

# **The Chronicles of Non- Randomized Studies: the Bias, the Fix, and the Propensity Score**

**SEMEL INSTITUTE FOR NEUROSCIENCE  
AND HUMAN BEHAVIOR**

**Center for Community Health  
January 17 , 2006**

**Scott Comulada, DrPH**

Thanks to Naihua Duan, PhD, for introducing me to propensity score methodology.



“Uh, oh, I think they just took a picture of me.”

# Example: Study on mothers with children



**Low income**



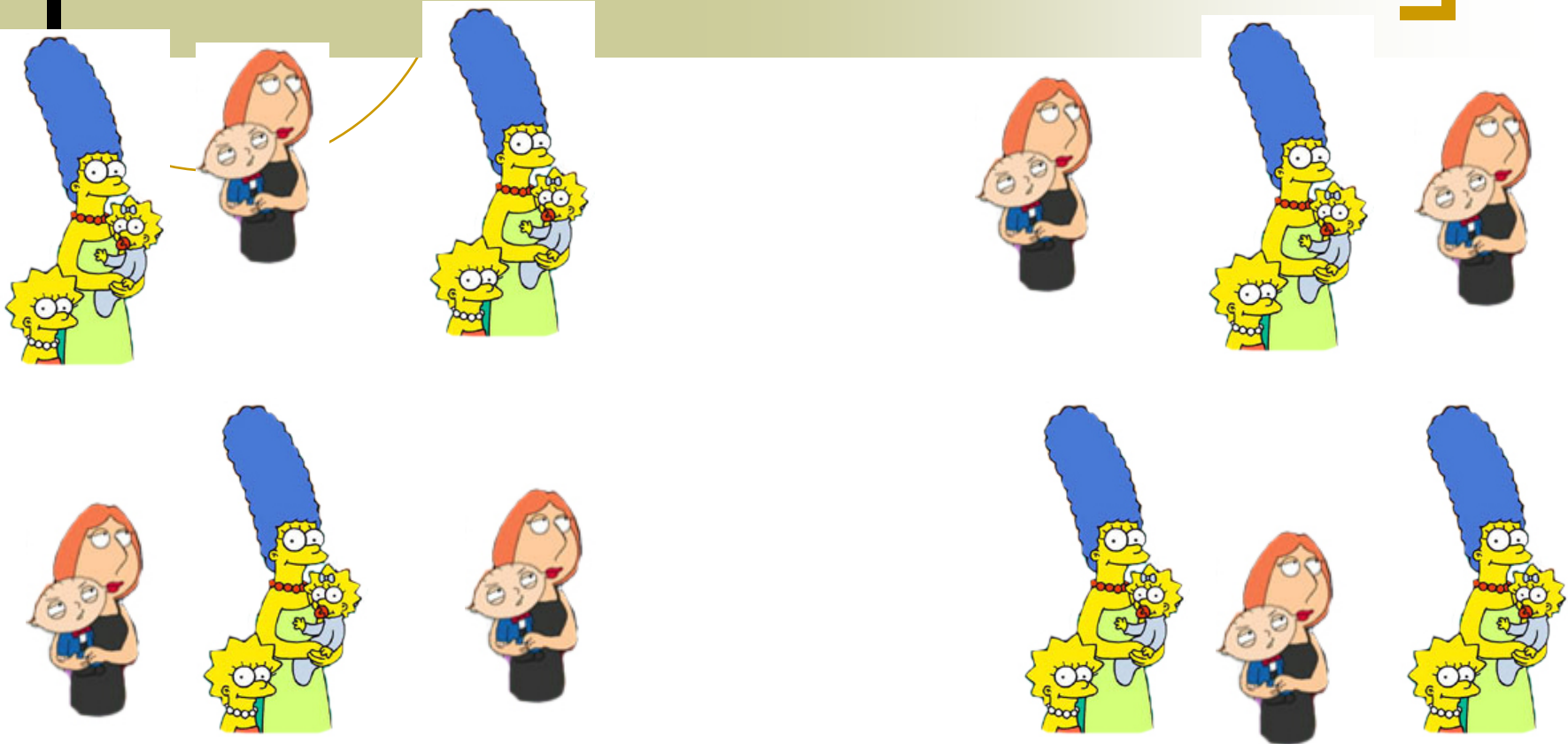
**High income**

# Randomized study

Sample characteristics evenly allocated across treatments

**Intervention**

**Control**

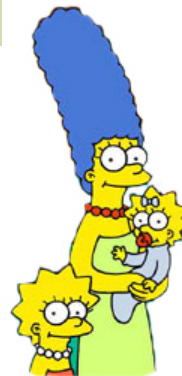


# Observational study

Even allocation of sample characteristics not possible


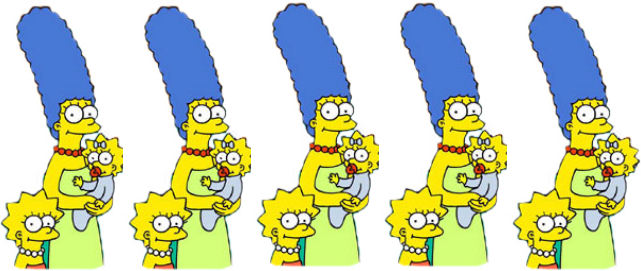
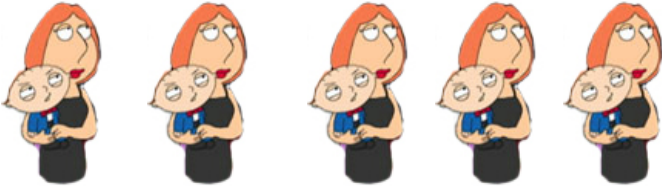

