Key Solutions to Model Longitudinal Natural History Data with Application in Ataxia Disease

Ralf-Dieter Hilgers

Institute of Medical Statistics, RWTH Aachen University

Intermediate Course (WP20)

This research is part of the EU-FP7 **IDeAl project** (GA No. 602552) embedded in DH's EU Horizon 2020 fundings from **EJP RD** (GA No. 825575) and **ERICA** (GA No 964908)

Typical Research questions

From longitudinal natural history data profiling (Explorative Trials) to inform randomized clinical trial designs (Confirmative Trials)

Background

- Reetz K, et al. Progression characteristics of the European Friedreich's Ataxia Consortium for Translational Studies (EFACTS): a 2 year cohort study. Lancet Neurol. 2016 Dec;15(13):1346-1354.
- Reetz K, et al.. Progression characteristics of the European Friedreich's Ataxia Consortium for Translational Studies (EFACTS): a 4-year cohort study. Lancet Neurol. 2021 May;20(5):362-372
- Traschütz A. et al. RFC1 Study Group. Natural History, Phenotypic Spectrum, and Discriminative Features of Multisystemic RFC1 Disease. Neurology. 2021 Mar 2;96(9):e1369-e1382.
- Traschütz A. et al. Responsiveness of the Scale for the Assessment and Rating of Ataxia and Natural History in 884 Recessive and Early Onset Ataxia Patients. Ann Neurol. 2023 Sep;94(3):470-485

- Reetz K, et al. Protocol of a randomized, double-blind, placebo-controlled, parallel-group, multicentre study of the efficacy and safety of nicotinamide in patients with Friedreich ataxia (NICOFA). Neurol Res Pract. 2019 Oct 15;1:33.

Outline

- Background and Objective
- Describing Disease Progression via Modelling longitudinal natural history data
- Application to Friedreich Ataxia
- Application to ATM Data with SAS Implementation
- Inform the design of an RCT
  - sample size justification
  - borrow control patients for use in randomised clinical trial
- Conclusion

Main message – Learning aims

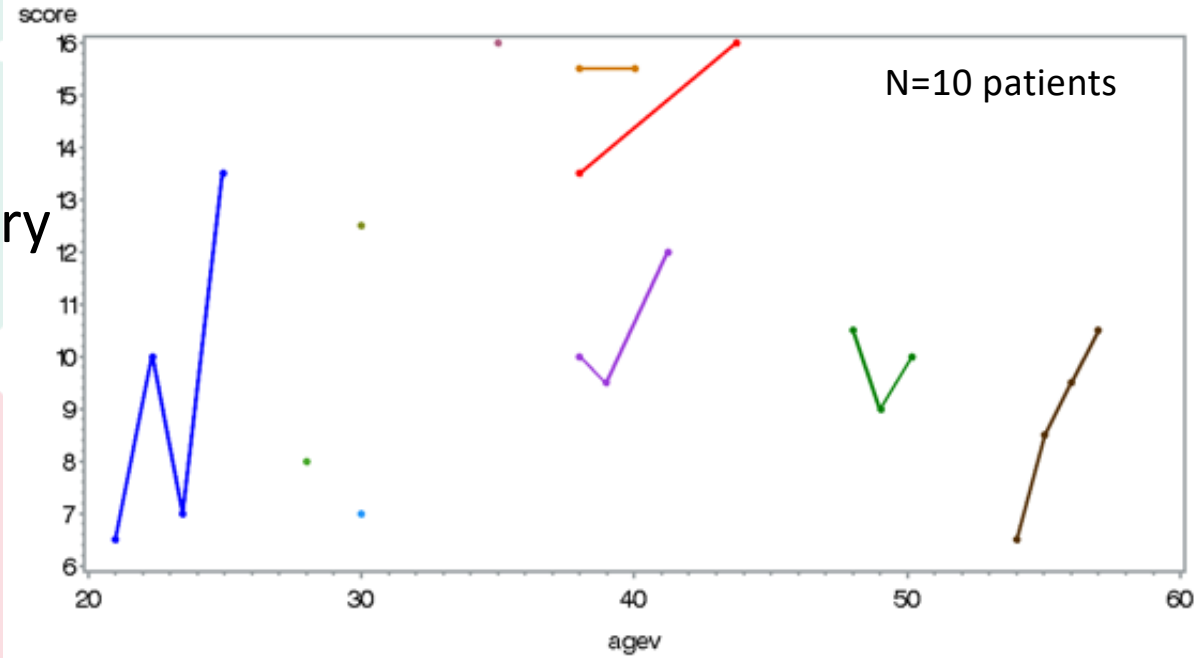
- Longitudinal natural history data are suitable for analysing linear trend of disease progression
- Tailored modelling of response with restricted value range (Tobit-Model) can improve model fit.
- Longitudinal natural history studies can to a certain extend inform about planning a comparative confirmatory trial
  - Arguments to justify sample size for RCT can be delivered
  - Borrow control patient data for use in a randomised clinical trial

