Extending the results of clinical trials using data from a target population

BROWN
School of Public Health

Issa Dahabreh
Center for Evidence-Based Medicine,
Brown School of Public Health

Disclaimer

- Partly supported through PCORI Methods Research Awards ME-1306-03758 & ME-1502-27794.

- All statements are solely mine and do not necessarily represent the views of the PCORI, its Board of Governors, or the Methodology Committee.

- Work in progress

Outline

- Applicability, external validity
  - Concepts & study design
- Extending trial findings to the target population
  - Structural assumptions
  - Statistical methods
  - Simulation study
- Applications
  - small data: revascularization for coronary artery disease
  - Big data: anti-hypertensives for chronic renal failure
- Extensions
- Concluding remarks

Applicability, external validity

Concepts and design-based solutions

Applicability and external validity

- "…essential for imparting significance to scientific research beyond the confines of the original investigation…"

- Evidence interpretation:

  obtain → assess validity → synthesize → assess applicability

- Unaided subjective judgment
  - Not transparent
  - Subject to cognitive "biases"
  - No conceptual framework

Key Question: How can we aid our judgment using causal assumptions & statistical methods?

Assessing applicability

- Based on similarity judgments
  - Populations
  - Interventions
  - Outcomes

- ***Applicability*** is tied to ***representativeness*** of trial participants

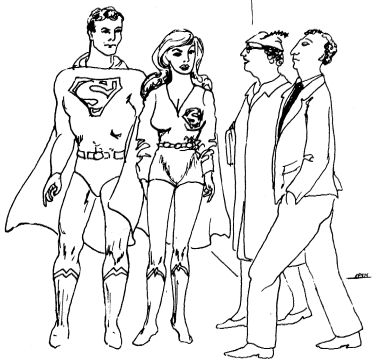
- ***Target*** population where the trial results will be applied

What is a representative sample?

Kruskal & Mosteller, Int Stat Review 1980

Meaning 1. *General, usually unjustified, acclaim for data: The emperor's new clothes.*

Meaning 2. *Absence of selective forces: Justice balancing the scales.*



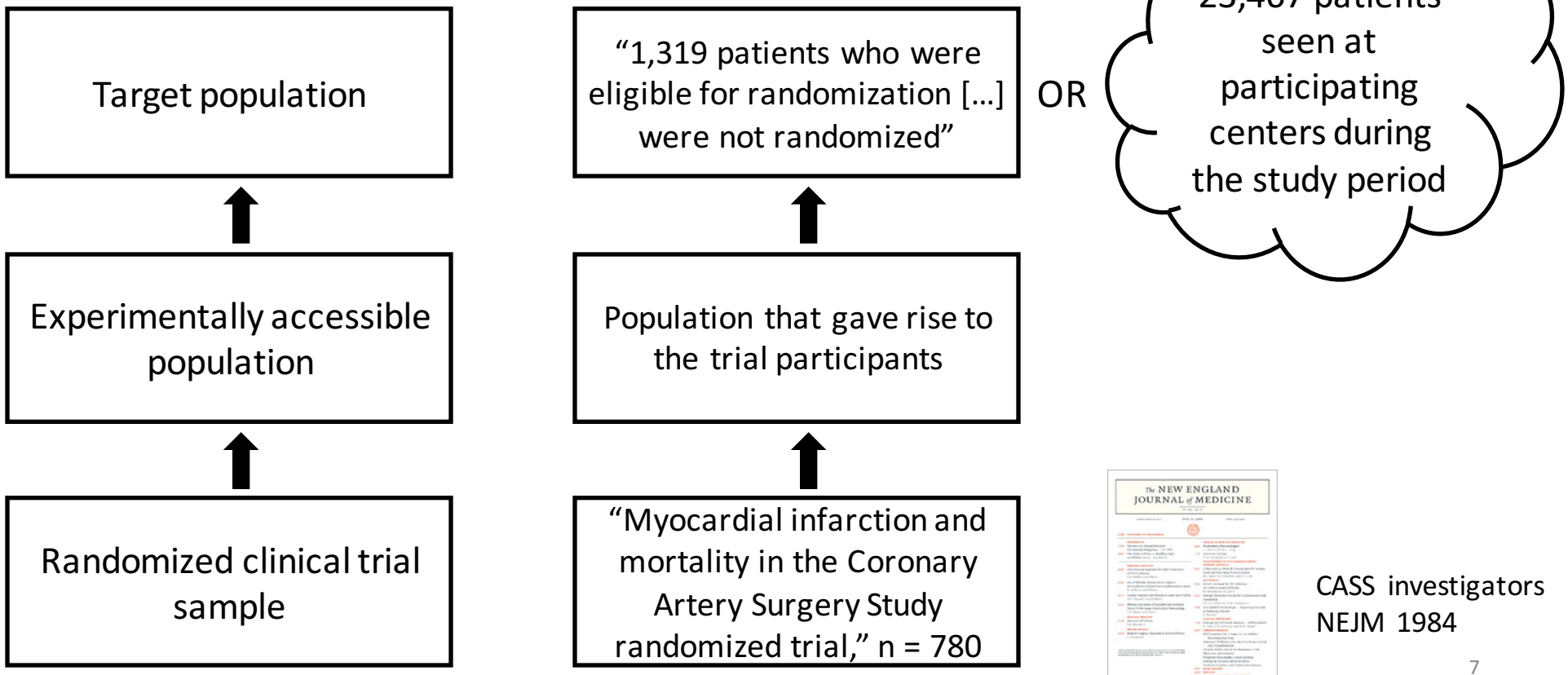
Meaning 4. *Typical or ideal case: Superman and Superwoman and Average man and Average woman.*

Meaning 5. *Coverage of the population: Noah's Ark.*

Meaning 7. *Some specific sampling method: The Sampling Department in action.*

Representativeness: main preoccupation in survey research

Clinical trials and target populations



CASS investigators
NEJM 1984

Should we worry about representativeness?

- Participants in the trials have different distributions of covariates than the target

- Treatment effects are heterogeneous over these covariates
  - Effect modification or "heterogeneity of treatment effects"

Key idea: applicability requires similar distribution of effect modifiers across populations

Is this a real problem?

