Goal Attainment Scaling, an individualized instrument with potential for outcome measurement in rare diseases

Hanneke van der Lee

Charlotte Gaasterland

Susanne Urach

Patient centered outcomes in rare diseases

- Generic outcome measures usually not responsive

- Development and validation of disease-specific outcome measures in rare diseases problematic

- Heterogeneity among rare disease trial participants

- Looking for an individual outcome measure: Goal Attainment Scaling (GAS)

    Kiresuk and Sherman, Community Mental Health Journal 1968; 4 (6): 443.

asterix

GAS in practice (1)

Heterogeneous patients, different goals

Adam

Brad

Chris

'I want to walk'

'I want to eat independently'

'I want to breathe independently'

How do we measure effect of intervention?

GAS in practice (2)

1. What are your goals?
2. Definition of 5 levels of attainment per goal
3. Which goals are most important to you (weights)?
4. *Intervention*
5. Independent assessment:

   At what level is each goal attained?

Dr. X

Adam

GAS in practice (3)

At baseline:

1. Selection of goals (1 or more)
2. Definition of attainment levels for each goal,

e.g.   -2   unable to walk

            -1   can take 3 steps

            0   can walk for 5 minutes

            +1   can walk for 15 minutes

   +2   can walk for a longer period

3. Goals may be weighted

Post intervention:

1. Assessment of goal attainment levels
2. Kiresuk T-score: weighted sum across all goals

$$T = 50 + \frac{10\,\Sigma w_i x_i}{\sqrt{(1-\rho)\Sigma w_i^{\,2} + \rho(\Sigma w_i)^2}}$$

Systematic review

- Has GAS been used in drug trials?
- For what (drug) interventions has GAS been used?
- What is known about the measurement properties?

31 drug studies without measurement properties

7 drug studies with measurement properties

20 non-drug studies with measurement properties
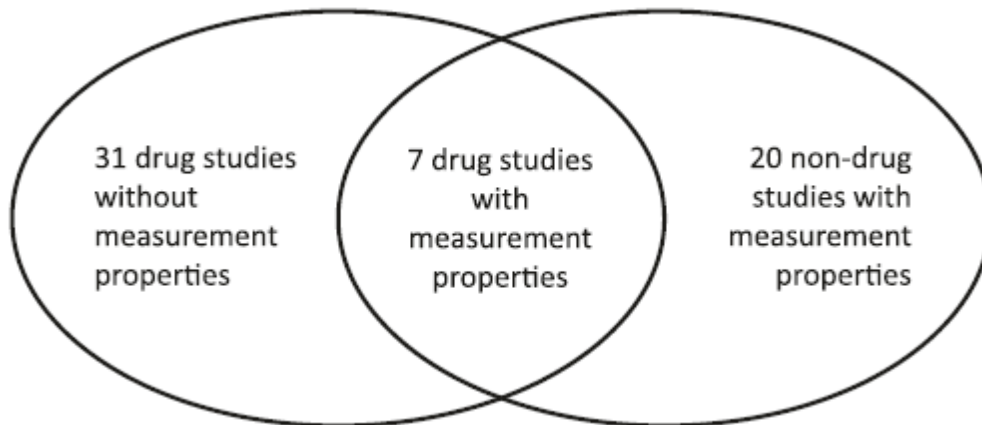
Mostly investigated:
Botox and Baclofen in patients with Cerebral Palsy
Donepezil and Galantamine in Alzheimer Disease patients

Conclusions SR

- Validation is mainly done in geriatrics/rehabilitation

- Usually in non-drug trials

- Insufficient information about validity

Gaasterland et al. BMC Medical Research Methodology (2016) 16:99.

When is GAS useful?

Useful:

- Chronic disease
- Effect of intervention expected on behavioral ability, that can be assessed independently
- Concurrent blinded controls

Not useful:

- Acute, episodic or unpredictable diseases
- Cross-over trials

A statistical approach for the efficient design of GAS studies

- How is a statistical analysis of GAS studies affected by
  - Maximum number of goals
  - Correlation between the goals
  - Proportion of goals affected by the treatment
  - Number of attainment levels

- How should the aggregated scores be analysed best?

- What kind of weights should be applied to the individual goals?

A model to simulate GAS data

- The treatment potentially affects several correlated goals of a patient.

- The observed ordinal score for each goal is the result of a discretization of a continuous normal variable.