**Intervention**

**Control**



# One solution

- Achieve even allocation of sample characteristics within strata
- Perform analyses within strata
- Combine analysis results from each strata

	Intervention	Control
Stratum 1		
Stratum 2		

# [ Create strata ]

---

- Use propensity score (Rosenbaum and Rubin 1983)
- Definition: Conditional probability an individual is in intervention group given a set of covariates
- Ex: Conditional probability mothers and children in intervention group given income status

# [ Calculating propensity score ]

- Fit logistic regression of indicator for treatment status to relevant covariates
- Obtain predicted probabilities of being in intervention group for each individual
- Rank individuals on predicted probabilities
- Create 5 strata with cut points at 20<sup>th</sup>, 40<sup>th</sup>, 60<sup>th</sup>, and 80<sup>th</sup> quantiles
  - Why 5 strata? Removes approx. 90% of bias from covariates (Rosenbaum and Rubin 1984)

# [ Perform analyses ]

- Obtain 5 stratum-specific parameter estimates
- Final estimate is weighted average of 5 stratum-specific estimates
- Weights =  $n_i / N$ ,
  - $n_i$  stratum-specific sample size for stratum  $i$
  - $N$  total sample size

# [ Example ]

---

- MD4Life study (N = 466)
- HIV+ English speaking participants, at least 18 yrs old
- Recruited from and received motivational-based intervention at one of 2 HMO specialty clinics or one of 4 community clinics (non-HMO)
- 3 intervention arms: Computer, Provider, Control
- Assessed every 3 months to 21 months

## Example: treatment allocation

Control	2 non-HMO clinics	n = 135
Intervention	2 non-HMO clinics	n = 156 (51%)
	2 HMO clinics	n = 151 (49%)

Intervention combines Computer and Provider arms.

## Example: Sociodemographics at baseline

	Control	Intervention
Mean age (SD)	40.2 (8.7)	43.2 (9.4)
Ethnicity		
A. Ameri.	27%	19%
Hispanic	39%	26%
White	35%	55%
Graduated HS	81%	92%
Number of HIV- partners	1.4 (3.2)	.8 (3.2)

P-value < .01 for all treatment group comparisons

## Example: stratification

- Collapsed Computer and Provider intervention arms
- Created 5 strata from propensity scores
- Covariates: Age, gender, ethnicity, own home (Y/N), graduated from high school (Y/N), self-reported financial status (1=poor, 0=can afford parking at UCLA), Number of HIV- sexual partners

# Example: Sociodemographics by stratum

Stratum	1		2		3		4		5	
Intv status	C	I	C	I	C	I	C	I	C	I
N	50	38	37	52	16	72	20	69	12	76
Age	37.8	37.2	38.1	37.6	40.6	41.7	44.4	44.7	49.3	50.0
Ethnicity										
A. Amer.	36%	34%	30%	25%	31%	29%	10%	14%	0%	0%
Hispanic	58%	47%	32%	37%	38%	49%	25%	13%	0%	0%
White	6%	18%	38%	38%	31%	22%	65%	72%	100%	100%
Grad HS	60	58	86%	88%	94%	96%	100 %	100 %	100%	100%
N. part.	2.6	2.3	1.0	1.1	1.1*	.6*	.4	.5	.2	.2

\* P < .01 for intervention status comparison

## Example: analyses

- Random-intercept logistic and Poisson regressions
- Sexual behavior and substance use outcomes
- Covariates: time, stratum indicator (5 categories), and intervention arm (Computer, Provider, or Control)

# Example: Comparison of stratified and unstratified results

Outcomes showing significant decreases over time

	Computer vs Control	Provider vs Control
Stratified and unstratified	# days using hard drugs	# unprotected sex acts # days using hard drugs
Stratified only	# unprotected sex acts	Probability of using hard drugs

# Alternative to propensity score method

- Covariate adjustment
  - Eliminate the middle man!



Adjustment covariates → ~~Propensity score~~ → Outcome

Disclaimer: Use of company logos for demonstrational purposes only, shopping at these stores may lead to buying frenzy.

# Benefits of each method

- Covariate adjustment
  - More likely to be understood by journal reviewers and anyone who has not watched the Lord of the Rings and Star Wars trilogies back to back
  - Fewer steps to implement
  - Propensity score might not remove bias on all covariates, will have to adjust anyways
- Propensity Score method
  - More computationally feasible in longitudinal analyses
    - Too many covariates → non-convergence



## [ Benefits of each method (cont) ]

- Simulation study comparing propensity score method vs covariate adjustment (Drake 1993)
  - Both methods “equally inadequate” at reducing bias if confounders are omitted
  - Propensity score better when model is misspecified

# [ Conclusions ]

---

- Best to deal with potential bias in design phase if possible
- If not possible, propensity score is powerful tool to reduce effect of bias on analyses
- Be prepared to do a lot of explaining to journal reviewers to explain methodology

# [ Propensity Score References ]

- D'Agostino Jr., R. (1998). Tutorial in biostatistics: Propensity Score methods for bias reduction in the comparison of a treatment to a non-randomized control group. *Statistics in Medicine*. 17: 2265-2281.
- Drake, C. (1993). Effects of misspecification of the Propensity Score on estimators of treatment effect. *Biometrics*. 49: 1231-1236.
- Rosenbaum, P and Rubin, D. (1983). The central role of the Propensity Score in observational studies for causal effects. *Biometrika*. 70:41-55.
- Rosenbaum, P and Rubin, D. (1984). Subclassification on the propensity score. *Journal of the American Statistical Association*. 79: 516-524.