of use for regulatory submissions.

Appreciating innovation



https://www.healthcare-brew.com/stories/2024/12/12/how-digital-twins-could-change-clinical-trials

Appreciating innovation



https://www.healthcare-brew.com/stories/2024/12/12/how-digital-twins-could-change-clinical-trials

https://mrctcenter.org/wp-content/uploads/2025/10/2025-09-30_AI-slides.pdf

Appreciating innovation

01	Digital Twins and Why They Are Useful
02	How to Build Digital Twins
03	How to Use Digital Twins

https://mrctcenter.org/wp-content/uploads/2025/10/2025-09-30_AI-slides.pdf

Appreciating innovation

01	Digital Twins and Why They Are Useful
02	How to Build Digital Twins
03	How to Use Digital Twins

https://mrctcenter.org/wp-content/uploads/2025/10/2025-09-30_AI-slides.pdf

View Webinar

Download Slides

UNLEARN + MULTI-REGIONAL CLINICAL TRIALS
THE MRCT CENTER of BRIGHAM AND WOMEN'S HOSPITAL and HARVARD

November 18, 2025

Digital Twins and Synthetic Data: Applications to Clinical Trials

Daniele Bertolini
Unlearn.AI

Agenda

01	Digital Twins Recap
02	PROCOVA and Bayesian PROCOVA
03	Digital Twins as Synthetic Controls
04	Subgroup Discovery
05	Design Optimization and in-silico Trials

Digital Twin of a patient

A Digital Twin of a patient is comprehensive computational model of that individual which makes predictions of future outcomes and can be dynamically updated with data from their physical twin.

It is typically built from multimodal patient-level and population-level data:

- Demographic
- Disease history
- Biomarkers
- Lab tests
- Genetic information
- Clinical assessments
- Images, etc

It can be used to:

→ Forecast disease progression and improve efficiency of RCTs

→ Evaluate interventions, identify long term health impacts, anticipate risks for significant health events, etc.

Predicted longitudinal outcomes

Time since baseline visit

Participant’s baseline data

Endpoints

A participant's digital twin is a comprehensive forecast of their predicted placebo outcomes generated from their baseline data.

Participant’s Digital Twin in ALS

Age (years): 55
Sex: Male
Race: Caucasian

Diagnosis: ALS F
Site of Onset: Bul
Symtom Onset: 6

Time (months)	Baseline	1	2
			
Alanine Aminotransferase	41	41.8 ± 12.3	42.3 ± 13.9
Albumin	42	42.2 ± 1.6	42.1 ± 2.1
Alkaline Phosphatase	76	77.8 ± 11.0	77.6 ± 13.7
ALSFRS Climbing	4	3.9 ± 0.4	3.7 ± 0.5
ALSFRS Cutting	4	3.9 ± 0.2	3.9 ± 0.4
ALSFRS Dyspnea	3	2.9 ± 0.7	2.9 ± 0.8
ALSFRS Handwriting	4	3.9 ± 0.3	3.8 ± 0.4
ALSFRS Hygiene	4	3.9 ± 0.3	3.8 ± 0.5
ALSFRS Insufficiency	4	4.0 ± 0.1	3.9 ± 0.2
ALSFRS Orthopnea	4	4.0 ± 0.2	3.9 ± 0.3
ALSFRS Salivation	3	2.8 ± 0.7	2.8 ± 0.8
ALSFRS Speech	2	1.9 ± 0.6	1.9 ± 0.8
ALSFRS Swallowing	3	2.9 ± 0.6	2.8 ± 0.8

Digital Twins and Synthetic Data

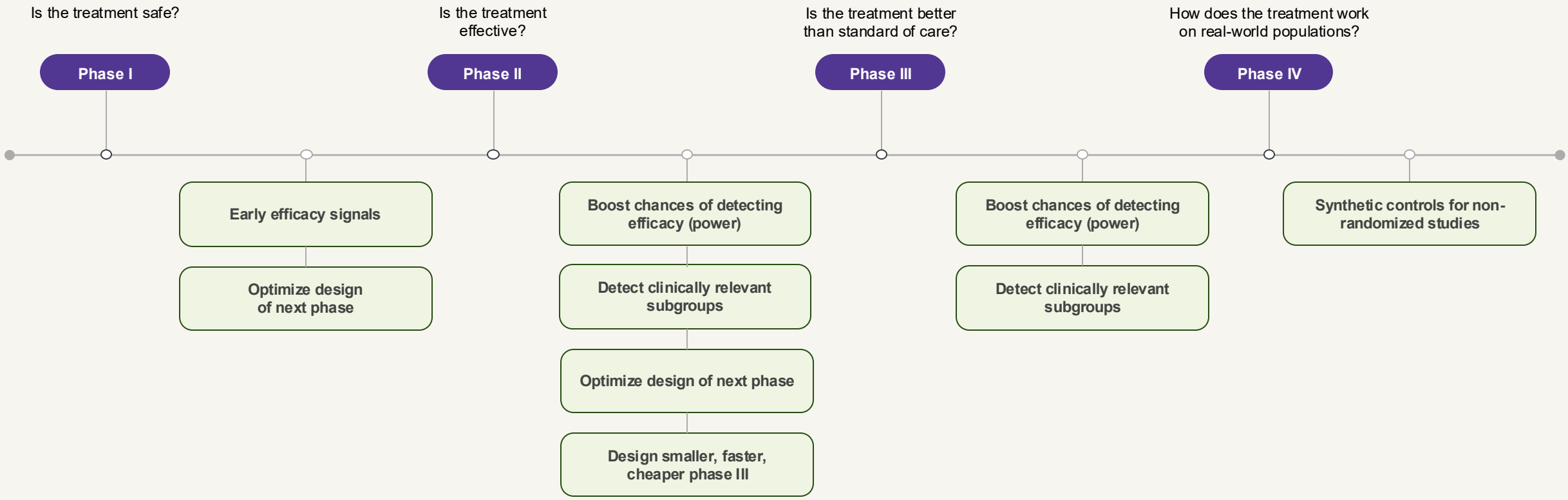
→ Digital Twins of patients are generated from their observed baseline and predict their likely progression under placebo. Predictions can be updated if new information about their physical twins becomes available.