- Systematic review, 52 studies (cardiology = 20; mental health = 17; oncology = 15)
- "*… a **high proportion of the general disease population was often excluded from trials.***"
- "*…highly selected and have a lower risk profile than real-world populations, with the **frequent exclusion of elderly patients and patients with co-morbidities**.*"
- "*71.2 % [of studies] concluded that **RCT samples were not broadly representative of real-world patients**…*"

Kennedy-Martin et al., Trials 2015

Design-based strategy (ideal) : survey experiments

- Randomly select individuals from the target population (survey)
- Randomly assign the sampled individuals (experiment)
- Analyze the results using survey weights for inference to the target population

Limitations of survey experiments

- The patients invited are different from those who agree to participate
- Exceedingly difficult to do in medical or health services research
  - E.g.: Oregon Health Insurance Experiment
- Applying the results of a survey experiment to different target population raises the same problem as for typical trials

Extending trial results to a target population

Structural assumptions and statistical methods

Counterfactual causality

- Causal effects = contrasts between *potential outcomes* under different interventions

- For the *i* th individual, i = 1,...,N
  Y^1_i is the *potential* outcome under treatment
  Y^0_i is the *potential* outcome under control

- Individual treatment effect = $Y^1_i - Y^0_i$

- Average treatment effect, $\Delta = E[Y^1_i - Y^0_i] = E[Y^1_i] - E[Y^0_i]$

Causal inference and missing data

- Foundamental problem of causal inference:

$$\text{we can never know } Y^1_i \textbf{\textit{ and }} Y^0_i$$

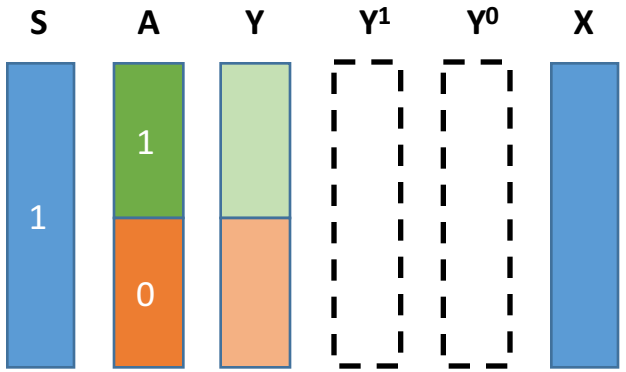
- Causal problem → missing data problem

- Deep connection
  - Causal inference problems are missing data problems
  - Missing data problems are causal inference problems

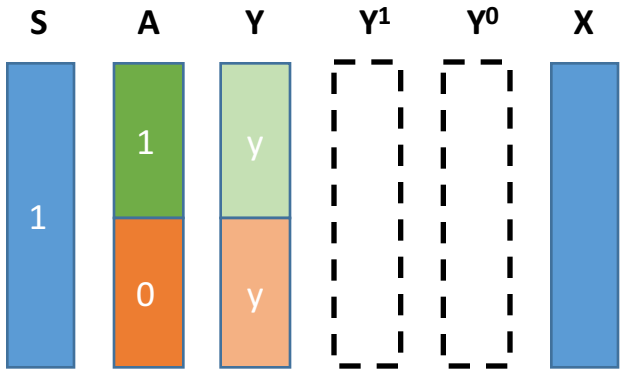
Rubin, J Ed Psych 1974; Rubin, JASA 2008

Missing data in the trial

- We have observed outcomes but know nothing about potential outcomes

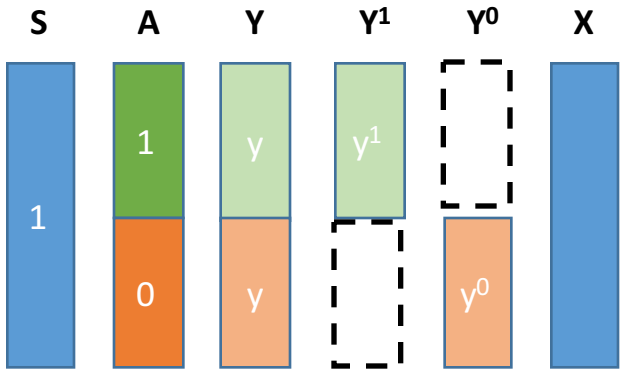


Identification in the trial



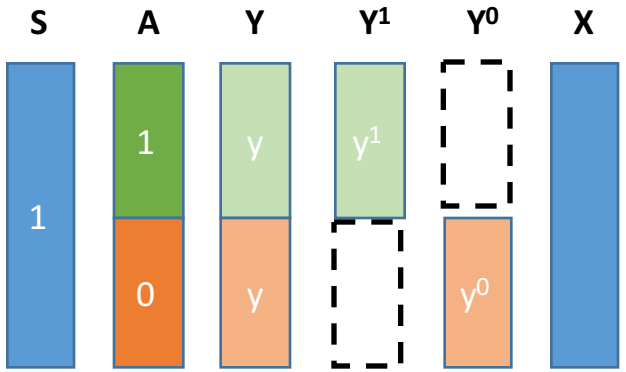
Identification in the trial

- Well defined interventions, consistency
  - $Y^{obs} = A\,Y^1 + (1 - A)\,Y^0$



Identification in the trial

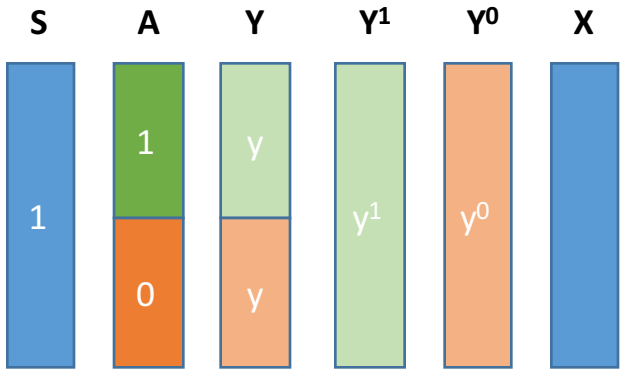
- Well defined interventions, consistency
  - $Y^{obs} = A\,Y^1 + (1 - A)\,Y^0$
- Positivity
  - $0 < P(A = 1 \mid X) < 1$



