# SAS® GLOBAL FORUM 2020

MARCH 29 – APRIL 1
WASHINGTON, DC

USERS PROGRAM

# Causal Effect Estimands: Interpretation, Identification, and Computation

**Reference:**

Thompson, C., Lamm, M., and Yung, Y.-F. (2019). “Causal Effect Estimands: Interpretation, Identification, and Computation.” In *Proceedings of the SAS Global Forum 2020 Conference*. Cary, NC: SAS Institute, Inc.

**Contact Information:**

clay.thompson@sas.com

**Causal Analysis in SAS/STAT 15.1:**

https://support.sas.com/rnd/app/stat/15.1/causal-analysis.pdf

Clay Thompson, Michael Lamm, Yiu-Fai Yung

Clay Thompson is a Senior Research Statistician Developer in the Multivariate Models Research Department at the SAS Institute, where he develops algorithms and software for the analysis of causal effects using graphical models. Prior to SAS, he worked as a quantitative systems pharmacologist in the pharmaceutical industry. He received a PhD in Applied Mathematics from North Carolina State University.

# Running Example

| Obs | ESL | PCGEd | ElmSize | MidSize | IncRatio | PreK | ElmProf | MidProf |
|---|---|---|---|---|---|---|---|---|
| 1 | 0 | 4 | 19.6 | 26.3 | 3.3 | 0 | 1 | 1 |
| 2 | 0 | 5 | 19.2 | 26.1 | 4.7 | 1 | 0 | 1 |
| 3 | 0 | 4 | 18.7 | 26.1 | 3.3 | 1 | 0 | 1 |
| 4 | 1 | 4 | 17.1 | 23.4 | 3.3 | 0 | 1 | 1 |
| 5 | 0 | 2 | 17.5 | 24.5 | 2.2 | 0 | 0 | 1 |
| 6 | 0 | 3 | 17.6 | 23.8 | 0.4 | 0 | 0 | 1 |
| 7 | 0 | 3 | 19.5 | 26.0 | 0.3 | 1 | 1 | 1 |
| 8 | 0 | 2 | 19.2 | 26.4 | 3.4 | 1 | 0 | 1 |
| 9 | 0 | 5 | 19.7 | 27.0 | 4.1 | 1 | 1 | 1 |
| 10 | 0 | 1 | 20.4 | 27.3 | 0.8 | 0 | 0 | 1 |

“Does enrollment in a PreK program improve school performance in later years?”

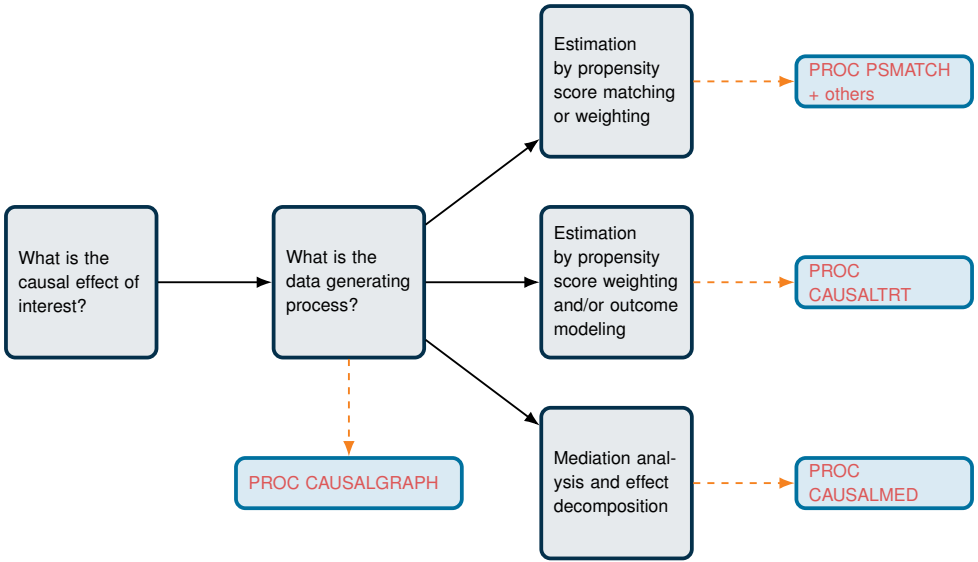
# Different research questions give rise to different causal estimands

