www.childhealthdata.org
Data Resource Center for Child & Adolescent Health
A project of the Child and Adolescent Health Measurement Initiative

MCHB

Making the Most Of National and State Child Surveys: Tips for the journey through the data analysis labyrinth

Christina Bethell
The Child and Adolescent Health Measurement Initiative

February 14, 2011

AMCHP, Washington DC

Christina Bethell

Has documented that she has no relevant financial relationships to disclose or COIs to resolve.

Presentation Goals

1) **Inspire more extensive and effective use** of national and state child surveys to inform and improve state policies and programs
2) Outline a **simple conceptual map** for thinking about within and across state analysis of national and state child survey data
3) **Build confidence** to incorporate new data analysis methods into your states standard analytic protocols

Making your way through the labyrinth of within and across state analysis of child survey data

Encouragement: There is a natural progression
Caution: You have to be able to get back out!

The Natural Progression Through the Data Analysis Labyrinth

First Steps

- **<u>My State</u>: What is the prevalence and what are disparities across child subgroups in my state?**

Venturing Further

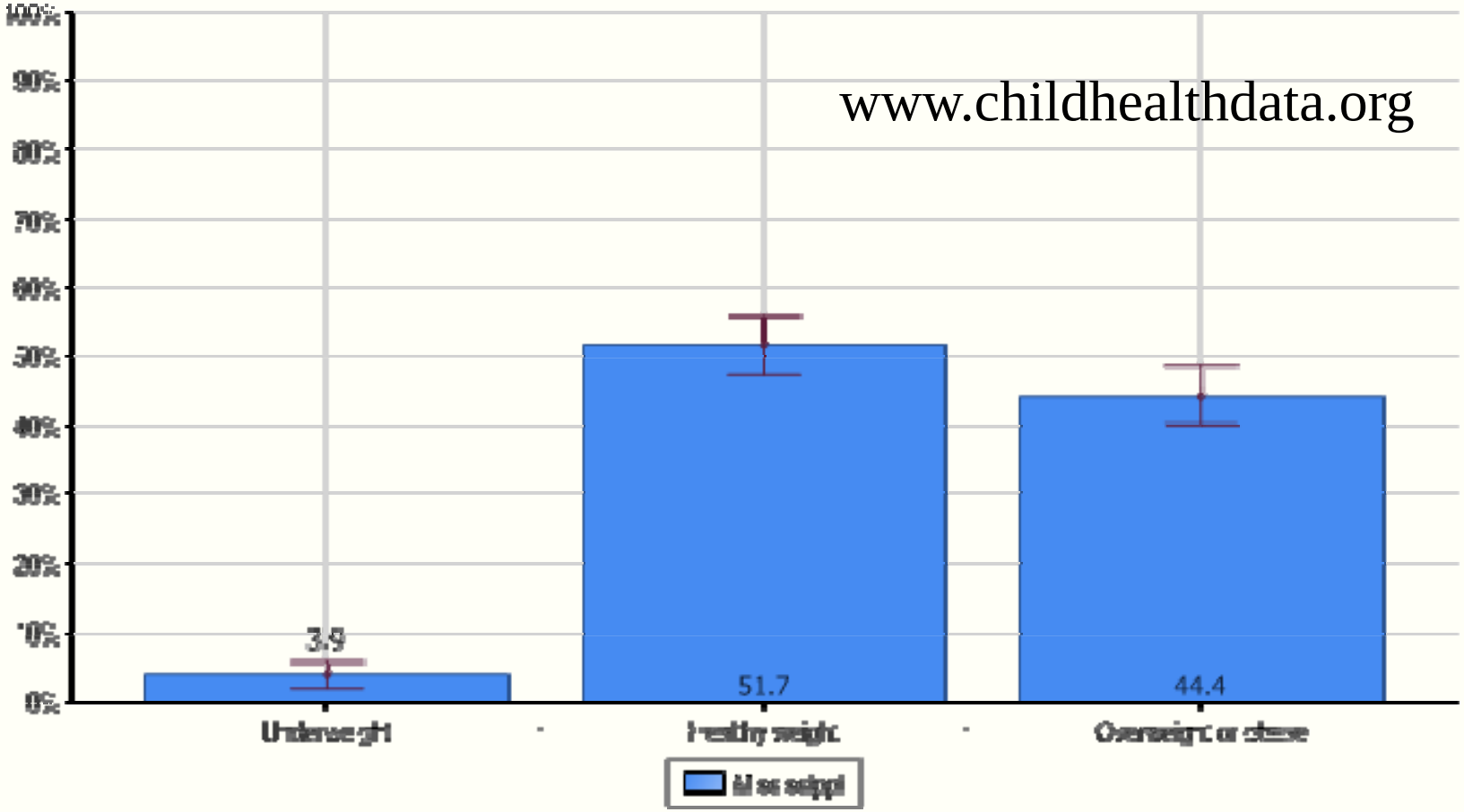
- **<u>Compare My State</u>: Where does my state rank? Is prevalence and are disparities in my state different?**

Approaching the Center

- **<u>Understand differences within and across states</u>: Are differences across states significant? What is associated with differences within and across states? Do these associations vary? Can they be explained by policy?**

