Hypothesis tests for Goal Attainment Scaling

Urach S.[1], Gaasterland C.M.W.[2], Roes K.[3], Rosenkranz G.[1], Jilma B.[1], Van der Lee J.H.[2], Posch M.[1], Ristl R.[1]

[1]Medical University of Vienna, [2]University of Amsterdam, [3]University Medical Center Utrecht

Background

Goal Attainment Scaling (GAS)

- Aim: Comparison of the treatment effect between two groups in a randomized clinical trial on the basis of individual, patient-specific goals which are measured on a prespecified common ordinal scale
- Advantages for research in small populations:
  - utilization of patient-centered outcomes
  - possibility to combine the information from patients in heterogeneous populations
  - inclusion criteria can be relaxed
  - larger sample sizes can be realized

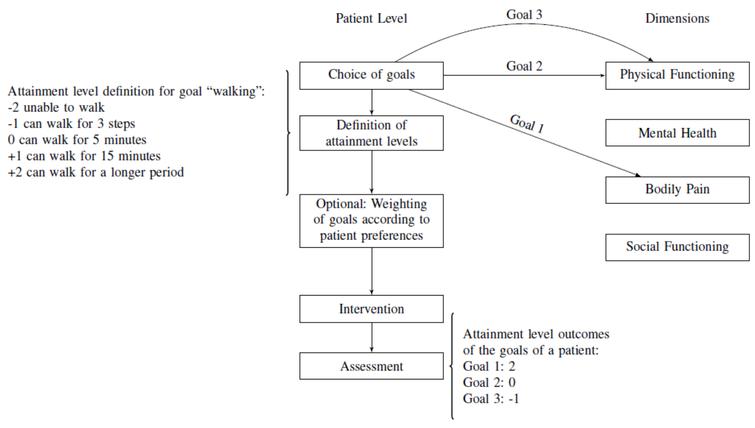


Figure 1: Illustration of the process of goal setting and measurement

Research objectives

Analyzing GAS data from randomized trials

- How to test for a treatment effect in an optimal way?
- What kind of weights should be applied to the individual goals?
- Interpretation of differences in patient specific effects?

Designing randomized trials with GAS outcome measure

How is a hypothesis test using a GAS endpoint affected by

- Maximum number of goals
- Correlation between goals
- Proportion of goals affected by the treatment
- Number of attainment levels

Methods

A data generating multilevel hierarchical model is set up based on discretized normally distributed latent variables modeling the variability between patients and goals. The attainment levels have to be aggregated and the resulting scores can be tested by a t test.

Null hypothesis

H_0 : The average goal attainment level of the experimental group and of the control group are equal.

Challenges

- **Clustered ordinal observations**: Goal attainment levels from within patients tend to be more alike than from different patients and provide less information about the overall treatment effect.
- **Different number and choice of goals per patient**: Less correlated or more goals of a patient provide more information about the overall treatment effect.

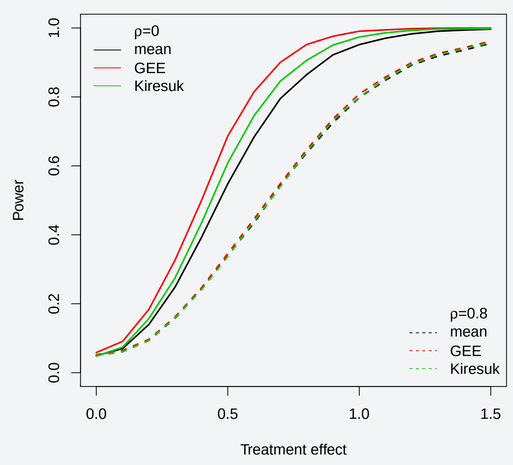
Aggregation methods

- **Mean of the goal attainment levels for each patient**
- **Kiresuk and Sherman method**:
  sum of the standardized mean goal attainment levels (OLS estimator)
- **Generalised estimation equation (GEE) method**:
  weights the goal attainment levels with the inverse of the covariance matrix (GLS estimator)

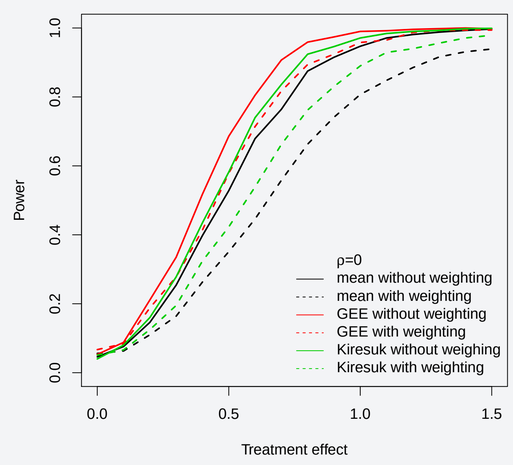
Results

Impact of analysis methods on the power

Power of methods for different correlations



Power of methods with patient preference weighting



- The GEE method weights the contribution of each patient depending on the number of goals and their correlation in a more optimal way than the Kiresuk method.
- Patient preference weighting leads to a substantial loss in power.

Impact of design parameters on the power

- The power increases with the number of goals affected by the treatment, but levels off: For weak correlation between goals, there can be a substantial power increase up to about 5 goals. If goals chosen by a patient are very similar, the gain in power by adding goals is small.
- Including goals that are not affected by the treatment can lead to a substantial loss in power.
- A scale with 5 levels appears to be sufficient. Further increasing the number of level has little influence on the power.

Conclusions

- Trade-off: Larger sample sizes and relevance of the endpoints for the patients, but no inferences for individual goals are possible, only for the difference in the overall treatment effect.
- Improvement in power is possible if a GEE approach instead of the suggested Kiresuk formula is used. Patient preference weighting of goals has a negative impact on the power.
- For an efficient application of GAS endpoints in clinical trials, the statistical implications of design choices should be considered.

References

[1] A. Agresti. *Analysis of ordinal categorical data*, volume 656. John Wiley & Sons, 2010.

[2] D. Hedeker and R. D. Gibbons. A random-effects ordinal regression model for multilevel analysis. *Biometrics*, pages 933–944, 1994.

[3] T. J. Kiresuk and M. R. E. Sherman. Goal attainment scaling: A general method for evaluating comprehensive community mental health programs. *Community mental health journal*, 4(6):443–453, 1968.

Contact Information

- susanne.urach@meduniwien.ac.at
- http://www.asterix-fp7.eu - https://cemsiis.meduniwien.ac.at/ms

Acknowledgements

This work has been funded by the FP7-HEALTH-2013-INNOVATION-1 project Advances in Small Trials Design for Regulatory Innovation and Excellence (ASTERIX). Grant Agreement No. 603160. Website: http://www.asterix-fp7.eu/