Improving Equivalent Scores for Clinical Neuropsychology:
A new method for regression model selection

Giorgio Arcara

IRCCS San Camillo Hospital, Venice, Italy
University of Padua, Padua, Italy

May 3rd, 2024

Test scores in clinical neuropsychology

In a working memory test a patient scores 5.

Test scores in clinical neuropsychology

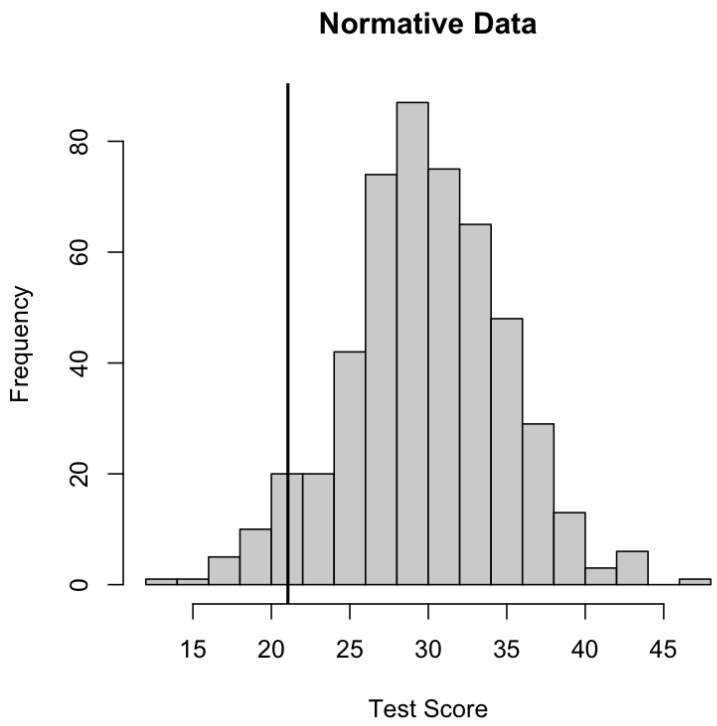
In a working memory test a patient scores 5.

In clinical neuropsychology, scores are almost always compared to *thresholds* to facilitate interpretation.

Typically, if the score is *below* the threhsolds (i.e. “cut-offs ”), it is intepreted as a probable impairment.

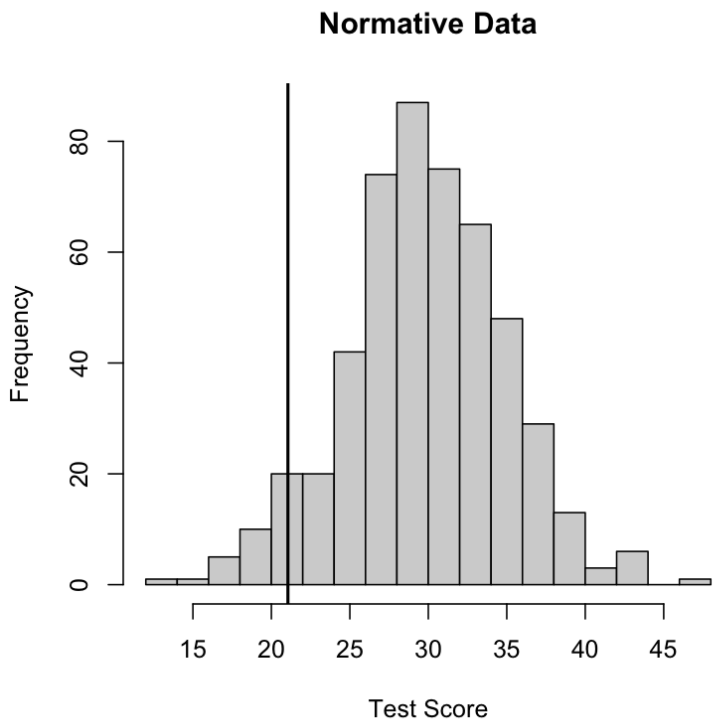
Demographic variables and normative data

Most methods for cognitive impairment are based on *normative data* that take into account age, education, and sex.



Demographic variables and normative data

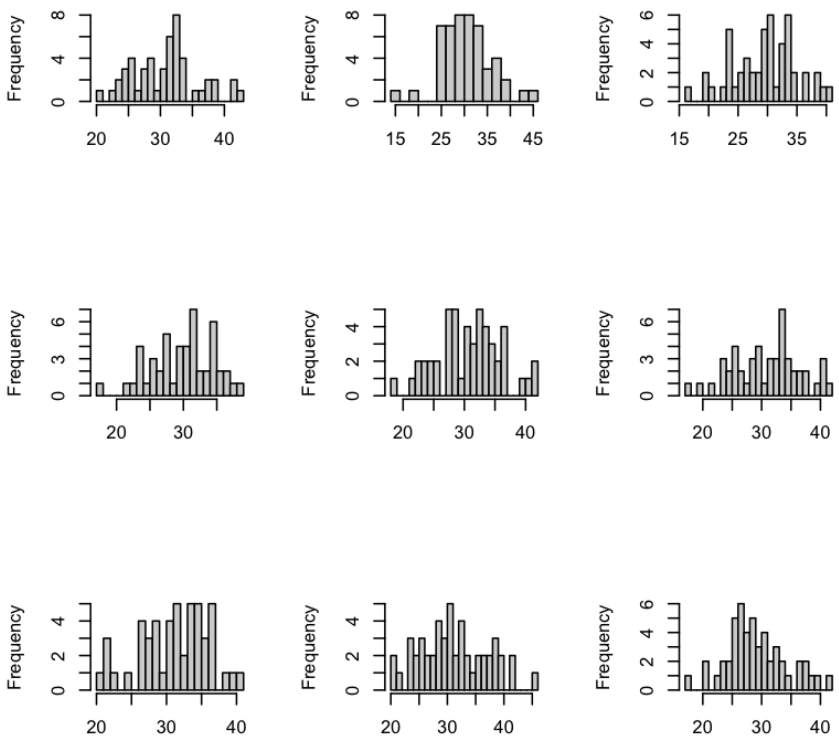
Most methods for cognitive impairment are based on *normative data* that take into account age, education, and sex.