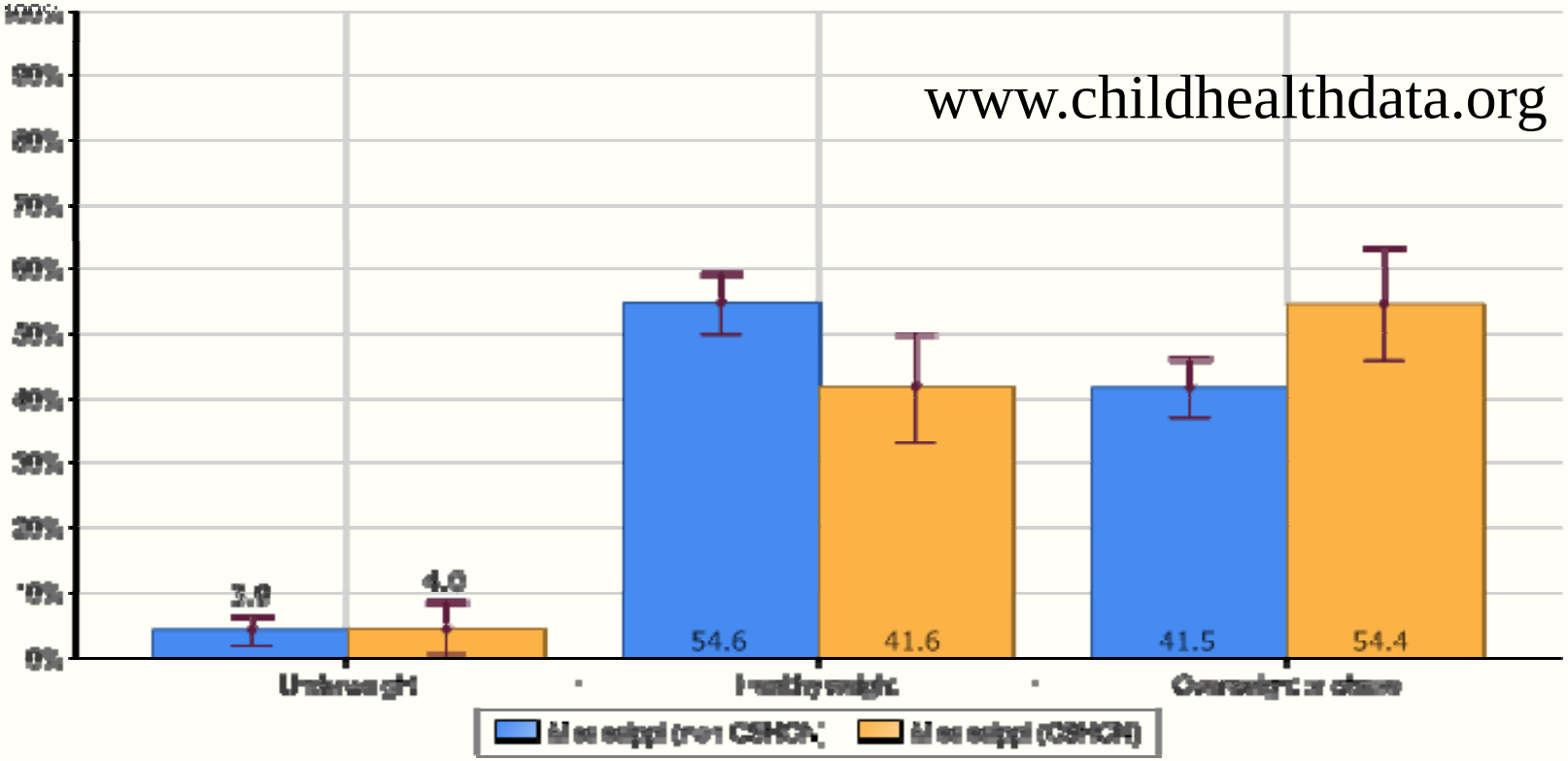
First Steps:
What is the prevalence in my state?

Weight status of children based on Body Mass Index for age (BMI-for-age)
Children age 10-17 years only
Mississippi



First Steps: Does prevalence vary for different subgroups of children in my state?

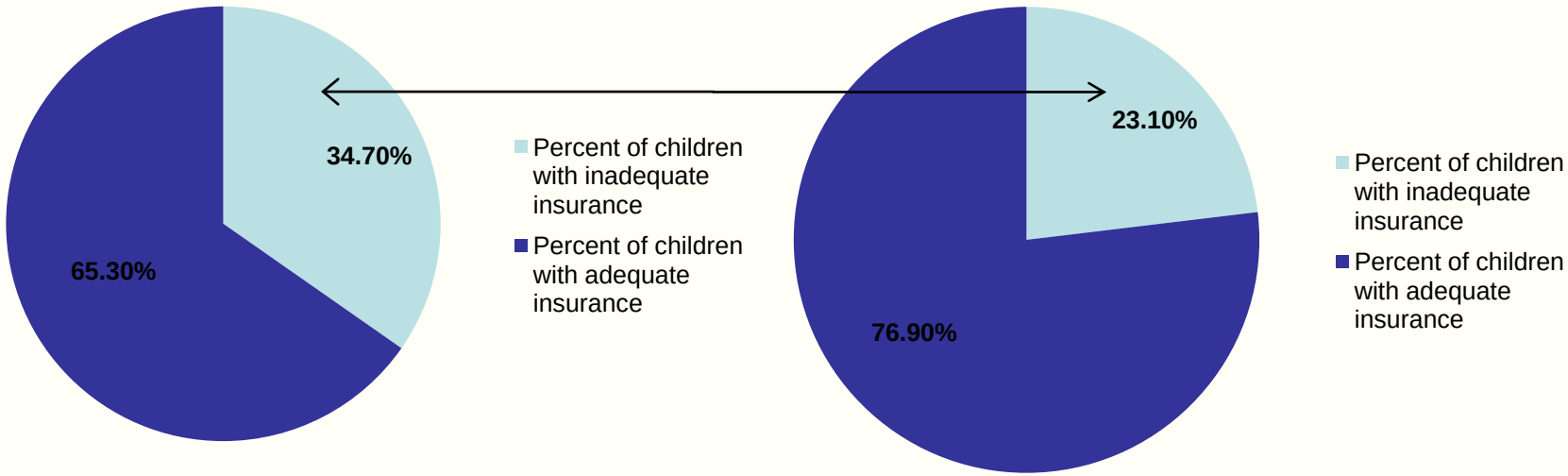
Weight status of children based on Body Mass Index for age (BMI-for-age)
Children age 10-17 years only
Mississippi
Special health care needs status



First Steps Analytic Questions
Are differences across subgroups of children in my state statistically significant and meaningful?