Identification in the trial

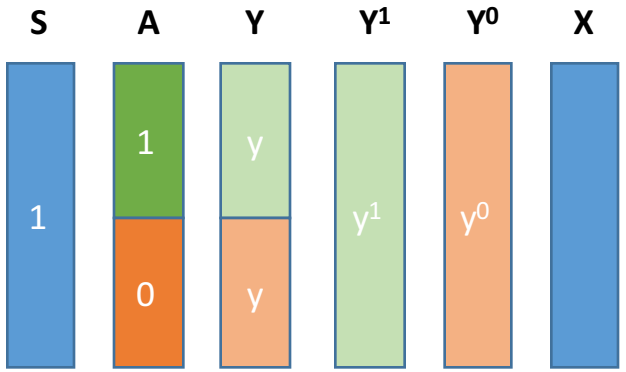
- Well defined interventions, consistency
  - $Y^{obs} = A\,Y^1 + (1 - A)\,Y^0$
- Positivity
  - $0 < P(A = 1 \mid X) < 1$
- Exchangeability
  - Marginal:

    $E[Y^a \mid A = a, S = 1] = E[Y^a \mid S = 1]$



Identification in the trial

- Well defined interventions, consistency
  - $Y^{obs} = A\,Y^1 + (1 - A)\,Y^0$
- Positivity
  - $0 < P(A = 1 \mid X) < 1$
- Exchangeability
  - Marginal:
    $E[Y^a \mid A = a, S = 1] = E[Y^a \mid S = 1]$
  - Conditional:
    $E[Y^a \mid A = a, X, S = 1] = E[Y^a \mid X, S = 1]$

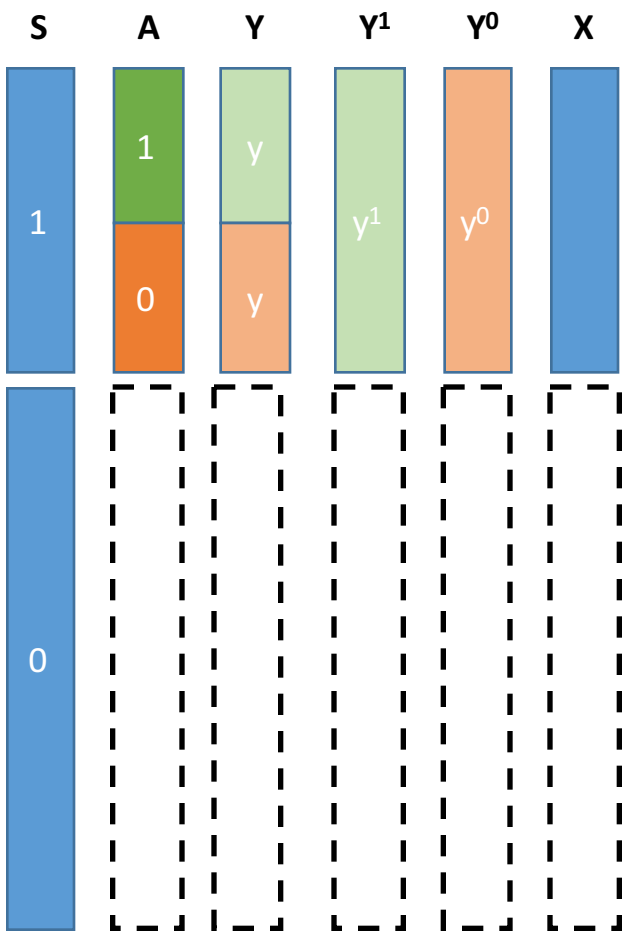


Will also assume that the trial is otherwise flawless: no dropouts, no missing data, no measurement error,…

This is only done for clarity of exposition

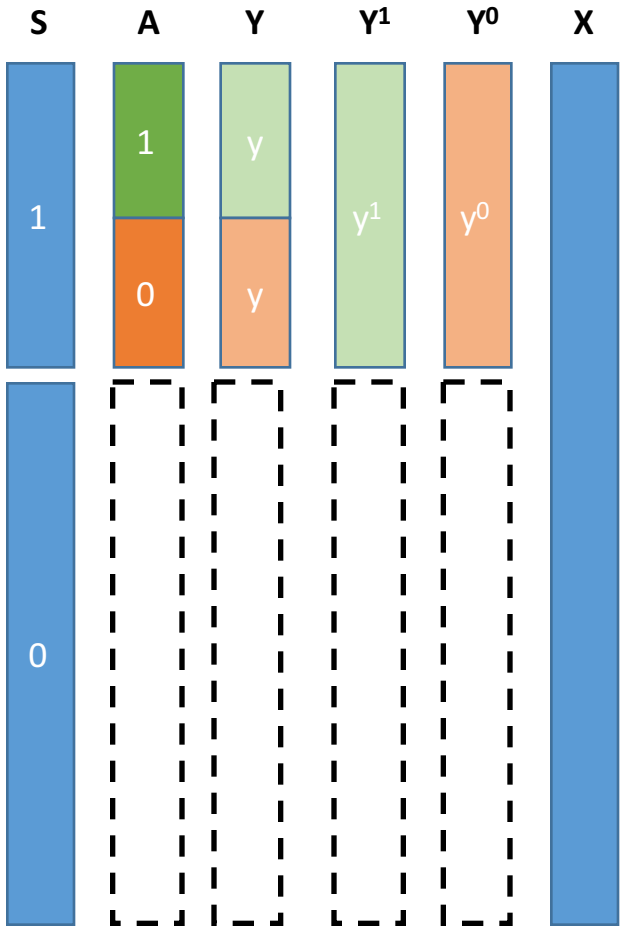
Extending the trial to the target population

