

# Propensity Scores

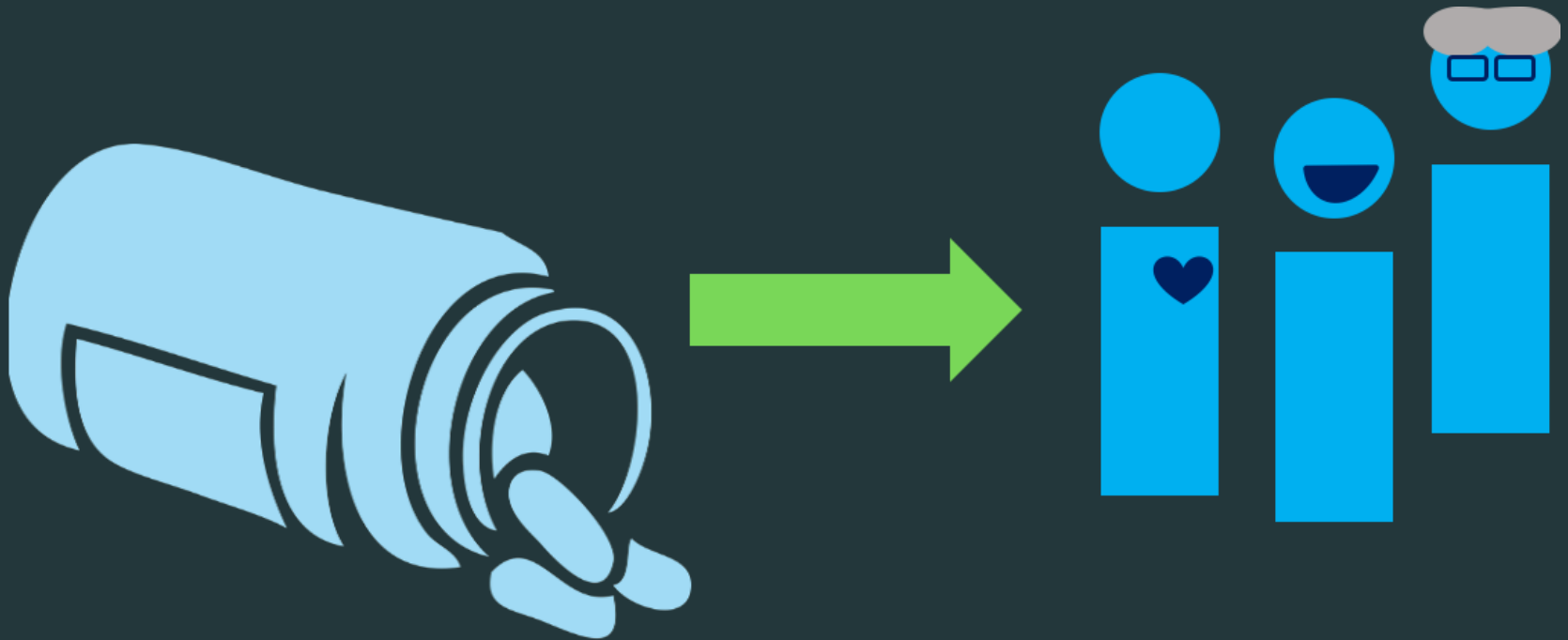
Lucy D'Agostino McGowan

Wake Forest University

2021-09-01 (updated: 2022-06-07)