→ Purely synthetic subjects (we will refer to as “synthetic data” for simplicity) do not take observed patient’s data as input. They represent hypothetical patient’s baselines and trajectories with the same statistical properties of actual patients. For example, one can generate a cohort of synthetic subjects satisfying certain inclusion/exclusion criteria to simulate a future trial.

→ As we discussed in the previous webinar, the same generative machine learning model can be used to generate both.

Examples of Digital Twins and Synthetic Data Applications

Digital twins and synthetic data help generate evidence and de-risk outcomes throughout the drug development life cycle.



Examples of Digital Twins and Synthetic Data Applications

1 Design

- → **Simulate trials**
  real-time patient trajectories, placebo response modeling, PTS simulations
- → **Optimize study designs**
  endpoints, timing, inclusion/exclusion criteria

2 Analysis

- → **Boost power and/or reduce sample size in RCTs**
  reduce variability of study analyses for accelerated decision making or reduce sample size required for target power
- → **Synthetic controls for single arm trials and open label studies**
  use digital twins as external control arms
- → **Refined endpoints and subgroups**
  uncover subgroups to better understand treatment effects

3 Commercial

- → **Support payor and market-access strategies** model real-world disease progression using digital twin-based simulations to project long-term outcomes and inform payer strategies, enabling better market access positioning.

Examples of Digital Twins and Synthetic Data Applications

1

Design

→ **Simulate trials**
real-time patient trajectories, placebo response modeling, PTS simulations

→ **Optimize study designs**
endpoints, timing, inclusion/exclusion criteria

2

Analysis

→ **Boost power and/or reduce sample size in RCTs**
reduce variability of study analyses for accelerated decision making or reduce sample size required for target power

→ **Synthetic controls for single arm trials and open label studies**
use digital twins as external control arms

→ **Refined endpoints and subgroups**
uncover subgroups to better understand treatment effects

3

Commercial

→ **Support payor and market-access strategies** model real-world disease progression using digital twin-based simulations to project long-term outcomes and inform payer strategies, enabling better market access positioning.

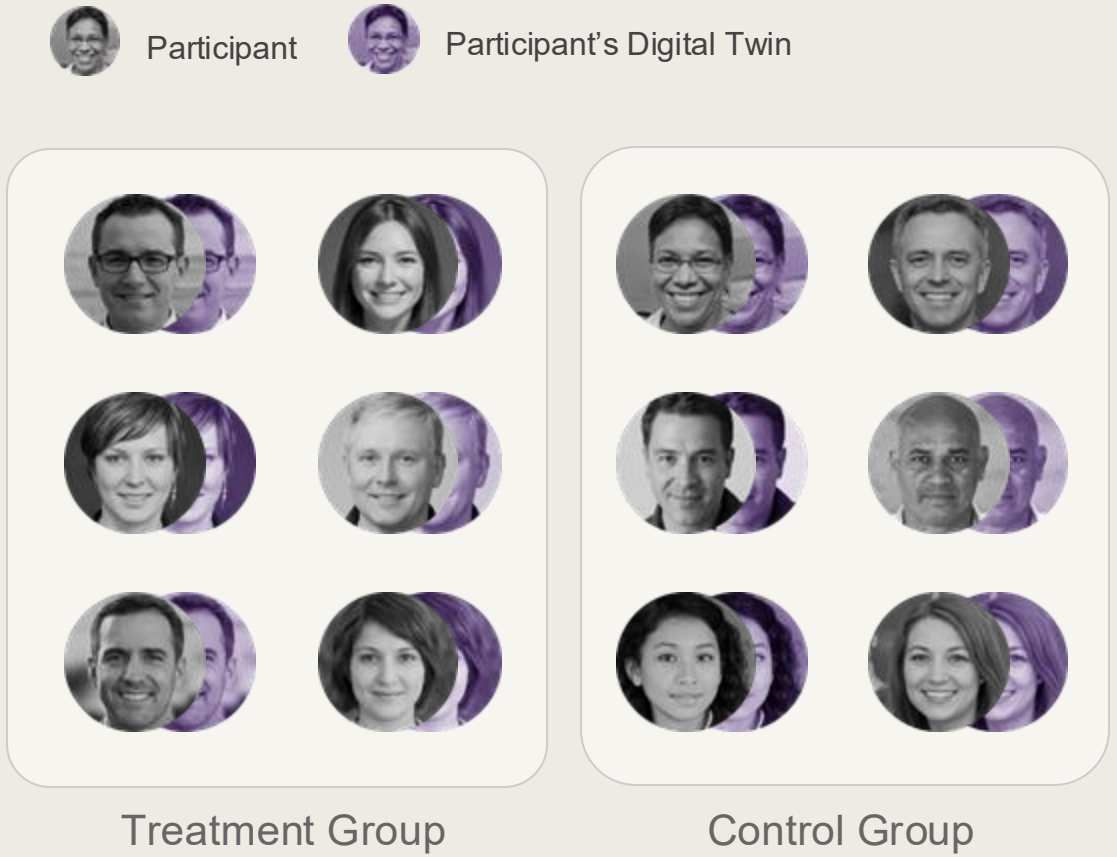
PROCOVA

Digital twins are used as covariates in a covariate-adjusted analysis. Instead of adjusting for, e.g., sex, age or other baseline covariates, you can adjust for the predicted outcome of each subject. This super-covariate can maximize efficiency. The method is called **PROCOVA**. For a continuous outcome for example, the analysis model will look like this:

$$y_i = \beta_0 + \beta_w w_i + \beta_m m_i + \epsilon_i$$

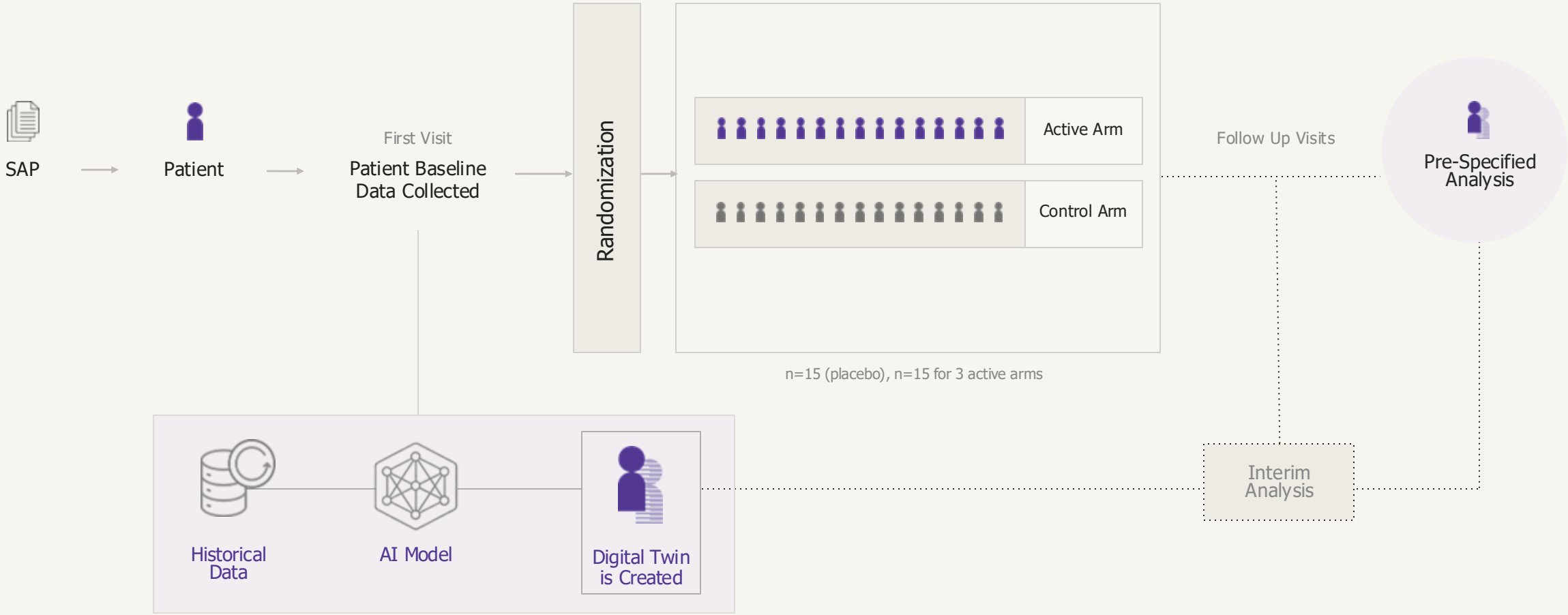
- y_i: outcomes
- β_w: trt effect
- w_i: trt assignment
- m_i: digital twins

Digital Twins are *not* new participants. They are *more information* about the participants in the study. They make every participant's data more rich by bringing information from past trials and observational studies into the analysis.



Schuler, Alejandro, et al. "Increasing the efficiency of randomized trial estimates via linear adjustment for a prognostic score." The International Journal of Biostatistics 18.2 (2021): 329-356.

