

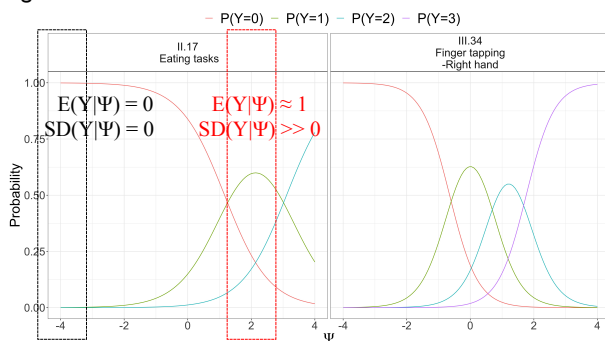
Introduction

Composite scales consist of many questions/items, that are summarised to a total score (TS) through an algorithm, where the TS is discrete and bounded. One example is the MDS-UPDRS scale for Parkinson's disease [1].

Adequately designed item response theory (IRT) models are the best way to handle composite scale data. These include item characteristic curves for each item and handle correlation between items through a latent variable (Ψ). However, IRT models may be cumbersome to develop, include many parameters and take long time to estimate. Also, they require that data on the item level are available. As an alternative, we propose TS-analyses, e.g. continuous variable (CV) and bounded integer (BI) [2] models which do not make use of item-level data, informed by an existing IRT model [3]. We investigated IRT-informed CV and BI models in terms of fit to both simulated and real data [4].

Methods

Figure 1. Item characteristic curves from IRT model

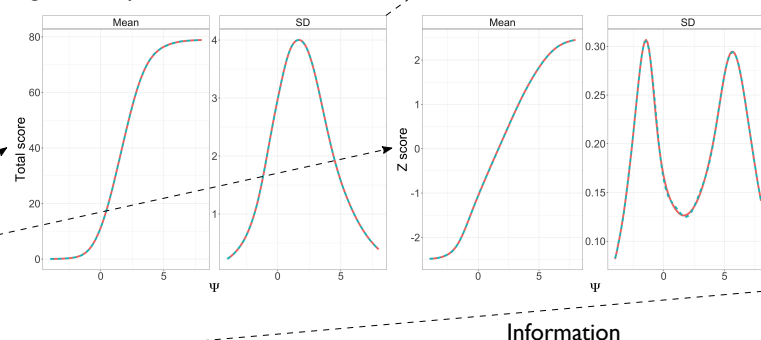


Two example items from the MDS-UPDRS motor subscale, which contains 30+ items in total. Each item is modelled with a separate function in the IRT model [3], but all are related to Ψ . Correlation between items can be handled via Ψ . The expected mean for the TS or Z score at any point of Ψ can be derived, which is called $E(Y|\Psi)$, and also the SD of TS or Z, called $SD(Y|\Psi)$. At very low values for Ψ , e.g. as the here indicated -4, the expected TS is 0 and the SD also approaches 0. As Ψ increases, so does the expected TS, as indicated when $\Psi=2$, the TS for this item is roughly 1. The SD shows the sensitivity to small changes in Ψ on the TS/Z score, which for $\Psi=2$ is nonzero but will tend to 0 again as Ψ increases to 4 in this case.

Models

The models investigated were 1) Standard CV model with homoscedastic error (S-CV), 2) Standard BI (S-BI) model, which respects the data nature and boundaries of the scale, 3) IRT-informed CV (I-CV) and 4) IRT-informed BI (I-BI) models. These are all models for TS-analyses, but the BI models are discrete data models with a latent variable on the Z score scale. The IRT-informed models included polynomial link functions for the mean and SD. Additionally, IRT-informed models with scaling parameters and also inter-individual variability (IIV) were estimated on top of the expected SD relations, to explain potential deviations from the expectations. Models were compared in terms of fit via Akaike information criterion (AIC).

Figure 2. Expected mean and standard deviation on TS and Z score scale



The relation between Ψ and TS was sigmoidal, but more linear between Ψ and Z score. The SD was unimodal for TS but bimodal for Z score, which is the latent variable in the BI model. Empirical Chebyshev polynomials (dashed blue lines) were used in the TS models, which approximated the exact relations (solid red lines). These polynomials are computed automatically in the R package piraid [5], which can output complete control streams for IRT-informed analyses.

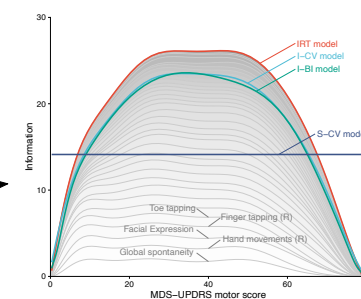
Results

Table 1. Fit to real data

Model	Disease progression	Standard deviation	θ	IIV (%CV)	Δ AIC	No. of estimated parameters	AIC
S-CV	Linear on TS	Homoscedastic (estimated θ)	5.3	-	-	10	18794 ^{Base}
		Heteroscedastic (fixed $SD(Y \Psi)$)	-	-	+259	9	19053
I-CV	Linear on Ψ	Heteroscedastic ($SD(Y \Psi) \cdot \theta$)	1.4	-	-206	10	18588
		Heteroscedastic ($SD(Y \Psi) \cdot \theta \cdot e^{\eta}$)	1.4	24	-356	14	18439 ^{Final}
S-BI	Linear on Z	Homoscedastic (estimated θ)	0.22	-	-	10	18706 ^{Base}
		Heteroscedastic (fixed $SD(Y \Psi)$)	-	-	+388	9	19094
I-BI	Linear on Ψ	Heteroscedastic ($SD(Y \Psi) \cdot \theta$)	1.4	-	-97	10	18610
		Heteroscedastic ($SD(Y \Psi) \cdot \theta \cdot e^{\eta}$)	1.3	38	-296	14	18411 ^{Final}

The IRT-informed models (I-CV, I-BI) had the best fit to the real data, although they required both a scaling factor and IIV on top of the expected relations. This is likely because real data has other sources of variability, such as Markovian features or correlations between items that can lead to both higher or lower than expected variability. For the simulated data (not shown), the IRT-informed functions alone were sufficient to provide the best fit.

Figure 3. Comparison of information content of different models



Through the polynomial link functions, also TS models can estimate the latent variable Ψ . The relative information content at a certain TS, calculated via Fisher information, can then be compared between models. The IRT model, which has item-level data, has the most information, while IRT-informed TS models can capture most of the same information. A homoscedastic CV model will overestimate the information at the boundaries but have overall less information.

Conclusions

Disease progression and residual unexplained variability in a total score analysis can be improved by using IRT-informed modelling. The approach requires no additional parameters to be added. Also, different model types can be directly compared.