Background and Objective

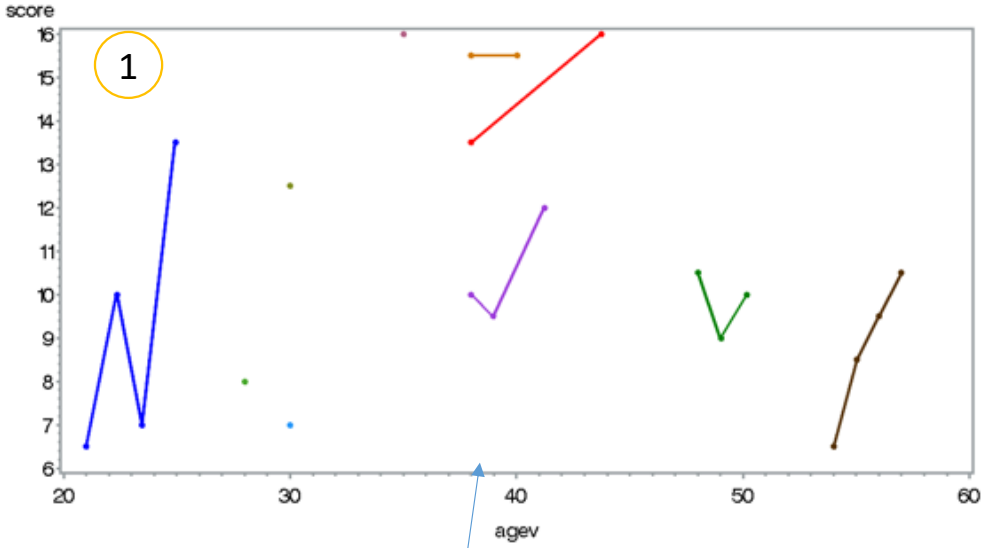
Background – Typical Data Layout

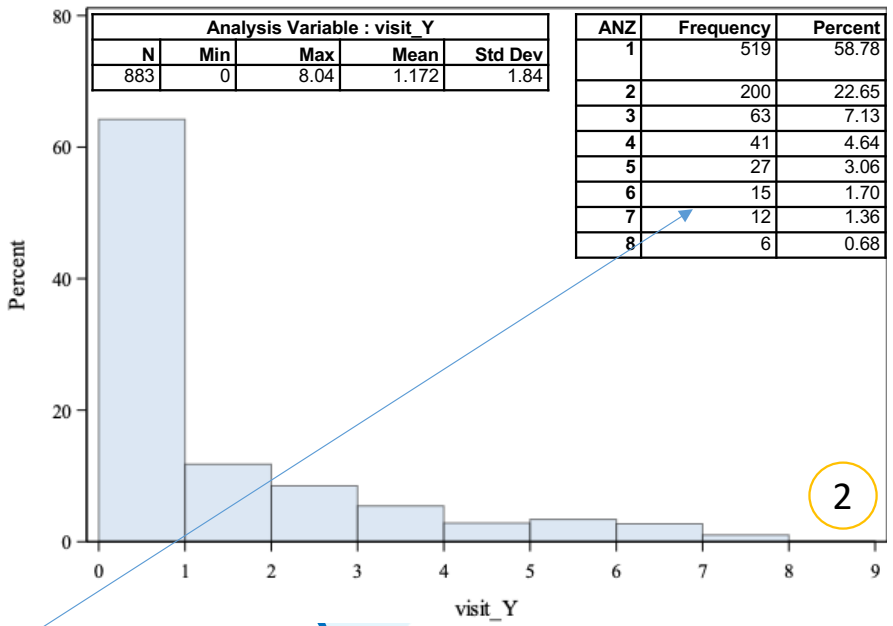
Situation:

If longitudinal data i.e. measurements of the same primary endpoint variable at one or more unscheduled time points are available, one primarily is interested in model disease progression.



Common Challenges with Longitudinal Data





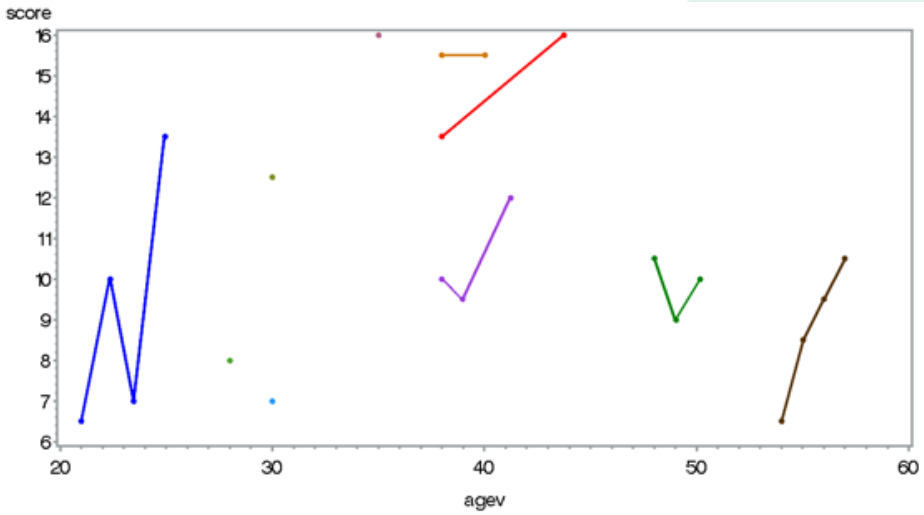
Typical challenges:

- Heterogenous time course
  - no of follow up visits
  - schedule of visits
- May result in missing observations
- Response with a restricted range (ceiled/floored)

ARCAS Registry:
Traschütz Annals of Neurology 2021

Heterogenous time course

Aim: **Modelling and Understanding Patients Disease Course (Progression)**



Data from ARCAS Register

Aspects, which have to be taken into account?

1. Left-Censoring
   We are unaware about the patient disease history before entering in the clinic/data base
2. Right-Censoring
   We are unaware about the patient disease progress beyond the follow up period
3. Unscheduled Visits
   patient disease status is mostly unknown at specific time points, e.g. between follow up visits
4. Distributional Assumptions on Scores
   Often primary endpoint variable have a limited range of values resulting in ceiled or floored effects
5. Functional form
   Often specific functional form of progression of primary endpoint variable over time is unknown