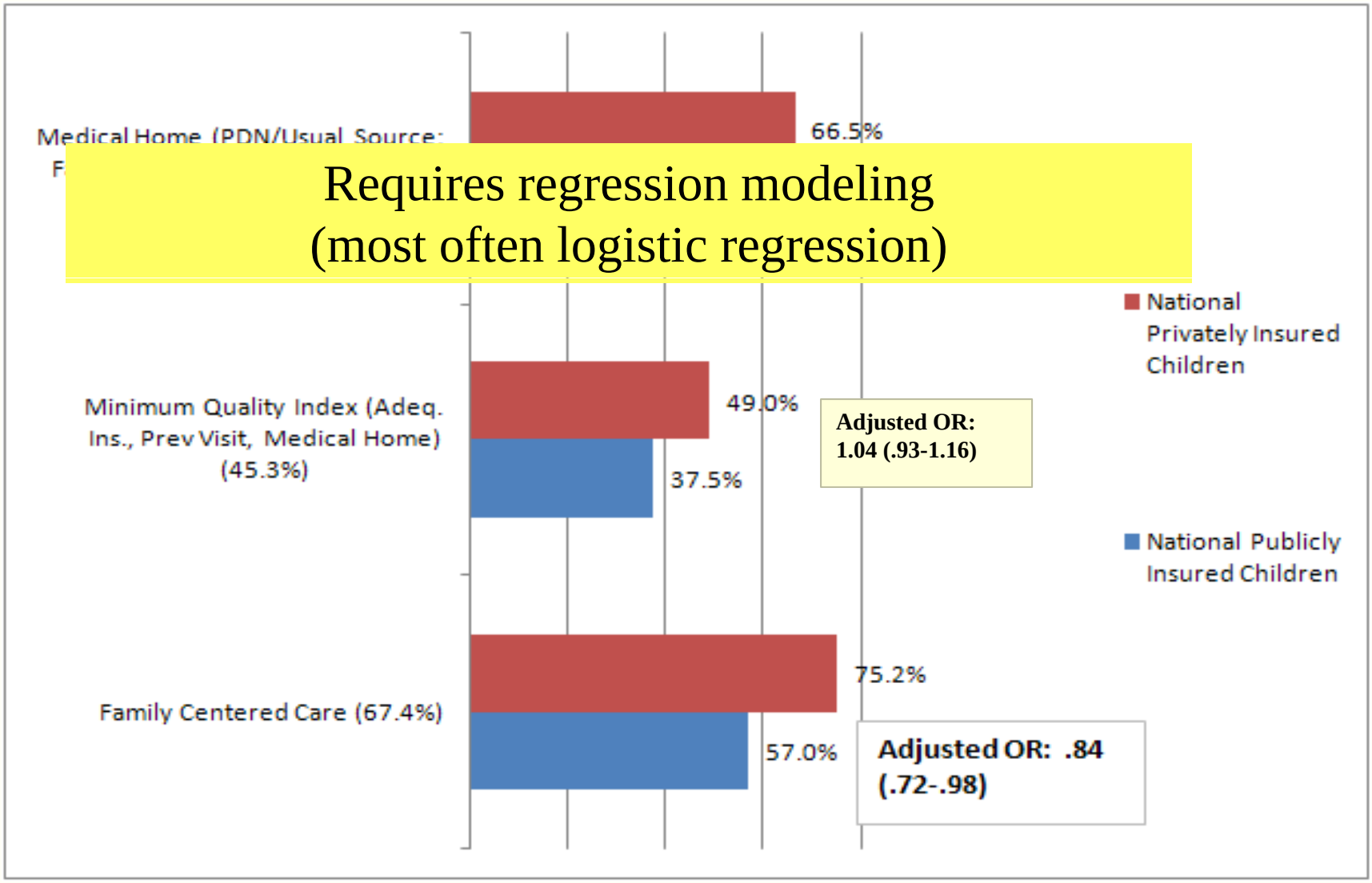
Example: Are differences in insurance adequacy between children with and without special health care needs significant?

Requires standardized t-tests or chi square tests—requires standard errors and adjustment for complex sampling



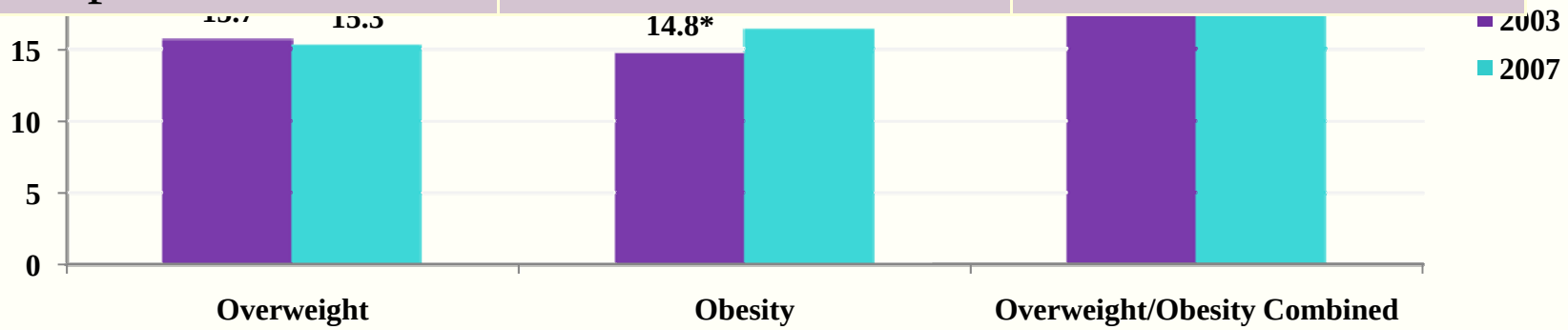
First Steps Analytic Questions:

Are child subgroups with higher prevalence still more likely to experience an outcome after adjusting for other factors?



First Steps Analytic Questions: Are my state's overall prevalence and child subgroup level prevalence rates different across years of survey data administration?

