Multiple Methods for Multiple Reporters of Child Maltreatment

Results from the Lehigh Study

Lehigh Measurement and Causal Inference Analyses Team

2025-06-02

Table of contents

1 Background

The Lehigh Study presents a unique opportunity. Data are collected on experiences of abusive discipline as reported by *administrative* reports, two *parental* reports at two different time points, and two *self* reports at two different time points. However, in the absence of a gold standard measure of abusive discipline, appropriately aggregating these multiple reports across multiple time points represents an analytic challenge.

In the manuscript below, we employ multiple strategies to estimate the relationship of these multiple reports from multiple reporters at multiple time points to a mental health outcome. We compare and contrast the advantages and disadvantages of these different methods, and conclude the manuscript with suggestions on optimal methodological approaches to confront the methodological challenges that are posed by having multiple reports from multiple reporters at multiple time points.

> **i** Note 1: Key Ideas or Questions
>
> - Many sources in the literature note that data can come from multiple reporters. It is often suggested that it is important to *triangulate* these different sources of information, yet what is meant exactly, or operationally, by such *triangulation* is often unclear.
> - How do the different methods highlight, or not highlight:
>   - The differences between prospective and retrospective reports, especially in predicting outcomes?
>   - The level of agreement or disagreement between different reporters?
>   - Issues of severity, chronicity, timing and developmental stages across reporters?

2 Basic Conceptual Model

We begin with a basic conceptual model of the reports and time points in the data, without at this point suggesting any associational or causal relationships.
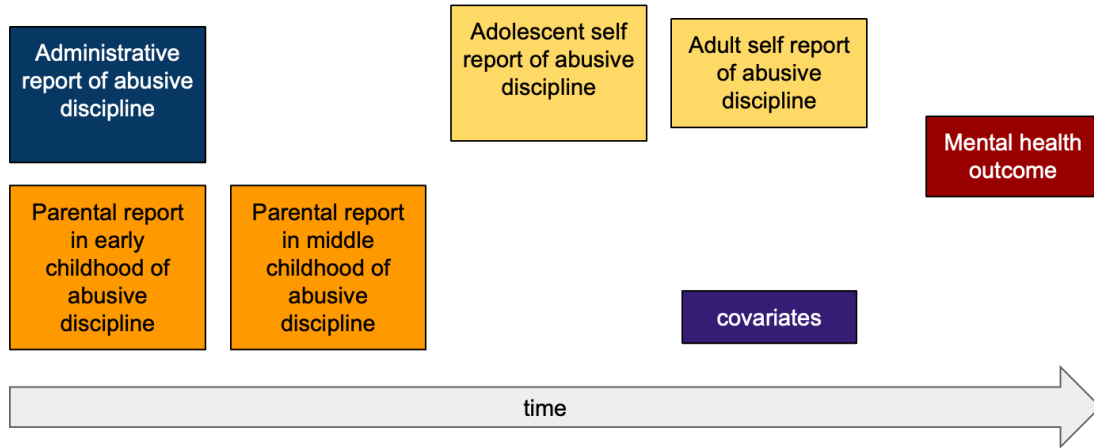

Figure 1: conceptual model

3 Variable Abbreviations

For parsimony, we use the following conventions for variable names in equations and statistical syntax.

Table 1: Variables and Variable Labels

Variable	Label
administrative	administrative report
PR1	parental report in early childhood
PR2	parental report in middle childhood
SR1	adolescent self report
SR2	adult self report
covariates	covariates (multiple variables)

Variable	Label
outcome	mental health outcome

4 Methods

4.1 OLS Regression

Our outcome is continuous. Therefore we here employ *ordinary least squares regression.* Were our outcome to be dichotomous, we could as easily employ *logistic regression.*

4.1.1 Diagram

4.1.2 Equation

$$\text{outcome} = \beta_0 + \beta\text{P1} + \beta\text{P2} + \beta\text{SR1} + \beta\text{SR2} + \beta\text{administrative} + \Sigma\beta\text{covariates} + e_i \quad (1)$$