- “What would be the effect of having every child enroll in PreK, compared to no child enrolled in PreK?”
  - Average Treatment Effect (ATE)
- “What is the effect of a PreK program among those children who were enrolled?”
  - Average Treatment Effect for the Treated (ATT)
- “What cost is incurred by students who are not enrolled in a PreK program?”
  - Average Treatment Effect for the Untreated (ATU)
- “How much of the total effect is (not) attributable to improved academic performance in early grades?”
  - Natural Indirect Effect (NIE), Natural Direct Effect (NDE)

# The SAS/STAT® product contains four procedures that are specifically designed for causal inference

| Procedure | Estimands | Primary Use | Release (Year) |
|---|---|---|---|
| PSMATCH | ATE, ATT, ATU | • Creation of matched data sets<br>• Assessing covariate balance in matched, weighted, or stratified data | SAS 9.4M4, SAS/STAT 14.2 (2016) |
| CAUSALTRT | ATE, ATT, ATU | • Direct estimation of a treatment effect | SAS 9.4M4, SAS/STAT 14.2 (2016) |
| CAUSALMED | ATE, NIE, NDE | • Causal mediation and related effects | SAS 9.4M5, SAS/STAT 14.3 (2017) |
| CAUSALGRAPH | Identification analysis | • Analysis of graphical causal models | SAS 9.4M6, SAS/STAT 15.1 (2018) |

# A simplified causal analysis workflow



# A well-considered workflow is essential to the clarity of your causal analysis

- Estimands in the Potential Outcomes Framework
  - Definitions and Assumptions
  - Causal Effect Estimands
- Using Data to Estimate Causal Effects
  - Randomization
  - Confounding
  - Identification
- Example: Estimating the Average Treatment Effect for the Treated
  - Identification Analysis with PROC CAUSALGRAPH
  - Effect Estimation with PROC PSMATCH
  - Effect Estimation with PROC CAUSALTRT
- Review of Key Themes

# Estimands in the Potential Outcomes Framework

# For simplicity, this presentation assumes a binary treatment variable

- **Treatment** $T$
  - Possible values $t$
  - Assume binary: $t \in \{0, 1\}$
- **Potential outcomes** $Y_t$
  - $Y_0$ and $Y_1$ for a binary treatment
  - Typically $Y_t = Y_t(Z)$, for some set of covariates $Z$

# Some assumptions are necessary in order to define the causal estimands

- **SUTVA**: the stable unit treatment value assumption
  - No hidden levels of treatment
  - No interference between subjects

- **Consistency**: $Y = Y_t$ if $T = t$
  - Treated: $T = 1 \Rightarrow Y = Y_1$
  - Untreated: $T = 0 \Rightarrow Y = Y_0$

# The SUTVA and consistency assumptions have practical implications

- No hidden levels of treatment
  - Are all PreK programs similarly effective?
  - Is the effect different for full-time vs. part time enrollment?
- No interference between subjects
  - To what extent are behaviors/norms collectivized?
  - Does this change with the density of enrollment?
- Consistent treatment assignment
  - Does the data reflect enrollment or attendance?

# Different research questions give rise to different causal estimands

(revisited)

- "What would be the effect of having every child enroll in PreK, compared to no child enrolled in PreK?"
  - Average Treatment Effect (ATE)
  - $\text{ATE} = \text{E}[Y_1 - Y_0] = \text{E}[Y_1] - \text{E}[Y_0]$
- "What is the effect of a PreK program among those children who were enrolled?"
  - Average Treatment Effect for the Treated (ATT)
  - $\text{ATT} = \text{E}[Y_1 - Y_0 \mid T = 1] = \text{E}[Y_1 \mid T = 1] - \text{E}[Y_0 \mid T = 1]$
- "What cost is incurred by students who are not enrolled in a PreK program?"
  - Average Treatment Effect for the Untreated (ATU)
  - $\text{ATU} = \text{E}[Y_1 - Y_0 \mid T = 0] = \text{E}[Y_1 \mid T = 0] - \text{E}[Y_0 \mid T = 0]$
- "How much of the total effect is (not) attributable to improved academic performance in early grades?"
  - Natural Indirect Effect (NIE), Natural Direct Effect (NDE)
  - $\text{NIE} = \text{E}[Y_{1M_1} - Y_{1M_0}] = \text{E}[Y_{1M_1}] - \text{E}[Y_{1M_0}]$
  - $\text{NDE} = \text{E}[Y_{1M_0} - Y_{0M_0}] = \text{E}[Y_{1M_0}] - \text{E}[Y_{0M_0}]$

