Joint multilevel modeling of a factor analytic and a covariance regression model

Emmanuel Lesaffre

Department of Biostatistics, Erasmus MC, Rotterdam

L-Biostat, KU Leuven

joint work with B. Li, L. Bruyneel and *Youngjo Lee*

LINSTAT 2014, Linköping August 24-28, 2014

Outline

- Motivating data set: RN4CAST project
- Review of covariance modelling
- Multilevel Covariance Regression (MCR) model
- Multilevel Higher-Order Factor (MHOF) model

The RN4CAST project

- Registered Nurse Forecasting FP7 project (Sermeus et al., 2011)
- Nurse survey across Europe (2009-2011)
- **Aim**: Study the impact of system-level features of nursing care on nurse wellbeing and patient safety outcomes on burnout, . . .
- **Swedish data removed** (no nursing unit information) & restriction to female nurses

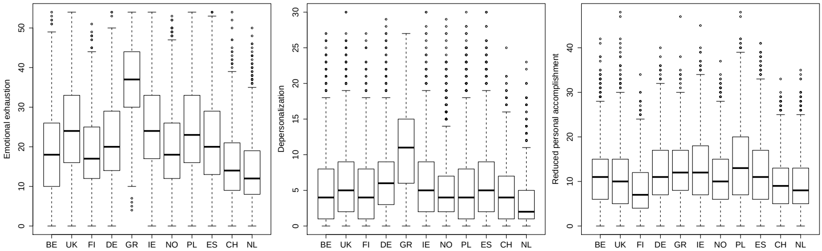
  $\Rightarrow$ 21,016 nurses, 2023 nursing units, 345 hospitals, 11 countries

RN4CAST project – outcomes of interest

- Three dimensions of burnout
  - Emotional exhaustion (EE)
  - Depersonalization (DP)
  - Reduced personal accomplishment (PA)
- Measured using the 22-item Maslach Burnout Inventory:
  - Q: "I feel emotionally drained from my work" (EE)
  - A: 0-never; 1-a few times a year or less; ...; 6-every day
- EE (9), DP (5) & PA (8) are sum scores within each dimension

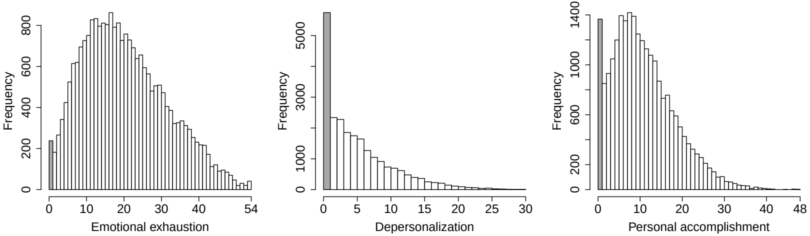
RN4CAST project – outcomes of interest

- Distribution of burnout per country



RN4CAST project – outcomes of interest

- Distribution of burnout across countries



- No classical transformation to normality

RN4CAST project – covariates of interest

- Working experience (yrs): working years being a registered nurse
- Work environment: average summary of practice environment scale of nursing working index
  - Item: ”Praise and recognition for a job well done”
  - Score: “Totally agree”=4, “Agree”=3, “Not agree”=2, “Totally not agree”=1
  - High values reflect a positive environment
- Teaching hosp (university hospital = 1, else = 0)
- Technical hosp ((heart/transplant) surgery present = 1, else = 0)
- Type of nursing unit (surgical = 1 or medical = 0)

RN4CAST project – covariates of interest

	Working experience(yrs)*	Work environment*	Size*†
Country	13.90 (9.05,18.84)	2.53 (2.25,2.87)	–
Hospital	14.29 (5.05,27.76)	2.52 (1.71,3.26)	483.60 (30,3213)
Nursing unit	13.92 (0.34,41.00)	2.54 (1.43,3.62)	11.36 (1,71)
Nurse	13.89 (0.05,50.00)	–	–

$*$: Mean (and range)

†: No. of beds at hospital level and No. of available nurses at nursing unit level

‡: Percentage

RN4CAST project – covariates of interest

	Working experience(yrs)*	Work environment*	Size*†
Country	13.90 (9.05,18.84)	2.53 (2.25,2.87)	–
Hospital	14.29 (5.05,27.76)	2.52 (1.71,3.26)	483.60 (30,3213)
Nursing unit	13.92 (0.34,41.00)	2.54 (1.43,3.62)	11.36 (1,71)
Nurse	13.89 (0.05,50.00)	–	–

∗: Mean (and range)
†: No. of beds at hospital level and No. of available nurses at nursing unit level
‡: Percentage

	Teaching hospital‡	Technical hospital‡	Surgery nursing unit‡
Country	–	–	–
Hospital	23.77%	28.99%	–
Nursing unit	–	–	49.88%
Nurse	–	–	–

RN4CAST project – research questions

- **Q1**: Are the **means** of 3 burnout dimensions associated with organizational & individual nurse characteristics?