Digital twins in a randomized clinical study



PROCOVA boosts trial efficiency

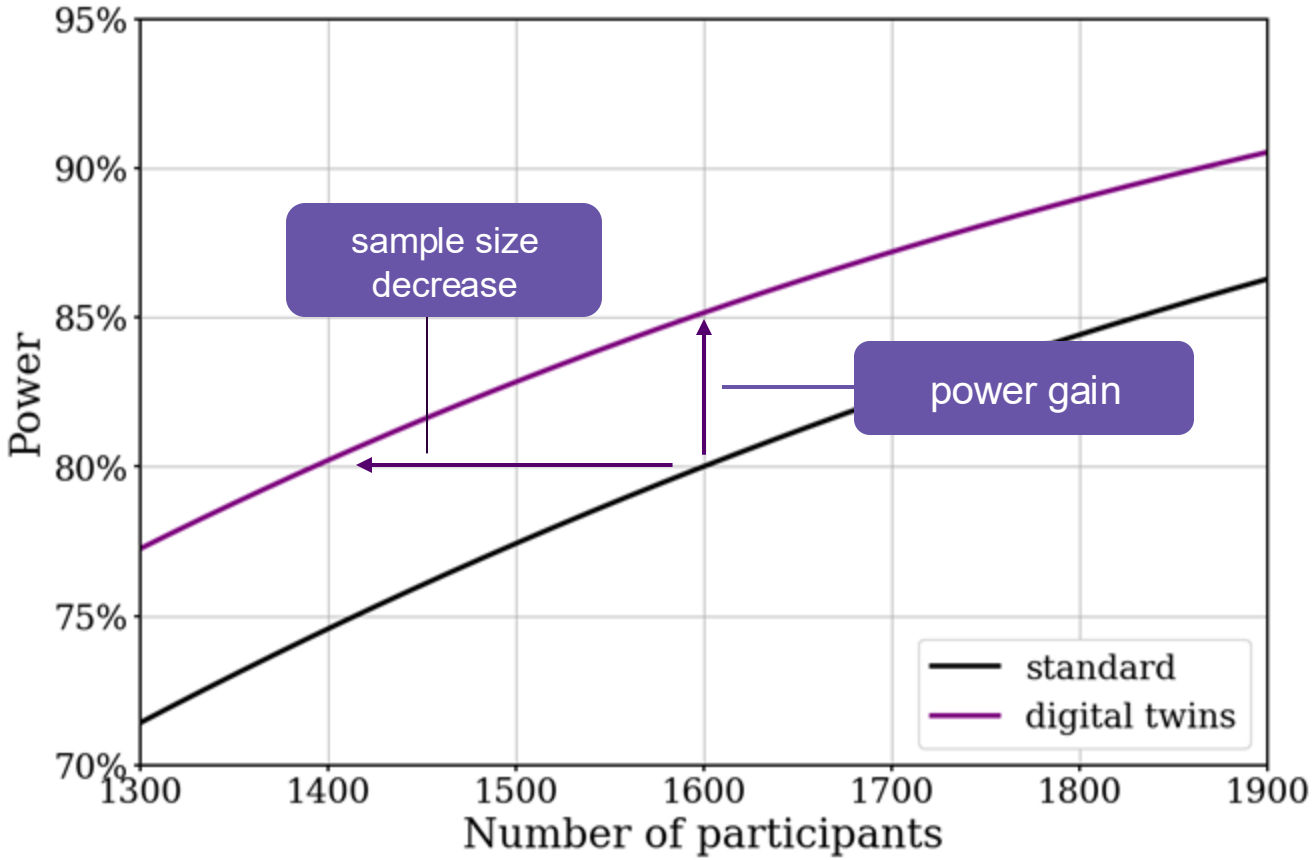
The prognostic value of the covariate can be used for:

Power add

→ Digital twins added at analysis time, this boosts the effective sample size, keeping enrollment constant

Sample size reduction

→ Expected prognostic value incorporated in the sample size calculation. Can decrease the placebo arm or both arms equally



Case studies

PROCOVA

Consider the reanalysis of past RCTs for AD, ALS, HD and compare PROCOVA to the efficiency of the pre-specified analysis.

	Bapineuzumab	Galantamine	Tilavonemab	DHA	Valproate	Resveratrol	Ceftriaxone	Coenzyme Q10
Indication	Alzheimer's	Alzheimer's	Alzheimer's	Alzheimer's	Alzheimer's	Alzheimer's	ALS	Huntington's
Sample Size	N = 1100	N = 1063	N = 453	N = 402	N = 313	N = 119	N = 513	N = 609
Phase	Phase 3	Phase 3	Phase 2	Phase 2	Phase 2	Phase 2	Phase 2	Phase 3
Timepoint	18 Months	12 Months	18 Months	18 Months	24 Months	12 Months	12 Months	24 Months
Primary Endpoint	Adas-Cog11	CDR-SB	CDR-SB	Adas-Cog11	NPI	Adas-cog14	ALSFRS-R	TFC
Met Primary Endpoint	No	Yes	No	No	No	No	No	No

Case studies

PROCOVA

	Prespecified Analysis Variance Reduction	Procova Variance Reduction	Procova Control Arm Reduction
Bapineuzumab	9.9%	20.7%	**34.3%**
Galantamine	0.7%	10.5%	**19.0%**
Tilavonemab	-0.1%	13.0%	**23.0%**
DHA	3.4%	18.1%	**30.3%**
Valproate	1.1%	13.2%	**23.2%**
Resveratrol	-0.5%	12.7%	**22.2%**
Ceftriaxone	0.3%	18.3%	**30.9%**
Coenzyme Q10	3.1%	13.5%	**23.8%**

Prespecified analysis included baseline scores, medication use, and stratification factors.

Digital twin methods demonstrated higher and consistent variance reduction.

Control arm reduction represents reduced sample size while maintaining the trial's original designed power.

Positive interactions with FDA and EMA on the use of PROCOVA in phase 2 and 3 clinical trials

PROCOVA is qualified by the EMA.

"CHMP qualifies PROCOVA as prognostic score adjustment and the proposed procedures, as described in a handbook for trial statisticians, could enable increases in power or precision of treatment effect estimates in controlled randomised clinical trials with continuous outcomes."


EUROPEAN MEDICINES AGENCY
SCIENCE MEDICINES HEALTH

20 September 2022
EMADOC-1700519818-907465
Committee for Medicinal Products for Human Use (CHMP)

Qualification opinion for Prognostic Covariate Adjustment (PROCOVA™)

FDA does not have an analog qualification process, but concurs that PROCOVA is a special case of covariate adjustment.