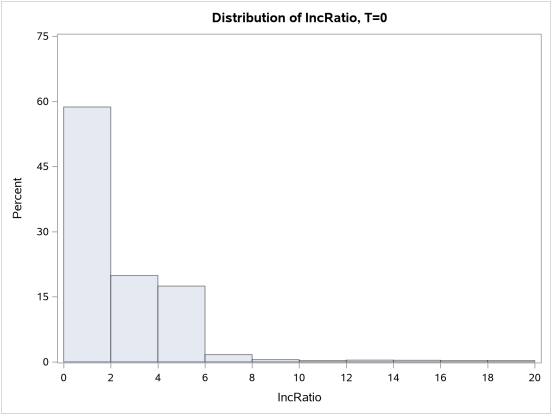
# The observed outcome is determined by the treatment received

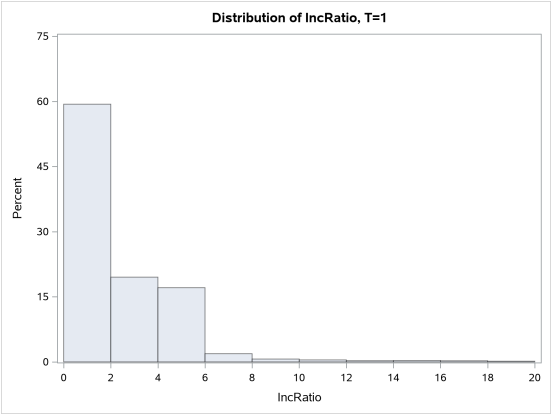
| Obs | $T$ | IncRatio | PCGEd | ESL | $Y_1$ | $Y_0$ | $Y$ |
|---|---|---|---|---|---|---|---|
| 1 | 0 | 3.3 | 4 | 0 | ? | 1 | 1 |
| 2 | 1 | 4.7 | 5 | 0 | 1 | ? | 1 |
| 3 | 1 | 3.3 | 4 | 0 | 0 | ? | 0 |
| 4 | 0 | 3.3 | 4 | 1 | ? | 1 | 1 |
| 5 | 1 | 2.2 | 2 | 0 | 1 | ? | 1 |
| 6 | 0 | 0.4 | 3 | 0 | ? | 0 | 0 |
| 7 | 0 | 0.3 | 1 | 0 | ? | 1 | 1 |
| 8 | 1 | 3.4 | 5 | 0 | 1 | ? | 1 |

- Consistency implies $Y = Y_t$ if $T = t$
  - Half of the potential outcomes are missing
  - "Causal inference is a missing data problem"
- How do you estimate a causal effect?
  - **Identification**: use observed data to construct an estimator

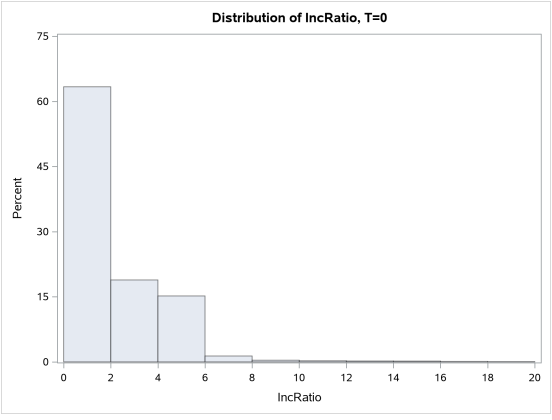
# Using Data to Estimate Causal Effects

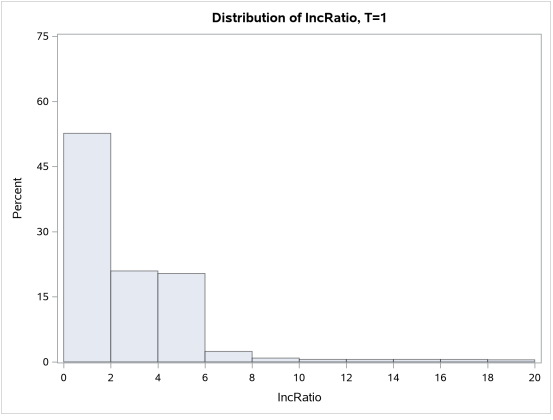
# Randomization creates groups that are exchangeable (on average)



