



Applications of Structural Equation Modeling in Epidemiology

Leila Amorim, Rosemeire Fiaccone, Carlos Teles, Darci Santos, Matildes Prado, Ana Marlúcia Assis, Silvano Oliveira, Maurício Barreto

Outline

- Motivating Applications in Epidemiology

- Structural Equation Models (SEM)
 - Aims
 - Terminology
 - Estimation and evaluation of goodness of fit

- Interpretation of results for applications

- Final Remarks

Motivating Applications in Epidemiology

- Nutrition: Definition of food patterns in 12-24 months aged children from cities in the country side of Bahia, Brazil.
- Psychological Development: Evaluation of risk factors for cognitive development in children from 20 to 42 months of age in Salvador, Bahia, Brazil.

Food Patterns in Children

- Goal: Examine hypothesized food patterns based on most commonly consumed items in the diet of children aged 12-24 months.

- Calories for 826 children from following items:
 - Sweets and sugar
 - Milk and related items (yogurt, cheese,)
 - Breast milk
 - Cereals
 - Beans
 - Rice
 - Meat

Food Patterns in Children

- Hypothesized food patterns:
 - “Milky”: composed by consumption of sweets and sugary foods, milk and related items, breast milk, and cereals.
 - “Traditional”: beans, rice and meat.

- Data Analysis:
 - Description of the contribution of items in the constitution of food patterns.
 - Evaluation of relationship between food patterns.

Cognitive Development in Small Children

- Goal: Evaluate effect of child's anthropometric measures, sex, age, and home quality status in her/his cognitive development.
- Data available on 320 children aged 20-42 months from Salvador, Bahia, Brazil.
- Cognitive development measure using *Bayley Scales for Infant Development*: a widely recognized, reliable, and valid standard instrument was used for assessing mental and motor development (Santos et al, 2008).

Cognitive Development in Small Children

- Anthropometric measures:
 - Birthweight (in kg)
 - Height-age z-score (HAZ)

- Quality of home environment: *Home Observation for Measurement of the Environment* (HOME) inventory, an instrument developed and validated by Caldwell & Bradley.

- Quality of home environment composed of two dimensions: “Organization of child’s environment” and “Parental styles and practices”.

Cognitive Development in Small Children

- Organization of child's environment (home environment):
 - Organization of physical and temporal environment
 - Availability of appropriate play materials and games
 - Variety of daily stimulation

- Parental styles and practices (relational environment):
 - Verbal and emotional responsiveness
 - Avoidance of punishment and restrictions
 - Maternal involvement with the child

- Data Analysis:
 - Definition of constructs to represent relational and home environments and to represent anthropometric measures.
 - Measure the impact of the constructs on the cognitive development scale.