Multivariate multilevel model for burnout means

RN4CAST project – research questions

- **Q1**: Are the **means** of 3 burnout dimensions associated with organizational & individual nurse characteristics?

  Multivariate multilevel model for burnout means

- **Q2**: Are the **variances/correlations** of 3 burnout dimensions stable across hospitals, nursing units and nurses, after taking into account a rich set of confounders at different levels?

  Multivariate multilevel model for burnout covariance matrix

RN4CAST project – proposed solutions

- **PART I:** sum scores as responses ⇒ multilevel covariance regression (MCR) model

RN4CAST project – proposed solutions

- **PART I:** sum scores as responses $\Rightarrow$ multilevel covariance regression (MCR) model

- **PART II:** original 22 items as responses $\Rightarrow$ multilevel higher-order factor (MHOF) model

  = combination of MCR model with multilevel factor analytic model

PART I: The Multilevel Covariance Regression (MCR) Model

Sum scores as responses (Li et al., 2013)

Brief review of covariance modelling

- Univariate multilevel (2-level) case
- Multivariate single level case
- Multivariate multilevel (2-level) case

Brief review of covariance modelling

Univariate multilevel case:

- Modeling variance with covariate x^*:

$$y_{ij} = \boldsymbol{x}_{ij}^T \beta + u_j + \varepsilon_{ij}$$
$$u_j \sim N(0, \sigma_u^2), \quad \varepsilon_{ij} \sim N(0, \sigma_{\varepsilon ij}^2), \quad \varepsilon_{ij} \perp u_j$$
$$\sigma_{\varepsilon ij}^2 = \rho(\boldsymbol{x}_{ij}^{*T} \beta^*)$$

- x^* could be continuous or categorical

Brief review of covariance modelling

Univariate multilevel case:

- Modeling variance further with random effects:

$$y_{ij} = \boldsymbol{x}_{ij}^T \beta + u_j + \varepsilon_{ij}$$
$$u_j \sim N(0, \sigma_u^2), \quad \varepsilon_{ij} \sim N(0, \sigma_{\varepsilon ij}^2), \quad \varepsilon_{ij} \perp u_j$$
$$\sigma_{\varepsilon ij}^2 = \rho(\boldsymbol{x}_{ij}^{*T} \beta^* + u_j^*), \quad u_j^* \perp u_j$$

- Foulley et al. (1992)
- DHGLM (Double hierarchical generalized linear model) (Lee and Nelder, 2006)

Brief review of covariance modelling

Multivariate single level case:

- Multiple (p) correlated responses

$$\mathbf{y}_i = \boldsymbol{B}\boldsymbol{x}_i + \varepsilon_i$$
$$\varepsilon_i \sim N(\mathbf{0}, \Sigma_\varepsilon)$$

- Σ_ε: $p \times p$ residual covariance matrix
- Let Σ_ε depend on covariates: $\Sigma_\varepsilon(\boldsymbol{x}^*)$
- **Problem**: Positive definiteness (pd) of $\Sigma_\varepsilon(\boldsymbol{x}^*)$

Brief review of covariance modelling

Multivariate single level case:

- Naive solution:
  - Model each covariance element directly
  - pd not guaranteed
- Overview of alternative solutions:
  - Logarithm transformation
  - Separation strategy
  - Modified Cholesky decomposition
  - Covariance regression

Brief review of covariance modelling

Matrix logarithm transformation of Σ_ε (Chiu et al., 1996):

- Definition matrix logarithmic transformation of C:
  - A symmetric, then $C = \exp(A) = \sum_{s=0}^{\infty} \frac{A^s}{s!}$ is pd $\Rightarrow$ A := log(C)
- Property matrix logarithmic transformation of C:
  - For each pd C, $\exists$ a A symmetric, such that $C = \exp(A)$

    $\Rightarrow$ Let upper triangular (unconstrained) elements of A depend on covariates

    $\Rightarrow$ **pd problem is solved**
- Interpretation: submatrix of $\Phi_\varepsilon \neq$ log(submatrix of Σ_ε)
- Too many parameters to estimate

Erasmus MC
University Medical Center Rotterdam

Brief review of covariance modelling

Separation strategy (Barnard et al., 2000):

- Separate covariance matrix into SD ($diag(S)$) and correlation (R) parts:
  $\Sigma_\varepsilon = diag(S)\, R\, diag(S)$
- Model each element in S with $\boldsymbol{x}^*$, but assume R constant $\Rightarrow$ **pd problem is solved**
- Not satisfying our needs here
- Too many parameters to estimate