# In observational studies, groups are no longer exchangeable



# Identification conditions are required for valid causal effect estimation

- **SUTVA**: the stable unit treatment value assumption
  - No hidden levels of treatment
  - No interference between subjects
- **Consistency**: $Y = Y_t$ if $T = t$
- **Positivity**: $P(T = t \mid \boldsymbol{X} = \boldsymbol{x}) > 0 \; \forall \; t, x$
- **Conditional Exchangeability**: $T \perp\!\!\!\perp Y_t \mid \boldsymbol{X}$

Conditional on $\boldsymbol{X} = \boldsymbol{x}$,
subjects are “as if randomized”

(**Conditional Ignorability**: Conditional Exchangeability + Positivity)

# Example: Estimating the Average Treatment Effect for the Treated

# What is the effect of a PreK program among those children who were enrolled?

| Obs | ESL | PCGEd | ElmSize | MidSize | IncRatio | PreK | ElmProf | MidProf |
|---|---|---|---|---|---|---|---|---|
| 1 | 0 | 4 | 19.6 | 26.3 | 3.3 | 0 | 1 | 1 |
| 2 | 0 | 5 | 19.2 | 26.1 | 4.7 | 1 | 0 | 1 |
| 3 | 0 | 4 | 18.7 | 26.1 | 3.3 | 1 | 0 | 1 |
| 4 | 1 | 4 | 17.1 | 23.4 | 3.3 | 0 | 1 | 1 |
| 5 | 0 | 2 | 17.5 | 24.5 | 2.2 | 0 | 0 | 1 |
| 6 | 0 | 3 | 17.6 | 23.8 | 0.4 | 0 | 0 | 1 |
| 7 | 0 | 3 | 19.5 | 26.0 | 0.3 | 1 | 1 | 1 |
| 8 | 0 | 2 | 19.2 | 26.4 | 3.4 | 1 | 0 | 1 |
| 9 | 0 | 5 | 19.7 | 27.0 | 4.1 | 1 | 1 | 1 |
| 10 | 0 | 1 | 20.4 | 27.3 | 0.8 | 0 | 0 | 1 |

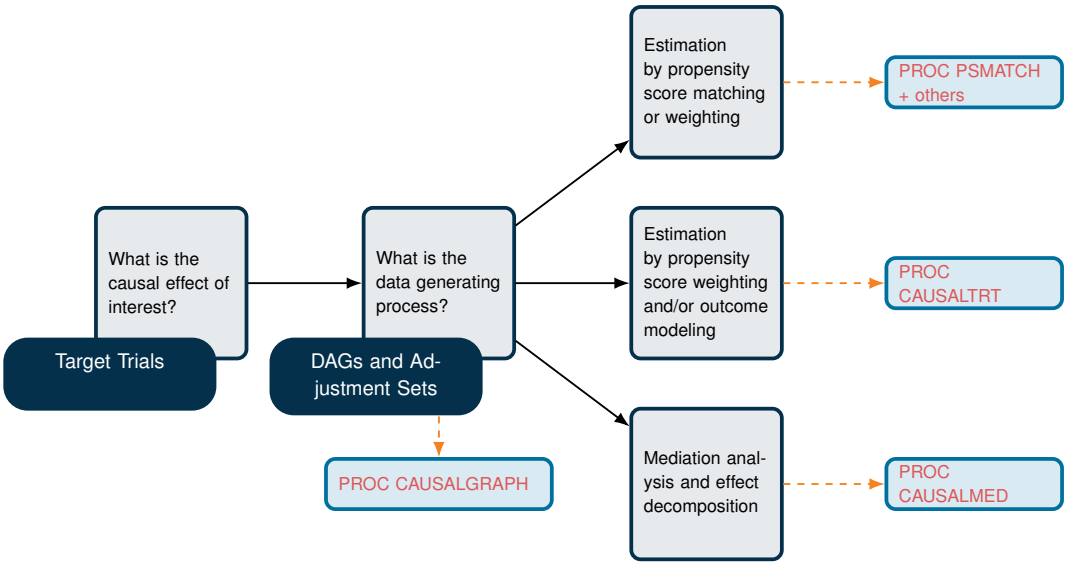
ATT = 0.0356

# This research question is answered by the ATT

- $\text{ATT} = \text{E}[Y_1 - Y_0 \mid T = 1] = \text{E}[Y_1 \mid T = 1] - \text{E}[Y_0 \mid T = 1]$
  - Subpopulation effect (those who enroll)
  - Useful to evaluate a pilot program

- Is the ATT identified by the available data?
  - SUTVA
  - Consistency
  - Conditional Ignorability

# Target trials and DAGs are useful to assess causal assumptions



# A “target trial” is a useful framework for creating well-defined causal questions

“Does enrollment in a PreK program improve school performance in later years?”

