

Evaluating Survey Data: Mediation Analysis and Propensity Score Matching

Emily Ricotta, ScM
Senior Research Assistant
Johns Hopkins Center for Communication Programs





Outline

- Mediation Analysis
 - What is mediation analysis? When do we use it?
 - Tanzania case study
- Propensity Score Matching
 - What is PSM? When do we use it?
 - Zambia case study





Mediation Analysis

- What does mediate mean?
 - 1) Bring about; 2) Be a means of conveying; 3) Form a connecting link between
 - Mediating variable is intermediate in the causal sequence relating an independent variable to a dependent variable
- What is mediation analysis and when do I use it?
 - How or why two things are related How do knowledge and beliefs lead to behavior
 - Example: Attitudes form intentions, intentions lead to actions
- "In an absence of a concern for such mediating or intervening mechanisms, one ends up with facts, but with incomplete understanding" – Morris Rosenberg, 1968





Example 1: Tanzania Survey

- The Communication and Malaria Initiative in Tanzania (COMMIT) project was a multi-tiered SBCC project
 - Aimed to foster positive social norms and increase confidence among the population about malaria prevention and interventions
 - Used messaging, community activities, and Community Change Agents (CCAs)
- Household survey conducted in October and November 2011 in three regions of Tanzania with different levels of reported CCA activity

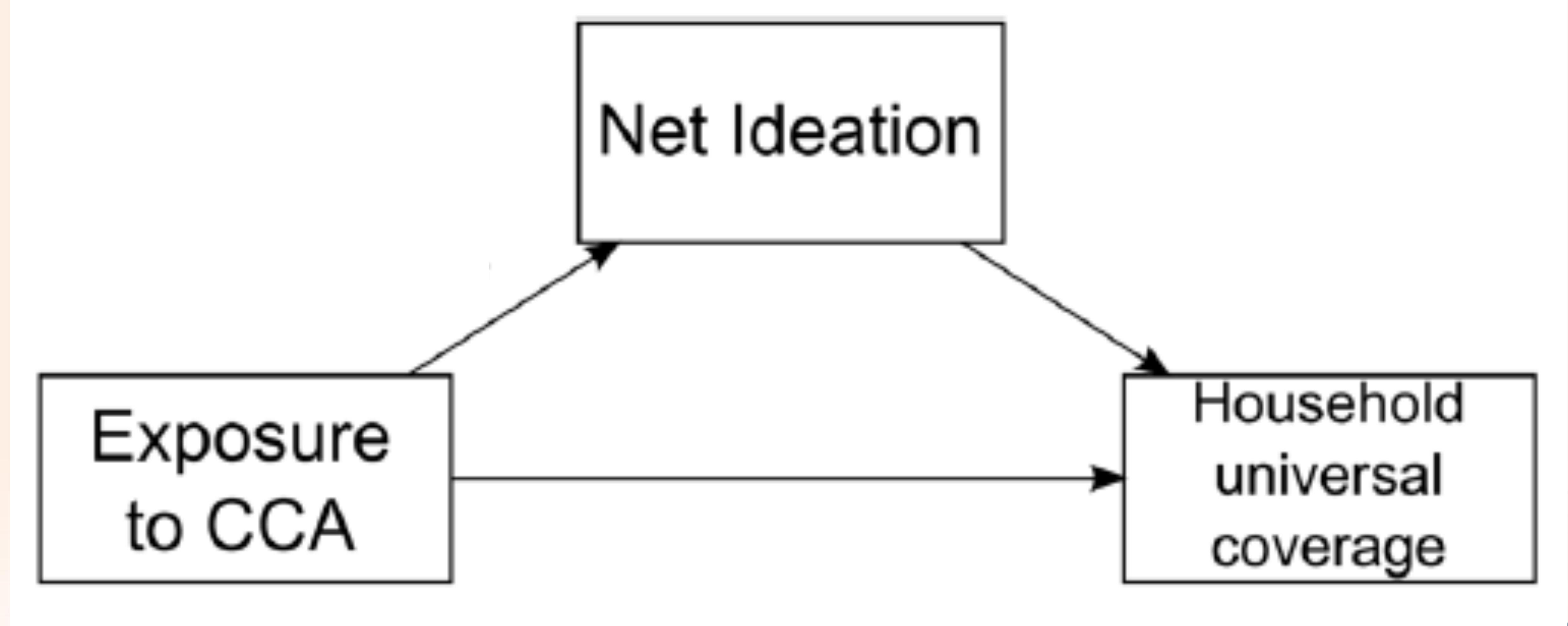




Variables

Exposure	Mediator	Outcome
 CCA only Media and community 	Net ideation	Universal coverage
messaging 3. Both CCA and M&C messaging		
(0 – Unexposed; 1 – Exposed)	(0 – 4 scale)	(0 – No; 1 – Yes)