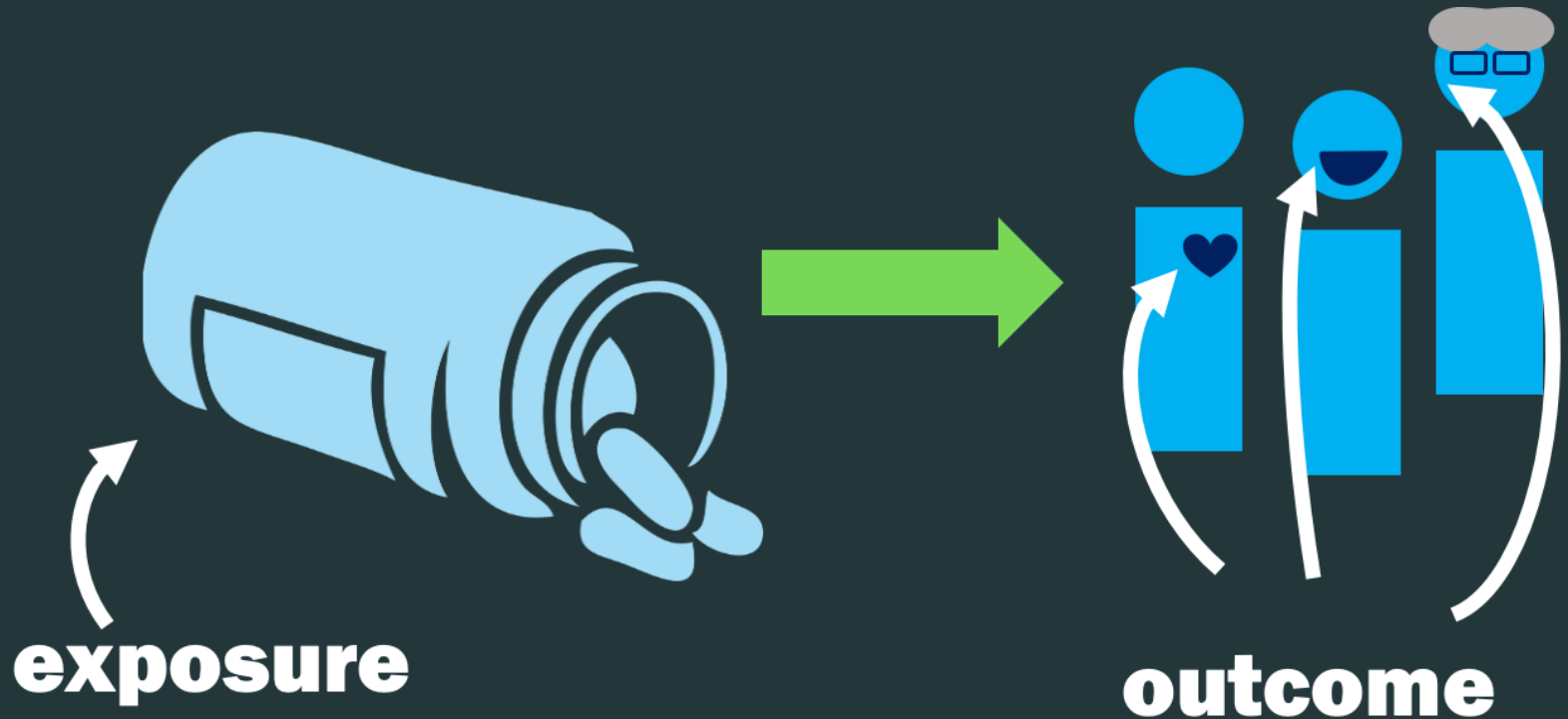
# Observational Studies

**Goal:** To answer a research question



# Observational Studies

**Goal:** To answer a research question



# Observational Studies

## Randomized Controlled Trial

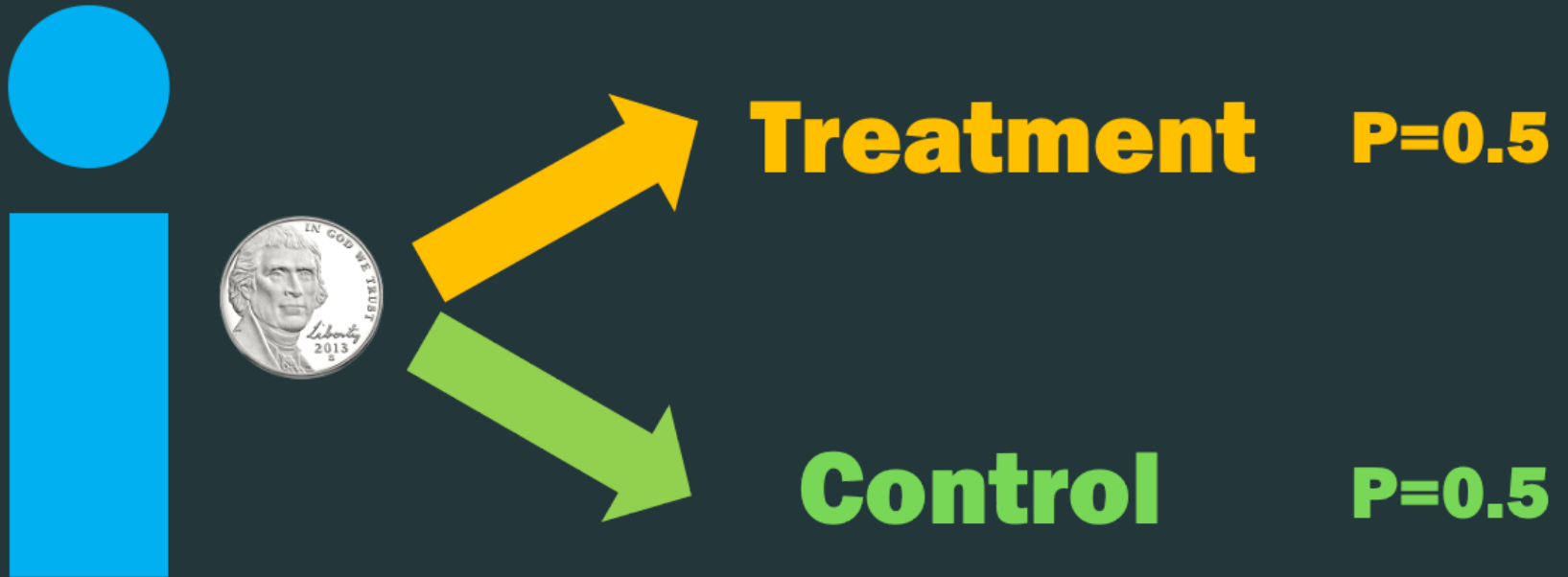