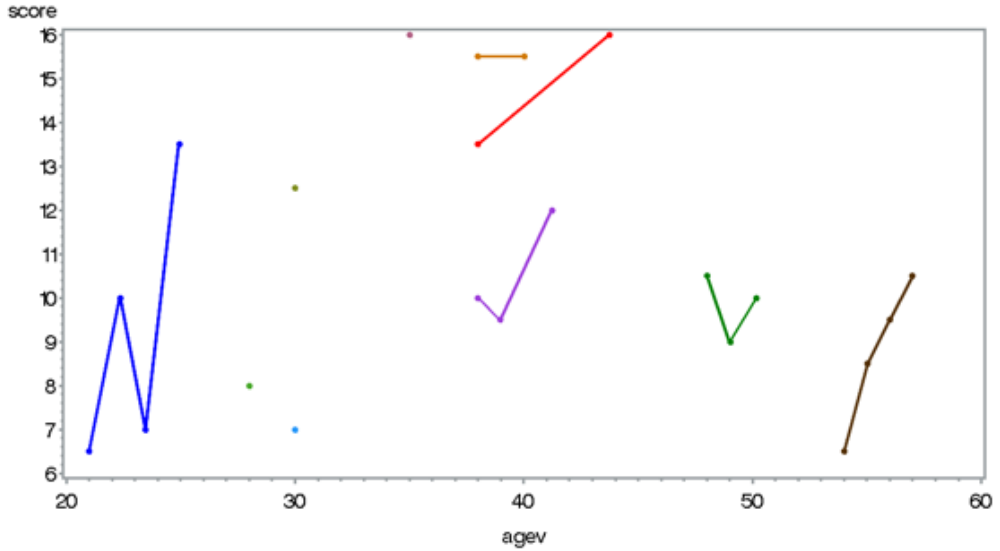
EUROPEAN JOINT PROGRAMME RARE DISEASES

ERICA

Objectives of Analysis

How to "measure" progression?

How varies progression by exploratory factors?



Data from ARCAS Register

Typical options for statistical implementation:

1. Estimate the effect at (single) point in time or (several) points in time
2. Progression modelling by change (typically change [difference] from Baseline)
3. **Progression modelling (Trend) by annual change (typically [linear] slope models)**
4. Progression modelling by functional fitting (typically assuming a common functional form – estimate model parameters)

EUROPEAN JOINT PROGRAMME RARE DISEASES ERICA

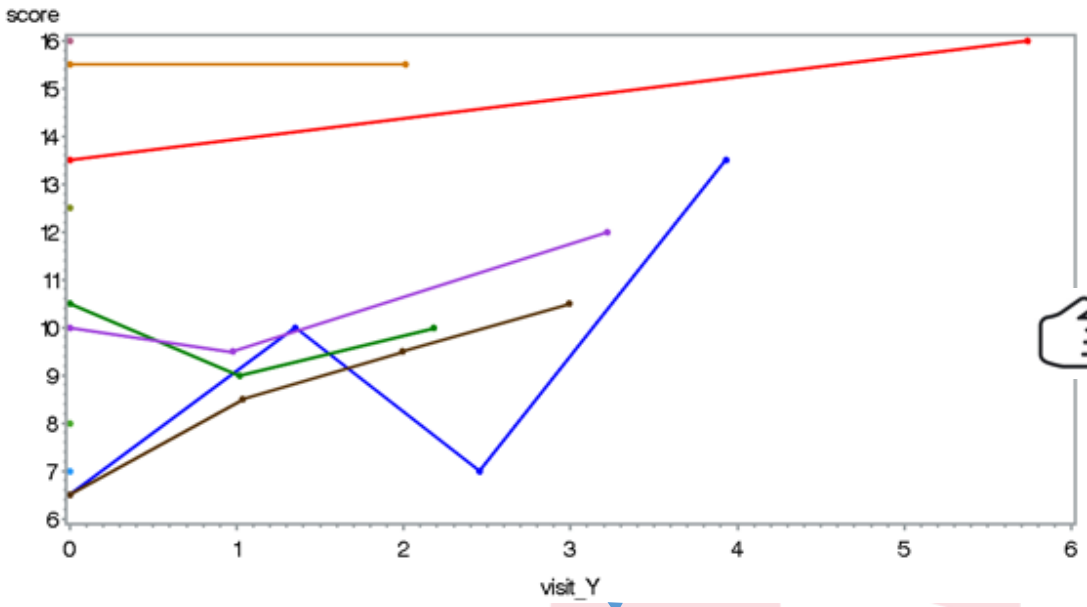
Describing Disease Progression

Describing Disease Progression

1. Estimate the effect at (single) point in time or (several) points in time

“Overcome” Left-Censoring

Set first visit as time zero



Data from ARCAS Register

Set Age of onset to Zero-Time

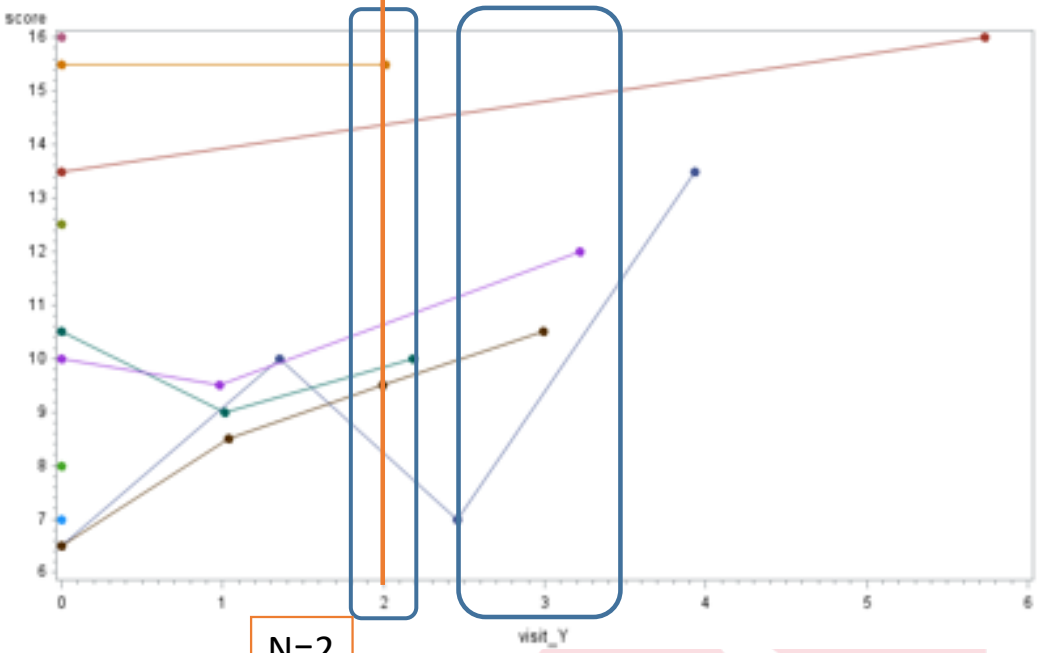
Solution: choose common zero time

Assumption 1: Course is independent from “onset”