FDA

"We concur with EMA that PROCOVA is a special case of ANCOVA. As such, CDER's current feedback is that PROCOVA does not appear to deviate from our guidance. Any future requests related to its use should be product or trial specific and directed by a sponsor to the appropriate review division."

– from FDA response to Unlearn's submission to ISTAND.

PROCOVA and Bayesian Analyses

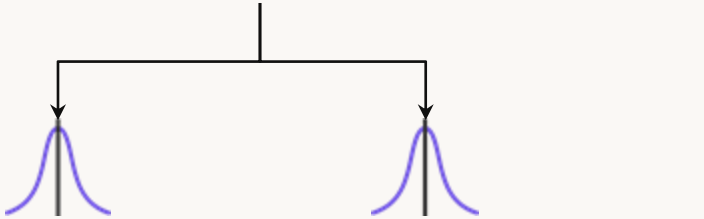
- → Both PROCOVA and bayesian analyses use historical data to improve efficiency.
- → PROCOVA is a covariate adjustment method. **Historical data is used to learn the best combination of baseline covariates and define prognostic scores for a future RCT**. It is a frequentist analysis and strictly controls type-I error. A prognostic score that's strongly correlated with the outcome improves the precision of the treatment effect estimate.
- → A bayesian analysis is a different framework for statistical inference. **Historical data can be used to define a prior distribution over parameters of interest** (e.g., the average control response) which is then updated with the observed data. It does not guarantee strict control of type-I error but can lead to larger gains in efficiency.

The two approaches are complementary and can be combined → Bayesian PROCOVA

Bayesian PROCOVA

Historical data and historical digital twins are used to set priors on analysis parameters. You trade off strict control on type-I error for larger variance reduction. ML metrics and flexible priors allow to minimize type-I error inflation. Typically more suitable for phase I/II trials where strict control of type-I error is not required.

historical data + their digital twins set priors on analysis parameters



$$y_i = \beta_0 + \beta_w w_i + \beta_m m_i + \epsilon_i$$

You now combine efficiency gains from a standard bayesian analysis with PROCOVA adjustment.

Efficiency substantially improved compared to PROCOVA.

Bayesian Prognostic Covariate Adjustment With Additive Mixture Priors

Alyssa M. Vanderbeek* †, Arman Sabbaghi †, Jon R. Walsh †, Charles K. Fisher †

February 29, 2024

Abstract

Effective and rapid decision-making from randomized controlled trials (RCTs) requires unbiased and precise treatment effect inferences. Two strategies to address this requirement are to adjust for covariates that are highly correlated with the outcome, and to leverage historical control information via Bayes' theorem. We propose a new Bayesian prognostic covariate adjustment methodology, Bayesian PROCOVA, that combines these two strategies. Covariate adjustment in Bayesian PROCOVA is based on generative artificial intelligence (AI) algorithms that construct a digital twin generator (DTG) for RCT participants.

Case studies

Bayesian PROCOVA

Consider the same case studies we reanalyzed with PROCOVA. Bayesian PROCOVA consistently provides substantially larger efficiency gains.

	Prespecified Analysis Variance Reduction	Procova Variance Reduction	Bayesian Procova Variance Reduction	Procova Control Arm Reduction	Bayesian PROCOVA Control Arm
Bapineuzumab	9.9%	20.7%	30.8%	34.3%	**47.1%**
Galantamine	0.7%	10.5%	12.2%	19.0%	**21.7%**
Tilavonemab	-0.1%	13.0%	20.5%	23.0%	**34.0%**
DHA	3.4%	18.1%	28.4%	30.3%	**44.2%**
Valproate	1.1%	13.2%	25.4%	23.2%	**40.5%**
Resveratrol	-0.5%	12.7%	24.1%	22.2%	**38.8%**
Ceftriaxone	0.3%	18.3%	30.3%	30.9%	**46.5%**
Coenzyme Q10	3.1%	13.5%	26.7%	23.8%	**42.1%**