Threshold typically delimits 5% of worse performances (at sample or population level)

Demographic variables and normative data

Demographic variables can be taken into acount by dividing into bins.



Equivalent Scores

Equivalent Scores is one of the most widespread method to calculate thresholds in clinical neuropsychology in Italy, originally developed by E. Capitani (Spinnler & Tognoni, 1987).

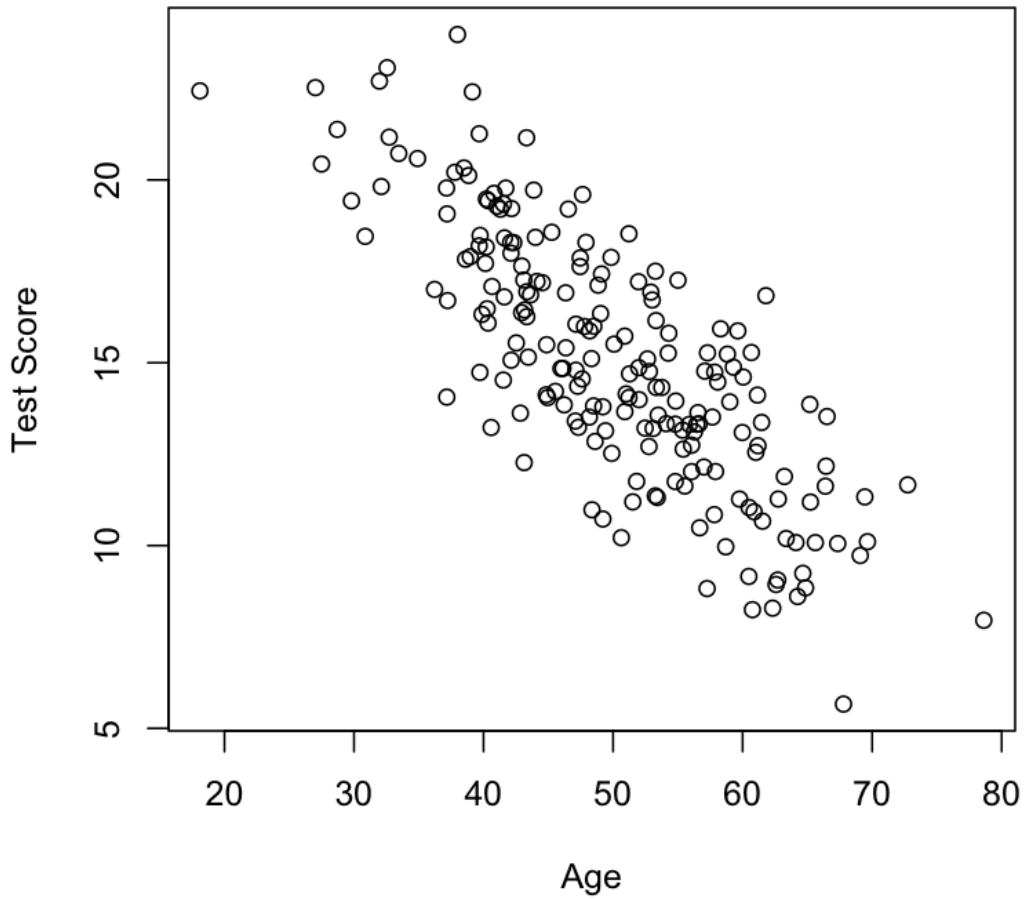
Equivalent Scores

Equivalent Scores is one of the most widespread method to calculate thresholds in clinical neuropsychology in Italy, originally developed by E. Capitani (Spinnler & Tognoni, 1987).

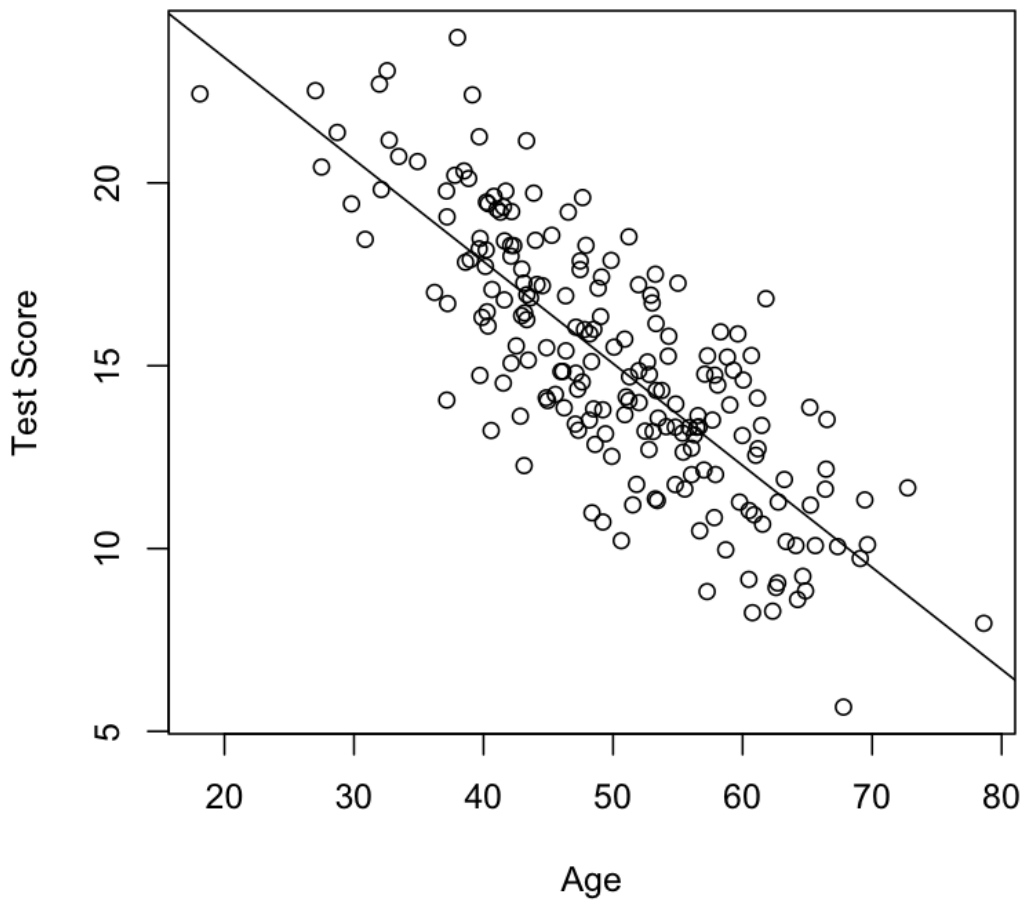
the method consists of:

- a first part using regression to calculate adjusted scores that consider age, education, and sex.
- a second part to derive thresholds for clinical inference (based on tolerance limits).