Structural Equation Models

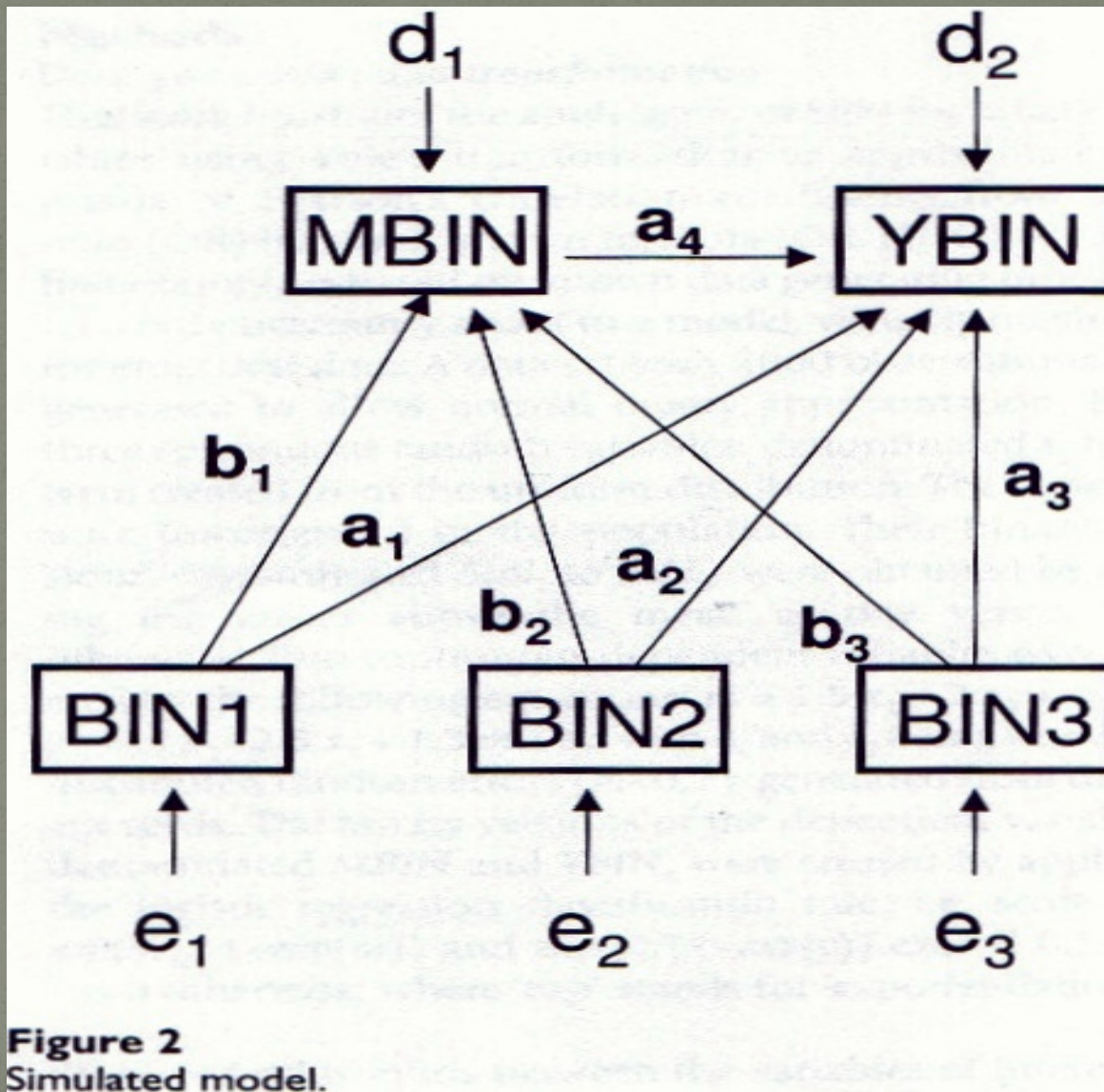
- Important analytical tool to evaluate complex causal relationships in many fields.

- Still of limited use in Epidemiology:
 - Evaluation of models for metabolic syndrome (European Journal of Epidemiology, 2006)
 - Test of a hypothesized model to the concepts of adherence and readiness in antiretroviral treatment (Patient Education and Counseling, 2007)