4.1.3 Syntax

```
regress outcome P1 P2 SR1 SR2 administrative covariates
```

For logistic regression, the appropriate syntax would be:

```
logit outcome P1 P2 SR1 SR2 administrative covariates, or
```

4.2 Summing Across Reporters

4.2.1 Diagram

4.2.2 Equation

First, we average parental reports:

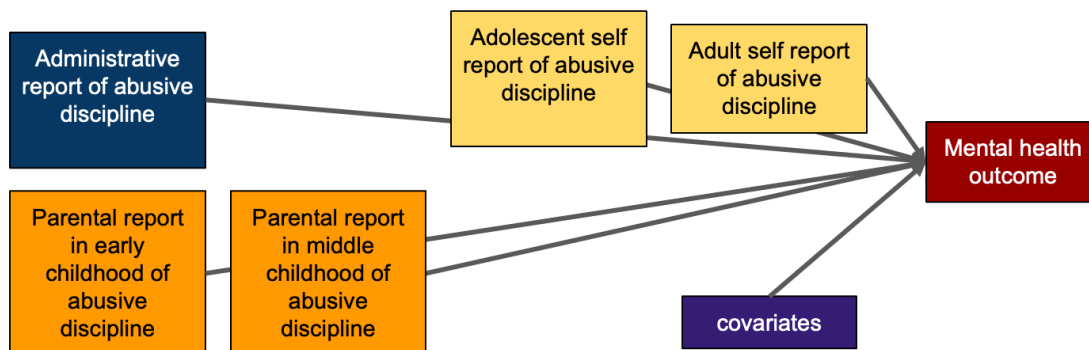$$P = \frac{P1 + P2}{2} \quad (2)$$
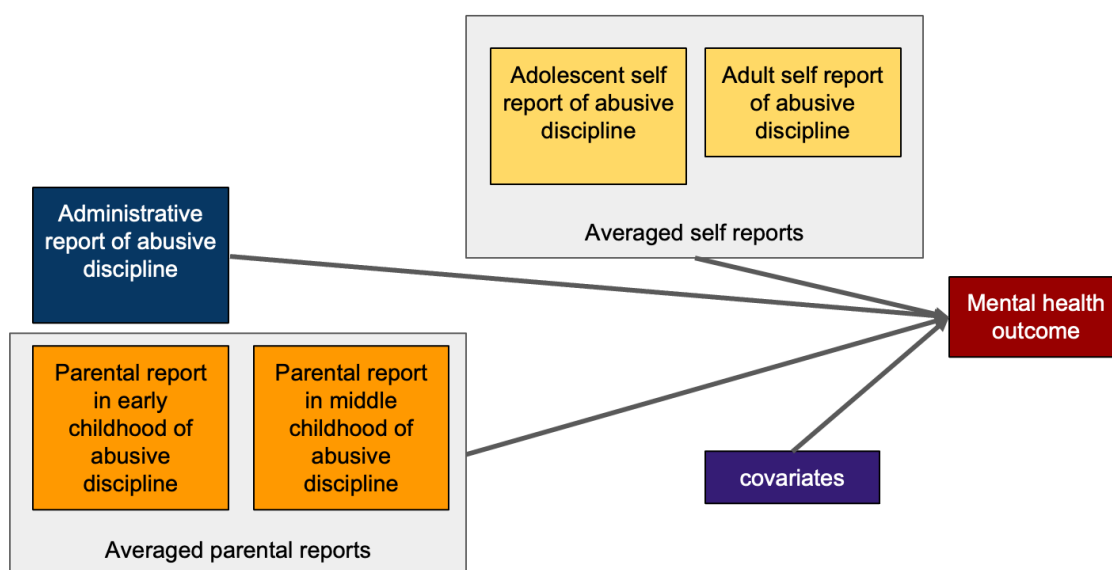
Then, we average self reports:

Figure 2: OLS

Figure 3: summing across reporters

$$SR = \frac{SR1 + SR2}{2} \tag{3}$$

Lastly, we estimate an OLS model in which averaged parental and self reports are variables in the model.