- Who will be studied?
  - What are the inclusion criteria?
  - What are the exclusion criteria?
- What intervention will be considered?
  - Will the exposure be similar for all treated subjects?
  - Is there a clear distinction between treated and untreated subjects?
- What are the study endpoints?
  - When will they be collected?
  - How do they relate to the outcome?

# The refined causal question is more likely to satisfy SUTVA and consistency

"Does full-time enrollment in a specific PreK program improve reading performance as measured by a standardized assessment of reading proficiency at the end of eighth grade?"

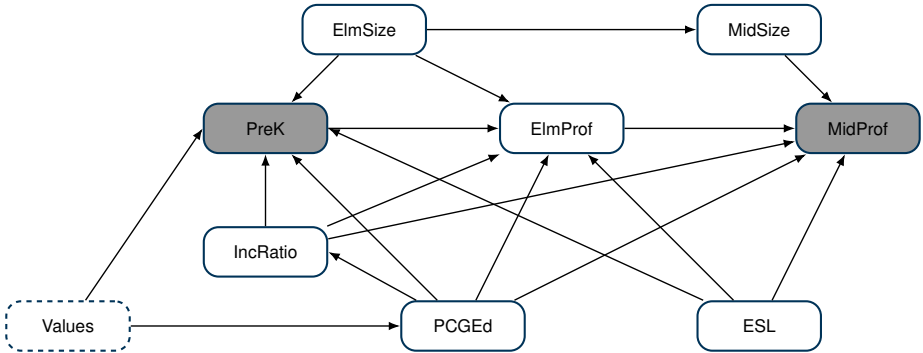
For more information about target trials, see Hernán and Robins (2016).

# Which set(s) of covariates satisfy the conditional ignorability assumption?

Conditional ignorability: $Y_t \perp\!\!\!\perp T | \boldsymbol{X}$

| Obs | ESL | PCGEd | ElmSize | MidSize | IncRatio | PreK | ElmProf | MidProf |
|---|---|---|---|---|---|---|---|---|
| 1 | 0 | 4 | 19.6 | 26.3 | 3.3 | 0 | 1 | 1 |
| 2 | 0 | 5 | 19.2 | 26.1 | 4.7 | 1 | 0 | 1 |
| 3 | 0 | 4 | 18.7 | 26.1 | 3.3 | 1 | 0 | 1 |
| 4 | 1 | 4 | 17.1 | 23.4 | 3.3 | 0 | 1 | 1 |
| 5 | 0 | 2 | 17.5 | 24.5 | 2.2 | 0 | 0 | 1 |
| 6 | 0 | 3 | 17.6 | 23.8 | 0.4 | 0 | 0 | 1 |
| 7 | 0 | 3 | 19.5 | 26.0 | 0.3 | 1 | 1 | 1 |
| 8 | 0 | 2 | 19.2 | 26.4 | 3.4 | 1 | 0 | 1 |
| 9 | 0 | 5 | 19.7 | 27.0 | 4.1 | 1 | 1 | 1 |
| 10 | 0 | 1 | 20.4 | 27.3 | 0.8 | 0 | 0 | 1 |

# A causal graph represents assumptions about the data generating process




For more information about using expert judgment to form a causal diagram, see Hanea et al. (2018).