Problem:
correct for constant or strict linear progression (with Covaribles)
Age of onset suppressed

Not meaningful in most situations

Unscheduled Visits



N=2

N=3
3 potential missing

Solution: Squeeze time points

Assumption 2: constant effects in selected time intervals

Assumption 3: linear effect over time

Problem:

Value of the score depend on random observational time

Definition of time intervals

➜ overlap

➜ missing values

Interpolation may be necessary

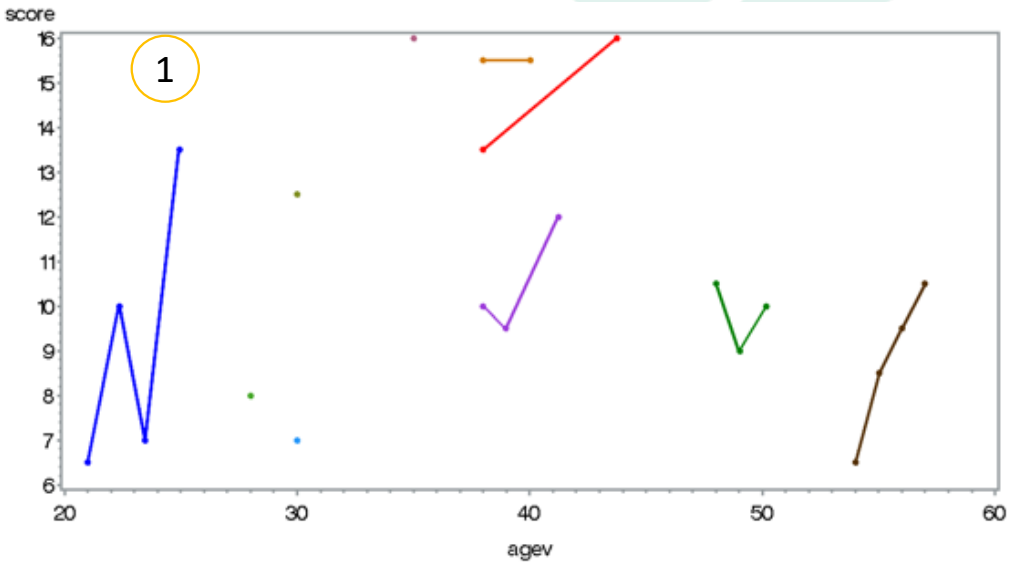
Not meaningful in most situations

Describing Disease Progression

2. Progression modelling by change, typically *difference from baseline*

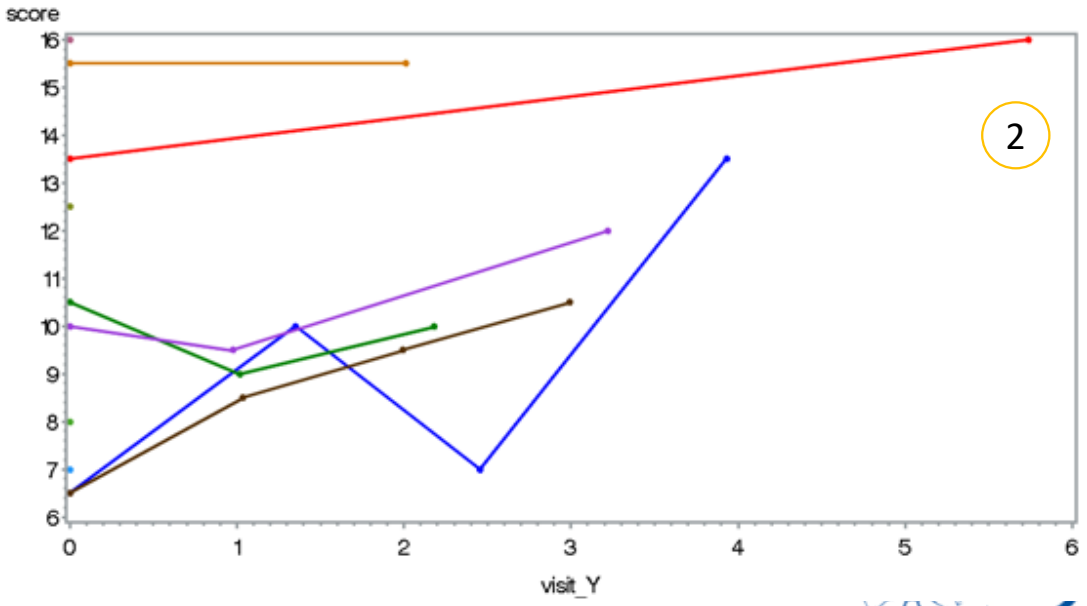
Graphical Presentation of Evaluation

Original data of the score: y_{ik}



i $patients$
k $repeated\ measures$
y_{ik} $response\ of\ patient\ i\ at\ time\ k$
y_{i0} $response\ of\ patient\ i\ at\ baseline\ (time\ 0)$