$$\text{outcome} = \beta_0 + \beta\text{P} + \beta\text{SR} + \beta\text{administrative} + \Sigma\beta\text{covariates} + e_i \tag{4}$$

4.2.3 Syntax

```
generate P = (P1 + P2) / 2 // is averaging appropriate?

generate SR = (SR1 + SR2) / 2 // is averaging appropriate?

regress outcome P SR administrative covariates
```

4.3 Path Model

4.3.1 Diagram

4.3.2 Equation

$$\text{outcome} = \beta_0 + \beta\text{P1} + \beta\text{P2} + \beta\text{SR1} + \beta\text{SR2} + \beta\text{administrative} + \Sigma\beta\text{covariates} + e_i \tag{5}$$

$$\text{SR2} = \beta_0 + \beta\text{SR1} + e_i$$
$$\text{P2} = \beta_0 + \beta\text{P1} + e_i$$

4.3.3 Syntax

```
sem (outcome <- covariates SR1 SR2 PR1 PR2 administrative) ///
  (SR2 <- SR1) ///
  (PR2 <- PR1) ///
  cov(e.outcome*e.SR2*e.PR2) // correlated errors
```
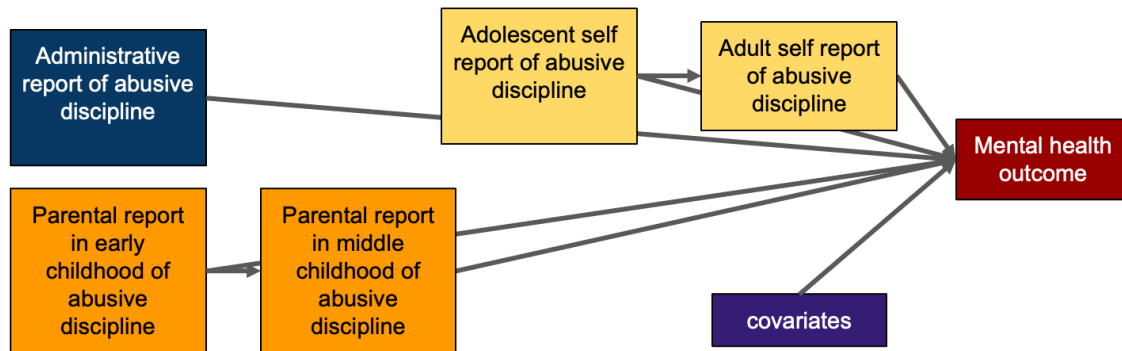
Figure 4: path model

