Development of a response shift algorithm detection in item response theory (ROSALI) in longitudinal patient-reported outcomes studies

---

Evaluation et analyse de la qualité de vie : ***nouveaux développements méthodologiques***
Cancéropôle Grand Sud-Ouest, 3-4 avril 2014

Myriam Blanchin, Alice Guilleux, Antoine Vanier, Jean-Benoit Hardouin, Véronique Sébille

EA 4275 "Biostatistics, Pharmacoepidemiology and Subjective Measures in Health Sciences"
University of Nantes
http://www.sphere-nantes.fr

INCA_6931

Longitudinal data and PRO

- Longitudinal data: evolution of patient-reported outcomes (PRO) over time

- Assumption: No change of patients' perceptions of the measured concept between the different times of measurement

  ⇒ Response shift

- Observed evolution= true change + response shift
- Measurement bias or measure of adaptation

Response shift (RS)

- Change in the meanings of one's self evaluation of a target construct as a result of (Sprangers and Schwartz, 1999):
  - **recalibration (RC):** a change in the respondent's internal standards of measurement
    - Uniform (RCu) or non-uniform (RCnu)
  - **reprioritization (RP):** a change in the respondent's values (i.e the importance of component domains constituting the target construct)
  - **reconceptualization:** a redefinition of the target construct

- Detecting and taking account of RS
  - Then-test (RC)
  - Relative importance measure (RP)
  - Random forest (RP)
  - Structural Equation Modelling (SEM) (All types of RS): Oort's procedure, Schmitt technique
  - Item Response Theory (IRT)?

PROGRESS project (INCA_6931)

- SEM: ignorable MD assumed, dimension level
- Item response theory: management of missing data even non-ignorable MD, item level

1. Assessment of RS with Oort's procedure at item level
2. Development of a response shift detection algorithm in item response theory (ROSALI)
   - Following Oort's procedure
   - At item level
3. Comparison of item-level Oort's procedure and ROSALI
   - Correct detection of response shift?
   - Correct management of missing data?

Oort's procedure (Oort, 2005)

- Step 1: Establishing a measurement model
  - Model 1: Measurement model (no invariance constraints)
- Step 2: Overall test of no response shift
  - Model 2: No response shift model (all invariance constraints)
  - LR test: Model 2 vs Model 1
- Step 3: Detection of response shifts (step by step removal of untenable invariance constraints)
  - Model 3: All apparent response shifts are accounted for
- Step 4: Assessment of true change
  - χ^2 test: statistical significance of true change
  - Optional: test for other types of invariance (common factor variances and correlations) -> Model 4

Best strategy to detect response shift using item-level Oort's procedure?

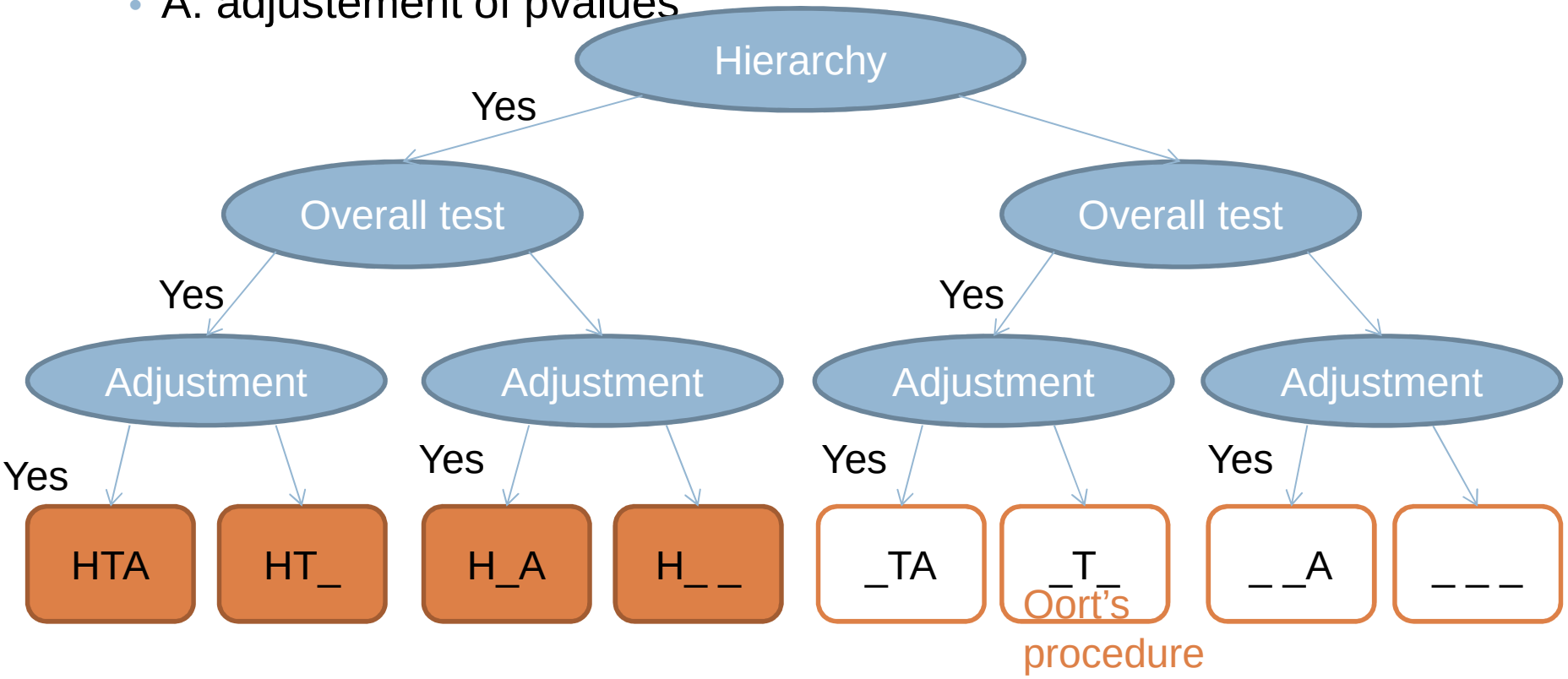
- Step 2: Overall test of response shift
  - Test + skip step 3 if not significant
  - No test + always step 3