- All we know is who is in and who is out of the trial



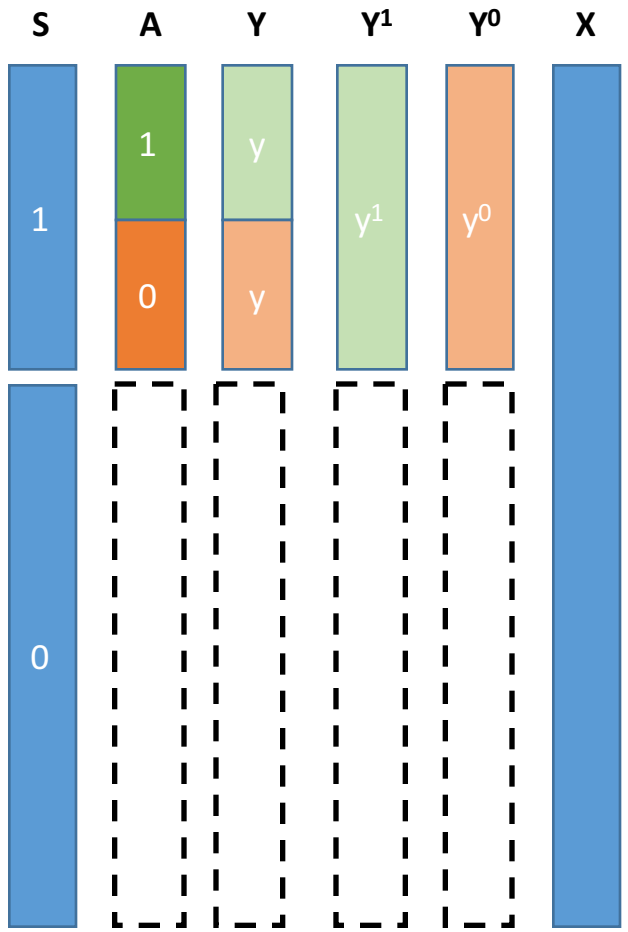
Identification in the target

- Often we have information on the covariate distribution in the target



Identification in the target

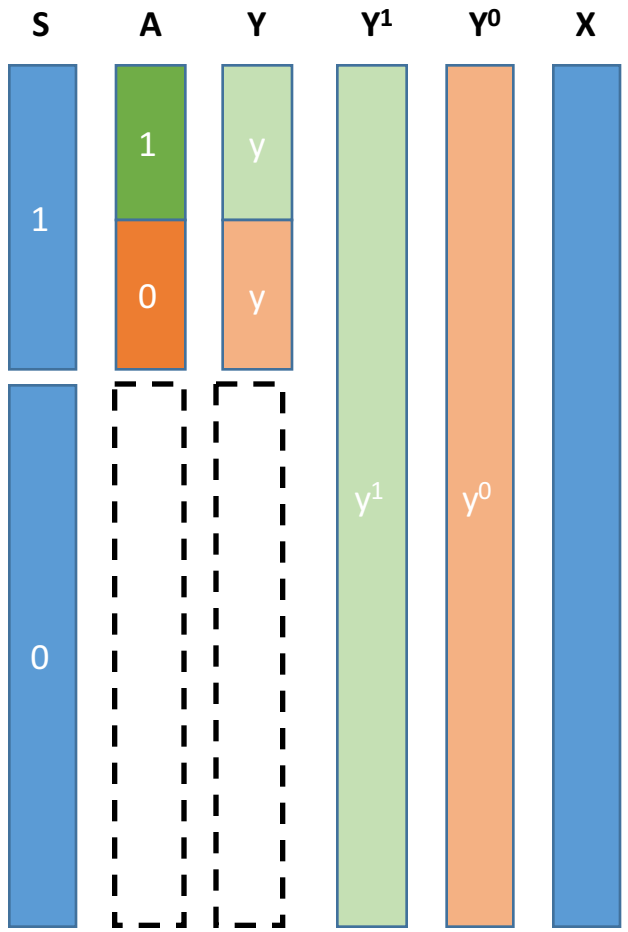
- Often we have information on the covariate distribution in the target
- Positivity of trial participation
  - P(S = 1 | X) > 0



Identification in the target

- Often we have information on the covariate distribution in the target
- Positivity of trial participation
  - $P(S = 1 \mid X) > 0$
- Conditional mean transportability

  $E[Y^a \mid X, S = 1] = E[Y^a \mid X]$



Positivity of trial participation

- No set of covariates predicts lack of participation perfectly
- No systematically excluded subgroups of patients
- Required for principled extrapolation
- Sometimes implausible, e.g.,
  - Tails of the age distribution
  - Disease severity
  - Socioeconomic status

Conditional mean transportability

- We know enough factors that determine the outcome so that trial participation itself is unimportant
- Trial participation itself does not have a direct effect on the outcome
- Often implausible
  - Hawthorne effect
  - Incomplete knowledge of the outcome mechanisms
  - Lack of measurements

Using a model for the outcome

- Treat the trial as an investigation into the mechanism of the outcome
- What statistics professors tell us not to do – "extrapolation"

- How to do it:
  1. Estimate 2 models for the outcome, one in each of the treatment arms
  2. Use the models to predict outcomes under treatment and no treatment for everyone in the target
  3. Compare the average values to estimate the causal effect

Imai & Ratkovic, Ann Applied Stat 2013

Outcome-model-based estimator

$$\hat{\mu}_{OM} = \frac{1}{N}\sum_{i=1}^{N} g_a(X_i;\hat{\gamma}_a)$$

$$g_a(X_i;\hat{\gamma}_a) \to g_a(X_i;\gamma_a^*) = \mathbb{E}[Y^{obs}|A=a,X,S=1]$$

$$\begin{aligned}
\hat{\mu}_{OM} &\to \mathbb{E}[g_a(X_i;\gamma_a^*)] \\
&= \mathbb{E}_X\left[\mathbb{E}[Y^{obs}|A=a,X,S=1]\right] \\
&\overset{(1)}{=} \mathbb{E}_X\left[\mathbb{E}[Y^a|A=a,X,S=1]\right] \\
&\overset{(2)}{=} \mathbb{E}_X\left[\mathbb{E}[Y^a|X,S=1]\right] \\
&\overset{(3)}{=} \mathbb{E}_X\left[\mathbb{E}[Y^a|X]\right] \\
&\overset{(4)}{=} \mathbb{E}[Y^a],
\end{aligned}$$

Using a model for the probability of participation

- Treat the trial as a survey where we lost the sampling weights
- Same structural, different modeling assumptions – "extrapolation"

- How to do it:
  1. Estimate a model for the probability of participation
  2. Weight the outcomes in each trial arm by the inverse of the probability of participation
  3. Compare the average weighted values to estimate the causal effect