Erasmus MC
University Medical Center Rotterdam

Brief review of covariance modelling

Modified Cholesky decomposition (Pourahmadi, 1999):

- Modified Cholesky decomposition: $T\Sigma_{\varepsilon}T^{T} = D$
- Interpretation:
  - T: conditional linear regression coefficients
  - D: conditional error variances
  - T and D can be expressed in covariates in an unconstrained manner $\Rightarrow$ **pd problem is solved**
- But only, when there is a natural ranking of responses
- And ... too many parameters to estimate

Brief review of covariance modelling

Covariance regression (Hoff and Niu, 2012):

$$\boxed{\Sigma_{\boldsymbol{x}^*} = A + \boldsymbol{B}\boldsymbol{x}^* {\boldsymbol{x}^*}^T \boldsymbol{B}^T}$$

- "Baseline" matrix A plus a matrix depending on $\boldsymbol{x}^*$, $\boldsymbol{B}$ is the coefficient matrix of $\boldsymbol{x}^*$ $\Rightarrow$ pd problem is solved
- Interpretation is intuitive: quadratic relationship in covariates
- Parsimonious representation effect of covariates

Brief review of covariance modelling

Covariance regression:

- A random-effects representation:

$$\boldsymbol{y}_i = \boldsymbol{B}\boldsymbol{x}_i + F_i \times \boldsymbol{B}^* \boldsymbol{x}_i^* + \varepsilon_i$$
$$\varepsilon_i \sim N(\mathbf{0}, \Sigma_\varepsilon), \quad F_i \sim N(0, 1), \quad F_i \perp \varepsilon_i$$

$\Rightarrow$ Factor model with loadings depending on covariates

$\Rightarrow$ Useful for modeling

Brief review of covariance modelling

Multivariate multilevel case:

- Very few publications on this subject
- We propose a solution through a factor model
- Extension of Hoff and Niu's covariance regression model = **Multilevel covariance regression (MCR) model**
- An example of a hierarchical (multivariate) generalized linear model with a factor structure: **HGLM factor model**

Multilevel covariance regression model

Outline:

- Model specification
- Implied marginal distribution
- Computational approaches
- Application to RN4CAST data set

Multilevel covariance regression model

A 2-level MCR model (i = subject, j = cluster):

$$\boldsymbol{y}_{ij} = \boldsymbol{B}\boldsymbol{x}_{ij} + \boldsymbol{u}_j + \boldsymbol{\delta}_{ij}$$
$$\boldsymbol{\delta}_{ij} = \boldsymbol{\lambda}_{ij} F_{ij} + \boldsymbol{\varepsilon}_{ij}$$
$$\boldsymbol{\lambda}_{ij} = \boldsymbol{B}^{*}\boldsymbol{x}_{ij}^{*} + \boldsymbol{u}_j^{*}$$

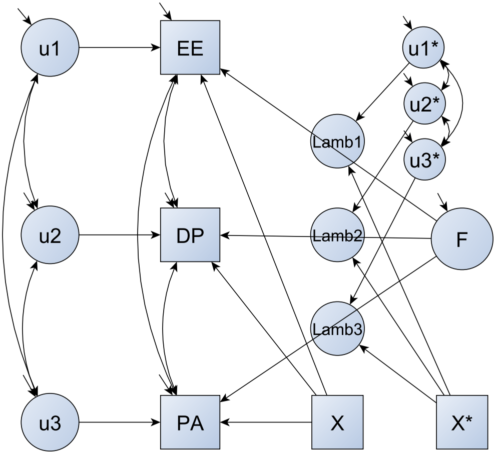
$$\boldsymbol{u}_j \sim N(\mathbf{0}, \Sigma_u), \quad \boldsymbol{u}_j^{*} \sim N(\mathbf{0}, \Sigma_u^{*})$$
$$F_{ij} \sim N(0, 1), \quad \varepsilon_{ij} \sim N(\mathbf{0}, \Sigma_\varepsilon)$$
$$\delta_{ij} \perp \boldsymbol{u}_j \quad \& \quad F_{ij} \perp \varepsilon_{ij}, \boldsymbol{u}_j^{*}$$

- The factor model guarantees a pd Σ_{ij} = covariance matrix of δ_{ij}

Erasmus MC
University Medical Center Rotterdam

Multilevel covariance regression model

Applied to RN4CAST study:



Multilevel covariance regression model

Properties 2-level model:

- Covariance matrix of response (conditional on random effects):