Set first visit as time zero



EUROPEAN JOINT PROGRAMME RARE DISEASES
ERICA

1. Difference from Baseline Model

Approach: difference from baseline to model progression

Change Model $D = y_{ik} - y_{i0}$

Assumption 4: baseline is uninformative

e.g. y_{ik} does not dependent on y_{i0}

Assumption 5: baseline's are comparable,

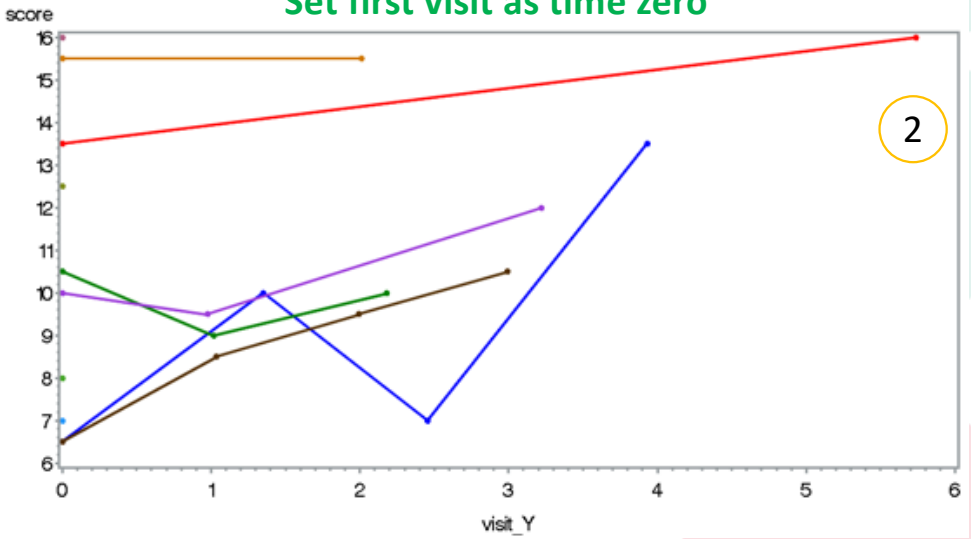
$D_1 = D_2$ although $y_{10} \gg y_{20}$

Remark: Increases variability compared to taking y_{ik} versus decrease power

Not meaningful in most situations, e.g. with different observational times

Graphical Presentation

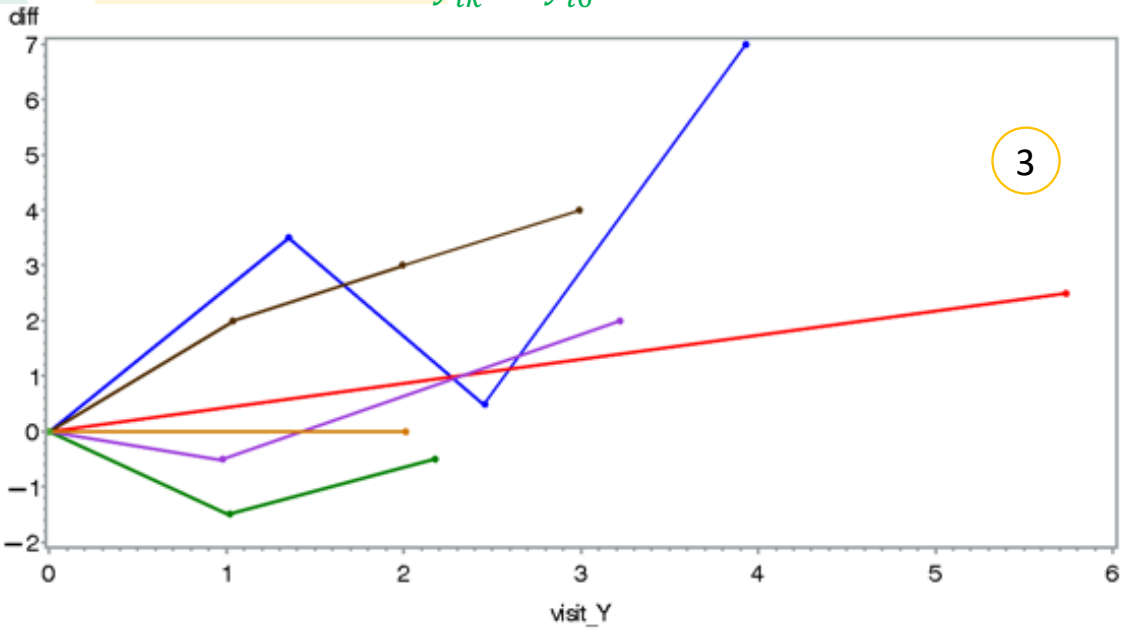
Set first visit as time zero



score: y_{i0}

Difference to first visit

$y_{ik} - y_{i0}$



2. Relative Difference from Baseline Model

Approach: Relative difference from baseline to model progression

Change Model $D = \frac{y_{ik} - y_{i0}}{y_{i0}}$

Assumption 6: change on a multiplicative scale

- linear in log scale
- only doubled difference $y_{ik} - y_{i0}$ and y_{i0} gives the same D

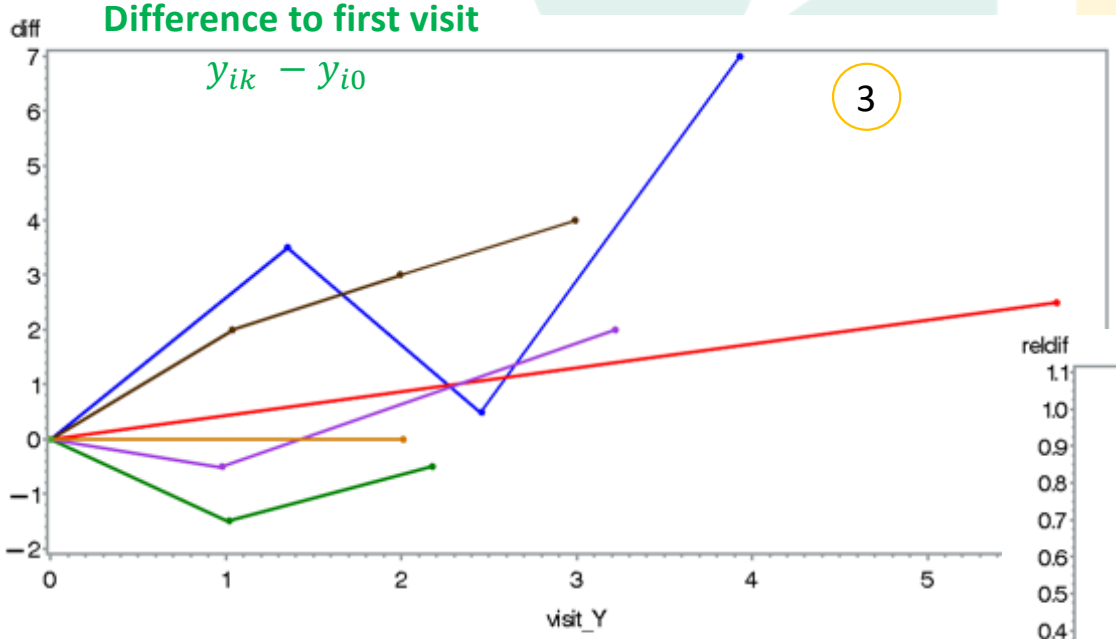
Remark : What happens with baseline obs. /Skip???
Reduces variability within patients with higher baseline scores
Increases variability within patients with lower baseline scores
Difficult to interpret