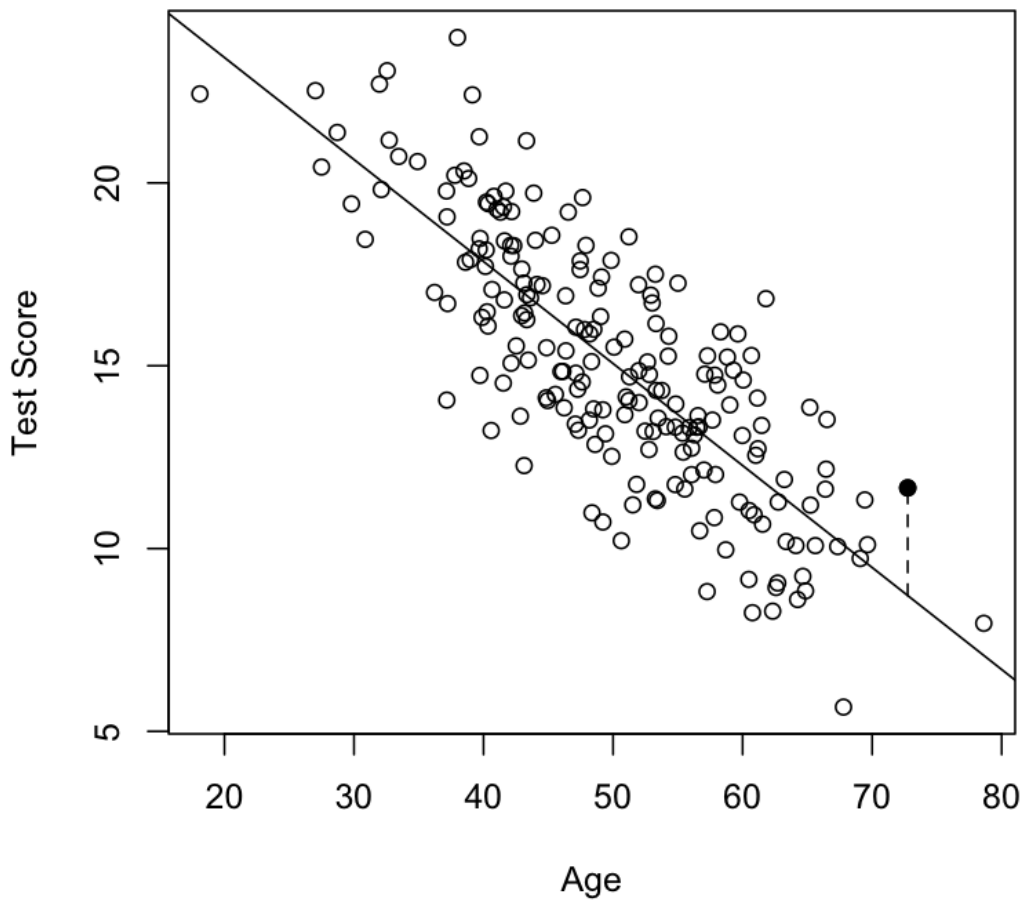
Regression Method



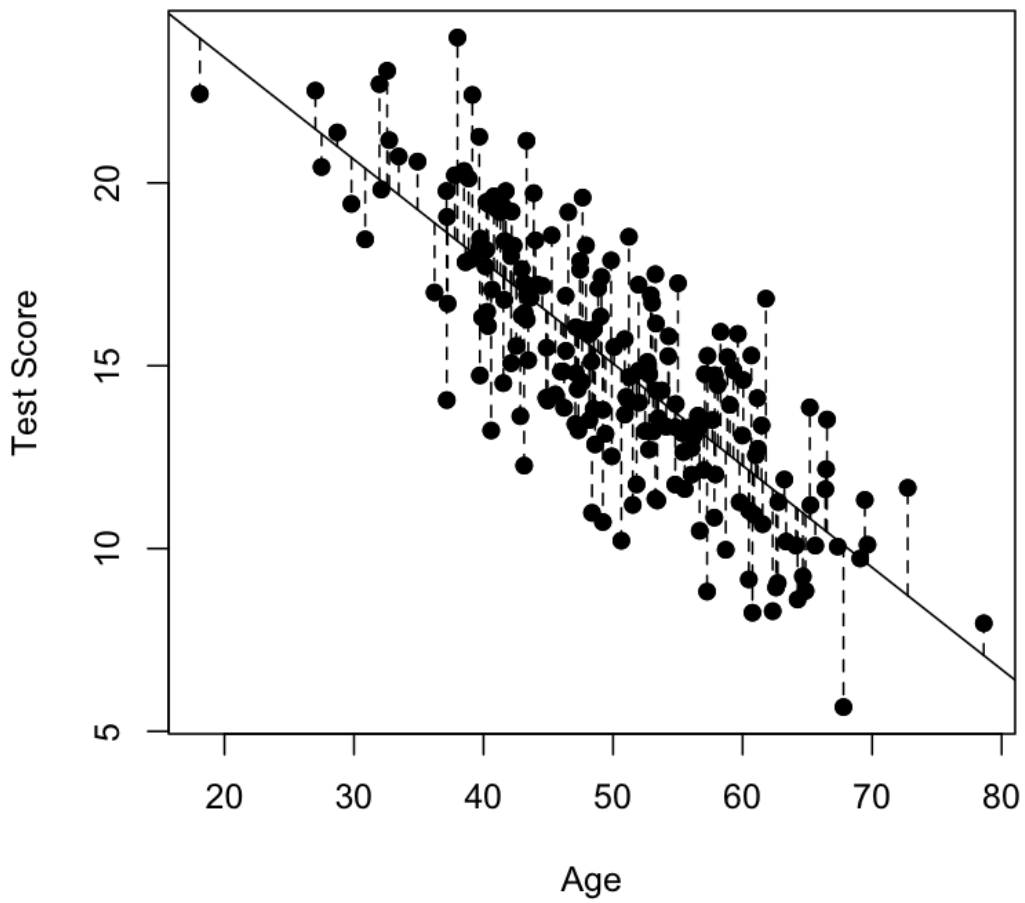
Regression Method



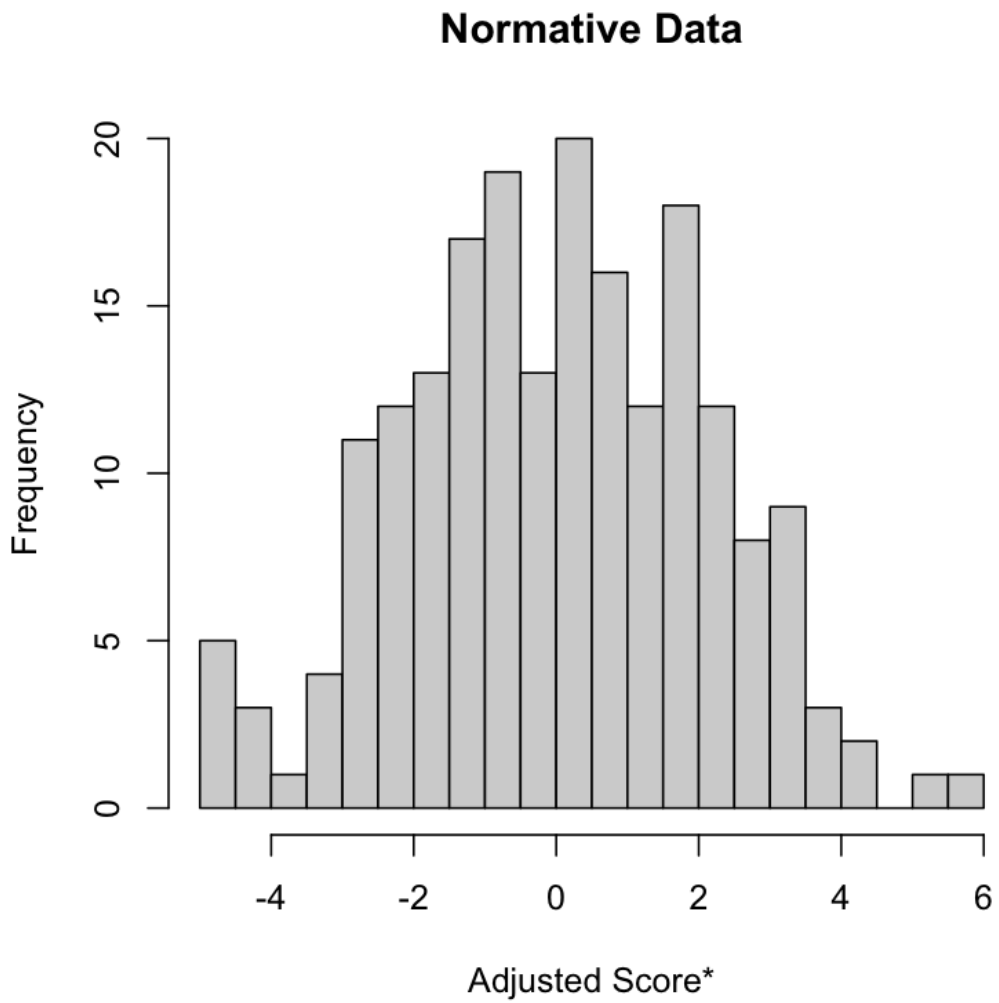
Regression Method



Regression Method

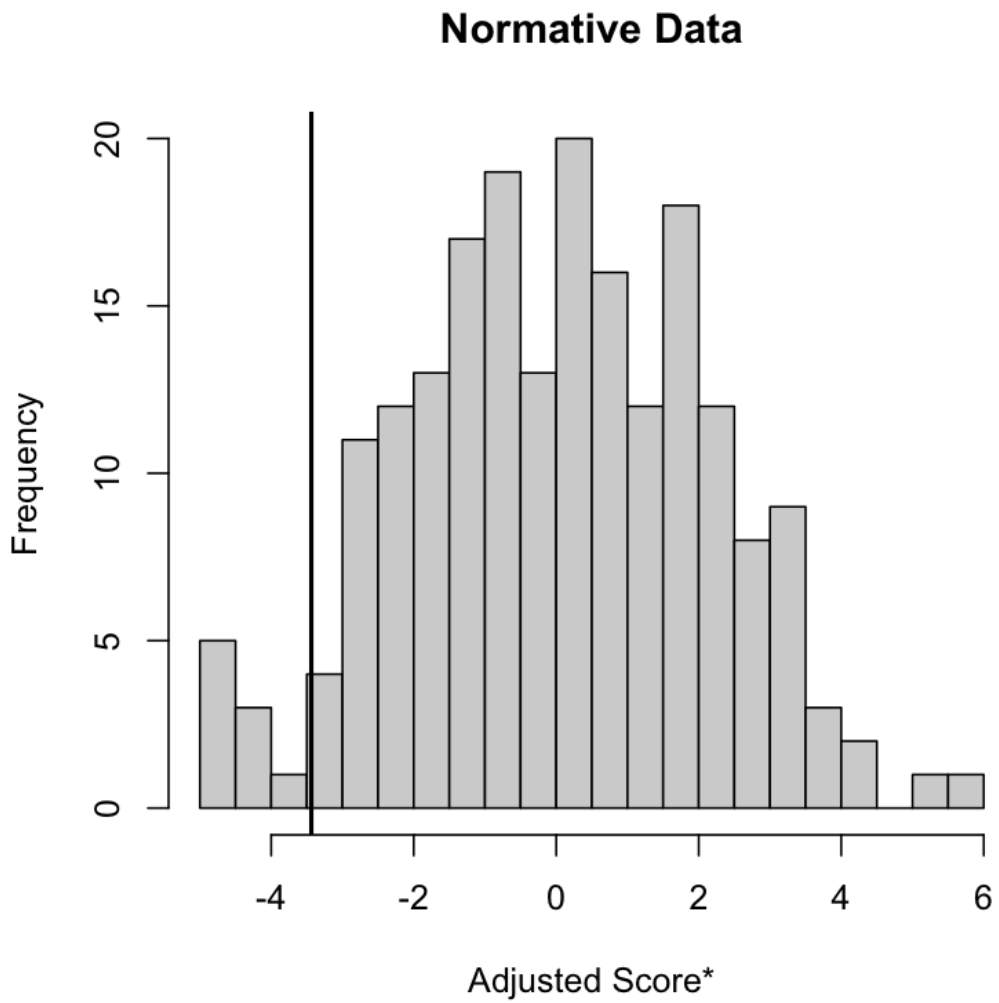


Regression Method



* Intercept is added to go back to original scale

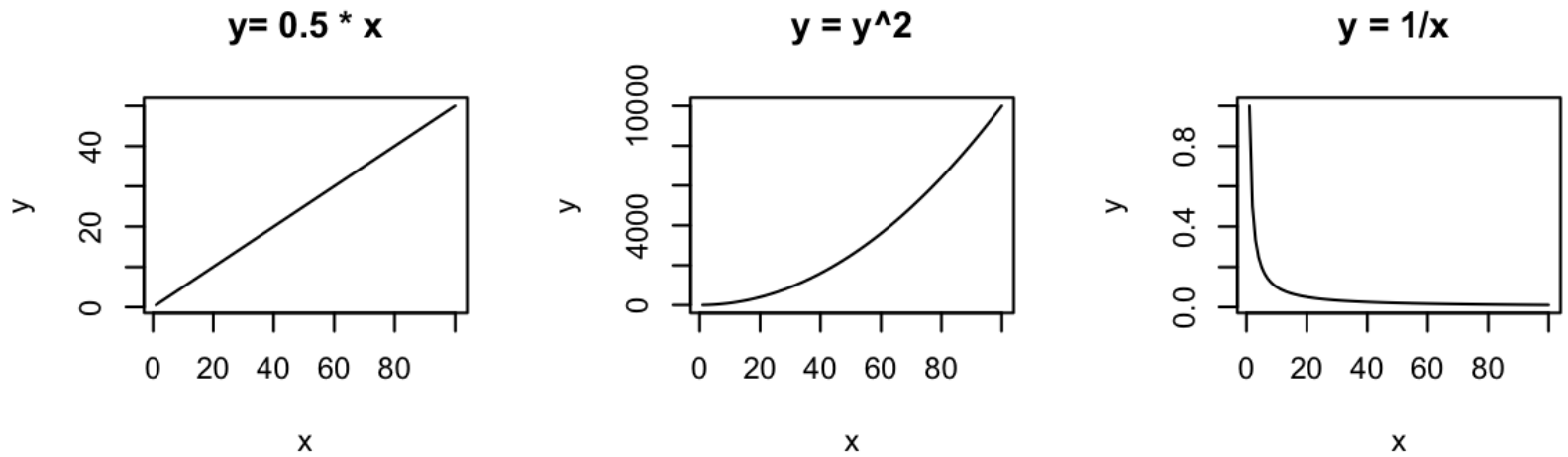
Regression Method



* Intercept is added to go back to original scale

Regression Modeling: two issues

1. what predictors should be included in the model? (Age, Education, Sex)
2. what is the better transformation of predictors? (log, quadratic, inverse etc.)