Covariates: Level of education, region of residence, SES, ward







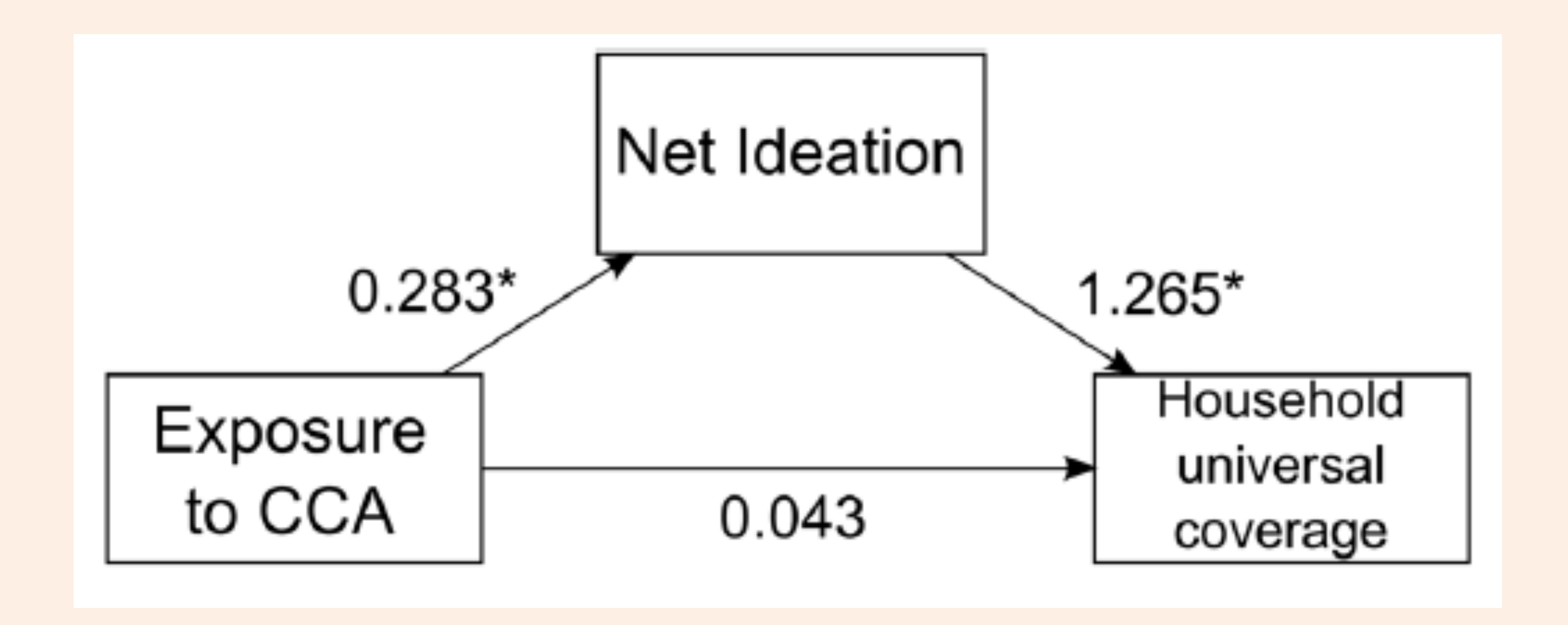
Mediation Analysis Steps

- Four requirements which need to be met (in most cases)
 - Step 1: Show that exposure is related to outcome
 - Step 2: Show that exposure is related to mediator
 - Step 3: Show that mediator is related to outcome
 - Step 4: Determine how much of the mediating variable accounts for the relationship between exposure and outcome
- Mediation analysis is done in three steps:
 - 1) Testing the direct relationship between the exposure on the outcome,
 - 2) Testing the effect of the exposure on the mediating variable,
 - 3) Telating the mediating variable to the outcome after controlling for the exposure and other variables