# Use PROC CAUSALGRAPH to find valid adjustment sets

```
proc causalgraph;
   model "ReadingProf"
      ElmProf => MidProf, MidSize => MidProf,
      ElmSize => ElmProf MidSize PreK,
      ESL IncRatio => ElmProf MidProf PreK,
      PCGEd => ElmProf IncRatio MidProf PreK,
      PreK => ElmProf, Values => PCGEd PreK;
   latent Values;
   identify PreK => MidProf;
run;
```

# Each valid adjustment set represents a possible identification strategy

| Covariate Adjustment Sets for ReadingProf | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Causal Effect of PreK on MidProf | | | | | | | | |
| | | | Covariates | | | | | |
| | Size | Minimal | ElmProf | ElmSize | ESL | IncRatio | MidSize | PCGEd |
| 1 | 4 | Yes | | * | * | * | | * |
| 2 | 5 | No | | * | * | * | * | * |

A valid adjustment set satisfies $T \perp\!\!\!\perp Y_t \mid \boldsymbol{X}$

For more information about PROC CAUSALGRAPH, see Thompson (2019).

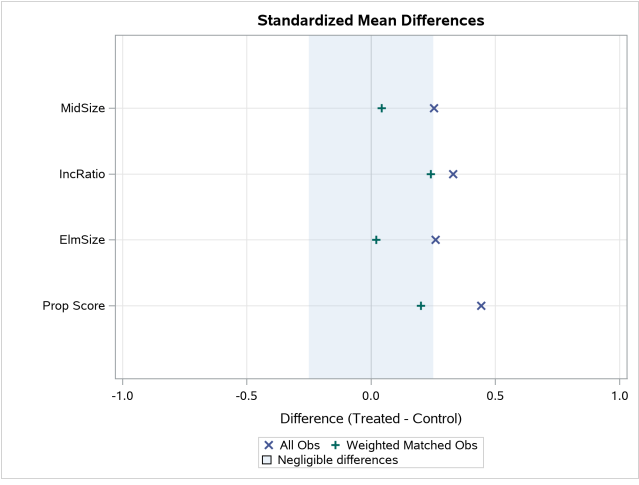
# You can use PROC PSMATCH to produce a matching analysis

```
proc psmatch data=ReadingObs region=treated;
   class ESL PCGEd PreK;
   psmodel PreK(treated="1") = ElmSize ESL IncRatio PCGEd;
   match method = optimal caliper=.;
   assess ps var=(ElmSize IncRatio MidSize) /
          plots(nodetails)=(stddiff box(display=(PS IncRatio)));
   output out=psATTMatchData weight=attWgt;
run;
```

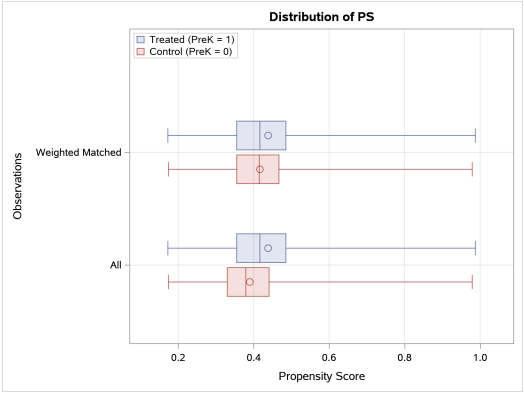
For more information about PROC PSMATCH, see:

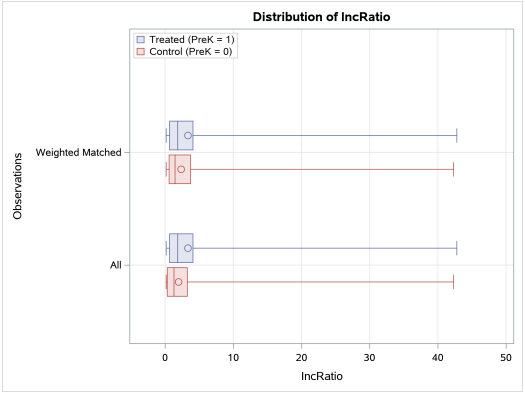
- Yuan, Yung, and Stokes (2017)
- Lamm, Thompson, and Yung (2019)

# Standardized mean differences diagnostics show a modest improvement in balance after matching

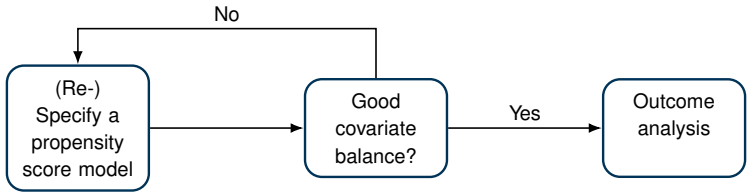


# Covariate balance diagnostics also show modest improvement



# Matching is an iterative process that excludes the outcome variable



# Use the PSMATCH output data set to estimate the causal effect

```
proc ttest data=psATTMatchData;
   class PreK;
   var MidProf;
   weight attWgt;
run;
```