Equivalent Scores - regression method *(Capitani1987)*

Capitani 1987 (Spinnler & Tognoni 1987).

1. fit separate simple regression models for *Age* and *Education* trying different transformations from a list (e.g. log, inverse, quadratic, etc.)
2. fit a multiple regression with best transformations for Age and Education identified in 1. (best = higher R^2), and Sex.
3. drop terms with $p > 0.05$, Bonferroni corrected (i.e., $p < 0.017$)

An alternative method *(Arcara2024)*

Arcara 2024 (*in prep.*)

1. fit all possible multiple regressions including *Age*, *Education* and *Sex* with all possible transformations for *Age* and *Education* (from a list). Calculate **AIC** for all models.
2. from each starting model in 1. , use a *stepwise* backward selection (`step` function in `R`) to find the best model (lowest AIC) and select among all reduced models, select the model with the lowest AIC.
3. drop iteratevely terms with $p > 0.05$, Bonferroni corrected (i.e., $p < 0.05/number\ of\ terms$)

Testing the methods

Methods can be compared by using the generative model assumed for ES and simulating normative data.

- simulate data with known transformations of Age and Education and calculate *Ground Truth* (GT) cut-offs.

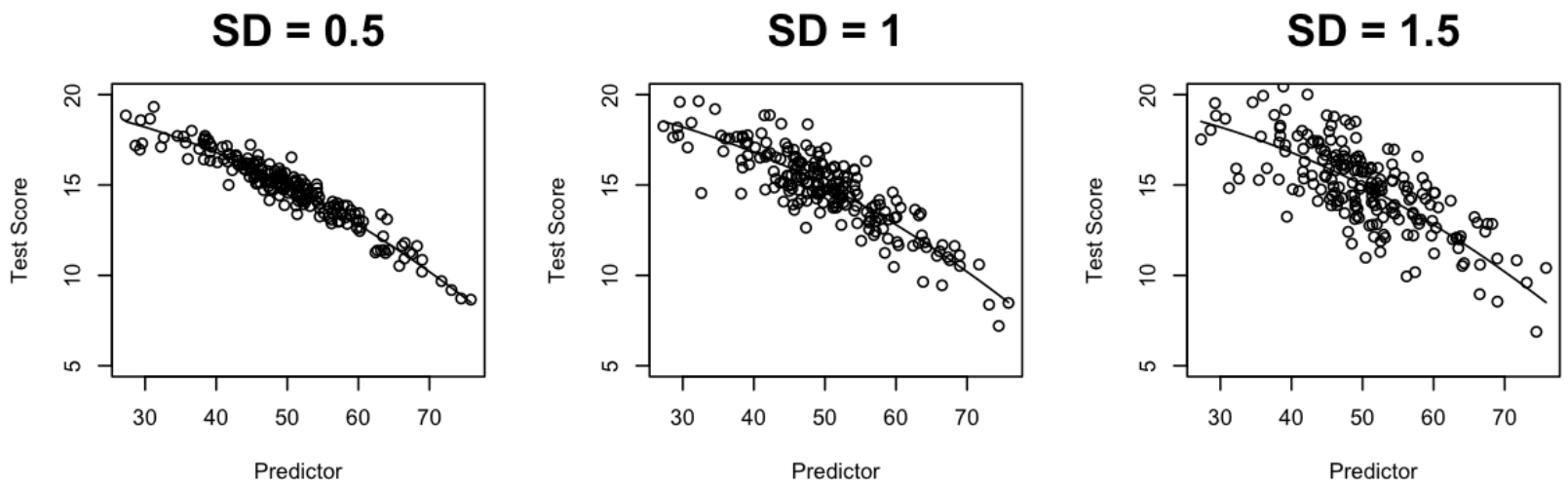
Parameters (1): Normative data Sample Size