4.4 Latent Construct(s)

4.4.1 Diagram

4.4.2 Equation

4.4.3 Syntax

```
sem ///
  (P1 P2 SR1 SR2 administrative <- X) /// measurement
  (outcome <- covariates X) // structural
```
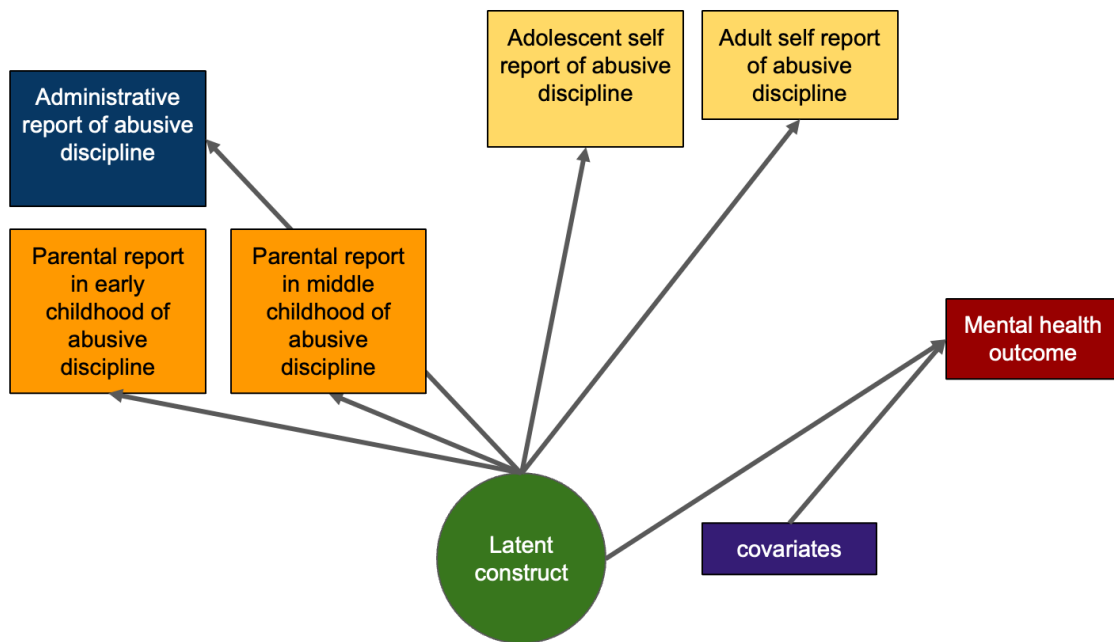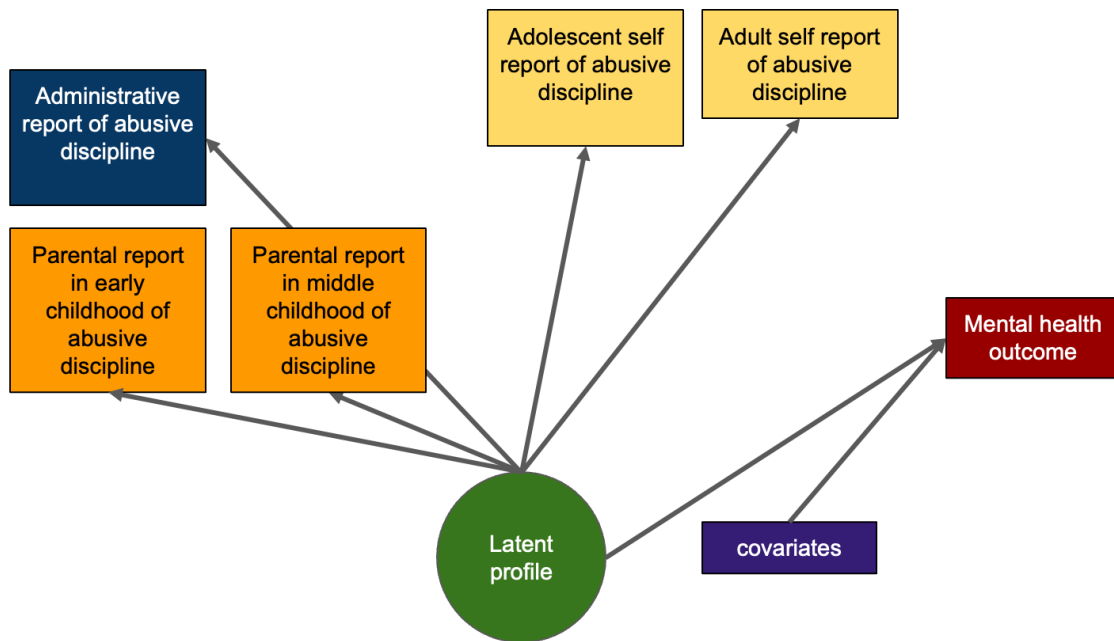
Figure 5: latent construct

Figure 6: latent profile

4.5 Latent Profile Analysis (Person Centered Approach)

4.5.1 Diagram

4.5.2 Equation

4.5.3 Syntax

We first run a latent class analysis to generate latent underlying classes based upon the reports of discipline from the different reporterers

In the syntax below, we estimate three latent classes. The actual number of latent classes is determined by running models with different numbers of latent classes, and comparing those models using *fit statistics*, and *likelihood ratio tests*.

```
gsem (P1 P2 SR1 SR2 administrative <-, gaussian), (lclass(C 3))
```