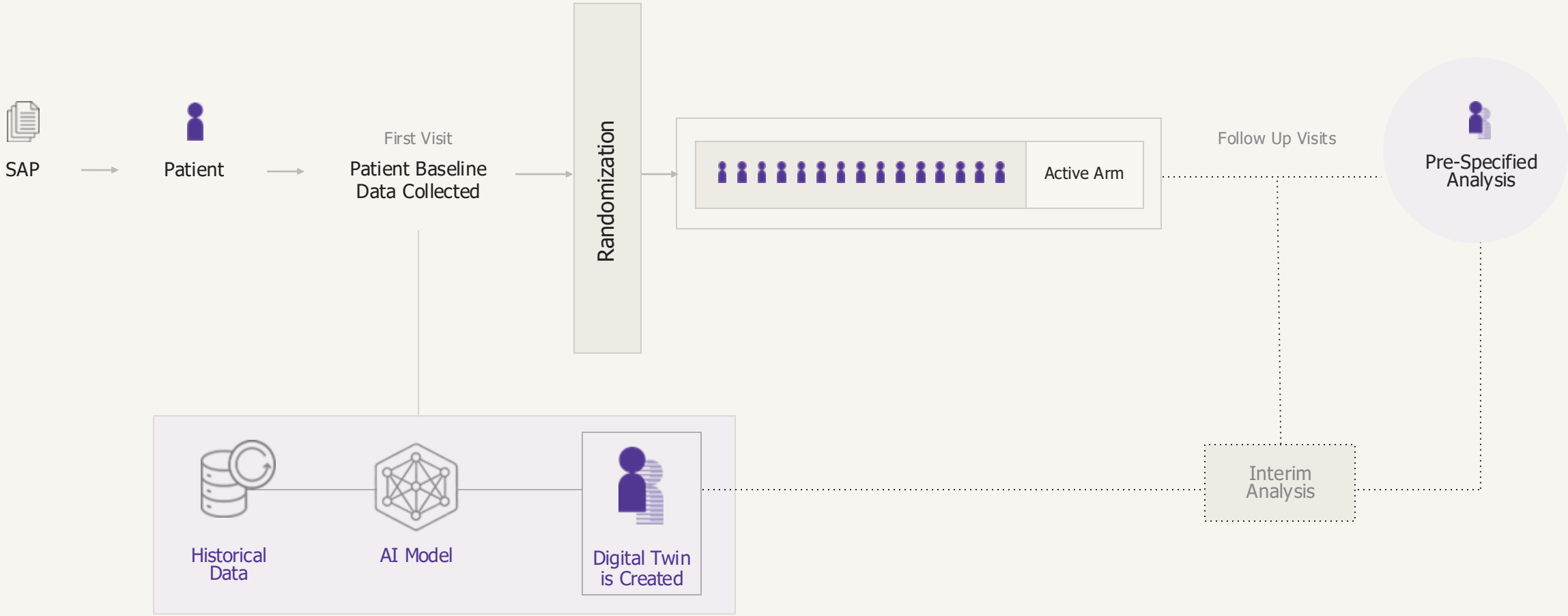
Single Arm Trials

- → Randomized control trials are the gold standard to establish efficacy of experimental treatments
- → In some cases, RCTs might not be feasible and single arm trials can be used. E.g.,
  - **Randomization might be unethical**
  - **Rare diseases**
  - **Pediatric trials**
  - **Early-phase trials, internal decision-making**
  - **Support accelerated approval**
- → We need an external control comparator

Digital twins can be used to construct synthetic controls.

For each treated subject, their digital twin provides a control counterfactual: it predicts how that particular subject would have progressed had they taken standard-of-care.

Digital Twins as Synthetic Controls



Digital Twins as Synthetic Controls

Digital twins provide a natural way to construct a synthetic control and can be **integrated with different statistical frameworks to improve robustness**. With any non-randomized method, we need stronger identifiability assumptions for historical data vs target trial: consistency of trt definitions, no unmeasured confounders, and population overlap.

(1) **Direct Comparison**

$E[Y - m \mid trial]$ Y: treated outcome, m: digital twin (predicted mean control outcome)

You rely on digital twins being unbiased.

---

(1) **Doubly Robust Estimator**

$E[Y - m \mid trial] - E[W(Y_C - m) \mid historical]$ Y_C: control outcome, W: propensity odds

The second term estimates bias of the model in the historical data and corrects for it. A propensity model is used to "transport" the estimated bias from the historical population to the trial population. This is called an Augmented Inverse Probability Weighting (AIPW) estimator and remains asymptotically unbiased if either digital twins model or propensity model are well-specified. It is also more efficient.

Digital Twins as Synthetic Controls

Why digital twins synthetic controls?

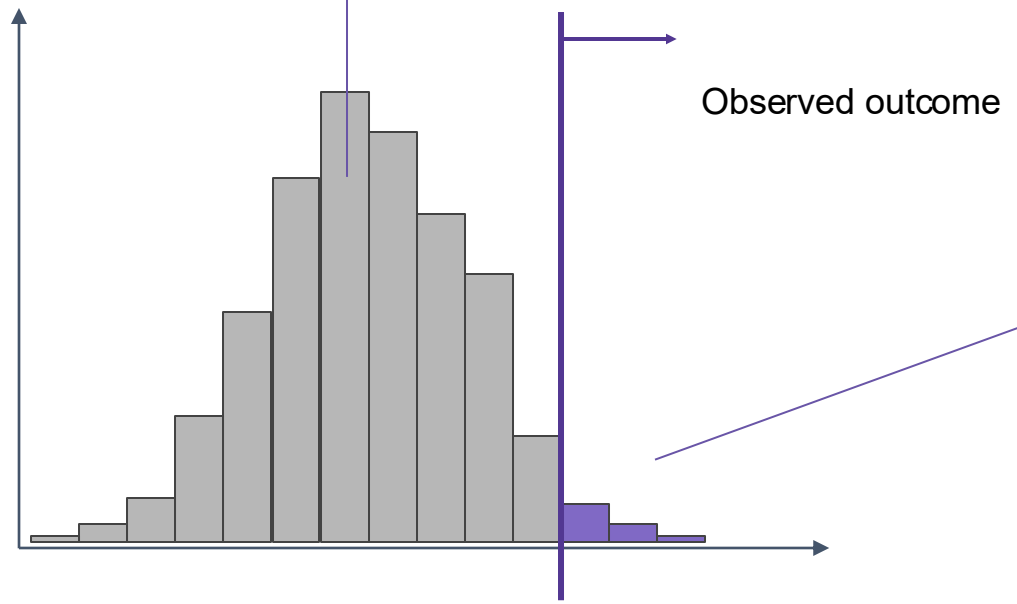
- → Digital twins are control predictions exactly matched to each individual subject, unlike pure data-based external controls (e.g., propensity score matching)
- → Digital twins leverages ML capabilities to learn from large, heterogeneous, often sparse datasets. Propensity score matching can be difficult in these cases
- → When used in a doubly robust framework, provides asymptotically the most precise estimate of treatment effect (if both digital twins and propensity models are well-specified)

Other applications of synthetic controls

- → Hybrid controls (internal+external)
- → Evidence for post-approval market access. Digital twins can help answer questions like:
  - Effectiveness on real-world populations
  - Comparative effectiveness
  - Long-term effectiveness
  - Cost-effectiveness

Responder Subgroup Discovery with Digital Twins

Digital twins provide a **distribution** of likely control outcomes per subject.