Normative Data Sample size

$$\begin{bmatrix} 100 & 300 & 500 & 700 \end{bmatrix}$$

Parameters (2): True Adjusted Score variability

True Adjusted score Standard Deviation
$\begin{bmatrix} 0.5 & 1 & 1.5 \end{bmatrix}$



Simulation

1000 simulation for each condition of the three parameters.

- *Normative data* (Train Set)
- *New participants* (i.e. Test Set of 1000 new participants for each Normative Data sample)

Simulation

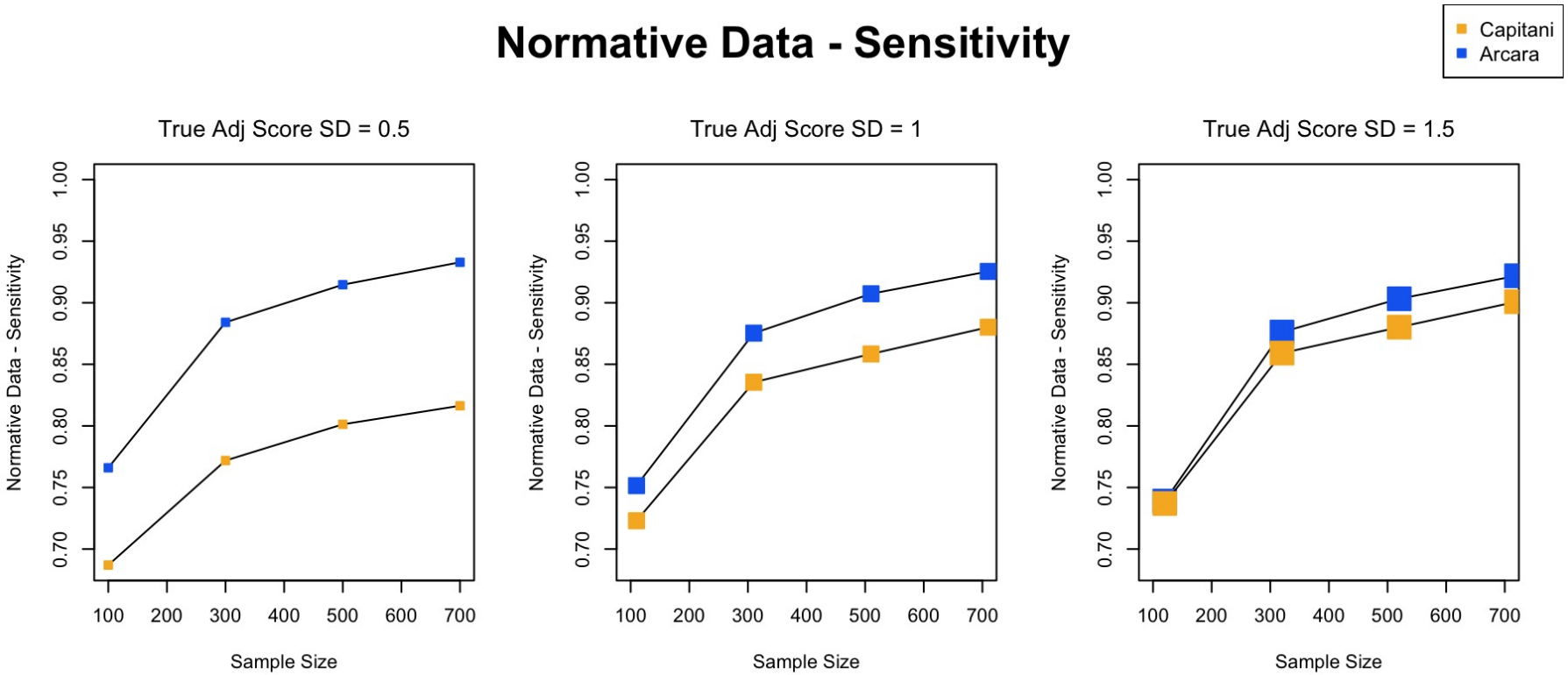
1000 simulation for each condition of the three parameters.

- *Normative data* (Train Set)
- *New participants* (i.e. Test Set of 1000 new participants for each Normative Data sample)

