Application of Item Response Theory to EDSS Modelling in Multiple Sclerosis

UPPSALA UNIVERSITET

Ana Kalezic[1], Radojka Savic[2], Alain Munafo[3], Sebastian Ueckert[1], Mats O. Karlsson[1]

[1] Dept of Pharmaceutical Biosciences, Uppsala University, Uppsala, Sweden [2] Dept of Bioengineering and Therapeutic Sciences, University of California San Francisco, San Francisco, USA [3] Modelling and Simulation, Merck Serono S.A., Geneva, Switzerland

Objectives

- To apply Item response theory (IRT) methodology to Expanded Disability Status Scale (EDSS) [1]
- To explore potential benefits of this novel methodology for characterizing disability caused by multiple sclerosis (MS) [2]

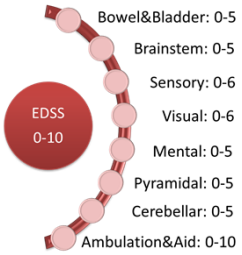
Background

- Despite numerous shortcomings of the EDSS to characterize MS patients and their disease progression, it is still considered a gold standard in assessing MS disability
- Time course of MS progression is slow and often unpredictable which makes it difficult to characterize and to detect a potential drug effect
- Traditional approaches to measurement scales generally disregard the underlying nature of the subcomponent data and usually model only summary score
- In contrast, IRT refers to a set of mathematical models that describe, in probabilistic terms, the relationship between a person's response to a survey question and its level of the "latent variable" being measured by the scale [3].

Methods

Patients and Data

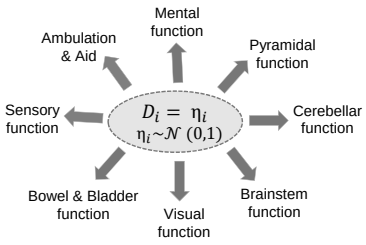
- EDSS is a 20 point scale (0=normal neurological exam and 10=death due to MS), based on the neurological assessment of 7 different functional systems and of walking ability [4].
- Unlike most measurement scales, EDSS total score does not result from simple addition of individual items, but instead results from individual components via a decision tree
- Data were collected from a 96-week Phase III clinical study, involving 117270 observations from 1319 patients with relapsing-remitting MS, treated with placebo or cladribine.



- EDSS score was assessed before the start of the treatment (at baseline) and every 12 weeks during the duration of the study.

Model building

- IRT model
- The main assumption is that the outcome of each item constituting EDSS depends on an unobserved variable "disability": D
- Two parameter logistic model was applied: each item was characterized by a location (b) and a slope (a) parameter
- Sets of parameters characterizing each item were modeled as fixed effects, while the MS disability was modeled as subject-specific random effect



- For each EDSS item, a model was developed in accordance with the nature of data, describing the probability of a given score as a function of disability variable (Eq. 1,2).

$$\mathbf{P}(Y_{ij} \geq k) = \frac{e^{a_j(D_i - b_j)}}{1 + e^{a_j(D_i - b_j)}} \quad (Eq.1)$$

$$P(Y_{ij} = k) = P(Y_{ij} \geq k) - P(Y_{ij} \geq k+1) \quad (Eq.2)$$

- Disease progression model
- Longitudinal data from placebo arms were used to describe the disease progression over time (Eq. 3)

$$Dis(t) = Dis_0 + SL * Time^{cst} \quad (Eq.3)$$

- Drug effect model
- IRT model was then extended to cladribine arms in order to charcterize the drug effect (Eq.4-6); and the symptomatic effect was considered exposure dependent

$$Dis(t) = Dis_0 + SL * Time^{cst} * (1 - EffDm_1) - EffDm_2 \quad (Eq.4)$$

$$EffS = \frac{Emax * Exps_i}{Exps_{50} + Exps_i} \quad (Eq.5) \qquad Exps_i = \frac{CumDose * Cl_{cr\,median}}{Cl_{cr}} \quad (Eq.6)$$

Dis_0: Baseline disability, SL: Slope, $Exps_i$: Exposure, EffDm: Disease modifying drug effect, Emax: Maximal obtainable effect, $Exps_{50}$: Exposure needed for half maximal effect based on cumulative dose administered, CumDose: Cumulative Dose administered, Cl_{cr}: Creatinine clairance, $Cl_{cr\,median}$: Median Cl_{cr} of the observed population sample

Item sensitivity

- Sensitivity with respect to patient disability was calculated as Fisher information for each item constituting the EDSS
- Items were ranked based on their expected Fisher information in the studied population

References

[1] Ueckert S . PAGE 2012 (Abstract 2318)

[2] Ueckert S. ACOP 2013 (Abstract T-019)

[3] Reeve BB, Fayers P.Oxford University Press, 2005. p. 55–73.