Not meaningful in most situations, e.g. with different observational times

Graphical Presentation

Difference to first visit

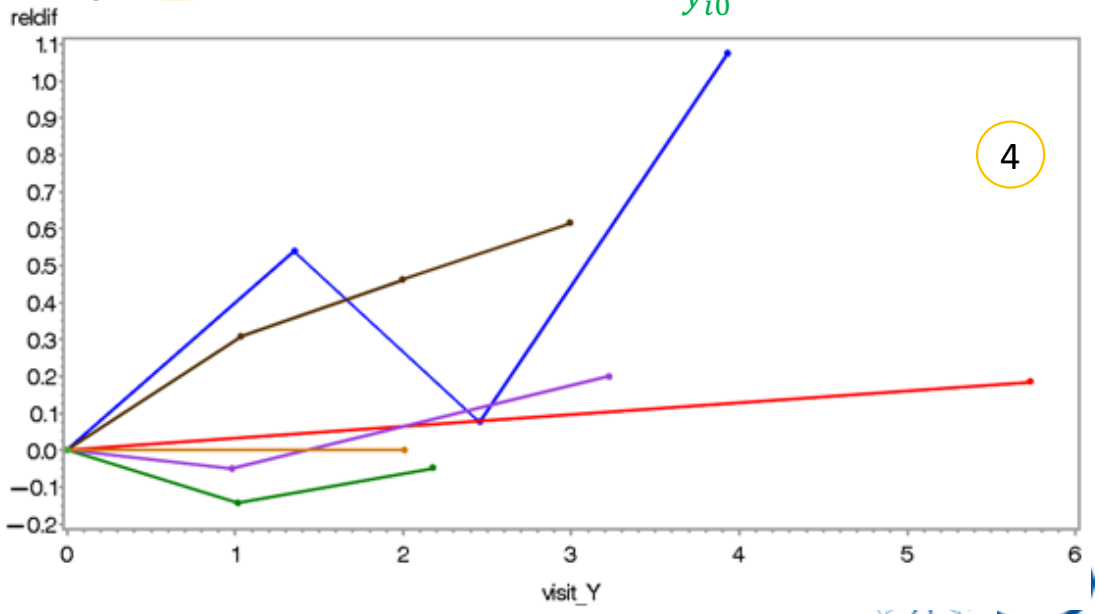
$y_{ik} - y_{i0}$



3

Relative Difference to first visit

$$\frac{y_{ik} - y_{i0}}{y_{i0}}$$



4

EUROPEAN JOINT PROGRAMME RARE DISEASES

ERICA

Summary

Not meaningful in most situations

- **choose common zero time**
- **queeze time points**

Less informative

- **difference from baseline to model progression**

Difficult to interpret

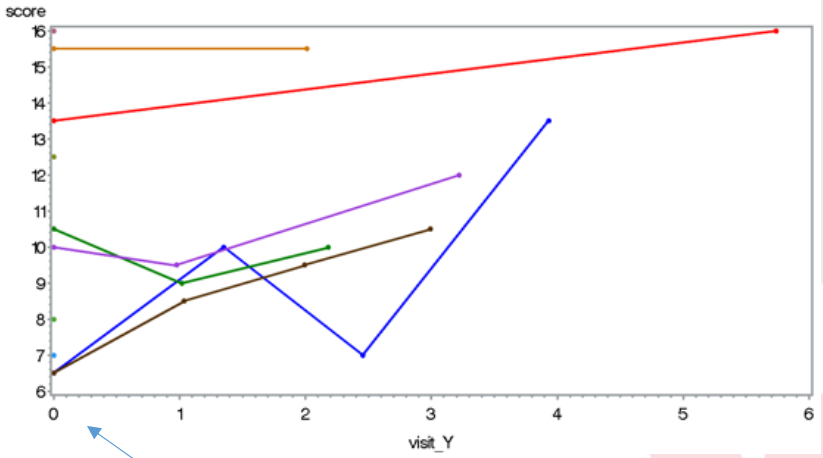
- **Relative difference from baseline to model progression**