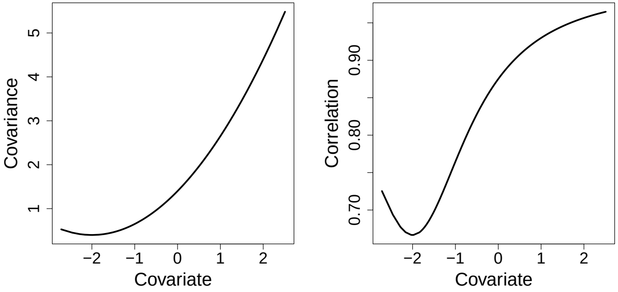
$$\Sigma_{ij} = \Sigma_{\varepsilon} + (\boldsymbol{B}^* \boldsymbol{x}_{ij}^* + \boldsymbol{u}_j^*)(\boldsymbol{B}^* \boldsymbol{x}_{ij}^* + \boldsymbol{u}_j^*)^T$$

- Covariates of each level can be included
- Single level case: Hoff and Niu's covariance regression model
- Easy interpretation: quadratic relationship as a function of covariates
- For 3 responses + no covariates: FA model reconstructs covariance matrix completely

Erasmus MC
University Medical Center Rotterdam

Multilevel covariance regression model

Relationship between covariance/correlation and covariate:



Multilevel covariance regression model

Implied marginal distribution:

- The marginal covariance matrix of the responses is:

$$\begin{aligned} \Psi_{ij} &= (\boldsymbol{B}^* \boldsymbol{x}_{ij}^*)(\boldsymbol{B}^* \boldsymbol{x}_{ij}^*)^T + \Sigma_u + \Sigma_u^* + \Sigma_\varepsilon \\ &= \boldsymbol{b} + \boldsymbol{a} + \boldsymbol{a}^* + \boldsymbol{c} \end{aligned}$$

- Marginal distributions of the responses are not normal
- Zero skewness for mutually independent random effects
- (excess) Kurtosis for the qth response:

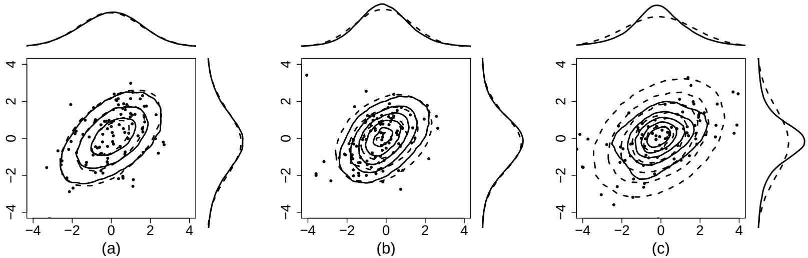
$$\text{kurtosis}_q = \frac{6a_q^{*2} + 12a_q^* b_q}{(a_q + a_q^* + b_q + c_q)^2}$$

a_q, a_q^*, b_q and c_q = qth diagonal element of $\boldsymbol{a}$, $\boldsymbol{a}^*$, $\boldsymbol{b}$ and $\boldsymbol{c}$

Multilevel covariance regression model

Implied marginal distribution:

- 3 bi-variate scenarios with kurtosis: (a) 0.24, (b) 1.50, (c) 3.60



Solid line (MCR model), dashed line (Gaussian model)

- MCR is capable of fitting heavier-tailed distributions

Computational procedures

- Classical likelihood approach – EM algorithm (Hoff & Niu's paper, but not here)
- Bayesian approach
- h-likelihood approach

Computational procedures

Bayesian approach:

- MCMC technique was used, since:
  - Large number of random effects & latent variables
  - Various distributions for random effects & latent variables

- Software: JAGS via R packages *rjags/dclone*
- Model selection: DIC and PSBF (Pseudo Bayes Factor)
- Convergence check: trace plots and Brooks-Gelman-Rubin checks
- Goodness of fit: PPC (Posterior Predictive Check) with χ^2 discrepancy function

Computational procedures

Bayesian approach – identification issue:

- **No random effects** in the loading part, factor part is:

$$(\beta_0^* + \beta_1^* x_{ij}^*) F_{ij}$$

Computational procedures

Bayesian approach – identification issue:

- **No random effects** in the loading part, factor part is:

$$(\beta_0^* + \beta_1^* x_{ij}^*)F_{ij}$$

- Since $F_{ij} \sim N(0, 1)$:
  $(\beta_0^* + \beta_1^* x_{ij}^*)F_{ij} \Longleftrightarrow (-\beta_0^* - \beta_1^* x_{ij}^*)(-F_{ij})$ – "flipping states"
- Different Markov chains result in symmetric (around 0) solutions
- Solution: flip negative chains over to positive

Computational procedures

Bayesian approach – identification issue:

- **With random effects** in the loading part:

$$(\beta_0^* + \beta_1^* x_{ij}^* + \boldsymbol{u}_j^*) F_{ij}$$

Erasmus MC
University Medical Center Rotterdam

Computational procedures

Bayesian approach – identification issue:

- **With random effects** in the loading part:

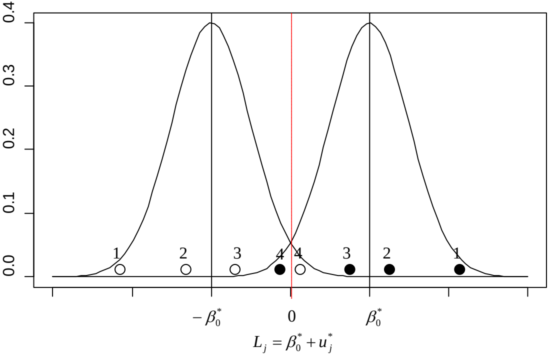
$$(\beta_0^* + \beta_1^* x_{ij}^* + \boldsymbol{u}_j^*)F_{ij}$$

- Flipping states issue is more complicated:
  - $\mathbf{L}_j = \beta_0^* + \boldsymbol{u}_j^*$
  - For cluster j, factor loading is $(\mathbf{L}_j + \beta_1^* x_{ij}^*)$ or $(-\mathbf{L}_j - \beta_1^* x_{ij}^*)$
  - For uni-modal distribution of $\mathbf{L}_j$ $(\boldsymbol{u}_j^*)$
    - Overall mean estimate of $\mathbf{L}_j$, i.e. $\hat{\beta}_0^*$ is close to zero
    - Σ_u^* is overestimated
  - Solution: Take
    - Vague prior on β_0^* and β_1^*
    - Bi-modal distribution for $\mathbf{L}_j$

Computational procedures

Bayesian approach – identification issue:

- Solution: $\mathbf{L}_j \sim 0.5N(-\beta_0^*, \Sigma_u^*) + 0.5N(\beta_0^*, \Sigma_u^*)$



$\Rightarrow$ All parameters in loading part can be identified up to a sign

Computational procedures

h-likelihood approach:

- **Extended likelihood**:

$$L_E(\beta, \boldsymbol{\sigma}, \boldsymbol{v} \mid \boldsymbol{y}, \boldsymbol{v}) = \prod_j \prod_i f_{\beta,\boldsymbol{\sigma}}(\boldsymbol{y}_{ij} \mid \boldsymbol{v}_{ij})\, f_{\boldsymbol{\sigma}}(\boldsymbol{v}_{ij})$$

with $\beta = \{\boldsymbol{B}, \boldsymbol{B}^*\}$, $\boldsymbol{\sigma} = \{\Sigma_u, \Sigma_u^*\}$, $\boldsymbol{v}_{ij} = \{\boldsymbol{u}_j, \boldsymbol{u}_j^*, F_{ij}\}$

Computational procedures

h-likelihood approach:

- **Extended likelihood**:

$$L_E(\beta, \boldsymbol{\sigma}, \boldsymbol{v} \mid \boldsymbol{y}, \boldsymbol{v}) = \prod_j \prod_i f_{\beta,\boldsymbol{\sigma}}(\boldsymbol{y}_{ij} \mid \boldsymbol{v}_{ij})\, f_{\boldsymbol{\sigma}}(\boldsymbol{v}_{ij})$$

  with $\beta = \{\boldsymbol{B}, \boldsymbol{B}^*\}$, $\boldsymbol{\sigma} = \{\Sigma_u, \Sigma_u^*\}$, $\boldsymbol{v}_{ij} = \{\boldsymbol{u}_j, \boldsymbol{u}_j^*, F_{ij}\}$

- Lee and Nelder (1996) proposed **hierarchical (h)- likelihood** approach
  - h-likelihood = extended likelihood when random effects combine additively with fixed effects in linear predictor
  - $\log(L_E)$ is called h-likelihood
  - Marginal likelihood computed by Laplace approximations

- Here h-likelihood approach generalized to HGLM factor models

Application to RN4CAST study

Outline:

- Description of data and research aims (again)
- Multilevel model for RN4CAST
- Modeling aspects
- Results: statistical and clinical

Description of data and research aims

- RN4CAST data set
  - 21,016 nurses, 2023 nursing units, 345 hospitals, 11 countries
  - Outcomes: EE, DP, PA
  - Covariates: working experience, work environment, teaching hospital, technology hospital, type of nursing unit
- Aims: Evaluate in a multi-level context relationship of covariates with
  - Means of burnout
  - Covariance matrix of burnout