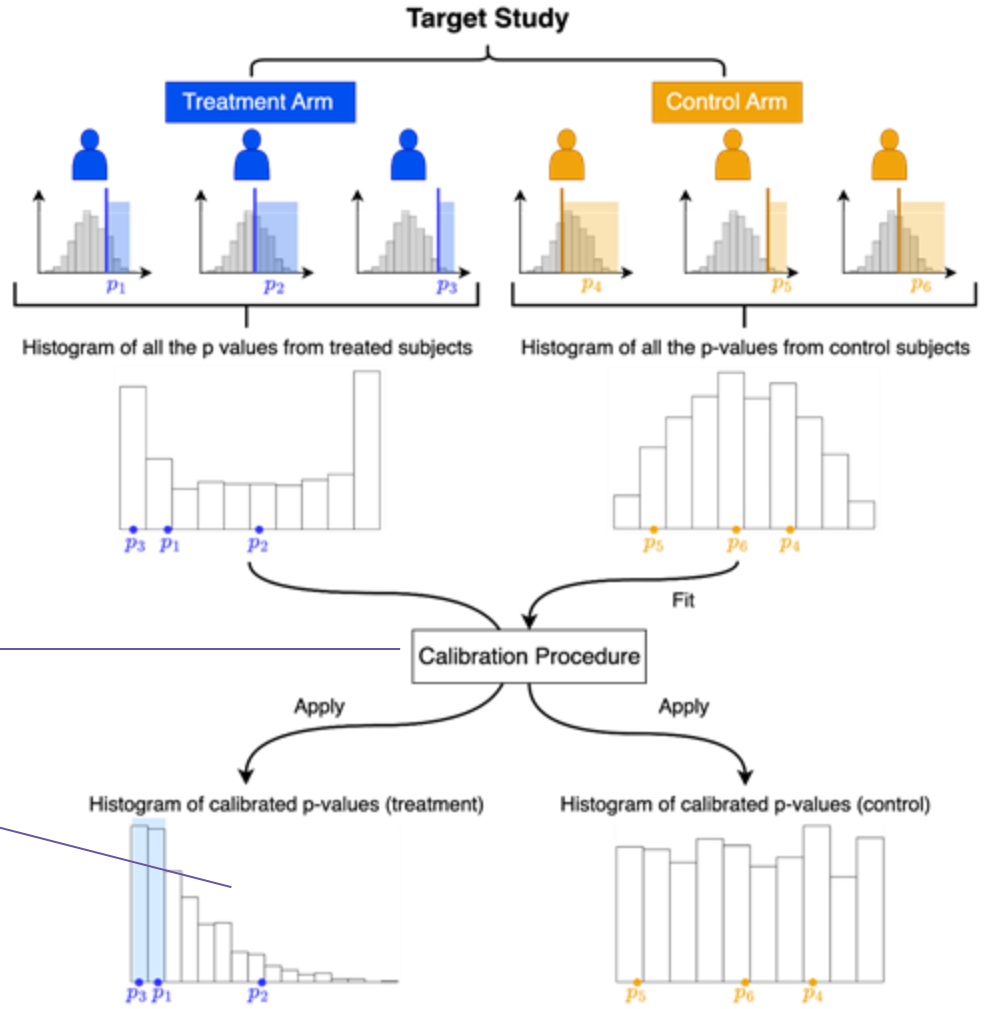


Per-subject **personalized p-value.** Any observed outcome can be compared to this distribution to evaluate how likely it is relative to the model's distribution.

Responder Subgroup Discovery with Digital Twins

In a study of interest we analyze the distribution of personalized p-values within the treatment and control groups, calibrating them to ensure that the distribution of control p-values is uniform.

If there is a treatment effect the treatment group will be skewed.



Step 1:
Compute original p-value with generated twins (gray bars) and observed outcome (blue and orange lines)

Step 2:
Re-calibrate with control subjects and apply to the treated subjects.

Step 3:
Identify responders in the treatment arm. (blue area on the left figure)

Examples of Digital Twins and Synthetic Data Applications

1

Design

- → **Simulate trials**
  real-time patient trajectories, placebo response modeling, PTS simulations
- → **Optimize study designs**
  endpoints, timing, inclusion/exclusion criteria

2

Analysis

- → **Boost power and/or reduce sample size in RCTs**
  reduce variability of study analyses for accelerated decision making or reduce sample size required for target power
- → **Synthetic controls for single arm trials and open label studies**
  use digital twins as external control arms
- → **Refined endpoints and subgroups**
  uncover subgroups to better understand treatment effects

3

Commercial

- → **Support payor and market-access strategies** model real-world disease progression using digital twin-based simulations to project long-term outcomes and inform payer strategies, enabling better market access positioning.