- Step 3: Detection of response shifts (step by step removal of untenable invariance constraints)
  - ∀ type of response shift (no hierarchy)
  - Hierarchy of measurement invariance (Nolte et al, 2009, Ahmed et al 2009)
    - Non-uniform recalibration (RCnu) / Uniform recalibration (RCu) / Reprioritization (RP)

  - Adjustment of pvalues for multiple comparisons?

Simulation study

- Comparison of 8 strategies (item-level Oort's procedure)
  - H: hierarchy of measurement invariance
  - T: overall test of no response-shift
  - A: adjustement of pvalues



Simulation study

- Longitudinal Partial-Credit model
- True change: γ=difference of the means of the latent variable at each time point: -0.2
- Covariance matrix of the latent variable: $\Sigma = \begin{pmatrix} 1 & 0.6 \\ 0.6 & 1 \end{pmatrix}$
- Sample size: N:100/200/300

- 4 or 7 polytomous items
  - 4 or 7 answer categories
- Response shift occurs on 0,1,2 or 3 items
  - If RS: RCNu or RCu or RP

102 combinations*500 replications=51000 datasets

- Items affectés par le response shift selon les cas :

- Recalibration uniforme et non-uniforme
- | | J=4 | J=7 |
- 1 item | 3 | 5
- 2 items | 3, 4 | 6,7
- 3 items | - | 4,6,7

- Décalage des difficultés d'items
- - pour la recalibration uniforme : -1
- - pour la non-uniforme si le nombre de modalités est de 7 [-0.5 -0.25 0 0 0.25 0.5]
- Si le nombre de modalités est de 4 [-0.5 0 0.5]

- Repriorisation
- Avec αj= 1 au temps 1 et αj= 1.5 au temps 2
- | | J=4 | J=7 |
- 1 item | 4 | 2
- 2 items | 2,4 | 3,6
- 3 items | - | 2,5,7

Results

- Erroneous detection of RS if no RS simulated

Strategy	Min	Max
H_ _	33.1%	50.7%
H_A	6.6%	29.6%
HT_	3.8%	18.9%
HTA	2.3%	14.6%

- Erroneous detection increases with:
  - Sample size
  - Number of answer categories

- Erroneous detection decreases when the number of items increases

Results

- Detection of the correct type of response shift on the correct item (potential other types of RS detected on all items)

Response shift on	1 item		2 items		3 items	
Strategy	Min	Max	Min	Max	Min	Max
H_ _	8.3%	22.6%	0.7%	14.5%	0.1%	3.9%
H_A	3.0%	16.8%	0.3%	8.9%	0%	0.9%
HT_	3.0%	14.4%	0.5%	13.3%	0.1%	3.8%
HTA	1.3%	10.8%	0.2%	8.3%	0%	0.9%

- RS detection increases with:
  - Sample size
  - Number of answer categories

Results

- Detection of the correct type of response shift on the correct item (no other type of RS detected)

Response shift on	1 item		2 items		3 items	
Strategy	Min	Max	Min	Max	Min	Max
H_ _	2.5%	8.5%	0.1%	3.1%	0%	0.2%
H_A	1.8%	7.9%	0.3%	1.4%	0%	0%
HT_	0%	1.7%	0%	2.0%	0%	0.1%
HTA	0.3%	2.2%	0.1%	1.6%	0%	0%

- RS detection increases with:
  - Sample size
  - Number of answer categories

Discussion

- High rate of erroneous detection
  - Highest for strategies without overall test of no RS
- Low rate of correct detection
  - Lowest for strategies with overall test of no RS

- Thing in common of the 4 strategies: hierarchy of the measurement invariance

- Bad performances due to
  - hierarchy? 4 other strategies
  - to item-level? other maximization method
  - to the size of the RS simulated?

- Strategies to avoid for RS detection at item-level
- First step of the PROGRESS project