Eleven ordered categories data: which modelling options?

Elodie L. Plan, Yang Sun and Mats O. Karlsson
Department of Pharmaceutical Biosciences, Faculty of Pharmacy, Uppsala University, Uppsala, Sweden.

Background

- Ordered categorical data:

If "low" # of categories ➔ ordered categorical model (OC),
if "high" # of categories ➔ continuous model (CO).

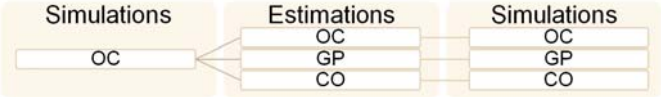
- 11 categories [from 0 (no pain) to 10 (worst pain)] data, like the Visual Analogue Scale (VAS) or the Likert Scale:

➔ Which model(s) should be used for the analysis?

➔ Is a generalized Poisson model (GP) also an option?

Objectives: - To evaluate models adapted for 11-point data
- To assess their capability of estimating a simulated drug effect

Methods



- **Step 1 = Original simulations**

OC:

231 individuals, ≅100 observations/ID (1/day over 18 weeks)

• Real placebo observations[1] analysed; obtained parameters used to perform 100 simulations of baseline data.

• Dose levels (0, 100, 200, 300) set up; drug effect (0.045 slope, 30% IIV) generated in 100 simulations of drug data.

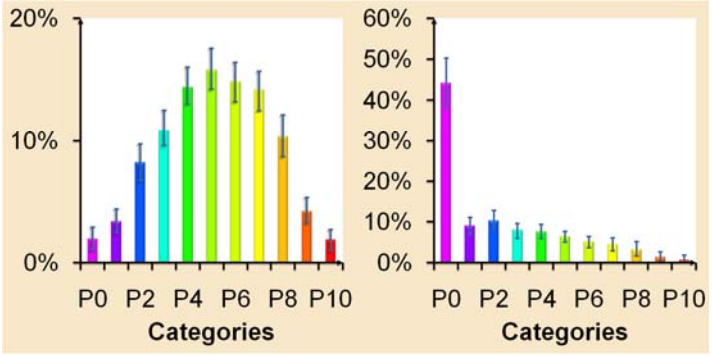


Figure 1: *Proportions of the 11 pain scores from the 100 simulations with the OC model (step 1) for baseline data (left) and drug data (right).*

- **Step 2 = Estimations**

OC:

• True proportional odds model: *10 θ, 1 η*

• Linear drug effect (in the logit): $Slope \times e^{\eta} \times Dose$

GP:

• 11-truncated generalized Poisson model: *1 λ, 1 δ, 2 η*

• Polynomial drug effect: $(Slope_1 \times Dose + Slope_2 \times Dose^2) \times e^{\eta}$

CO:

• 10-fold logit transformed continuous model: *1 λ, 1 η, 1 ε*

• Polynomial drug effect (in the logit)

- **Step 3 = Final simulations**

Each vector of estimated parameters used to generate a simulated dataset for evaluation purposes.

Results

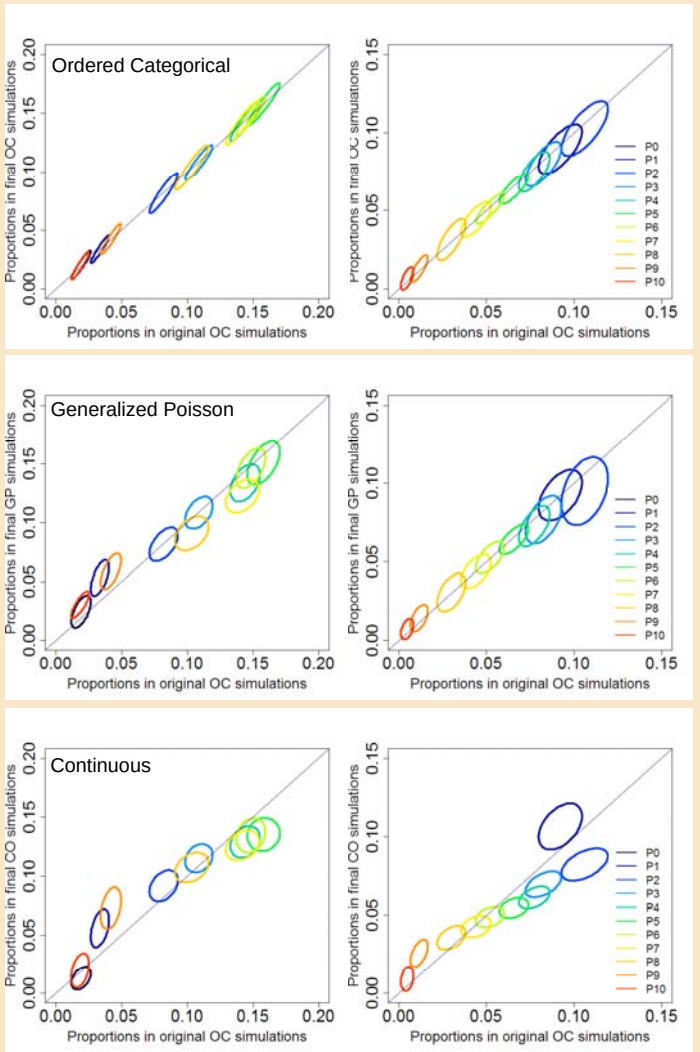


Figure 2: *Ellipses of the proportion (%) of each category simulated from final estimates (y axis) versus simulated with original values (x axis) for the 3 studied models OC, GP and CO represented for baseline data (left) and drug data (right).*

Table 1: *Average scores (±SD) from steps 1 and 3 for baseline data (left) and drug data (right).*

Placebo	OC	GP	CO	Drug	OC	GP	CO
5.2±2.3	5.2±2.3	5.1±2.4	5.2±2.2	2.2±2.6	2.2±2.6	2.2±2.7	2.3±2.6

• An ordered categorical model accurately analyses 11-point data, at least with rich datasets and no Markovian pattern,

• A truncated generalized Poisson model performs well,

• A logit-transformed continuous model presents less accurate results when a drug effect is included.

Conclusion

Although seldom used for Likert data, the OC model performs well. The other models may be alternatives for sparser data sets and in the presence of serial correlations.

Reference

[1] Plan *et al.* New models for handling correlated underdispersed Likert pain scores. *PAGE* 2009.

Contact: Elodie.Plan@farmbio.uu.se