Tanzania Results Example







Mediation Analysis – Pro & Con

 Allows assessment of the process underlying the observed relation between an intervention and a behavior

 Can get statistically complicated if multiple mediators are involved, a multilevel design is used, or certain assumptions about the data cannot be made





Propensity Score Matching

- What is Propensity Score Matching?
 - Statistical technique that attempts to evaluate whether exposure to an intervention was actually responsible for the changes in observed behavior
 - Uses background population characteristics to create a statistically identical control group to which we can then compare our intervention group
- When would you use it?
 - Ideally, evaluation done with a Randomized Control Trial which randomly assigns people to receive or not receive a treatment or intervention
 - We cannot easily do this with SBCC programs or health interventions almost impossible to randomize who will or will not be exposed
 - This is why we use PSM, to simulate exposed and unexposed groups based on certain characteristics (their propensity to be exposed)





Example 2: Zambia Survey

- Using data from the 2010 Zambia Malaria Indicator Survey (MIS)
- This dataset was selected for this analysis due to the quality of the data and the strength of the BCC activities prior to the data collection period
- Zambia had an ongoing national mass and print media campaign for net use, as well as community-level campaign using interpersonal and community-based approaches





Variables

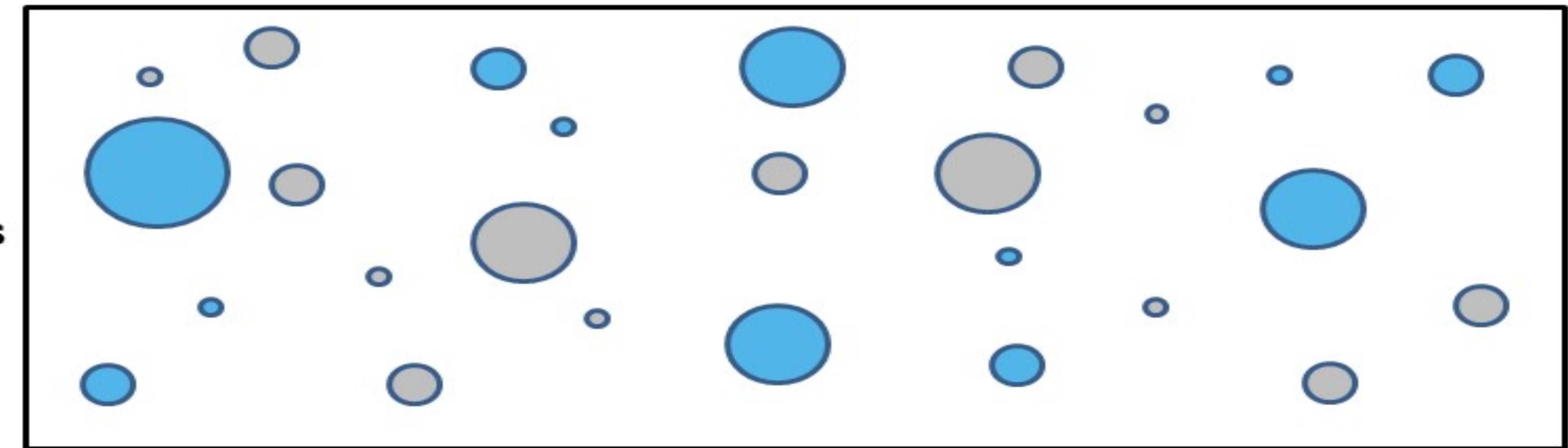
Exposure	Matching variables	Outcome
BCC messaging	 Age of respondent Level of education Child < 6 years old Wealth quintile Residence type Scheduled for IRS 	Woman slept under ITN
(0 – Unexposed; 1 – Exposed)	7. Province of residence	(0 – No; 1 – Yes)

