Machine-Learning Methods to Investigate Heterogeneity in Longitudinal Patient-Reported Outcome Measures

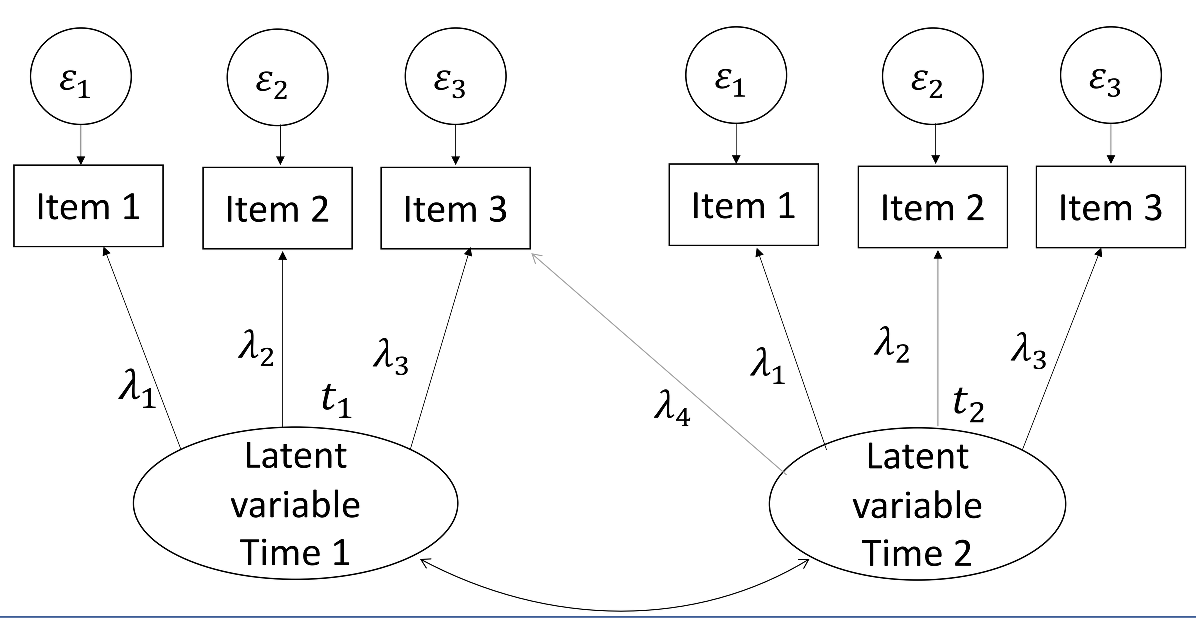
University of Manitoba

The VADA Program
Visual and Automated Disease Analytics Graduate Training Program

Muditha Bodawatte Gedara, Lisa M. Lix
Department of Community Health Sciences, University of Manitoba, Winnipeg, Manitoba

INTRODUCTION

- **Patient-reported outcome measures (PROMs)**
  - Appraisals from patients about their well-being and quality of life
  - Are comprised of multi-item self-report scales that capture perceptions about physical, mental, and social health
  - Provide the patient's perspective on the effectiveness of healthcare interventions, such as surgeries or health programs
  - Are sensitive to differences in interpretation by sub-groups within the population
- **What is Response Shift in PROMs?**
  - Patients may change their interpretation of the items that comprise a multi-item scale over time (i.e., at multiple measurement occasions)
  - Results in inconsistencies in measurement properties of multi-item scales collected over time. Types of response shift: recalibration, reprioritization, and reconceptualization



Note: Item 1, 2, and 3: observed variables, $\varepsilon_1, \varepsilon_2$, and ε_3: error terms, λ_1, λ_2, λ_3, and λ_4: factor loadings

- **Recalibration**: Change in measurement structure leads to changes in error variances over time
- **Reprioritization**: Change in the relative importance of items and latent variables contribute to change in factor loadings over time
- **Reconceptualization**: Change in the way the latent variable is conceptualized leads to changes in factor loading patterns over time

Figure 1: Latent variable model at two measurement occasions

- Observed variables may be responsible for variation in response shift and result in response shift patterns.

- **Item Response Theory Models for Response Shift**
  - Item Response Theory (IRT) models have been used to identify individual PROM scale items that are sensitive to response shift
  - Conventional IRT models assume homogeneity in parameter estimates over time may result in imprecise response shift, if heterogeneity exists
  - Unsupervised machine-learning methods can aid in identifying clusters of individuals with similar response shift patterns

PURPOSE & OBJECTIVES

Purpose: To develop new methods to detect heterogeneity (i.e., variation) in response shift in longitudinal PROMs data

Objectives:

- Develop a longitudinal IRT model that uses unsupervised machine-learning techniques to detect response shift in PROMs data
- Compare this new model with existing statistical models (i.e., conventional IRT model) to detect response shift
- Apply this newly-developed model to real-world clinical data about PROMs for patients having joint replacement surgery

Research Questions

- Do unsupervised and conventional longitudinal IRT models differ in their sensitivity to detect variation in response shift in PROM item responses?
- Are longitudinal unsupervised and conventional IRT equally sensitive to detect different types of response shift in PROMs item responses?