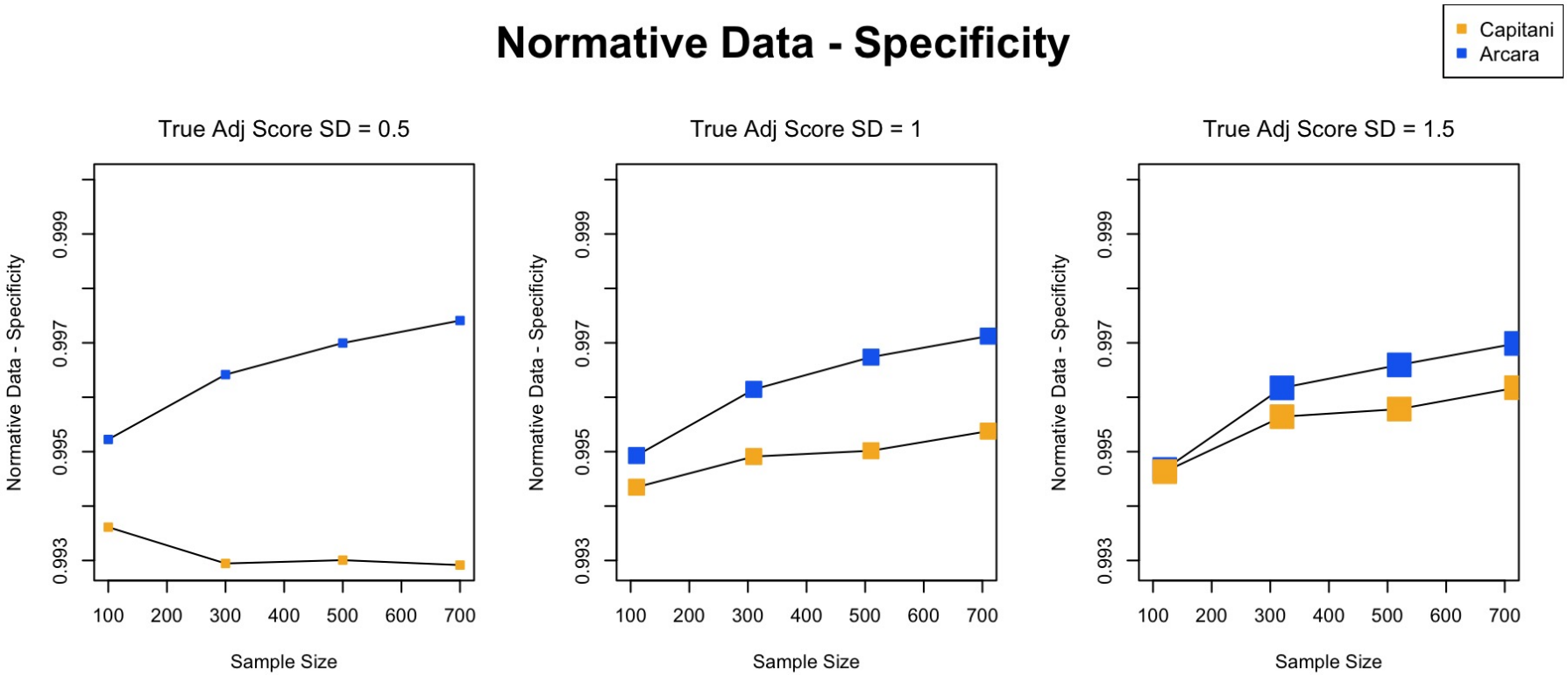
For each simulation performance is tested in classifying correctly below/above cut-off considering GT classifications (Sensitivity, Specificity, and others)

NOTE: as data are simulated with healthy participants, about 0.05 (at population level) are expected to fall below cut-off.