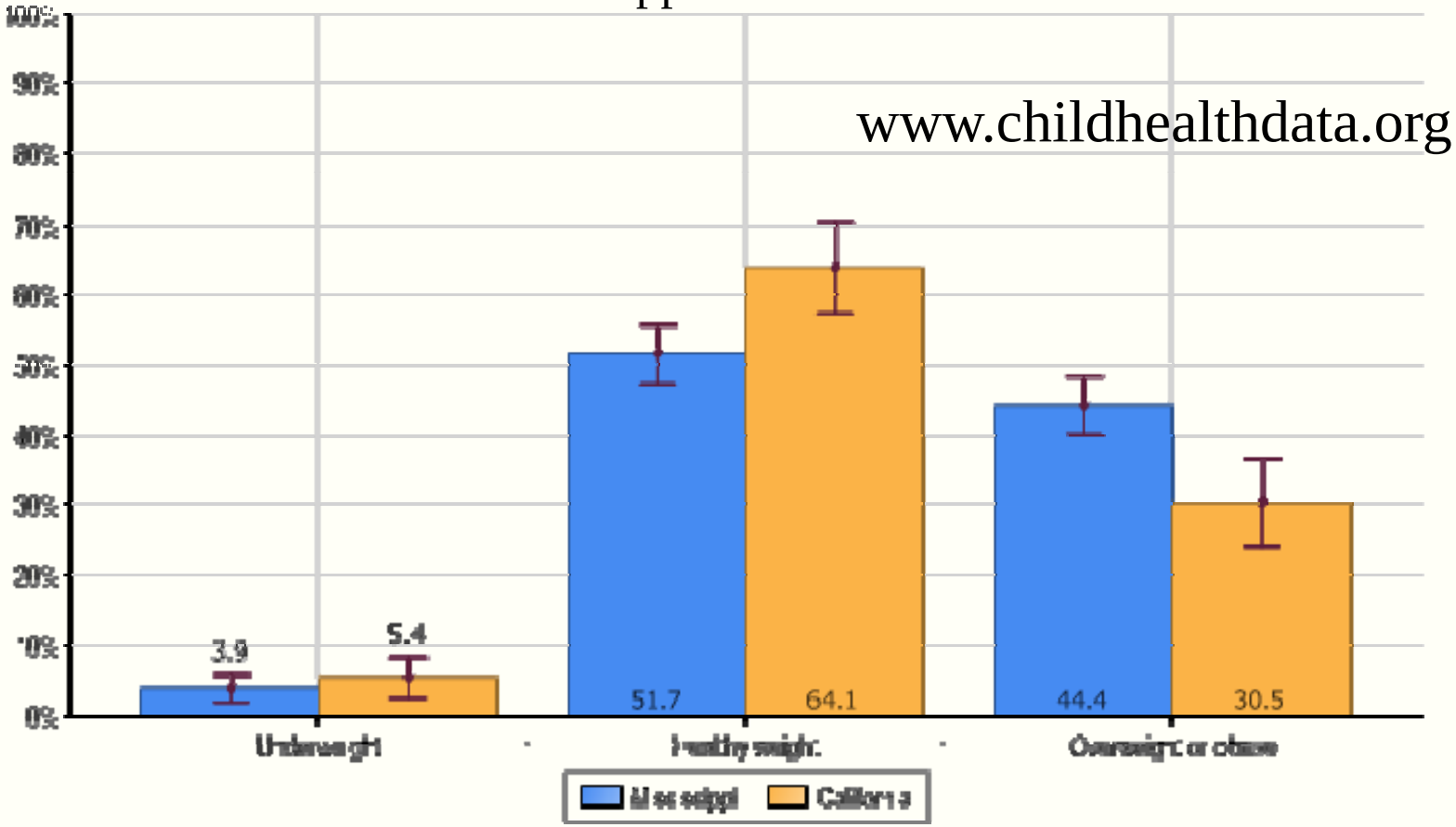
	2003	2007
Publicly Insured	39.6%	43.2%*
Poor (< 100% FPL)	39.8%	44.8%*
Hispanic	37.7%	41.0%**



*2003 versus 2007 rates of obesity are significantly different (P<0.05). ** p < .07.

Venturing Further: How does the prevalence in my state compare to other states or the nation as a whole?

Weight status of children based on Body Mass Index for age (BMI-for-age) Children age 10-17 years only
Mississippi vs. California



Venturing Further: Where does my state rank compared to other states or the nation as a whole?

RESULTS Share this Link | Add to Briefcase | Print

Selected Criteria:

- All States
- 2007
- Physical and Dental Health
- 1.4a: Weight status - BMI-for categories, age 10-17 (detail

Go Back

Question: Indicator 1.4a: How many children are c Index for age (BMI-for-age)? (details...)

Notes: Click on the Column Header to sort th
To get a detailed explanation of the c

Region	Underweight (less than 5th percentile) %			
1. Mississippi	3.9			
2. Arkansas	4.3			
3. Georgia	4.0			
4. Kentucky	4.7			
5. Tennessee	4.2			
6. Alabama	4.0			
7. Louisiana	3.6			
8. West Virginia	4.5			
9. District of Columbia	4.8			
10. Illinois	3.7			
11. Nevada	3.9			
12. Alaska	4.9			
13. South Carolina	7.4			
14. North Carolina	8.5			
15. Ohio	4.7	62.0	33.3	100.0
16. Delaware	6.2	60.7	33.2	100.0

Question: ator 1.4a: How many children are currently overweight or obese, based on Body Mass r age (BMI-for-age)? (details...)