Stuart & Cole, Am J Epi 2010

Probability of participation estimator

$$\hat{\mu}_{IPPW} = \frac{1}{N}\sum_{i=1}^{N}\frac{S_i I(A_i = a)Y_i^{obs}}{w_a(X_i;\hat{\beta}_a)}$$

$$w_a(X_i;\hat{\beta}_a) \rightarrow w_a(X_i;\beta_a^*) = P(S=1|X)P(A=a|X,S=1)$$

$$\begin{aligned}
\hat{\mu}_{IPPW} &\rightarrow \mathbb{E}\left[\frac{SI(A=a)Y^{obs}}{w_a(X;\beta_a^*)}\right] \\
&\overset{(1)}{=} \mathbb{E}_X\left[\mathbb{E}\left[\frac{SI(A=a)Y^{obs}}{w_a(X;\beta_a^*)}|X\right]\right] \\
&= \mathbb{E}_X\left[\frac{1}{w_a(X;\beta_a^*)}\mathbb{E}\left[SI(A=a)Y^{obs}|X\right]\right] \\
&= \mathbb{E}_X\left[\frac{1}{w_a(X;\beta_a^*)}\mathbb{E}\left[Y^{obs}|A=a,X,S=1\right]P(S=1|X)P(A=a|X,S=1)\right] \\
&\overset{(2)}{=} \mathbb{E}_X\left[\mathbb{E}\left[Y^{obs}|A=a,X,S=1\right]\right] \\
&\overset{(3)}{=} \mathbb{E}[Y^a],
\end{aligned}$$

Extreme model reliance

- Need for correct specification of the outcome or participation models
  - Otherwise estimators are inconsistent
- Ideally, we want a method that
  - Protects us from getting the wrong answer if we could get one of the models right (and did not know which)
  - Is as efficient as the outcome model-based approach
  - Provides a way to assess model specification

Doubly robust estimation

- Such a method exists!
- Smart way to combine the outcome model and the probability of participation
  - Make the weighted estimator more efficient and gain robustness
- Analogous to de-confounding methods for combining weighting and regression modeling

Scharfstein et al. JASA 1999; Robins & Rotnitzky, Stat Sinica 2001;
Robins et al., Stat Science 2007

Doubly robust estimators

Unbounded, Horwitz-Thompson

1. Model participation in the trial and calculate inverse probability of participation weights

2. Estimate a model for the outcome using regression and predict outcomes

3. *Combine observed outcomes, predicted outcomes, and the probability of participation* in a smart way

$$\hat{\delta}_{\mathrm{DR}} = \frac{1}{N}\sum_{i=1}^{N}\left\{\frac{S_i A_i Y_i}{w_1(X_i,\hat{\beta})} + \frac{w_1(X_i,\hat{\beta}) - S_i A_i}{w_1(X_i,\hat{\beta})} g_1(X_i,\hat{\gamma}_1) - \frac{S_i(1-A_i)Y_i}{w_0(X_i,\hat{\beta})} - \frac{w_0(X_i,\hat{\beta}) - S_i(1-A_i)}{w_0(X_i,\hat{\beta})} g_0(X_i,\hat{\gamma}_0)\right\}$$

Bounded, regression-based

1. Model participation in the trial and calculate inverse probability of participation weights

2. Estimate a model for the outcome using *weighted regression* and predict outcomes

3. Use the predictions to estimate the causal effect

3 claims

- Consistent if one of the two models (or both) are correctly specified
- In large samples, if both models are correctly specified, as efficient as the outcome model based estimator
- Reasonable performance in finite samples

Key idea: 2 opportunities to make a modeling mistake for the price of 1

Simulation study

(selected results)

Simulation setup

$N = 1{,}000{,}000$

$X_{j,i} \sim \mathcal{N}(0,1)$; $j = 1,2,3$; $i = 1,\ldots,N$; $\mathbf{X}_i^T = (1, X_{1,i}, X_{2,i}, X_{3,i})$

$S_i \sim \text{Bernoulli}(\pi_i)$

$A_i \sim \text{Bernoulli}(0.5)$ if $S_i = 1$

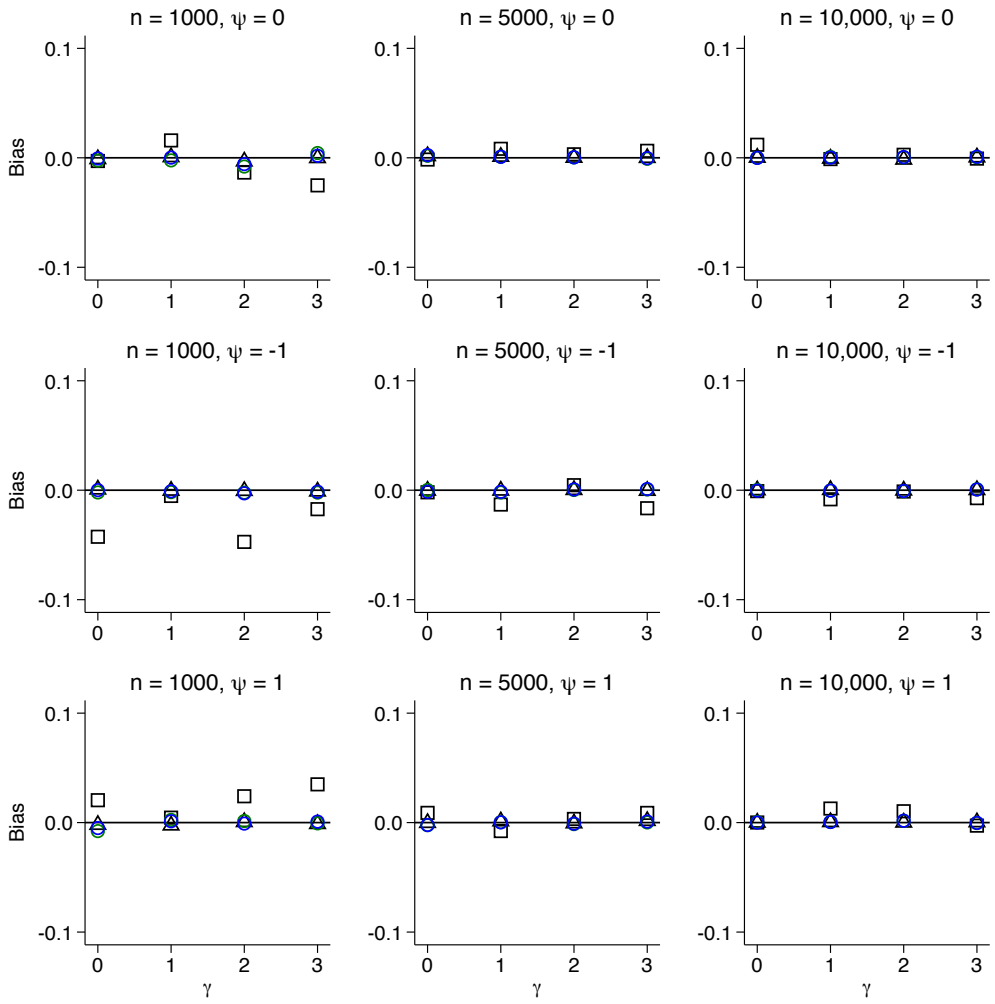
$\pi_i = \exp(\mathbf{X}_i^T\boldsymbol{\alpha}) / (1 + \exp(\mathbf{X}_i^T\boldsymbol{\alpha}))$; $\boldsymbol{\alpha}^T = (\alpha_0, 1,1,1)$