Structural Equation Models

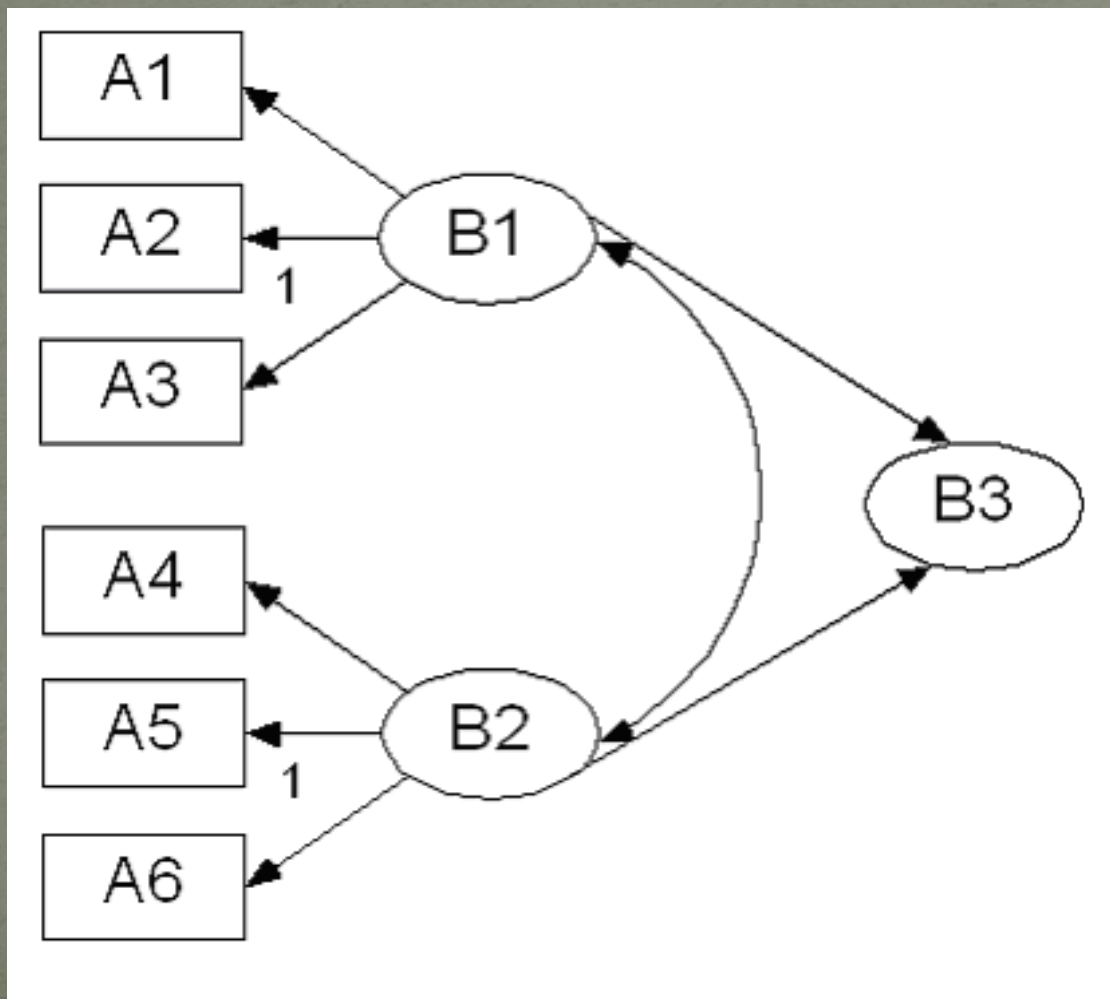
- ❑ Statistical methodology that combines factor analysis and path analysis.
- ❑ SEM is a confirmatory technique rather than exploratory.
- ❑ It allows a hypothesized theoretical framework to be tested statistically.
- ❑ It can incorporate not only observed but also unobserved (latent/construct) variables in the analysis.

Figure 1 – Example of a Path Diagram



Source: Kupek E (2005). In: BMC Medical Research Methodology, 6(13).

Representation of Relationships in SEM



Observed x Latent
variable

Endogenous x
Exogenous variable

Models in SEM

- Two models are considered for estimation:
 - *Measurement model*: definition of latent variables (CFA).
 - *Structural model*: estimation of relationships between latent variables and outcome(s).

Structural Equation Models

- Several indexes can be used to measure the model fit in SEM.

- The statistics usually reported in SEM are:
 - Chi-square [$p < 0.05$]
 - Goodness-of-fit index (GIF) [over 0.90]
 - Comparative fit index (CFI) [over 0.90]
 - Root mean square error of approximation (RMSEA) [below 0.05]

Structural Equation Models

- ❑ Currently several statistical software implement SEM.