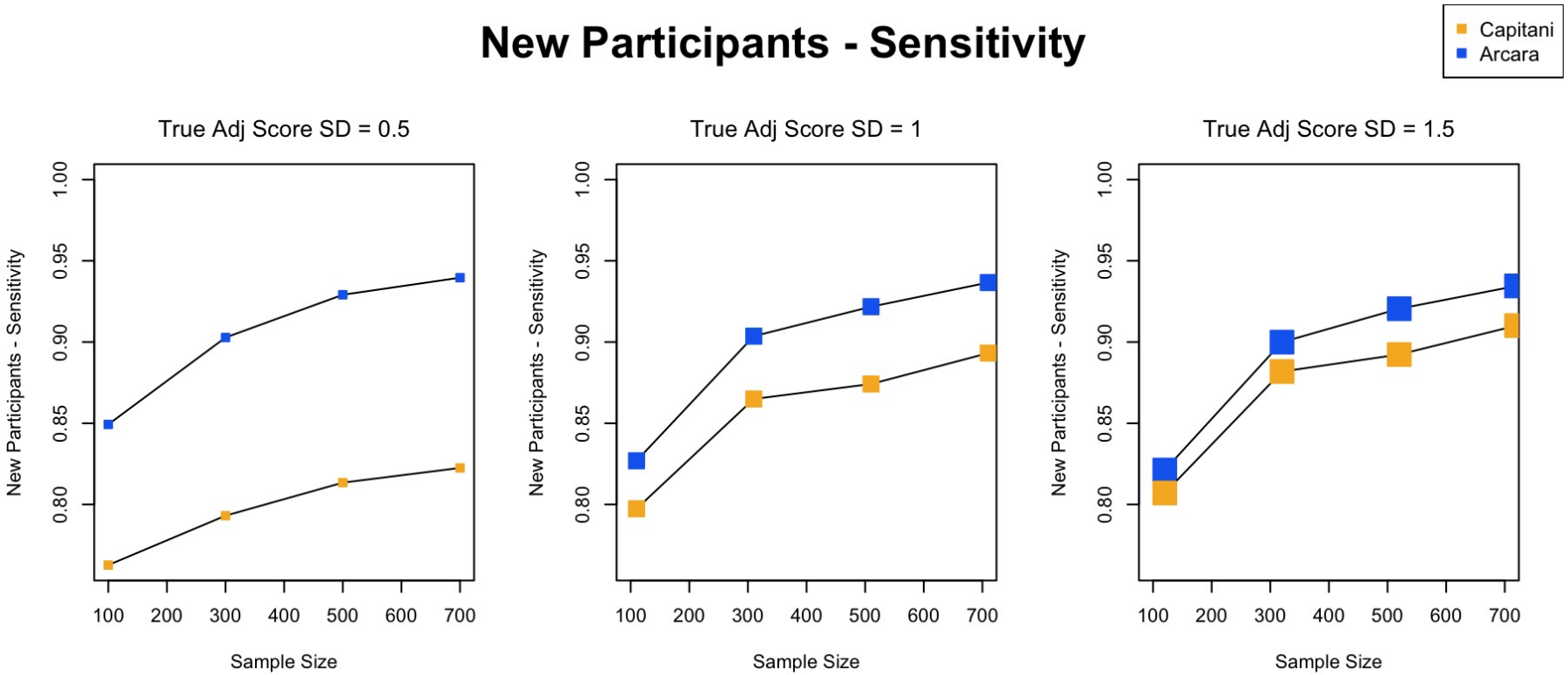
Sensitivity - Normative Data



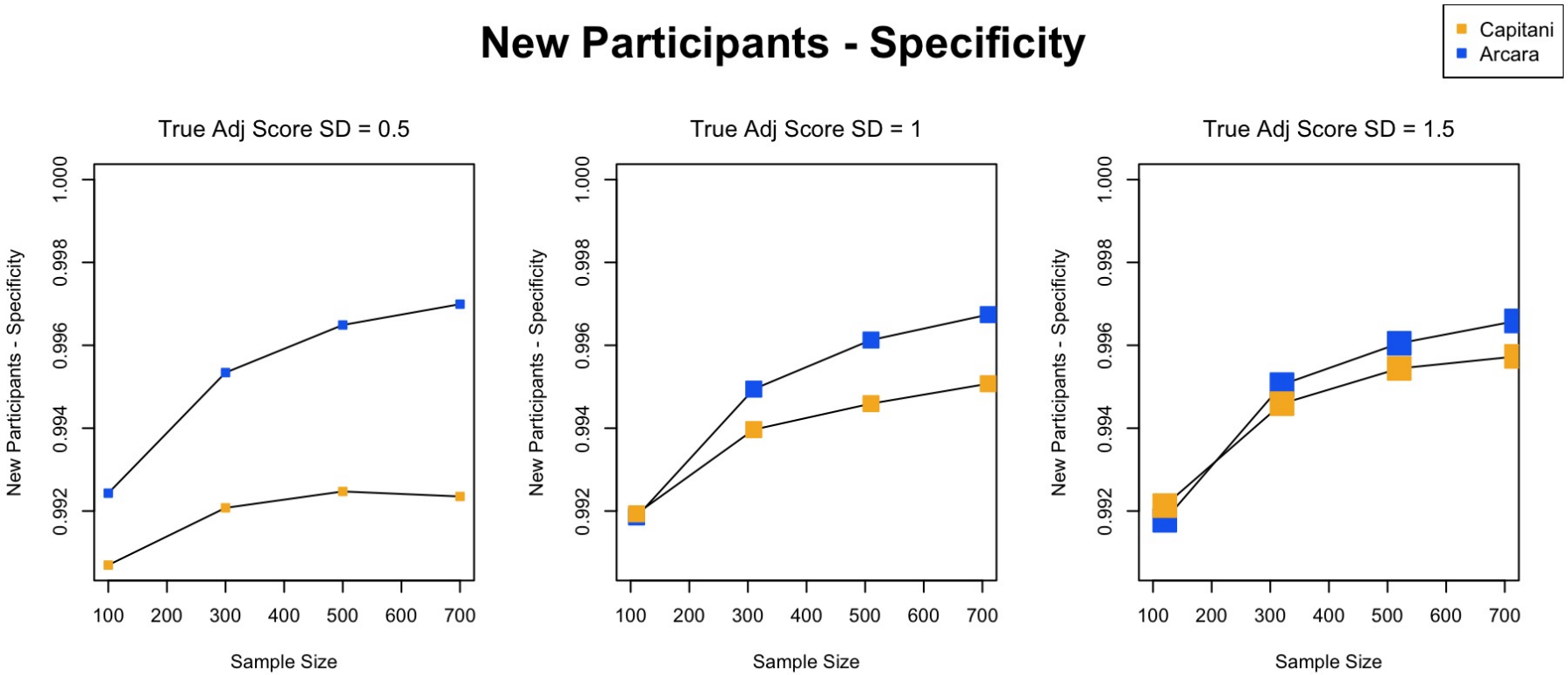
Specificity - Normative Data



Sensitivity - New Participants



Specificity - New Participants



Additional checks

(not shown, but trust me)

- better performance in guessing correctly the Ground Truth transformation.
- better performance with different conditional distributions of age and education.
- better performance with different proportion of participants below/above cut-off.
- the better method includes all three steps.

Discussion

A simulation approach can help in develop better methods for clinical neuropsychology.

The new proposed method outperforms the existing method across a wide range of parameters and conditions

The Equivalent Score method (as many other methods in Clinical Neuropsychology) can be improved.

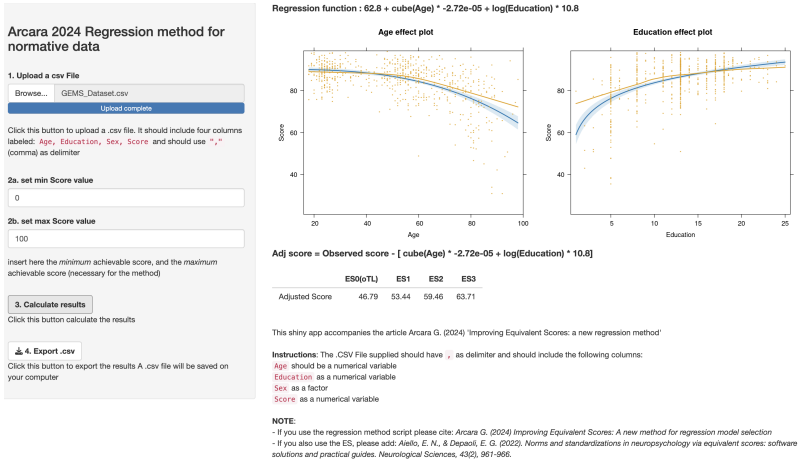
Future Directions

Many aspects left:

- what about reliability of scores? (with Andrea Spoto)
- what about heteroschedasticity and other realistic scenarios?
- rising awareness into issues and limitations of cut-offs.

Resources

- All code and scripts are available at `https://github.com/giorgioarcara/Reg-Method-Norm-Data-Sim`.
- All supplementary materials available at `https://osf.io/yma69/`.
- Manuscript soon on Psyarxiv.
- ShinyApp for easy application of the method (with cautionary notes)



Thanks!

Any suggestion (on additional analayses or other) is more than welcome.

giorgio.arcara@gmail.com
giorgio.arcara@hsancamillo.it
github.com/giorgio.arcara