Description of data and research aims

- RN4CAST data set
  - 21,016 nurses, 2023 nursing units, 345 hospitals, 11 countries
  - Outcomes: EE, DP, PA
  - Covariates: working experience, work environment, teaching hospital, technology hospital, type of nursing unit
- Aims: Evaluate in a multi-level context relationship of covariates with
  - Means of burnout
  - Covariance matrix of burnout
- Implies fitting a 3-variate 4-level model in mean and covariance
  - Here: results of Bayesian analysis
  - H-likelihood approach gave basically the same results

Multilevel model for RN4CAST

3-variate 4-level model in mean & covariance:

$$\boldsymbol{y}_{ijkl} = \boldsymbol{B}\boldsymbol{x}_{ijkl} + \boldsymbol{u}_{jkl} + \boldsymbol{u}_{kl} + \mathbf{u}_{l} + \delta_{ijkl}$$

$$\delta_{ijkl} = \boldsymbol{\Lambda}_{ijkl} F_{ijkl} + \varepsilon_{ijkl}, \quad \boldsymbol{\Lambda}_{ijkl} = \boldsymbol{B}^{*}\boldsymbol{x}^{*}_{ijkl} + \boldsymbol{u}^{*}_{jkl} + \boldsymbol{u}^{*}_{kl} + \boldsymbol{u}^{*}_{l}$$

$$\boldsymbol{u}_{jkl} \sim N(\mathbf{0}, \Sigma_{u}), \quad \boldsymbol{u}_{kl} \sim N(\mathbf{0}, \Sigma_{h}), \quad \boldsymbol{u}_{l} \sim N(\mathbf{0}, \Sigma_{c})$$

$$\boldsymbol{u}^{*}_{jkl} \sim N(\mathbf{0}, \Sigma^{*}_{u}), \quad \boldsymbol{u}^{*}_{kl} \sim N(\mathbf{0}, \Sigma^{*}_{h}), \quad \boldsymbol{u}^{*}_{l} \sim N(\mathbf{0}, \Sigma^{*}_{c})$$

$$F_{ijkl} \sim N(0, 1), \quad \varepsilon_{ijkl} \sim N(\mathbf{0}, \Sigma_{\varepsilon})$$

All random parts independent

Modeling aspects

- **Covariates aggregation**: partition the covariate into each level (Neuhaus and Kalbfleisch, 1998)

$$\begin{aligned} x_{ijkl} &= (x_{ijkl} - \bar{x}_{jkl}) + (\bar{x}_{jkl} - \bar{x}_{kl}) + (\bar{x}_{kl} - \bar{x}_l) + \bar{x}_l \\ &= x_n + x_u + x_h + x_c \end{aligned}$$

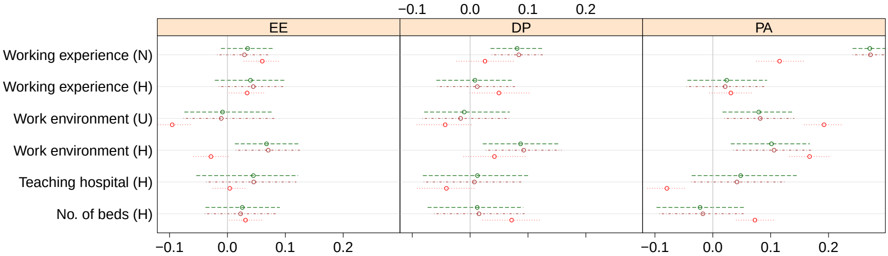
- **Non-normal burnout**: apply BOS approach (Lesaffre et al., 2007)
- **Missing data**:
  - Missing response: treat sum scores with missing items as interval censored
  - Missing covariates: assume stochastic + jointly sample from posterior predictive distribution

Results

- **Statistical**: MCR model is a significant improvement (DIC, PSBF) over considering equal covariance matrices
- **Clinical** - Mean part of burnout:
  - Longer working experience ⇒ less burnout at all levels
  - Better work environment ⇒ less burnout at nursing unit & hospital level
  - Nurses working in surgical nursing unit have more burnout
- **Clinical** - Covariance part of burnout:
  - Correlations rather stable across models
  - Experienced nurses have a larger variance of burnout
  - Random effects: variance of burnout differs across units

Results: statistical and clinical

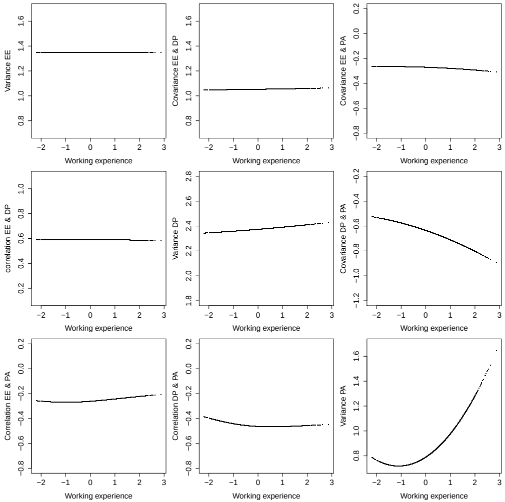
Fixed effect estimates in factor loadings



brown dashed-dotted line (our model)