| PreK | Method | Mean | 95% CL Mean | | Std Dev | 95% CL Std Dev | |
|---|---|---|---|---|---|---|---|
| 0 | | 0.7574 | 0.7443 | 0.7706 | 0.4287 | 0.4196 | 0.4382 |
| 1 | | 0.7987 | 0.7864 | 0.8110 | 0.4010 | 0.3925 | 0.4099 |
| Diff (1-2) | Pooled | -0.0412 | -0.0592 | -0.0233 | 0.4151 | 0.4088 | 0.4215 |
| Diff (1-2) | Satterthwaite | -0.0412 | -0.0592 | -0.0233 | | | |

# You can use CAUSALTRT to estimate the ATT by inverse probability weighting

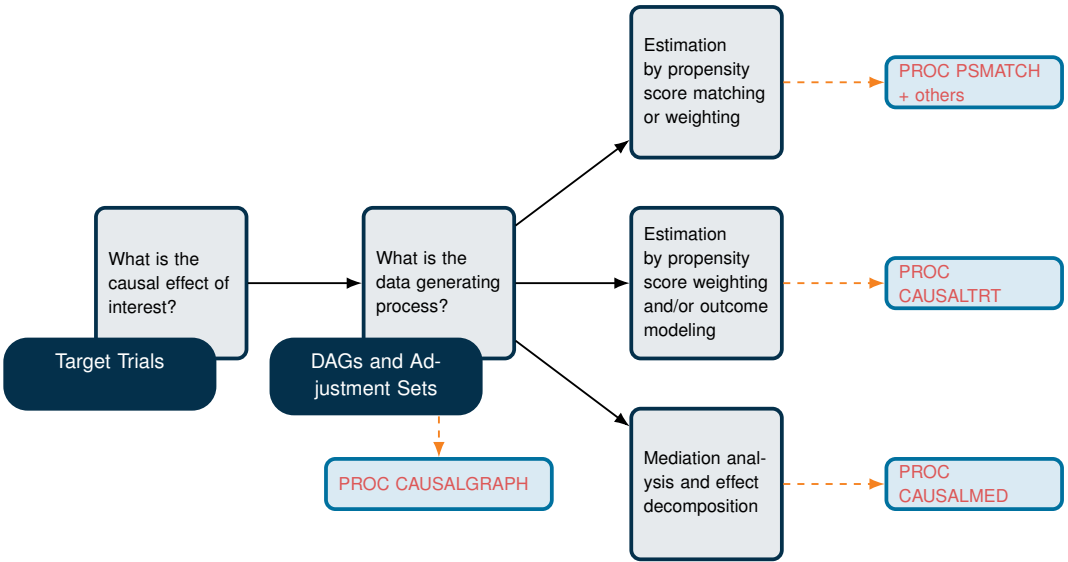
```
proc causaltrt data=ReadingObs att;
   class ESL MidProf PCGEd PreK / desc;
   psmodel PreK = ElmSize ESL IncRatio PCGEd;
   model MidProf;
   bootstrap seed=1976;
run;
```

**Analysis of Causal Effect**

| Parameter | Treatment Level | Estimate | Robust Std Err | Bootstrap Std Err | Wald 95% Confidence Limits | | Bootstrap Bias Corrected 95% Confidence Limits | | Z | Pr > \|Z\| |
|---|---|---|---|---|---|---|---|---|---|---|
| POM | 1 | 0.7987 | 0.00626 | 0.00611 | 0.7864 | 0.8110 | 0.7862 | 0.8106 | 127.51 | <.0001 |
| POM | 0 | 0.7644 | 0.00636 | 0.00643 | 0.7519 | 0.7768 | 0.7517 | 0.7774 | 120.26 | <.0001 |
| ATT | | 0.03430 | 0.00883 | 0.00889 | 0.01700 | 0.05160 | 0.01540 | 0.05075 | 3.89 | 0.0001 |

# Review of Key Themes

# Target trials and DAGs are useful to assess causal assumptions



**Reference:**

Thompson, C., Lamm, M., and Yung, Y.-F. (2019). “Causal Effect Estimands: Interpretation, Identification, and Computation.” In *Proceedings of the SAS Global Forum 2020 Conference*. Cary, NC: SAS Institute, Inc.