Describing Disease Progression

3. Progression modelling (Trend) by annual change (typically [linear] slope models)

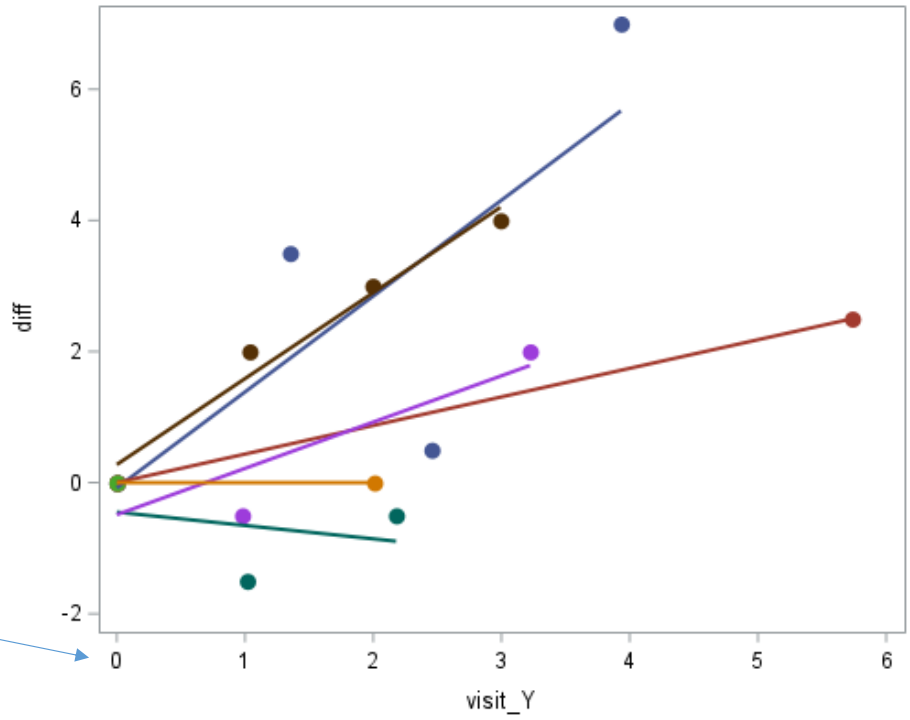
Linear Slope Model

Set first visit as time zero



score: y_{i0}

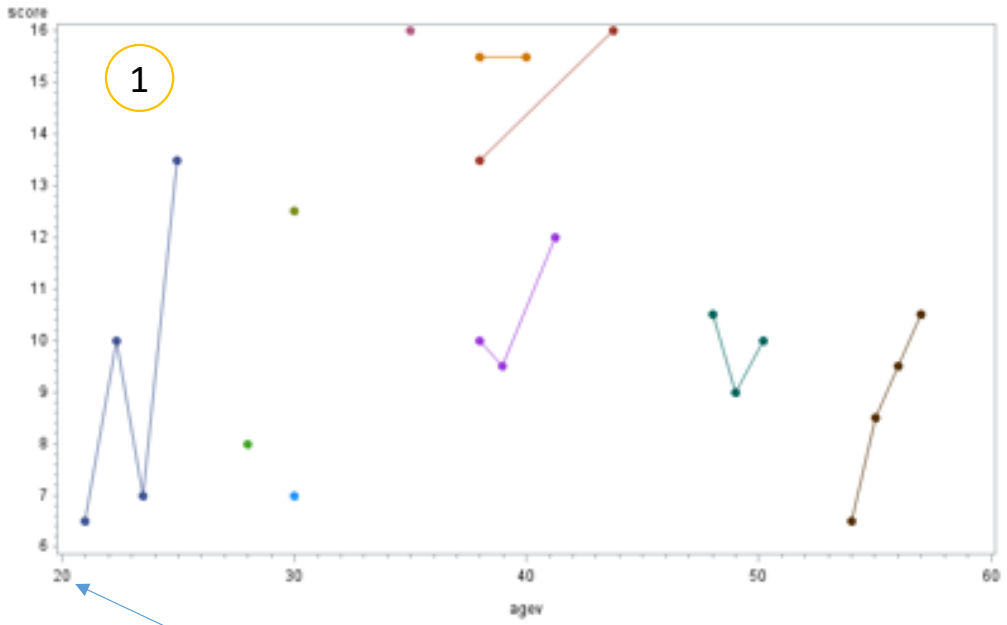
Individual linear regression
i.e. linear trend



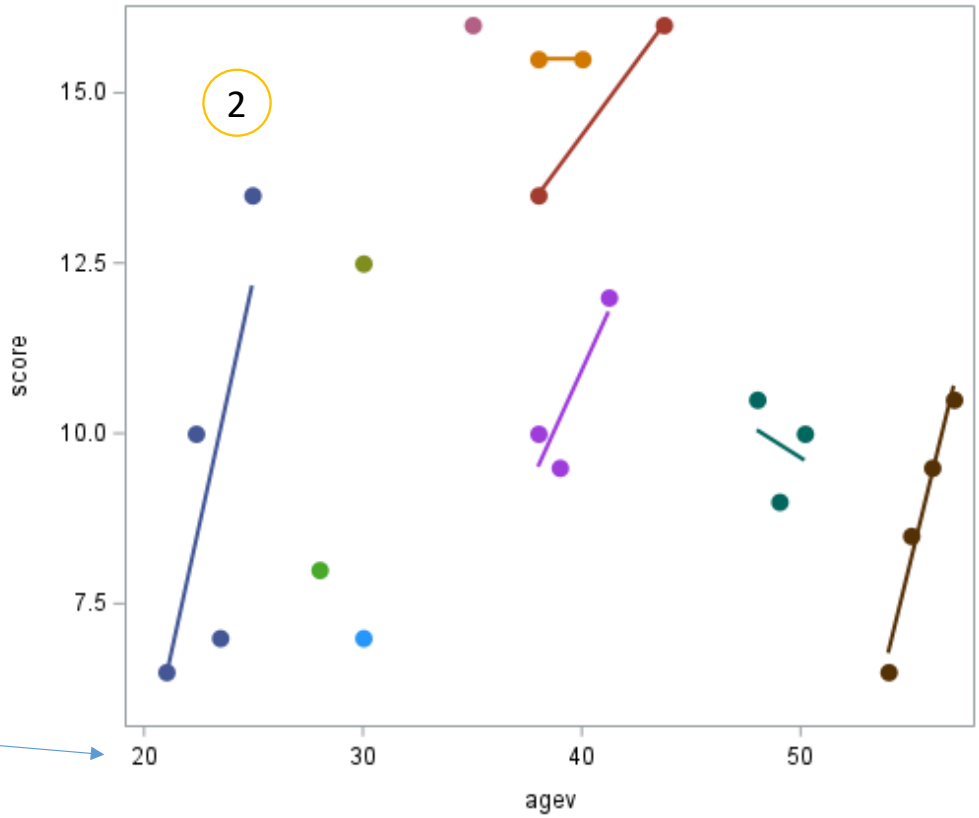
Group	Int	Slope
687	-0.079	0.2757
523	22E-17	-0.489
751	-0.45	0
222	0.2757	0
026	-0.489	0
685	0	0
725	0	0
598	0	1.4623
714	0	0.436
466	0	-0.204
687	1.4623	1.3114

