

RTI(h)(s) RTI HEALTH SOLUTIONS®

Innovative Methods to Identify and Explore Heterogeneity

Introduction and Methodological Overview

Presented by Donald Stull
Director, Retrospective Data Analysis,
RTI Health Solutions, Manchester

RTI Health Solutions
Research Triangle Park, NC, US
Ann Arbor, MI, US
Barcelona, Spain
Lund, Sweden
Manchester, UK
Sheffield, UK
Waltham, MA, US
www.rtihs.org
e-mail: rtihealthsolutions@rti.org

LEADING RESEARCH...
MEASURES THAT COUNT

What Is Heterogeneity?

- Individuals respond differently to treatment or life experiences
 - Not all respondents in the same group respond in the same way or to the same extent
 - Example: treatment nonresponse, placebo response
- Can be due to differences in life situations, experiences, stressors, and genetic make-ups
- Common in clinical and observational studies

Confidential Information 2 **RTI(h)(s)**

Why Is Heterogeneity Important?

- Understanding that particular patients are more, or less, likely to benefit from treatment can ensure optimal treatment administration and health care
- This information can
 - Aid in designing clinical trials that are more appropriately powered and include the most relevant patients
 - Help minimize exposure to potentially toxic treatments for those least likely to respond
 - Reduce the cost and burden of administering treatment to such patients

Confidential Information 3 **RTI(h)(s)**

How Can Heterogeneity Be Identified and Measured?

- An indicator that heterogeneity is present within data is the presence of variability in scores for endpoints during, or at the end of, a study
- Some heterogeneity may be attributable to identifiable (observed) factors
 - Example: genetics, age, gender, comorbidities
- Other heterogeneity must be inferred from the data (unobserved)
 - Example: baseline levels and change over time, different standards of care

Confidential Information 4 **RTI(h)(s)**

Methods for Identifying Sources of Heterogeneity

Classification of Commonly Used Methods for Heterogeneous Populations

Method	Source of heterogeneity	Outcome variables	Latent variable model
Discriminant analysis	Observed	Continuous	No
Logistic regression	Observed	Categorical	No
MANOVA	Observed	Continuous	No
Multigroup CFA	Observed	Continuous and/or categorical	Yes
K-means clustering	Unobserved	Continuous	No
Latent class analysis	Unobserved	Categorical	Yes
Latent profile analysis	Unobserved	Continuous	Yes
Factor mixture modeling	Unobserved	Continuous and/or categorical	Yes

Note. MANOVA = multivariate analysis of variance; CFA = common factor analysis.

Source: Lubke GH, Muthen B. Investigating Population Heterogeneity With Factor Mixture Models. *Psychological Methods* 2005;10(1):21-39.

Unobserved Heterogeneity: Traditional Approach

- In the absence of a predefined responder, an analyst must make multiple comparisons, trying different cut-off values in variables of interest and consequently violating assumptions of independence
 - May require dozens of data slices of different definitions of high, medium, and low on outcome variables

Unobserved Heterogeneity: Innovative Approach

- Methods based on structural equation modeling (SEM) deal more efficiently with measuring unobserved heterogeneity
- Model-based methods have the advantage of using more rigorous approaches to compare alternative models
- Model-based methods can uncover subsets of patients who exhibit within-class homogeneity yet are themselves different from the larger class of patients from which they were drawn

Innovative Approach: Mixture Models

- Cross-sectional: Factor mixture models
- Longitudinal: Growth mixture models
- Known class and/or latent class
 - Allows for both continuous (factor model) and categorical (latent-class model) group assignment
 - Can help detect subclasses such as “placebo responders,” “hyporesponders,” or “hyperresponders” in clinical trials

Mixture Modeling

- What is it?
 - It is *NOT* mixed-effects regression
 - Combines (i.e., is a *mixture* of) common factor model and latent-class model

Confidential Information

9



Mixture Modeling

- Factor models serve to cluster items
 - Latent variables in factor models are continuous
- Latent-class models cluster participants
 - Latent variables in latent-class models are categorical
 - The number of categories (latent classes) represent the number of clusters in the sample

Confidential Information

10



Factor-Mixture Modeling

- Cross-sectional
- Also known as finite-mixture modeling

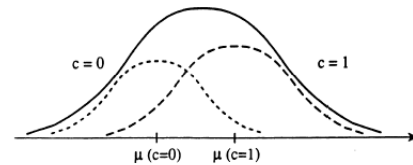
Confidential Information

11



Factor-Mixture Modeling

- Overall population distribution, two subpopulation distributions



Example: Overall height or weight; male and female height or weight.
Example: Overall treatment response; placebo and treatment group responses.