α_0 to obtain expected trial sample sizes of 1000, 5000, or 10,000

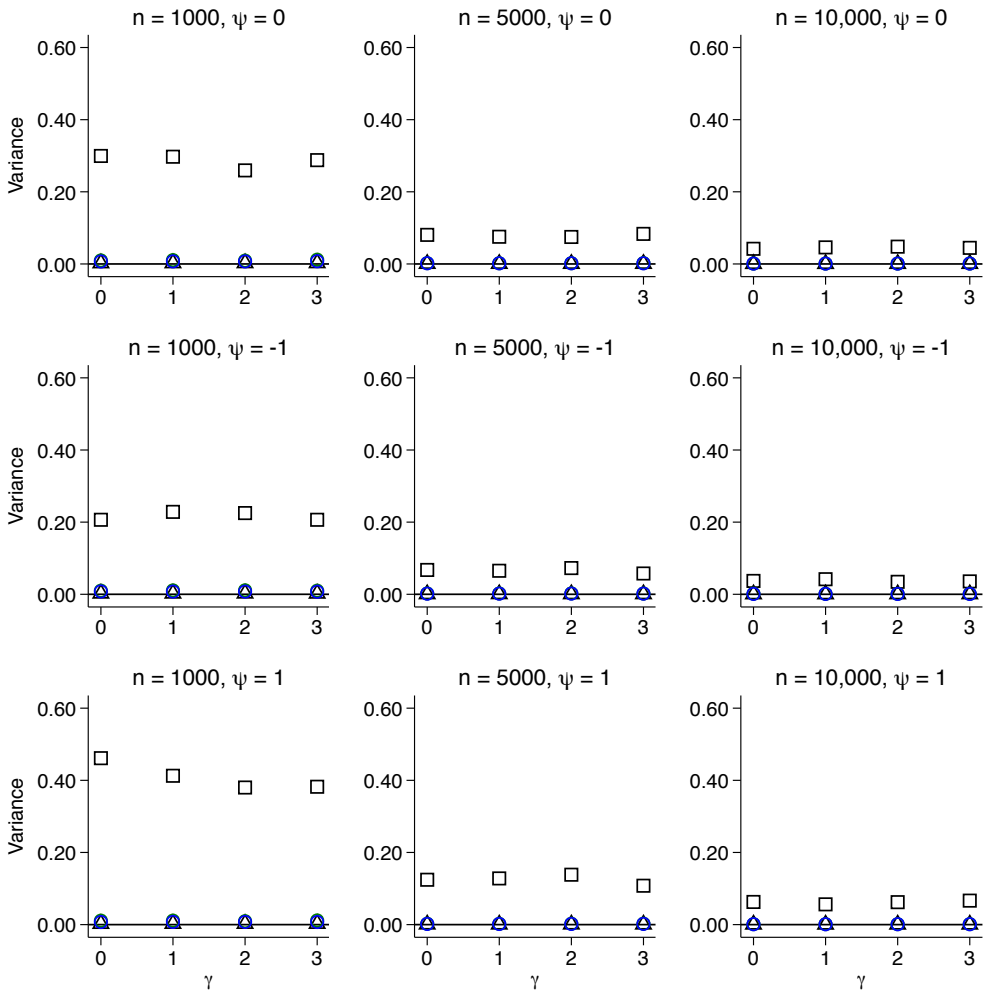
$\mu_i = \beta_0 + \sum\beta_j X_{j,i} + \psi X_{1,i}A_i + \gamma A_i$

$Y_i \sim \mathcal{N}(\mu_i,1)$

Bias


n = 1000, ψ = 0
n = 5000, ψ = 0
n = 10,000, ψ = 0
n = 1000, ψ = -1
n = 5000, ψ = -1
n = 10,000, ψ = -1
n = 1000, ψ = 1
n = 5000, ψ = 1
n = 10,000, ψ = 1
Bias
γ
0.1
0.0
-0.1
0
1
2
3

Variance


n = 1000, ψ = 0
n = 5000, ψ = 0
n = 10,000, ψ = 0
n = 1000, ψ = -1
n = 5000, ψ = -1
n = 10,000, ψ = -1
n = 1000, ψ = 1
n = 5000, ψ = 1
n = 10,000, ψ = 1
Variance
0.00
0.20
0.40
0.60
0
1
2
3
γ