Results

Impact of covariates on (co)variances and correlations



PART II: The Multilevel Higher-Order Factor (MHOF) model

Original 22 items as responses (Li et al., 2014)

Dealing with high-dimensional response

- Burnout was originally measured through 22 items
- Three dimensions proposed by Maslach and Jackson (1981) were obtained from a different population
- These dimensions may be different in RN4CAST study
- Alternative: model the original 22 items directly

Dealing with high-dimensional response

- Burnout was originally measured through 22 items
- Three dimensions proposed by Maslach and Jackson (1981) were obtained from a different population
- These dimensions may be different in RN4CAST study
- Alternative: model the original 22 items directly

- Possible analysis strategies:
  - Higher rank model:

$$\boldsymbol{y}_i = \boldsymbol{B}\boldsymbol{x}_i + F_i \times \boldsymbol{B}^{*}\boldsymbol{x}_i^{*} + G_i \times \boldsymbol{B}^{**}\boldsymbol{x}_i^{*} + \varepsilon_i$$

$$\varepsilon_i \sim N(\mathbf{0}, \Sigma_\varepsilon), \quad F_i \sim N(0,1), \quad F_i \perp \varepsilon_i, \quad G_i \sim N(0,1), \quad G_i \perp F_i, \quad G_i \perp \varepsilon_i$$

Dealing with high-dimensional response

- Burnout was originally measured through 22 items
- Three dimensions proposed by Maslach and Jackson (1981) were obtained from a different population
- These dimensions may be different in RN4CAST study
- Alternative: model the original 22 items directly

- Possible analysis strategies:
  - Higher rank model:

$$\boldsymbol{y}_i = \boldsymbol{B}\boldsymbol{x}_i + F_i \times \boldsymbol{B}^* \boldsymbol{x}_i^* + G_i \times \boldsymbol{B}^{**} \boldsymbol{x}_i^* + \varepsilon_i$$

$$\varepsilon_i \sim N(\boldsymbol{0}, \Sigma_\varepsilon), \quad F_i \sim N(0,1), \quad F_i \perp \varepsilon_i, \quad G_i \sim N(0,1), \quad G_i \perp F_i, \quad G_i \perp \varepsilon_i$$

  - First multilevel factor model (MFA) to find 'intrinsic' burnout dimensions (MFA), then MCR model

Dealing with high-dimensional response

- Burnout was originally measured through 22 items
- Three dimensions proposed by Maslach and Jackson (1981) were obtained from a different population
- These dimensions may be different in RN4CAST study
- Alternative: model the original 22 items directly

- Possible analysis strategies:
  - Higher rank model:

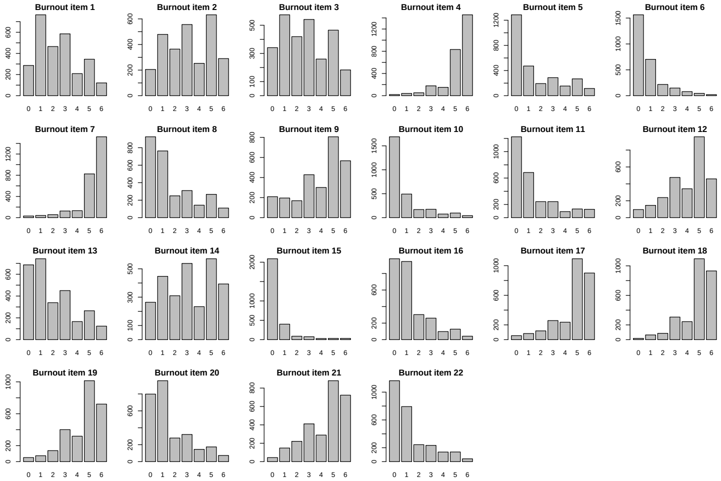
$$\boldsymbol{y}_i = \boldsymbol{B}\boldsymbol{x}_i + F_i \times \boldsymbol{B}^* \boldsymbol{x}_i^* + G_i \times \boldsymbol{B}^{**} \boldsymbol{x}_i^* + \varepsilon_i$$

$$\varepsilon_i \sim N(\mathbf{0}, \Sigma_\varepsilon), \quad F_i \sim N(0,1), \quad F_i \perp \varepsilon_i, \quad G_i \sim N(0,1), \quad G_i \perp F_i, \quad G_i \perp \varepsilon_i$$