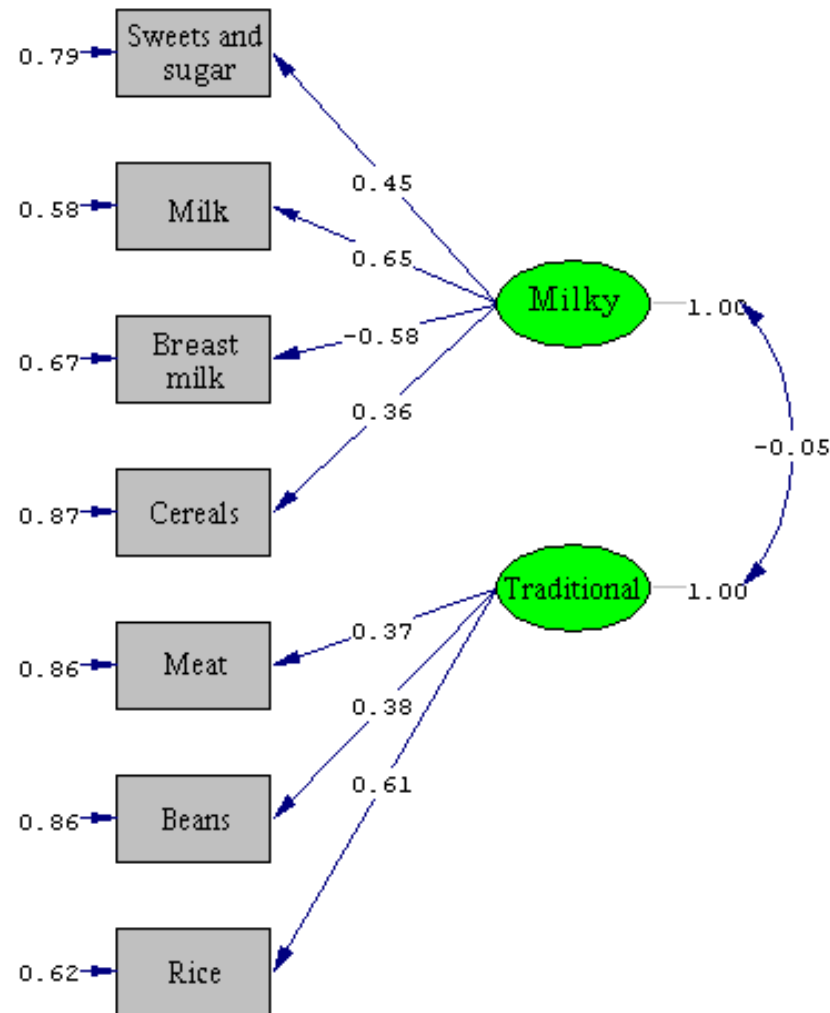
- ❑ The most commonly software used are:
 - LISREL (Linear Structural Relationships)
 - AMOS (Analysis of Moment Structures)
 - MPLUS
 - EQS
 - STATA (GLAMM)
 - R (SEM package)

Applications in Epidemiology

□ Food patterns in children:

- Use of confirmatory factorial analysis (CFA)
- Definition of latent variables: consumption of milky and traditional diet.
- Analysis performed using software LISREL v.8.80.

Figure 2 – Standardized Coefficients in the Path Diagram Analysis of Food Patterns



Chi-Square=27.56, df=13, P-value=0.01042, RMSEA=0.037

LISREL - v.8.80

Cognitive Development in Small Children

- ❑ Use of CFA to define three constructs: relational, environmental and nutritional.
- ❑ SEM adjusted to evaluate the impact of these constructs on IDM (index of cognitive development), while adjusting for age and sex of the child.
- ❑ Analysis performed using AMOS v.6.