**Treatment**

**Control**

# Observational Studies

## Randomized Controlled Trial



# Observational Studies

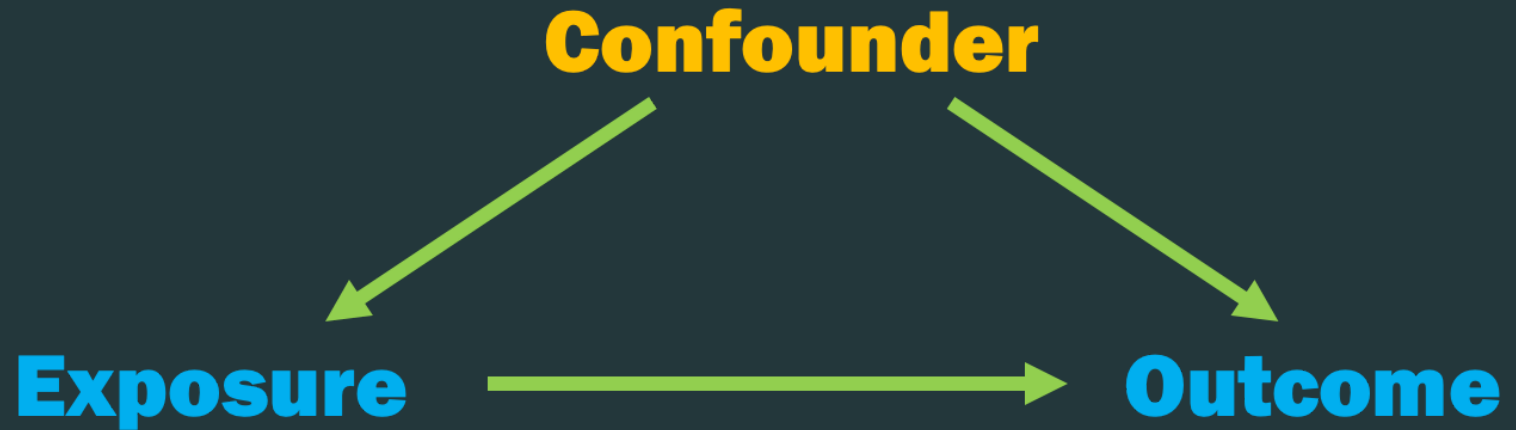




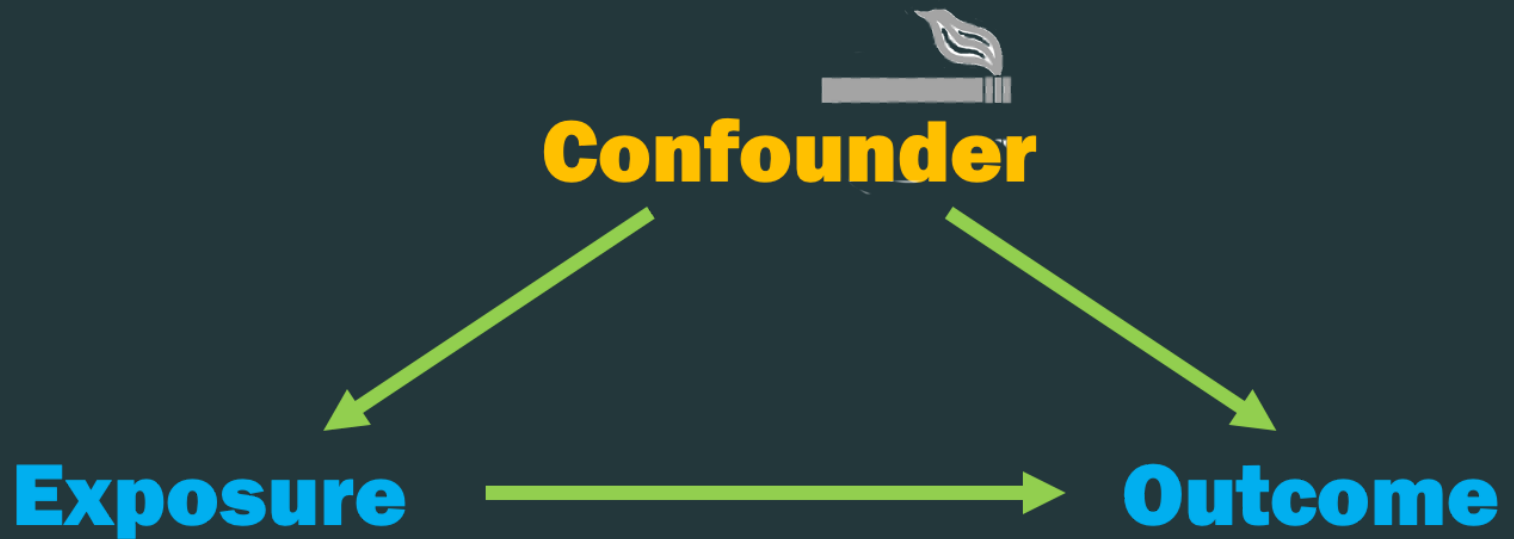




# Confounding



# Confounding



# Propensity scores

Rosenbaum and Rubin showed in observational studies, conditioning on **propensity scores** can lead to unbiased estimates of the exposure effect

- 1 There are no unmeasured confounders**
- 2 Every subject has a nonzero probability of receiving either exposure**