  - First multilevel factor model (MFA) to find 'intrinsic' burnout dimensions (MFA), then MCR model
  - Jointly estimate MFA model and MCR model – **MHOF model**

Erasmus MC
University Medical Center Rotterdam

Distribution of the 22 original burnout items



Multilevel factor analytic model

- Find the latent factors underlying a group of variables in a multilevel context
- A two level MFA model is:

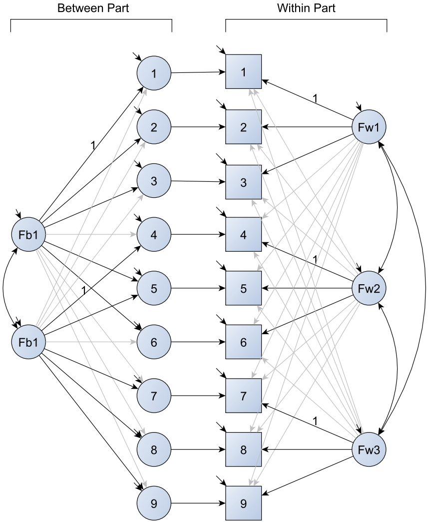
$$\boldsymbol{y}_{ij} = \boldsymbol{\mu} + \boldsymbol{L}_B \boldsymbol{f}_j + \boldsymbol{u}_j + \boldsymbol{L}_W \boldsymbol{f}_{ij} + \varepsilon_{ij}$$
$$\boldsymbol{f}_j \sim N(\mathbf{0}, \Sigma_{fB}), \quad \boldsymbol{u}_j \sim N(\mathbf{0}, \Sigma_u)$$
$$\boldsymbol{f}_{ij} \sim N(\mathbf{0}, \Sigma_{fW}), \quad \varepsilon_{ij} \sim N(\mathbf{0}, \Sigma_\varepsilon)$$
$$i = 1, 2, ..., n_j;\ j = 1, 2, ..., k$$

All random parts independent

- Implied covariance matrix for the MFA model is:

$$\Sigma = \boldsymbol{L}_B \Sigma_{fB} \boldsymbol{L}_B^T + \Sigma_u + \boldsymbol{L}_W \Sigma_{fW} \boldsymbol{L}_W^T + \Sigma_\varepsilon$$

Multilevel factor analytic model



Multilevel higher-order factor model

Multilevel higher-order factor (MOHF) model =

- Is a combination of MFA model with covariance regression model
- If there are **3** intrinsic factors, then MCR model for covariance regression model is a good option
- If there are > 3 intrinsic factors, then MCR model is less optimal but can still be considered

Multilevel higher-order factor model

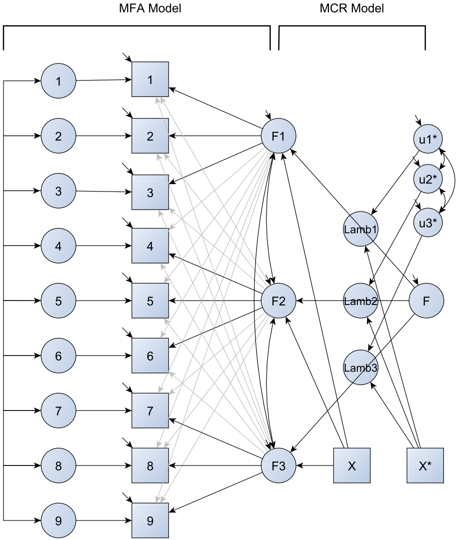
The MFA part:

- Factor structure at the lowest level only
- Estimate the whole covariance matrix at higher levels

The MCR part:

- Use the lowest factor scores as responses
- Include covariates at each level

Model specification



Application to RN4CAST study

- Applied to Belgian part of RN4CAST study
- A 3-level MFA model based on 22 items and a 3-variate 3-level MCR model are jointly estimated
- Same modeling aspects as for MCR model
- Basically same clinical results as before

Conclusion + discussion

- MCR and MHOF inspired by clinical questions
- Alternative: SEM software (Mplus), but cannot handle this problem
- Informative priors would not change the outcome of the study
- A simulation study showed good frequentist performance of our approach

Conclusion + discussion

- MCR and MHOF inspired by clinical questions
- Alternative: SEM software (Mplus), but cannot handle this problem
- Informative priors would not change the outcome of the study
- A simulation study showed good frequentist performance of our approach

- h-likelihood approach is interesting alternative + yielded same results as Bayesian approach, in a **much shorter computation time**:
  - Bayesian approach (rjags, dclone): 15 hours $\Leftrightarrow$ h-likelihood approach: 3 hours
  - MCMC software is quite flexible and models can relatively easy be extended
  - h-likelihood approach: software package needed to be extended, but it is a serious competitor to INLA

That's it!

FINALLY