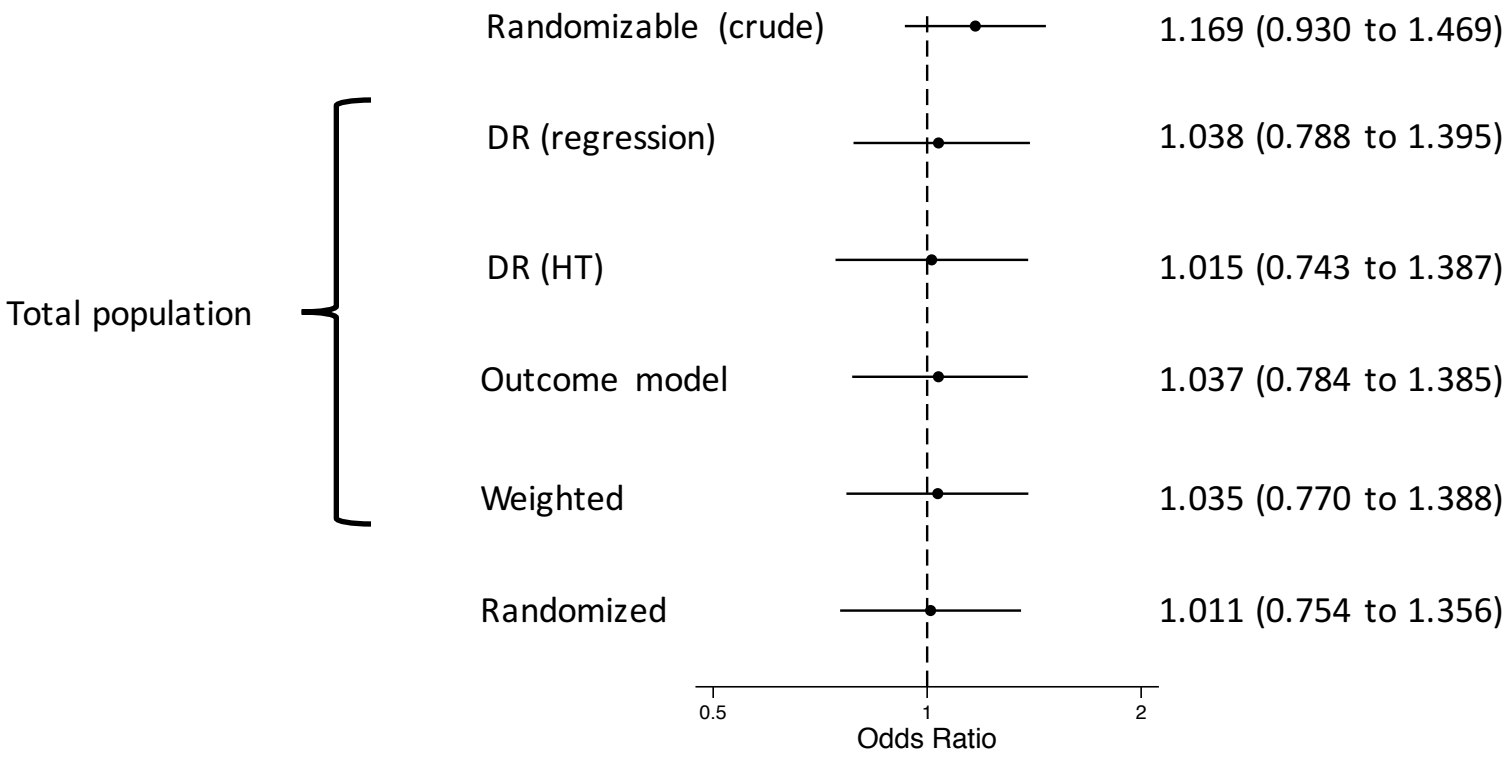
Applications

Data, Big and small

Small data: randomized preference design
Chronic coronary artery disease

- Coronary artery bypass grafting (CABG) vs. medical therapy
- 2099 met trial criteria:
  - 780 randomized (390 CABG; 390 medical)
  - 1319 self-selected (570 CABG; 745 medical)
- Outcome: death by year 24 (no dropout)

Treatment effects – CASS study

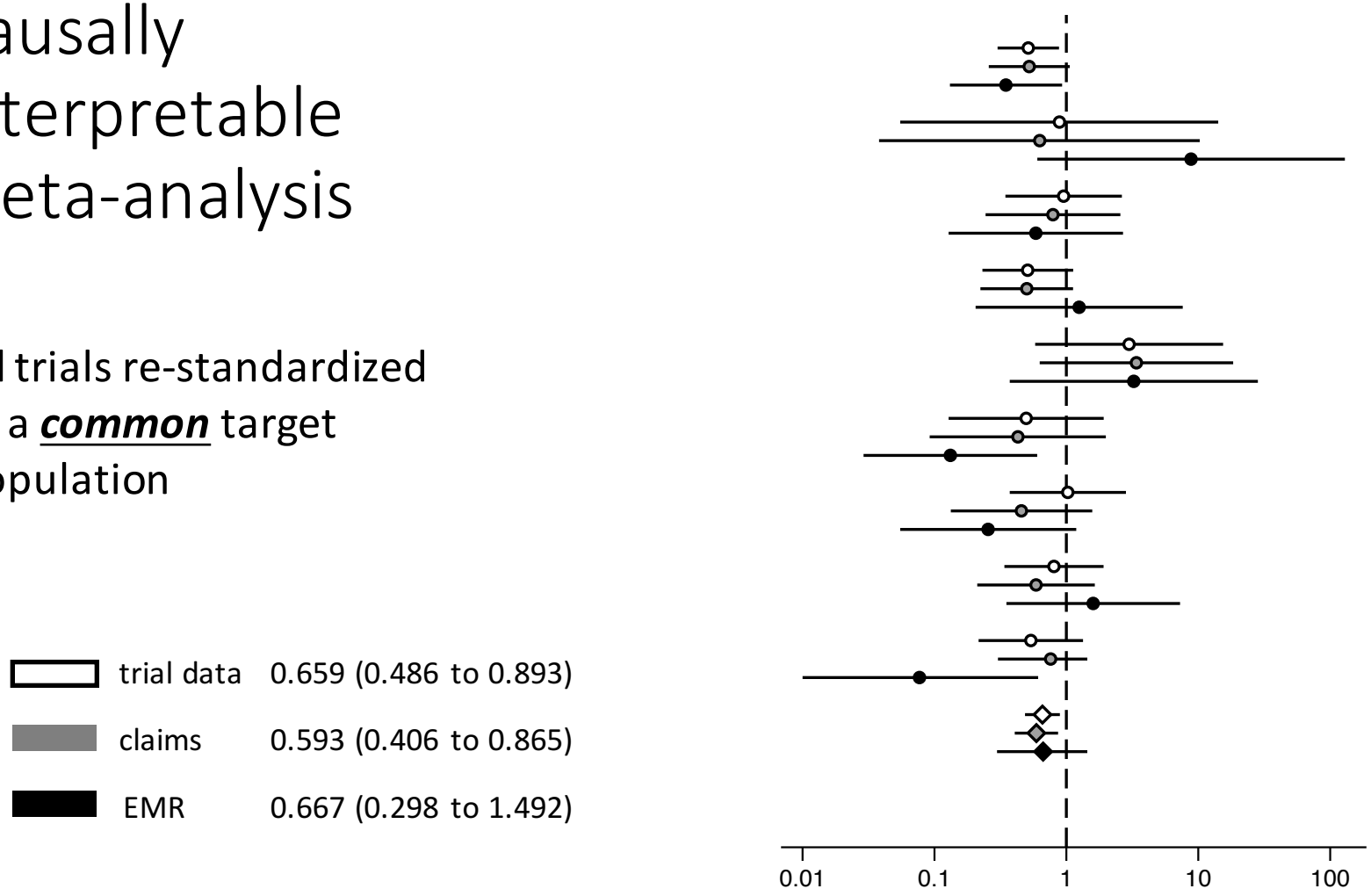


Big data: combined claims and EMR
Non-diabetic chronic kidney disease

- Angiotensin-converting inhibitors (ACE) vs. other anti-hypertensives (control)
- 9 trials, 791 ACEi vs 765 control
- Outcome: dialysis
- Optum Labs, chronic kidney disease cohort
  - 93,160 non-diabetic kidney disease patients
  - met trial inclusion-exclusion criteria (e.g., no diabetes, cancer, dementia)
  - Data on
    - Claims: age, sex, race, hypertension + comorbidities
    - EMR: age, sex, race, hypertension, BMI, height, weight, creatinine values, systolic + diastolic blood pressure, estimated glomerular filtration rate + comorbidities

Causally interpretable meta-analysis

All trials re-standardized to a ***common*** target population



How seriously should we take these results?