Select a Response: Overweight or obese

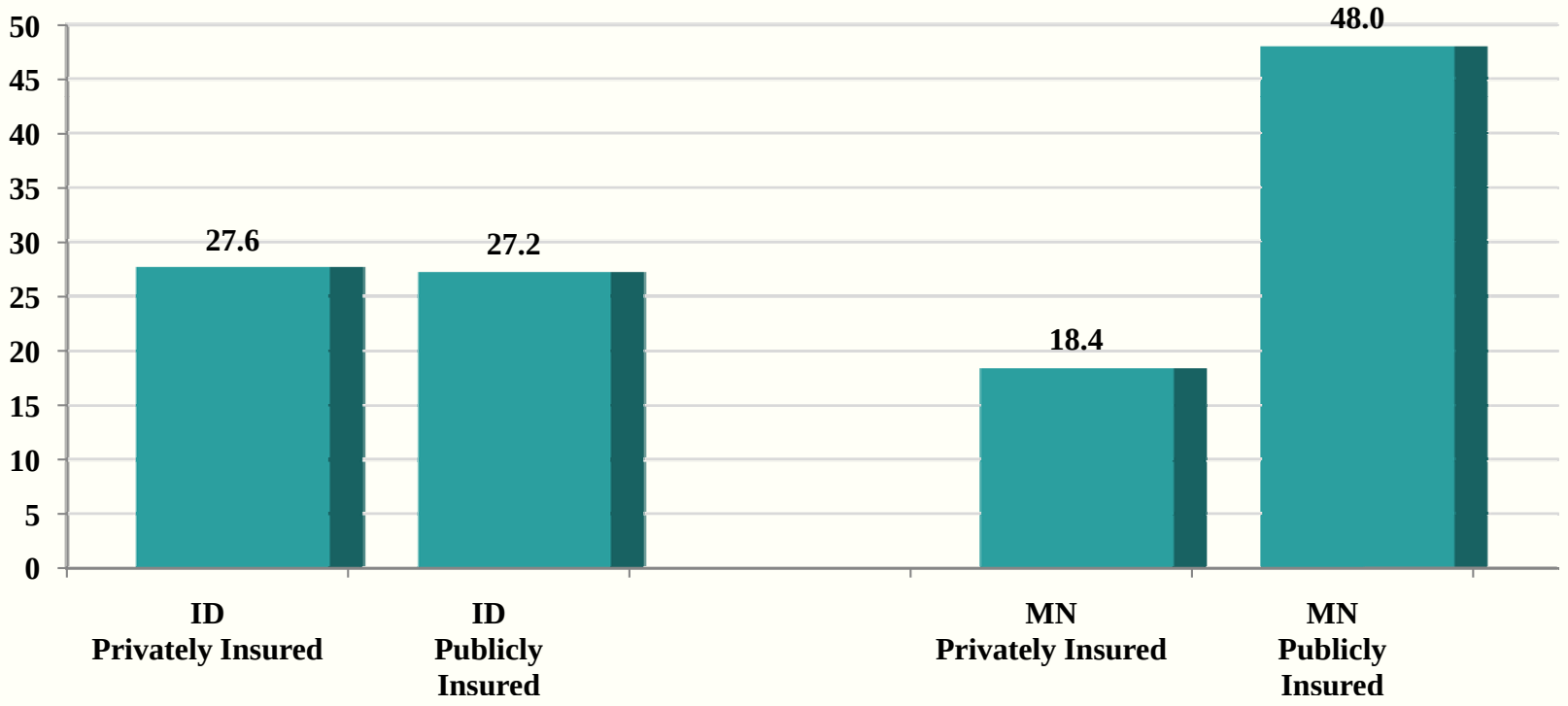
Notes: Column Header to sort the results by ascending or descending order.
To iled explanation of the data HOVER over the text in the table.

Region	One or more emotional behavioral or developmental issues %	No qualifying emotional behavioral or developmental issues %	No qualifying special health needs (non-CSHCN) %
1. Illinois	62.4	33.2	33.3
2. Mississippi	57.6	52.9	41.5
3. New Jersey	56.3	30.5	28.9
4. Missouri	52.2	32.3	28.0
5. New York	51.1	32.6	31.5
6. Georgia	49.8	39.3	35.9
7. Arkansas	49.2	46.0	34.0
8. District of Columbia	48.6	45.3	31.4
9. Florida	48.5	21.7	33.7
10. Nebraska	47.5	20.7	32.1
11. Maine	46.5	32.7	25.1
12. Alabama	46.4	37.6	34.5
13. Massachusetts	46.3	26.2	28.8
14. Indiana	45.5	26.2	28.5
15. Pennsylvania	45.4	39.1	25.8

Venturing Further: How do disparities in my state compare to those in other states?

Disparities in the Disparities!