METHODS

Statistical Models

- The implementation of the **conventional IRT models** will follow the steps below.
  - Establish the measurement model for the latent construct (e.g., mental health, physical health)
  - Test for an overall response shift effect
  - Test for response shift on each PROM item
  - Fit the final response shift model and use it to estimate change in the latent variable scores for each individual
- The implementation of the **unsupervised machine-learning IRT model** will follow the steps in **Figure 2**.

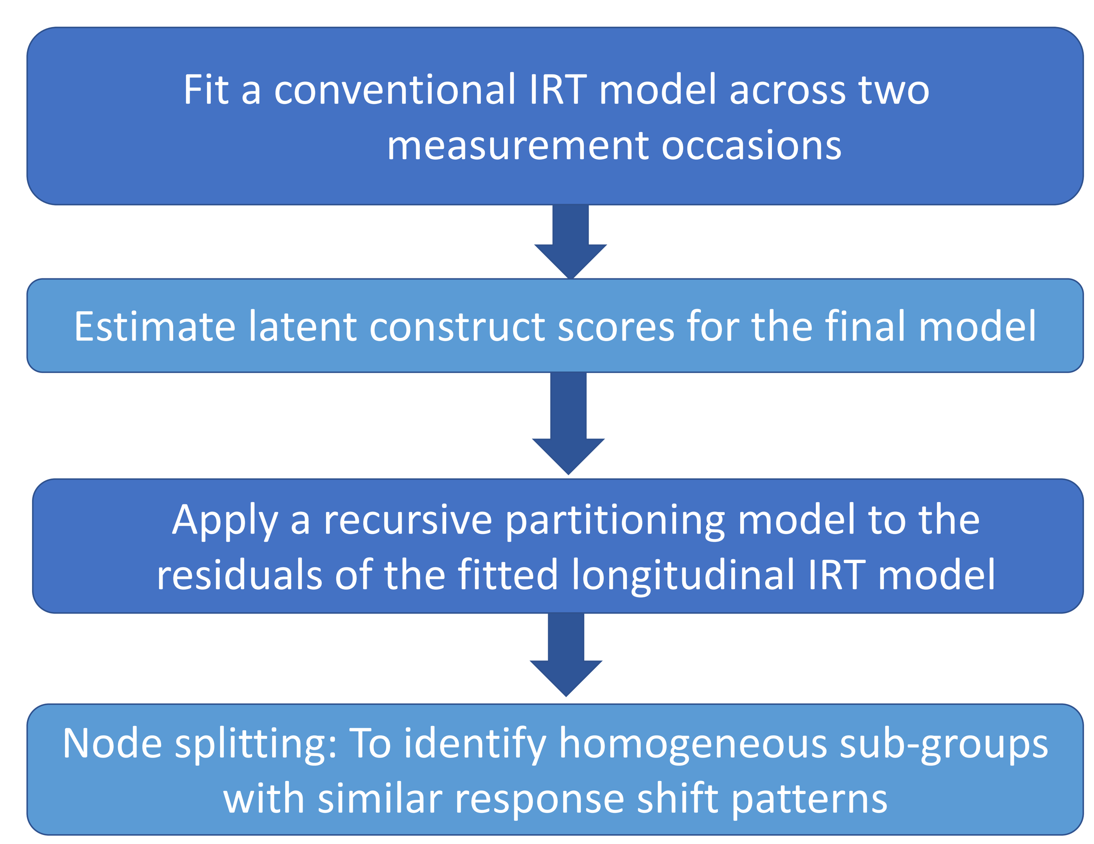


Figure 2: Implementation of unsupervised IRT model

Computer Simulation Studies

- Models will be compared using computer simulations to detect response shift.
- **Simulation study characteristics:**
  - ***Research design characteristics***: total sample size
  - ***Response characteristics***: number of PROMs items measured using an ordinal scale, number of item response categories, type of response shift effect
  - ***Covariate effects:*** number of covariates associated with response shift (i.e., number of clusters with similar response shift patterns)
  - ***Covariance/correlation***: magnitude of correlation among the measurement occasions, variance of the covariates
  - ***Effect size***: magnitude of the true difference between measurement occasions; response shift effect size.

We will then apply two models to real-world data from WRHA

Clinical Data

- Winnipeg Regional Health Authority (WRHA) Joint Replacement Registry
- Covariates: socio-demographic and clinical characteristics
- Both general-purpose and condition-specific PROMs are captured in the data (Table 1).
- Longitudinal Data: PROMs are collected one month prior to surgery and one year following surgery

General-purpose PROMs	Condition-specific PROMs
Short Form Health Survey (SF-12): used to assess both mental and physical health	**Oxford Hip Score:** assess function and pain before and after hip replacement surgery
	Oxford Knee Score: assess function and pain before and after knee replacement surgery

SIGNIFICANCE

- WRHA Joint Replacement Registry contains a rich set of covariates
- Contribute to the development of valid and sensitive PROMs methods
- Help to advance the use of unsupervised machine-learning methods for PROMs data
- Valid analytic techniques will contribute to better interpretation of the patient's perspective

ACKNOWLEDGEMENT

We would like to acknowledge the support and funding provided by the Visual and Disease Analytics (VADA) program and the Natural Sciences and Engineering Research Council of Canada (NSERC).

Contact: Muditha Lakmali - bodawatm@myumanitoba.ca