Synthetic Data to Optimize Trial Design

→ In the last webinar we discussed how having a generative model of populations and disease progression allows to sample under different conditions. If we condition on the baseline characteristics of a specific patient, we get a predicted probability distribution for the future outcomes of that patient over time. This is their **digital twin**.

$$p(X_{\text{longitudinal}}(t > 0), X_{\text{events}}(t) \mid X_{\text{baseline}})$$

→ If we condition on just inclusion/exclusion criteria we can sample purely synthetic baselines and predict their future outcomes. These synthetic cohorts have the same statistical properties of real subjects and can be used to **simulate trials** (i.e., “in-silico trial”). One can imagine different scenarios where the treatment effect itself is simulated based on assumptions or learned from relevant data.

$$p(X_{\text{longitudinal}}(t > 0), X_{\text{events}}(t) | X_{\text{baseline}} \in \text{IE criteria})$$

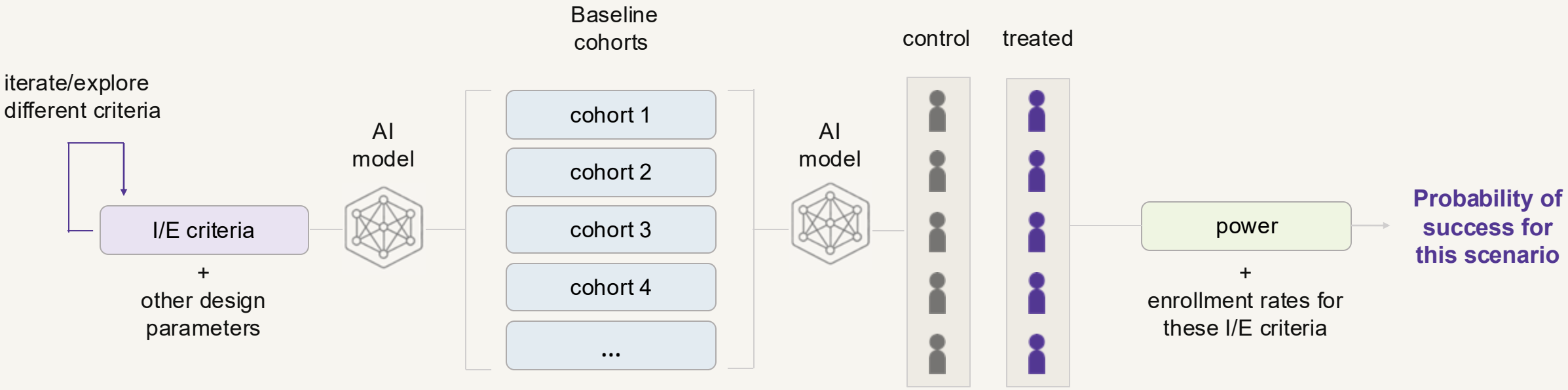
Synthetic Data to Optimize Trial Design

High-fidelity, subject-level, synthetic data can be used to answer key clinical trial design questions and test assumptions.

For example:

- → For a given set of inclusion/exclusion criteria, what's the expected **progression and variability** across different outcomes?
- → How long should a trial be to **observe sufficient change** in target outcomes?
- → What set of inclusion criteria give me the highest **trial power to detect an effect while balancing enrollment rate potential**?
- → What's the relationship between **biomarkers and clinical outcomes** at different time points and different hypothetical populations enrolled?

Synthetic Data to Optimize Trial Design



Synthetic Data to Optimize Trial Design

Some of the advantage of using a synthetic data generator include:

- → Ability to simulate scenarios where historical data density is low or outcomes are sparse
- → Ability to handle multiple data sources (e.g., subject-level + population-level) via model training
- → Assumptions about treatment effects can be easily integrated
- → Synthetic data can be shared more easily as there are fewer privacy concerns
- → In some cases one can design hybrid approaches where both observed data and synthetic data can be used to simulate trials, leveraging strengths of both

Regulatory Interactions for AI applications Should Follow a Risk-Based Approach

1. Define the question of interest
2. Define the context of use
3. Asses AI model risk
4. Develop a plan to establish credibility

Early interactions, iterative

1. Execute the plan
2. Document results
3. Determine adequacy for context of use

Execution and evaluation

GUIDANCE DOCUMENT

Considerations for the Use of Artificial Intelligence To Support Regulatory Decision-Making for Drug and Biological Products

Draft Guidance for Industry and Other Interested Parties

JANUARY 2025

In its January 2025 draft guidance FDA outlines a **seven steps risk-based framework** when considering the use of AI/ML applications.

All applications described here can be used following the FDA framework.

Summary

- PROCOVA: Digital Twins can be used as super covariates in an RCT analysis to reduce sample size while maintaining the same power or boost power with same sample size. When strict control of type-I error is not required, a Bayesian version of PROCOVA can lead to even larger efficiency gains by combining the advantages of PROCOVA with those of a bayesian analysis.
- In single arm trials, digital twins can be used to construct matched synthetic controls and combined with different statistical framework to enhance robustness (e.g., doubly robust).
- Digital twins can be used to discover responder subgroups by comparing their observed outcomes to a predicted control distribution (their digital twin).
- Synthetic data generated from the same probabilistic models can be used for high-fidelity simulations of future trials and design optimization.
- These are just few example of digital twins / synthetic data applications. Context of use, risk assessment, and tailored risk mitigation should always be taken into consideration. Especially for regulatory interactions.

UNLEARN + MULTI-REGIONAL CLINICAL TRIALS
THE MRCT CENTER of BRIGHAM AND WOMEN'S HOSPITAL and HARVARD

Thank you.