Covariates: # ITNs owned by household, discussion of malaria with a HCW



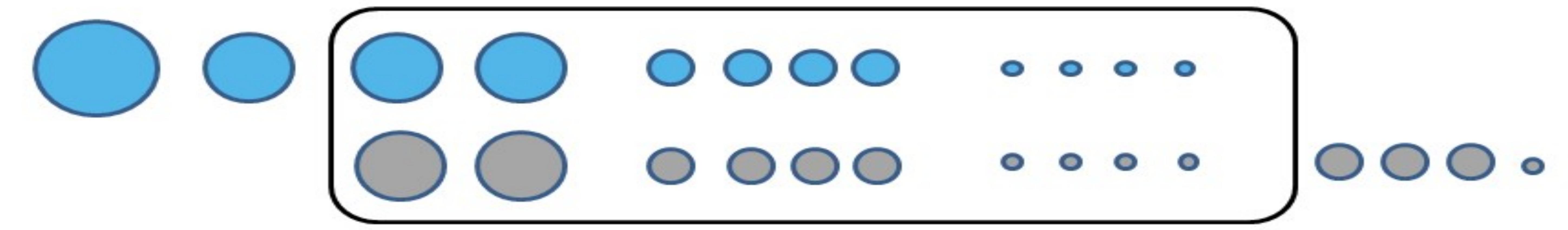


Population with varying characteristics





Study Group with Matching









Propensity Score Matching Steps

- Two requirements for PSM
 - Representative survey data collecting as many descriptive variables of the population as possible
 - After matching, individuals with similar propensity scores do not have any other differences except being exposed or unexposed to the intervention
- PSM is done in four steps
 - Step 1: Determine the relevant background covariates (a priori, stepwise selection, etc.)
 - Step 2: Obtain propensity scores for population using a logit/probit regression of the background covariates on intervention status (Exposed/Unexposed) – the result of being exposed to the intervention is called the "propensity score"
 - Step 3: Match similar exposed and unexposed individuals using the propensity scores
 - Individuals whose background variables or propensity score are unable to match with another in the comparison group are dropped from the analysis
 - Step 4: Compare difference in proportion of practicing the outcome behavior between matched treatment and control groups (e.g. 70% of those exposed to SBCC messages slept under an ITN, compared to only 30% of those who were not exposed to SBCC messages)





Zambia Results Example

- Of the 3,380 women owning at least one ITN, 3,263 had complete information and were included in the analysis
 - Of these, 729 exposed respondents were matched with replacement to 2,498 unexposed respondents
- Without matching on propensity, there was a 13-percentage point difference in ITN use between exposed and unexposed respondents
- After matching on propensity score, there was a 29.4percentage point difference in ITN use between exposed and unexposed respondents
 - Created statistically equivalent groups matched on their probability of being exposed to the BCC messages (similar to a RCT)





Propensity Score Matching – Pro & Con

 Propensity scores are useful when the intervention was not randomly assigned, but you want to draw causal conclusions

 Requires large samples, good data, complete capturing of covariates (village-level, householdlevel, etc.)





Resources

- Resources (Mediation Analysis)
 - Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic and statistical considerations. *Journal of Personality and Social Psychology, 51*, 1173-1182.
 - Imai, K., Keele, L., & Tingley, D. (2010). A general approach to causal mediation analysis. Psychological Methods, 15, 309-334.
 - MacKinnon, D. P. (2008). *Introduction to statistical mediation analysis*. New York: Erlbaum.
 - Ricotta, E., et al. (2015). The use of mediation analysis to assess the effects of a behaviour change communication strategy on bed net ideation and household universal coverage in Tanzania. Malaria Journal, 14:15.
- Resources (Propensity Score Matching)
 - Boulay, M., Lynch, M., Koenker, H. (2014). Comparing two approaches for estimating the causal effect of behavior-change communication messages promoting insecticide-treated bed nets: an analysis of the 2010 Zambia malaria indicator survey. *Malaria Journal*, 13:342.
 - Heinrich, C., Maffioli, A., Vazquez, G. (2010). A primer for applying propensity-score matching. IBD Impact-Evaluation Guinelines, Technical Notes.
 - Austin, P. (2011). An introduction to propensity score methods for reducing the effects of confounding in observational studies. Multivariate Behavioral Research, 46:399-424.
 - Health Communication Capacity Collaborative. (2014). Propensity Score Matching: An HC3 Research Primer.