[4] Kurtzke JF, Neurology 33 (11):1444-52, 1983

Results

- The final model contained 8 ordered categorical submodels for a total of 56 parameters.
- All estimated item characteristic curves were well defined, with a person with higher disability having higher probability of high scores. (Figure 1)
- The slope estimates (a) ranged from 0.39 to 3 and location parameters (b) from -1.61 to 4.02, indicating considerable variation across items
- Proportions of simulated scores were in agreement with observations across items (Figure 2).

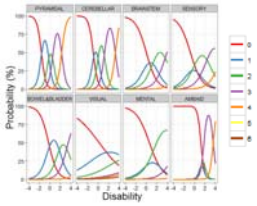


Figure 1. ***Item characteristic curve:*** *probability of each score as a function of disability.*

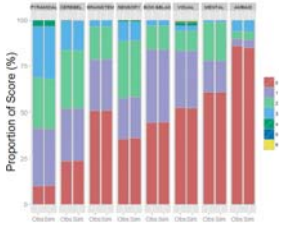


Figure 2. Comparison of *proportions of scores at baseline* *in the original data and simulations for each item .*

Longitudinal model

- Typical individual receiving placebo will progress 0.17 EDSS points over 2 years.
- Disease-modifying effects of cladribine, defined as change in the rate of increase of EDSS, and temporal offset of the increase in EDSS respectively, were estimated (Figure 3, Table 1).

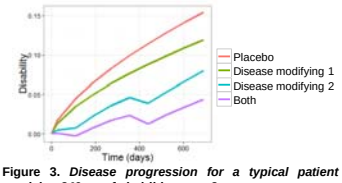


Figure 3. ***Disease progression for a typical patient receiving 240mg of cladribine over 2 years***

Table 1. Final parameter estimates

	Parameter estimates (RSE%)
Slope*	0.101
Cst*	0.669
Emax	0.272 (3.6%)
IIV Emax (CV%)	128.1% (2.8%)
Exp50	591.24 (1.9%)
Dis.modif. effect	0.229 (2.7%)
Corr Dis./SL*	0.147

*: Parameters estimated on placebo data only

- Simulations from the IRT model were in good agreement with the observed EDSS and item-level data (Figure 4-5).

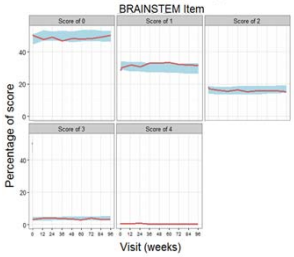


Figure 4. ***Visual predictive checks (VPCs) describing the time-courses of each score.*** *Median (solid line) of the observed data is compared to the 95% prediction interval (shaded areas) for the simulated data*

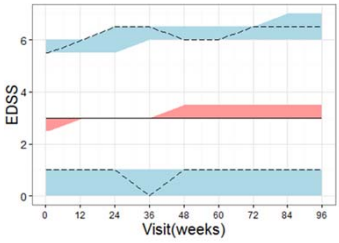


Figure 5. ***VPC for the placebo model describing the change in EDSS vs time.*** *Median (solid line), 2.75th and 97.25th percentiles (dashed lines) of the observed data are compared to the 95% confidence intervals (shaded areas) for the simulated data's 2.75th, 97.25th percentiles and median.*

Item sensitivity

- In the studied population, the cerebellar item was found to be the most sensitive one
- 80% of the information was contained in 4 out of 8 items (Table 2)

Table 2. Item information ranking

ITEM	Information	% Total Inf.
Cerebellar	2.18	33.98
Pyramidal	1.62	25.28
AmbAid	1.14	17.8
Bowel&Bladder	0.45	7
Brainstem	0.37	5.81
Cerebral	0.29	4.54
Sensory	0.28	4.39
Visual	0.08	1.21

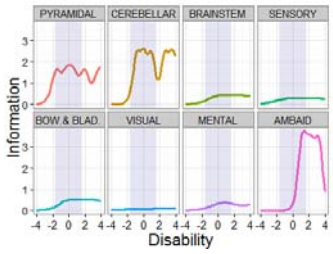


Figure 6. ***Sensitivity of EDSS items*** *and disability range for 95% of studied population*

Conclusions

- This is the first time that the IRT methodology has been applied to the MS area and to a score that is not a summation of items.
- Disease progression in MS patients was succesfully described in patients with and without tretament
- The IRT model allowed to identify considerable variation in item sensitivity, providing a basis for a potential simplification of the EDSS assessment

The research leading to these results has received support from the Innovative Medicines Initiative Joint Undertaking under grant agreement n° 115156, resources of which are composed of financial contributions from the European Union's Seventh Framework Programme (FP7/2007-2013) and EFPIA companies' in kind contribution. The DDMoRe project is also supported by financial contribution from Academic and SME partners. This work does not necessarily represent the view of all DDMoRe partners.

Following the feedback received from regulatory agencies, Merck Serono made the decision to withdraw the agent from the regulatory approval process.