We then use class membership to predict the outcome.

```
regress outcome i.class covariates
```

4.6 Network Analysis

4.6.1 Diagram

4.6.2 Equation

$$
\begin{array}{ccccccc}
 & \text{P1} & \text{P2} & \text{SR1} & \text{SR2} & \text{administrative} & \text{outcome} \\
\text{P1} & 1 & r_{\text{P1, P2}} & r_{\text{P1, SR1}} & r_{\text{P1, SR2}} & r_{\text{P1, administrative}} & r_{\text{P1, outcome}} \\
\text{P2} & & 1 & r_{\text{P2, SR1}} & r_{\text{P2, SR2}} & r_{\text{P2, administrative}} & r_{\text{P2, outcome}} \\
\text{SR1} & & & 1 & r_{\text{SR1, SR2}} & r_{\text{SR1, administrative}} & r_{\text{SR1, outcome}} \\
\text{SR2} & & & & 1 & r_{\text{SR2, administrative}} & r_{\text{SR2, outcome}} \\
\text{administrative} & & & & & 1 & r_{\text{administrative, outcome}} \\
\\
\text{outcome} & & & & & & 1
\end{array}
\tag{6}
$$

4.6.3 Syntax

```
corr P1 P2 SR1 SR2 administrative outcome
```
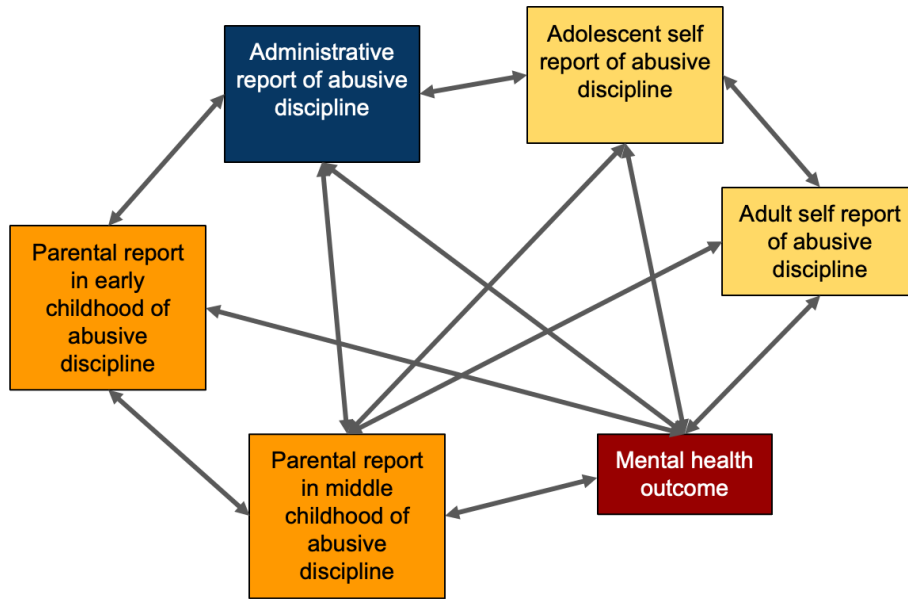
9

Figure 7: network model

4.7 ~~Multilevel Modeling~~

4.8 ~~Classification and Regression Tree (CART) (Machine Learning)~~

4.9 ~~Random Forest (Machine Learning)~~

References

Machlin, L., Sheridan, M. A., Tsai, A. P.-T., & McLaughlin, K. A. (2025). Research review: Assessment of early-life adversity and trauma – cumulative risk and dimensional approaches. *Journal of Child Psychology and Psychiatry, n/a.* https://doi.org/https://doi.org/10.1111/jcpp.14170

McLaughlin, K. A., Sheridan, M., Humphreys, K. L., Belsky, J., & Ellis, B. J. (2020). *The value of dimensional models of early experience: Thinking clearly about concepts and categories.* PsyArXiv. https://doi.org/10.31234/osf.io/29fmt