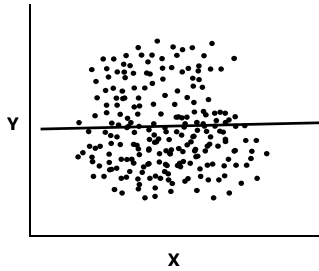
Confidential Information

12



Factor-Mixture Modeling

- Analysis of the overall sample suggests that there is no relationship between X and Y



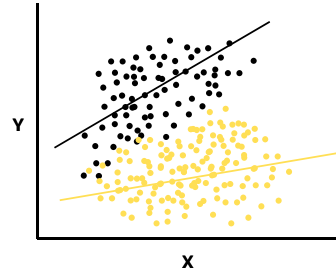
Confidential Information

13



Factor-Mixture Modeling

- Using mixture models, it becomes clear that there are two underlying distributions, each with its own intercept and slope



Confidential Information

14



What Is Different About These Two Latent Classes?

- Each of these latent classes can be examined to see how they differ
- Latent-class assignment for each individual is merged with the original study data, and post hoc comparisons can be made

Confidential Information

15



Growth-Mixture Modeling

- Longitudinal
 - Helps identify subgroups of patients who respond differentially to treatment (or who respond when not expected, i.e., placebo responders)
 - Helps uncover heterogeneity in responses to treatment

Confidential Information

16



Mixture Models

- These analytic methods can uncover a subset (or subsets) of patients who exhibit within-class homogeneity yet as a class are heterogeneous with respect to other classes of patients in the larger sample
- Such subsets of patients can be examined post hoc to see what variables may account for class membership, thus differentiating these patients from the remaining patients

Confidential Information

17



What Are the Benefits of These Innovative Approaches?

- Flexible growth-curve shape
- Individually varying times of observation
- Random effects (intercepts, slopes) integrated with other latent (unobserved) variables
- Regressions among random effects
- Multiple processes (i.e., changes in multiple outcome variables modeled simultaneously)
- Multiple populations (can compare growth processes among groups)
- Multiple indicators (can model measurement error)
- Embedded growth models
- Categorical latent variables allowing for growth mixtures

Confidential Information

18



Example of Hidden Subgroup of Nonresponder

- Treatment (A) and placebo (B) groups
- No difference at baseline
- No significant treatment difference

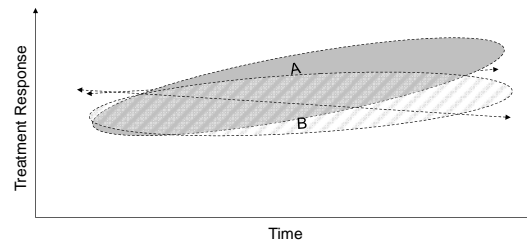
Confidential Information

19



Mixture Models

- Hidden subgroup of treatment nonresponders (C) emerges
 - Consequently, difference between “true” responders and placebo patients increases



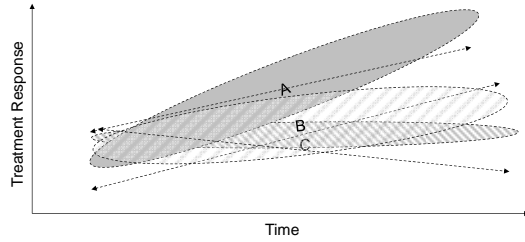
Confidential Information

20



Mixture Models

- Hidden subgroup of treatment nonresponders (C) emerges
 - Consequently, difference between “true” responders and placebo patients increases



Confidential Information

21



When Hidden Subgroup of Treatment Nonresponders Emerges

- Main treatment group (A) shows significant increase in outcome resulting from treatment
 - Thus, modeling nonresponders separately allows for examination of treatment effects on those patients who respond
- Characteristics of nonresponders can be examined in post hoc analyses

Confidential Information

22