**Contact Information:**

clay.thompson@sas.com

**Causal Analysis in SAS/STAT 15.1:**

https://support.sas.com/rnd/app/stat/15.1/causal-analysis.pdf

# References I


Hanea, A. M., McBride, M. F., Burgman, M. A., and Wintle, B. C. (2018). "Classical Meets Modern in the IDEA Protocol for Structured Expert Judgement." *Journal of Risk Research* 21:417–433.

Hernán, M. A., and Robins, J. M. (2016). "Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available." *American Journal of Epidemiology* 183:758–764.

Lamm, M., Thompson, W. C., and Yung, Y.-F. (2019). "Building a Propensity Score Model with SAS/STAT Software: Planning and Practice." In *Proceedings of the SAS Global Forum 2019 Conference*. Cary, NC: SAS Institute Inc. `https://www.sas.com/content/dam/SAS/support/en/sas-global-forum-proceedings/2019/3056-2019.pdf`.

Thompson, W. C. (2019). "Causal Graph Analysis with the CAUSALGRAPH Procedure." In *Proceedings of the SAS Global Forum 2019 Conference*. Cary, NC: SAS Institute Inc. `https://www.sas.com/content/dam/SAS/support/en/sas-global-forum-proceedings/2019/2998-2019.pdf`.

# References II

Yuan, Y., Yung, Y.-F., and Stokes, M. (2017). “Propensity Score Methods for Causal Inference with the PSMATCH Procedure.” In *Proceedings of the SAS Global Forum 2017 Conference*. Cary, NC: SAS Institute Inc. `http://support.sas.com/resources/papers/proceedings17/SAS0332-2017.pdf`.

# The SAS/STAT® product contains four procedures that are specifically designed for causal inference

| Procedure | Estimands | Primary Use | Release (Year) |
|---|---|---|---|
| PSMATCH | ATE, ATT, ATU | • Creation of matched data sets<br>• Assessing covariate balance in matched, weighted, or stratified data | SAS 9.4M4, SAS/STAT 14.2 (2016) |
| CAUSALTRT | ATE, ATT, ATU | • Direct estimation of a treatment effect | SAS 9.4M4, SAS/STAT 14.2 (2016) |
| CAUSALMED | ATE, NIE, NDE | • Causal mediation and related effects | SAS 9.4M5, SAS/STAT 14.3 (2017) |
| CAUSALGRAPH | Identification analysis | • Analysis of graphical causal models | SAS 9.4M6, SAS/STAT 15.1 (2018) |


MARCH 29 – APRIL 1 | WASHINGTON, DC | #SASGF

# PROC CAUSALGRAPH can be used to find valid adjustment sets

- List valid identification strategies or test a user-specified strategy
- Multiple identification criteria:
  - Constructive backdoor criterion
  - Backdoor criterion
  - Conditional instrumental variable criterion
- Specify unmeasured, latent, or excluded variables
- Explore causal and noncausal paths between treatments and outcomes
- Enumerate observationally testable assumptions in a model
- Analyze multiple models simultaneously
- Analyze any static treatment regime (multiple treatments, multiple outcomes)

# PROC PSMATCH can build propensity score models for weighting or matching

- Create propensity scores by fitting a logistic regression model
- Input propensity scores created by some other method
- Compute ATE or ATT weights
- Multiple matching strategies, including caliper restrictions
  - Greedy nearest neighbor
  - Optimal matching (fixed ratio, variable ratio, full)
  - Matching with replacement
- Many tabular and graphical assessments of covariate balance

# PROC CAUSALTRT can estimate total causal effects directly

- Same theoretical foundation and similar functionality to PSMATCH
- Create propensity scores by fitting a logistic regression model
- Use computed propensity scores to estimate causal effects directly
- Supports outcome regression to estimate causal effects
- Supports doubly robust estimation

# PROC CAUSALMED can perform a causal mediation analysis

- Maximum likelihood estimation of causal mediation effects by regression adjustment methods
- Bootstrap and asymptotic estimation of standard errors and confidence intervals
- Support of various data types:
  - continuous or binary outcome, treatment, and mediator variables
  - continuous or categorical covariates
- Several two-, three-, and four-way decompositions of total effects
- Flexible evaluation of controlled direct effects and conditional mediation effects
- Analysis with case-control design