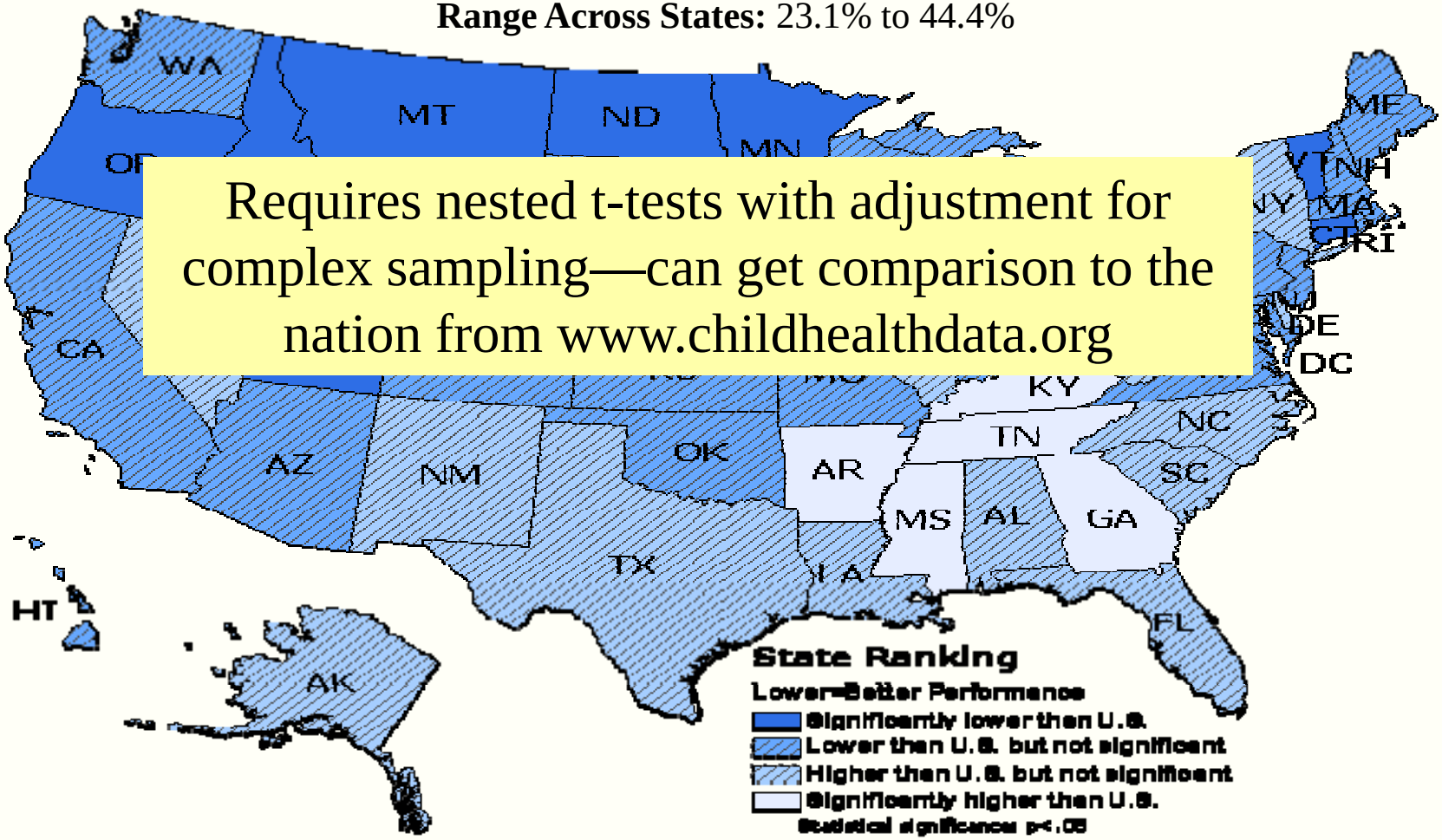
State with the lowest overall overweight or obese rate (MN) had the highest insurance disparity



www.childhealthdata.org

Analytic Questions When Venturing Further:
Does the prevalence in my state differ significantly from other states or the nation as a whole?

Overweight/Obese Children
(age 10-17) 2007 National Survey of Children's Health
Nationwide: 31.6% of children met indicator
Range Across States: 23.1% to 44.4%



Requires nested t-tests with adjustment for complex sampling—can get comparison to the nation from www.childhealthdata.org

Analytic Questions When Venturing Further: Are variations across states statistically significant?

Measure	Hi State	Low State
Overweight/Obese	44.4% (MN)	23.1% (UT/MN)
Developmental Screening	47.0% (NC)	10.7% (PN)
Participates in Activities	73.0% (CA)	90.5% (MN)
Repeated a Grade in School	25.4% (LA)	1.8 (UT)

Analytic Questions When Venturing Further: Are variations across states statistically significant?

Measure	Hi State	Low State
Ove		
Dev Scre		
Part Acti		
Repeated a Grade in School	25.4% (LA)	1.8 (UT)

The Simple Option

-Independent samples t-test accounting for complex sampling

-Will tell you whether the lowest and highest states vary significant (or any other two states)

But…..don't you really want to know…

1. The effect of child level demographic and other factors and whether differences across states in these factors explain differences observed across states in your outcome of interest (e.g. age, race/ethnicity, poverty…)

AND, BETTER YET….

2. Whether associations with these demographic and other child (e.g. CSHCN status), family (e.g. household income) and system level factors (e.g. insurance type) are the same or different across states (e.g. is having public sector health insurance in one state is the same as in another state?)

AND, EVEN MORE IDEALLY….

3. Whether any variation remains within and across states after accounting for these child level demographic and other factors

AND, FINALLY….

4. Whether any across state variation in your outcome of interest that remains might be associated with differences in programs and policies across states (e.g. policies impacting nutrition labeling, school food policies, availability of parks and fitness programs, etc.)

Approaching the Center: Is there real variation across states, even after adjusting for population differences? Do associations with population characteristics differ across states? Any variation after that? Policy impact?

That natural progression….

Enter Multi-Level Modeling (MLM)

MLM: The simple explanation for a complex tool helping you approach the center….

- MLM involves two levels of assessment to evaluate associations and predictors of outcomes of interest (*analogy:* students within classrooms—effect of student vs. effect of the teacher/class)
  - **<u>Level 1: Individual child, family and other factors</u>**
    - Estimates the association between child level factors and outcome, allowing the magnitude of these associations to vary across states (median odds ratio-MOR)
    - Assesses the significance of any across-state differences in the association between an outcome and a child level factor (e.g. age, insurance status, etc.)
    - Tells you how much variation remains across states after accounting for level 1 factors—is there a "state effect"?

MLM: The simple explanation for a complex tool helping you approach the center….

- **Level 2 predictors: State level variables**
  - Examines whether a state level variable (e.g. contextual variable) has a large, small or no effect relative to the variation left unexplained by level 1 factors (Interval Odds Ratio –IOR)
  - **Example State Policy Variables**:
    - school nutrition policies
    - payment levels for developmental screening
    - eligibility rules for early intervention
    - violence education and tolerance policies in schools
    - public insurance benefits standards
    - case management for CSHCN in schools

MLM EXAMPLE: CSHCN Repeating a Grade In School—Does State Policy for Case Management for Children with Disabilities Make a Difference?

State Policy Variable Background

1. In 2006, 26 states had a policy requiring the provision of case management for children with disabilities who need it.
   - Obvious questions: What constitutes case management? What constitutes "needing case management"? Who decides a child needs case management and how do they decide it? Putting this aside...
2. Average rate of grade repetition for CSHCN in states with a case management policy: **15.72%**
3. Average rate of grade repetition for CSHCN in states without a case management policy: **20.32%**
   - Obvious question: Is the average across states driven by any outliers? Are states with a policy more advantaged in other ways that might explain this difference (poverty levels, severity of special health care needs, etc.)?