- The means of the continuous normal variables shift due to the treatment effect.
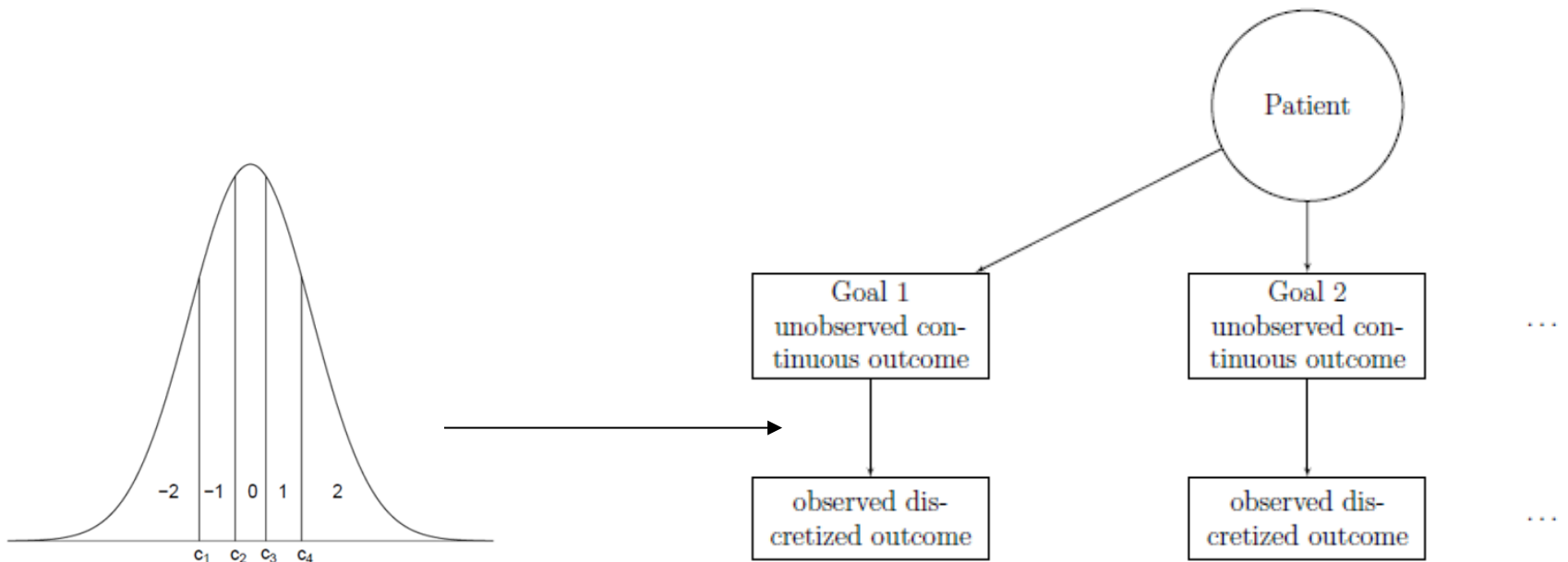
The use of GAS to demonstrate treatment effects

- Some aggregation is needed because the number of goals varies and the goals are not directly related to one another.

- To demonstrate treatment effects, mean Kiresuk T scores between treatment groups can be compared. The interpretation of mean T scores in single arm trials is challenging.

- Treatment effects can only be estimated on the scale of the Kiresuk T scores. For the clinical interpretation, the goals and weights chosen by the patients have to be taken into account.

- The use of parametric test procedures to compare mean Kiresuk T scores is justified because of the robustness to non-normality of tests of central tendency such as the t-test and the sum is taken across different goals.

Designing trials with GAS outcome

- The power increases with the number of goals affected by the treatment per patient, but levels off.

- For weak correlation between goals, there can be substantial power increase up to about 5 goals.

- 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.

Analysis of GAS data

- Improvement in power is possible if a GEE approach is used instead of the suggested Kiresuk formula.

- Weighting of goals
  - If the weights are not correlated with the treatment effect on the goals, weighting may lead to a substantial loss in power.

  - We are investigating to which extent power can be gained by choosing weights that are correlated with the treatment effect on each goal.

Discussion

- Validation of GAS faces specific challenges: generic validation across diseases/interventions not possible

- Randomization and blinding is of paramount importance to address potential sources of bias, as, e.g., the patient's and investigator's choice of goals.

- For an efficient application of GAS endpoints in clinical trials, the statistical implications of design choices (as, e.g., the maximum number of goals) should be considered.

- GAS is a promising instrument for heterogenous patient groups. We propose to develop it as an endpoint for applications in the regulatory context.

References

Kiresuk, T. J. (1973). Goal attainment scaling at a county mental health service. *Evaluation: The International Journal of Theory, Research and Practice*.

Agresti, A., & Kateri, M. (2011). *Categorical data analysis* (pp. 206-208). Springer Berlin Heidelberg.

Hedeker, D., & Gibbons, R. D. (1994). A random-effects ordinal regression model for multilevel analysis. *Biometrics*, 933-944.

www.asterix-fp7.eu