- Mean transportability and positivity are strong assumptions
  - Results can be thought as attempts to approximate the causal effect
  - Tools for informing judgments
  - Input to the design of a future study
- Transparent assumptions
- No "no confounding" assumption in the target (triangulation)
- Better covariate collection in trials (and observational studies)

Extensions

- Assessing applicability
  - Assessing whether the trial and target distribution are close enough
  - High dimensions (visualization + testing)
- Methods for extending trial findings
  - Additional data available in the target population
  - Behavior under model misspecification
  - Sensitivity analysis
  - Implications for study design and economic modeling

Concluding remarks

- Elements of a framework for assessing applicability and extending trial findings
  - Heroic assumptions needed (made explicit)
  - Transparency
- Methods for extending trial findings, with a view towards robustness
  - 2 real world applications
  - first causally interpretable meta-analysis ever
- Causal inference and evidence interpretation: “extrapolations are normally components of a complex web of interrelated evidence that must be considered together in assessing a hypothesis”

"…a highly standardized experiment supplies direct information only in respect to the narrow range of conditions achieved by standardization. ***Standardization, therefore, weakens rather than strengthens our ground for inferring a like result, when, as is invariably the case in practice, these conditions are somewhat varied.***"

R.A. Fisher
The design of experiments, 1935

Thank you!

- BSPH
  - Chris Schmid
  - Sarah Robertson
  - Iman Saeed
- HSPH
  - Miguel Hernan
  - James Robins
- John's Hopkins
  - Liz Stuart
- Tufts MC
  - David Kent
  - John Wong
  - Jason Nelson
  - Lesley Inker
- Duke
  - Liz Delong
- ICER
  - Moira Kapral
- PCORI
  - Emily Evans

Misspecified outcome model

$$
\begin{aligned}
\hat{\mu}_{a,DR} \quad &\rightarrow \quad \mathbb{E}\left[\frac{SI(A=a)Y}{w_a(X;\beta_a^*)}\right] + \mathbb{E}[g_a(X;\gamma_a^*)] - \mathbb{E}\left[\frac{SI(A=a)g_a(X;\gamma_a^*)}{w_a(X;\beta_a^*)}\right] \\
&\overset{(1)}{=} \quad \mathbb{E}[Y^a] + \mathbb{E}[g_a(X;\gamma_a^*)] - \mathbb{E}\left[\frac{S_i A_i g_a(X;\gamma_a^*)}{w_a(X;\beta_a^*)}\right] \\
&\overset{(2)}{=} \quad \mathbb{E}[Y^a] + \mathbb{E}[g_a(X;\gamma_a^*)] - \mathbb{E}_X\left[\frac{g_a(X;\gamma_a^*)}{w_a(X;\beta_a^*)}\mathbb{E}\left[SI(A=a)|X\right]\right] \\
&\overset{(3)}{=} \quad \mathbb{E}[Y^a] + \mathbb{E}[g_a(X;\gamma_a^*)] - \mathbb{E}\left[\frac{g_a(X;\gamma_a^*)}{w_a(X;\beta_a^*)}P(S=1|X)P(A=a|X,S=1)\right] \\
&= \quad \mathbb{E}[Y^a] + \mathbb{E}[g_a(X;\gamma_a^*)] - \mathbb{E}[g_a(X;\gamma_a^*)] \\
&= \quad \mathbb{E}[Y^a],
\end{aligned}
$$

Misspecified probability of participation model

$$
\begin{aligned}
\hat{\mu}_{a,DR} \quad \rightarrow \quad & \mathbb{E}\left[\frac{SI(A=a)Y^{obs}}{w_a(X;\beta_a^*)}\right] + \mathbb{E}[g_a(X;\gamma_a^*)] - \mathbb{E}\left[\frac{SI(A=a)g_a(X;\gamma_a^*)}{w_a(X;\beta_a^*)}\right] \\
\overset{(1)}{=} \quad & \mathbb{E}\left[\frac{SI(A=a)Y^{obs}}{w_a(X;\beta_a^*)}\right] + \mathbb{E}[Y^a] - \mathbb{E}\left[\frac{SI(A=a)g_a(X;\gamma_a^*)}{w_a(X;\beta_a^*)}\right] \\
\overset{(2)}{=} \quad & \mathbb{E}_X\left[\mathbb{E}\left[\frac{SI(A=a)Y^{obs}}{w_a(X;\beta_a^*)}|X\right]\right] + \mathbb{E}[Y^a] - \mathbb{E}_X\left[\mathbb{E}\left[\frac{SI(A=a)g_a(X;\gamma_a^*)}{w_a(X;\beta_a^*)}|X\right]\right] \\
= \quad & \mathbb{E}_X\left[\frac{P(S=1|X)P(A=a|X,S=1)}{w_a(X;\beta_a^*)}\mathbb{E}[Y^{obs}|A=a,X,S=1]\right] + \mathbb{E}[Y^a] \\
& -\mathbb{E}_X\left[\frac{g_a(X;\gamma_a^*)}{w_a(X;\beta_a^*)}\mathbb{E}\left[SI(A=a)|X\right]\right] \\
\overset{(3)}{=} \quad & \mathbb{E}_X\left[\frac{P(S=1|X)P(A=a|X,S=1)}{w_a(X;\beta_a^*)}\mathbb{E}[Y^{obs}|A=a,X,S=1]\right] + \mathbb{E}[Y^a] \\
& -\mathbb{E}_X\left[\frac{g_a(X;\gamma_a^*)}{w_a(X;\beta_a^*)}P(S=1|X)P(A=a|X,S=1)\right] \\
= \quad & \mathbb{E}[Y^a],
\end{aligned}
$$