CSHCN Repeating a Grade In School in 2007 (NSCH): Does a state policy requiring case management for children with disabilities make a difference after accounting for complexity and type of special health care need?

Level 1: Percentage of CSHCN Repeating a Grade in 2007: by Service Need

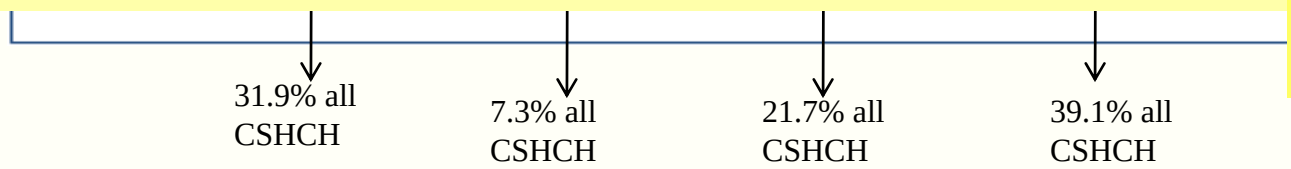
CSHCN with more complex needs and EBD

- on average across states
- have nearly 5 times the odds of having repeated a grade in school.
- the effect of having more complex needs and an EBD varies significantly across states

If similar CSHCN moved from the state with the lowest to the highest prevalence of grade repetition among CSHCN…

- …they would still have yet greater odds of repeatir grade—even after adjusting for all other individual characteristics assessed (1.30)

Other Adjustment Variables:
Age, Sex, Race/Ethnicity, Household Income, Insurance Adequacy



MLM EXAMPLE:

CSHCN Repeating a Grade: Does State Policy Make a Difference?

<u>Level 2 Question:</u> Is the remaining variation across states in CSHCN who repeated a grade in school partly accounted for by whether or not a state has a policy requiring case management for students with disabilities?

- **Interval Odds Ratio for Case Management: .65[s]**

- **CSHCN are less likely to repeat a grade in school** if they live in states with a case management policy, even after adjusting for a range of individual child level characteristics

Stay tuned for more findings and examples such as….

- ✓ Actual percentage of variation explained by level 1 and level 2 factors assessed
- ✓ Impact of violence prevention and prohibition policies on CSHCN participation in activities (Sneak Peek---IOR: 1.22)
- ✓ Whether school case management policies impact CSHCN engagement in school and doctor-school communication

Cautions before venturing in!

1. State level policy variables that are comparable across states and sensitive to the outcome of interest are **HARD TO FIND!**

   - MLM is **still useful** to learn whether there is any state contextual effect at all—even if we can't explain it!
   - MLM is **still useful** to learn if associations with child and family variables vary across states
   - Even generalized effects using imperfect variables **can point in the right direction**

2. **Don't begin** using multi-level modeling **unless you are prepared** for the journey and have required software, computer systems/memory and staff….even the bravest among us have gotten lost in the labyrinth!

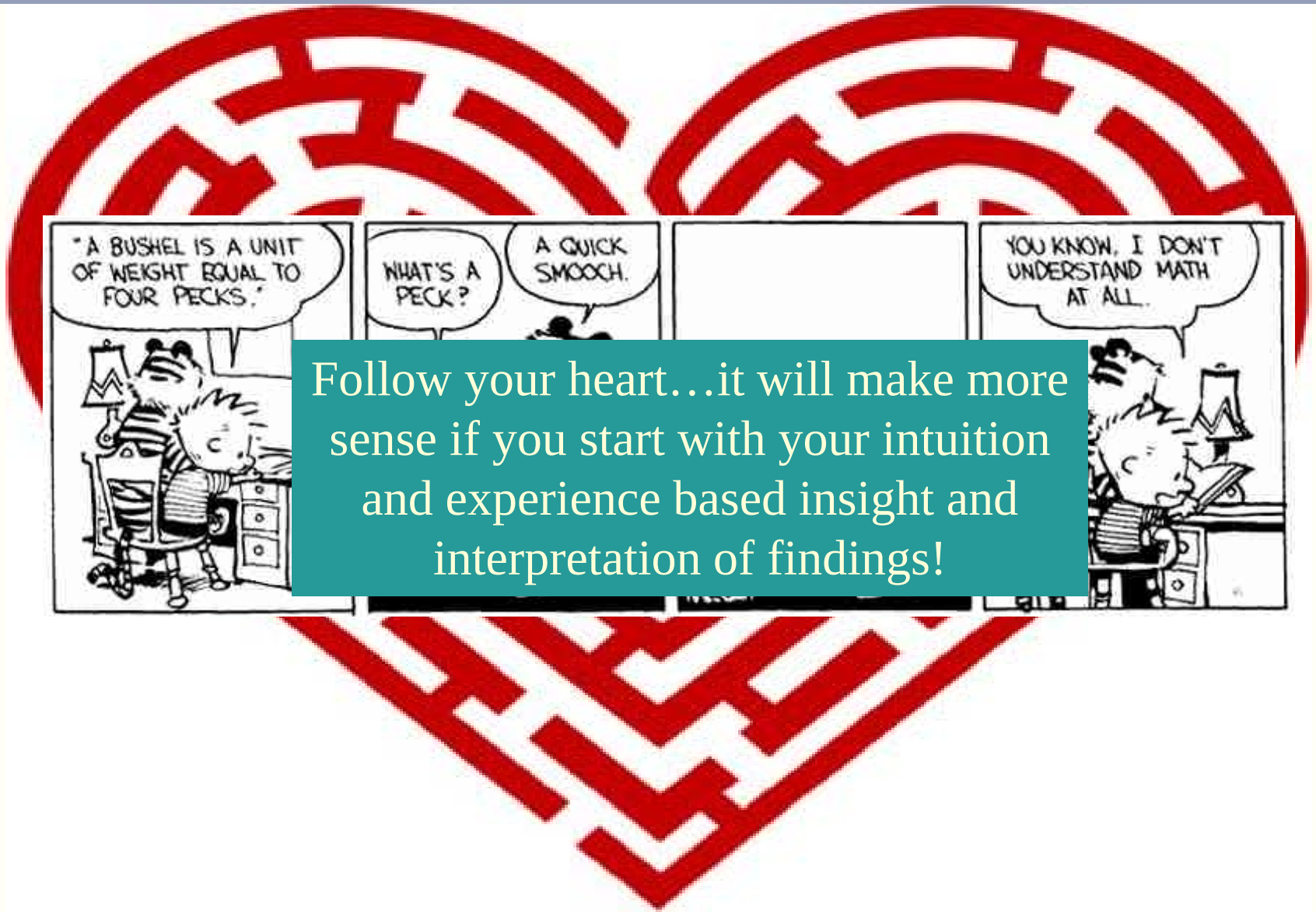
   - There are plenty of compelling analyses for you to do in understanding state findings without doing MLM