Figure 3 – Standardized Estimates from CFA for analysis of cognitive development data

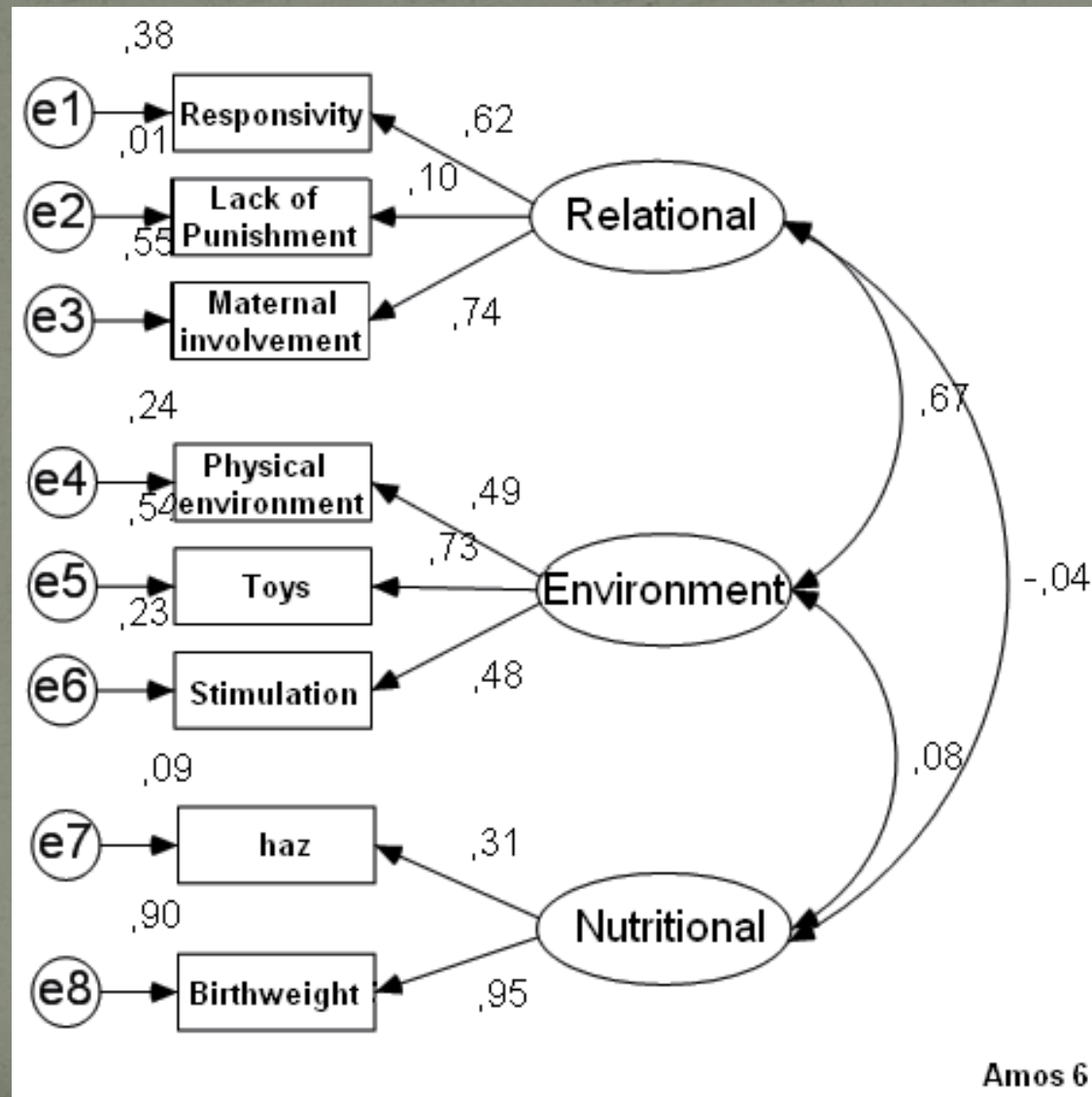
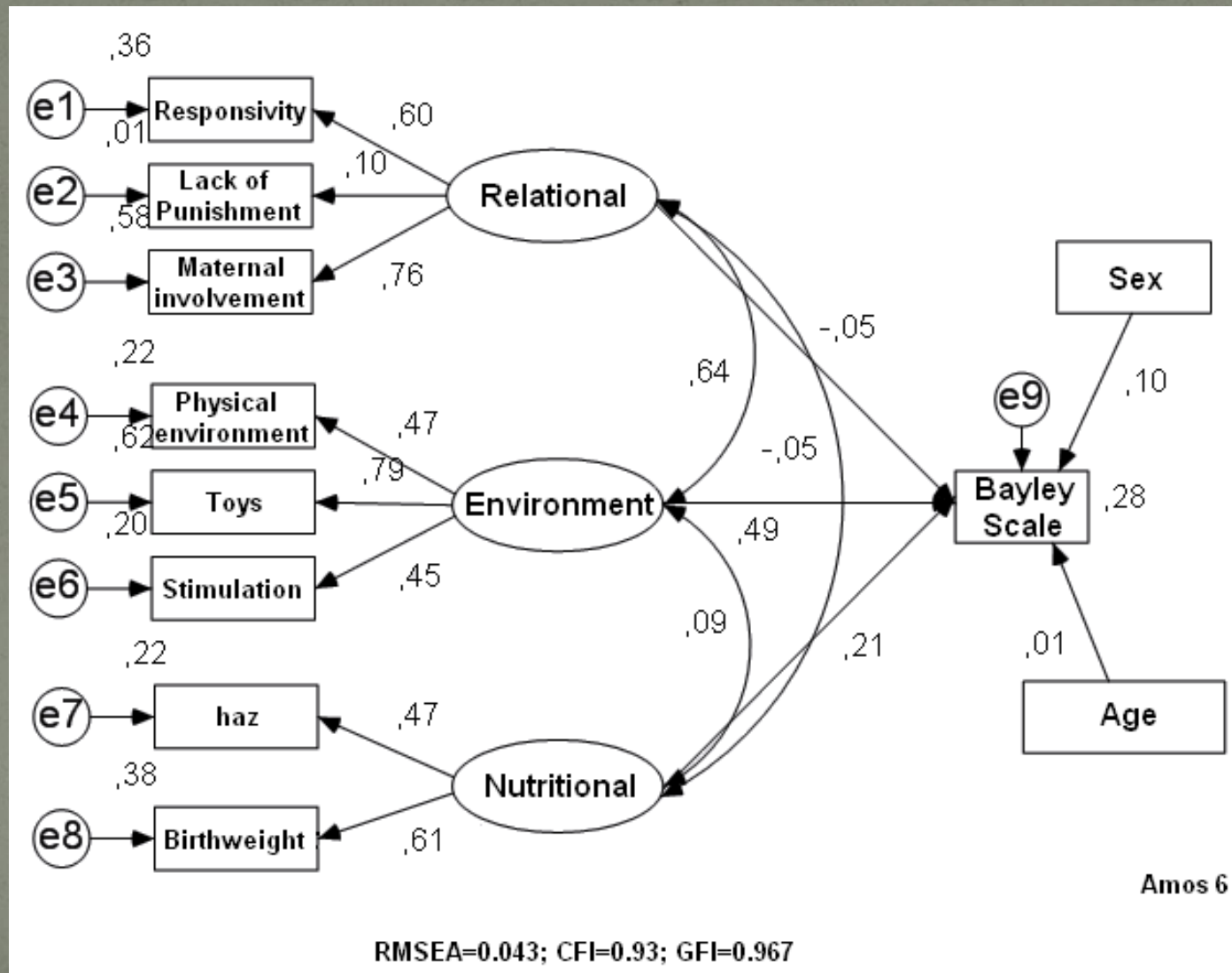


Figure 4 – SEM standardized estimates for analysis of cognitive development data



Final Remarks

- ❑ Hypothesized models for motivating applications fit the data well.
- ❑ However, SEM does not allow us to rule out the possible models/theories.
- ❑ The SEM analysis is based on a confirmatory approach.

- Advantages of SEM approach include:
 - Possibility of modeling several regression equations simultaneously;
 - Inclusion of mediated effects and potential confounding latent variables;
 - Possibility to estimate direct and indirect effects.

- Methodology is available for evaluation of models with binary outcomes as well as complex data structure, such as clustered or longitudinal data.

- Dissemination of this methodology for epidemiological analysis may contribute to evaluate intricate research questions using a new perspective.