# Propensity scores

Fit a **logistic regression** predicting exposure using known covariates

$$Pr(exposure = 1) = \frac{1}{1 + \exp(-X\beta)}$$

Each individuals' predicted values are the **propensity scores**

# Propensity scores

```
library(tidyverse)  
library(broom)
```

# Propensity scores

```
glm(exposure ~ confounder_1 + confounder_2 + confounder_3 + ...,  
     data = df,  
     family = binomial())
```

# Propensity scores

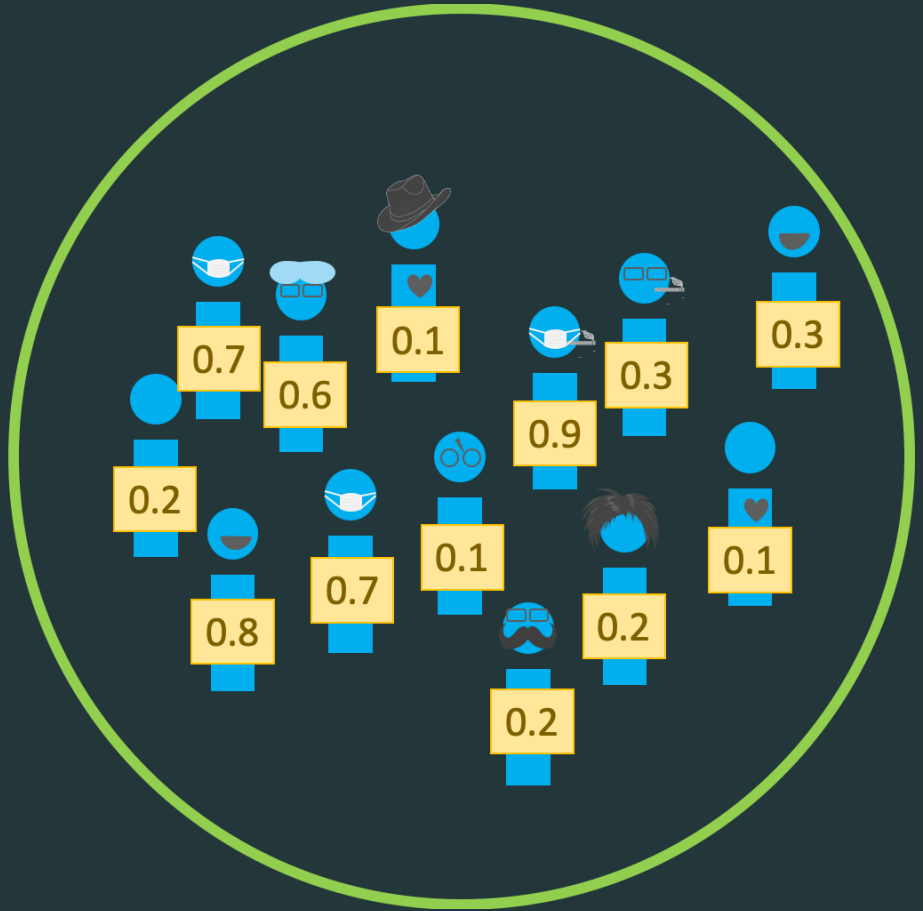
```
glm(exposure ~ confounder_1 + confounder_2 + confounder_3 + ...,  
     data = df,  
     family = binomial()) %>%  
augment(type.predict = "response", data = df)
```

# Propensity scores

```
glm(exposure ~ confounder_1 + confounder_2 + confounder_3 + ...,  
     data = df,  
     family = binomial()) %>%  
augment(type.predict = "response", data = df)
```



# Propensity scores





# Your turn

- 1 Using the **confounders** identified in the previous DAG, fit a propensity score model for qsmk
- 2 Stretch: Create two histograms, one of the propensity scores for those that quit smoking and one for those that do not

10:00