Cautions for making your way back out!

Don't lost sight of the big picture in there!

- For example, even though children without access to a park/recreation center or who are publicly insured are more likely to be overweight or obese
  - nearly 60% of all overweight and obese children are privately insured and
  - 85% have access to a park or recreation center.
- A significant state level effect may still be small compared to the impact of child level factors—assess what will make the most difference (a policy or an improved pediatric practice environment, etc.)
- At the same time….very small effects can translate into a large number of children. Interpretation requires your knowledge and wisdom!

FINAL WORD



Follow your heart…it will make more sense if you start with your intuition and experience based insight and interpretation of findings!

Special Notes About MLM with Categorical Outcomes

- A common case example using national and state child surveys:
  - e.g., Does a child have access to a medical home or not?
- Categorical (yes/no) variables increase MLM complexity and require special interpretation.
  - More difficult to examine and interpret variance components in MLMs with categorical outcomes.
  - Due to nonlinear relationship between covariates and outcome and difficulty partitioning variance.

Adjusted Odds Ratios, MLM Style

- Level 1: Analogy to Adjusted Odds Ratio (logistic regression) is the Median Odds Ratio (MLM)
- MOR attenuates problems related to categorical variables by quantifying the variance among clusters.
  - Essentially compares two randomly chosen individuals with the same values on all covariates, but from two different clusters (i.e., contexts).
    - Conceptually, repeat this for all possible pairs.
      - Take the median odds ratio from all of these comparisons.
  - MOR gives the median odds ratio between individuals with higher propensity compared to people with lower propensity.
    - Always greater than or equal to .
    - If equal to 1, no variation exists between contexts.
    - If large, considerable variation exists.
    - Directly comparable to fixed-effects odds ratios.
      - Can make relative statements.

Adjusted Odds Ratios, MLM Style

- **<u>Level 1:</u>** Analogy to Adjusted Odds Ratio (logistic regression) is the Median Odds Ratio (MLM)
  - MOR encapsulates the increased risk that would occur if similar individuals moved from one context to another (in the median/on average).
  - MOR attenuates problems related to categorical variables by quantifying the variance among clusters.
  - Essentially compares two randomly chosen individuals with the same values on all covariates, but from two different clusters (i.e., contexts).
    - Conceptually, repeat this for all possible pairs.
      - Take the median odds ratio from all of these comparisons.
      - Why it can take hours to run an MLM and need sufficient sample in each conceivable cluster

Adjusted Odds Ratios, MLM Style

- **<u>Level 2:</u>** Analogy to Adjusted Odds Ratio is the Interval Odds Ratio (IOR)
  - IOR essentially quantifies the effect of contextual variables relative to variance across level 1 clusters.
  - How does the odds ratio for the contextual variable compare to the amount of variance across contexts after accounting for the contextual variable?

Interval Odds Ratio (IOR)

- Consider two random individuals with different values of a cluster-level covariate but same individual-level covariate values.
  - Compute odds for individual in context with higher propensity vs. lower propensity.
  - Consider all possible pairs of individuals.
    - Results in distribution of odds ratios (ORs).
- IOR = interval that contains 80% of these values.
  - If IOR contains 1, contextual variability large *compared to* effect of cluster-level variable.
  - If IOR does not contain 1, large cluster-level variable effect *compared* to unexplained contextual variation.

Categorical Outcomes

- No clear distinction between individual and cluster-level variance.
  - If we know the prevalence of the outcome in each cluster, we know the variance within a cluster.
    - Not true with a continuous variable.
      - If we know cluster's mean, we cannot infer cluster's variance.

- Thus, cannot simply partition the variance like we do in a continuous model.
  - Yet, need remains to quantify variance across clusters and interpret in line with odds ratio interpretations.
    - See Merlo, et al. (2006) for more discussion.

- What can we do?

Adjusted Odds Ratios, MLM Style

- MOR gives the median odds ratio between individuals with higher propensity compared to people with lower propensity.
  - Always greater than or equal to .
  - If equal to 1, no variation exists between contexts.
  - If large, considerable variation exists.
  - Directly comparable to fixed-effects odds ratios.
    - Can make relative statements.

Median Odds Ratio (MOR)

- MOR summarizes variance across contexts among people with the same values on the covariates.
  - It encapsulates the increased risk that would occur if an individual moved from one context to another.
    - (In the median).
  - With no covariates in model, describes extent to which outcome depends on context.
- However, likely still wish to examine whether a contextual variable has large effect relative to unexplained variation between contexts.

MLM EXAMPLE: CSHCN Repeating a Grade In School—Does State Policy for Case Management for Children with Disabilities Make a Difference?

- Examines relationship between a child repeating a grade in school and child level factors
- Examines variation in the effect of these child factors across states
- Examines variation across states after controlling for within state demographic and child health status differences (e.g. CSHCN status)
- Examines whether state prevalence in CSHCN who repeated a grade in school differs significantly across states after controlling for differences in demographic and other factors?

Answers whether state level contextual level variable predicts variance in the relationship between CSHCN status and repeating a grade in school