Set first visit as time zero

Modelling linear Trend (calender Time)



Individual linear regression
i.e. linear trend



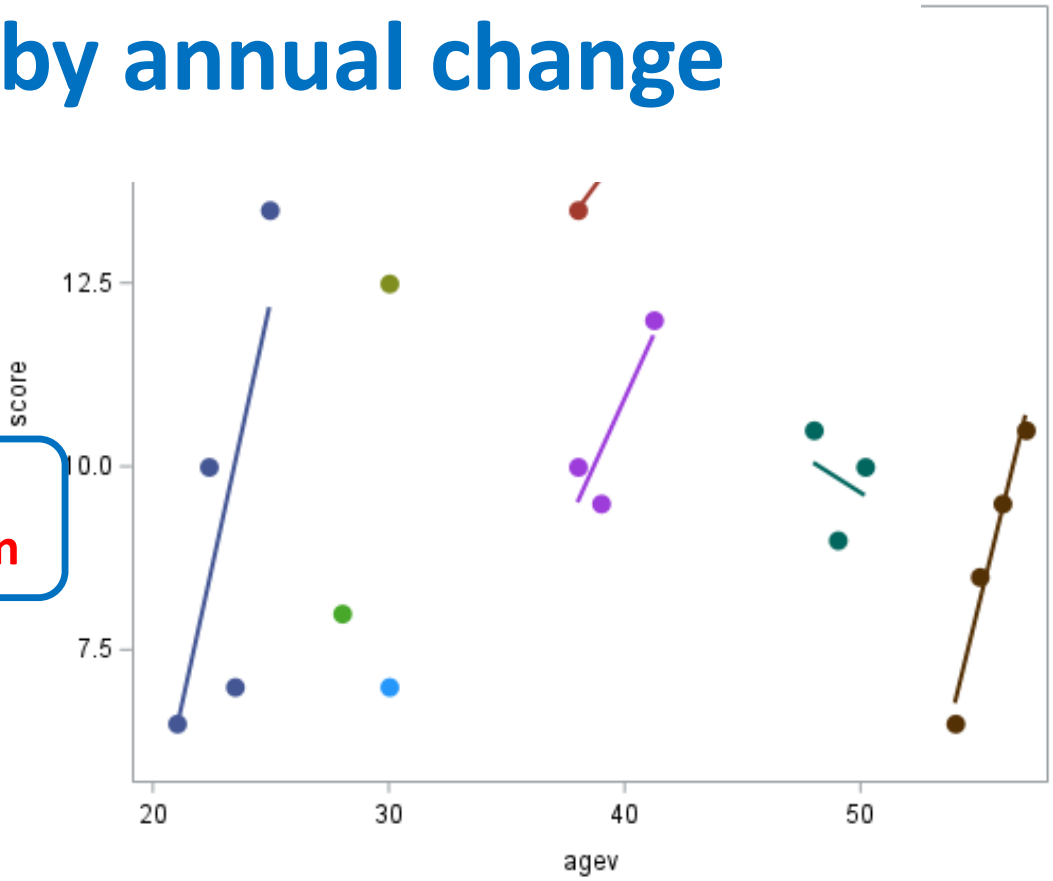
Time as “Age of Onset”

Progression modelling by annual change

$$y_{ik} = \beta_i \; y_{i0} + error_{ik}$$

Solution: reflects trends in data

Assumption 8: linear (monotone) trend
Assumption 9: distributional assumption



Data Form EFACTS Register

Extension:

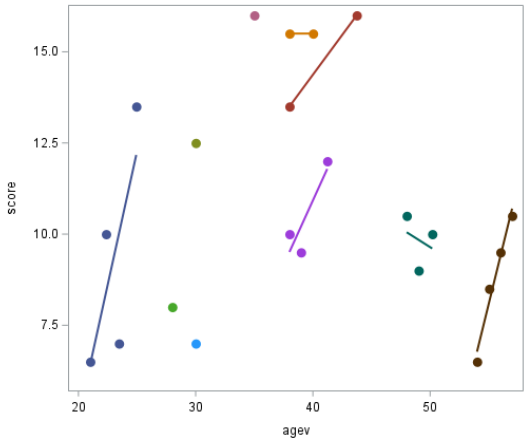
- Implemented via LMEM (rand. Intercept and Slope)
- Tobit models account for ceiled and floored effects
- Covariables

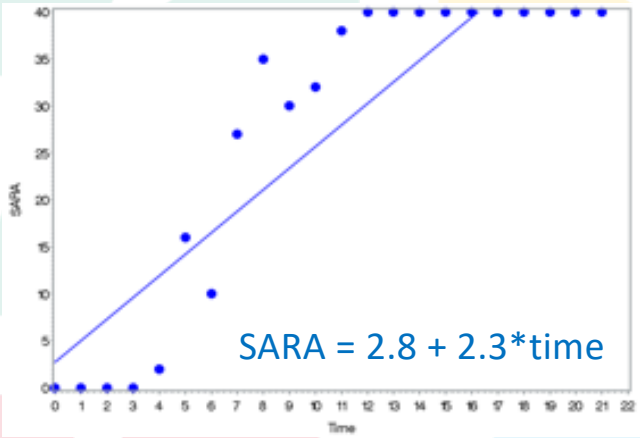
Restricted Score

Ceiled/Floored Scores

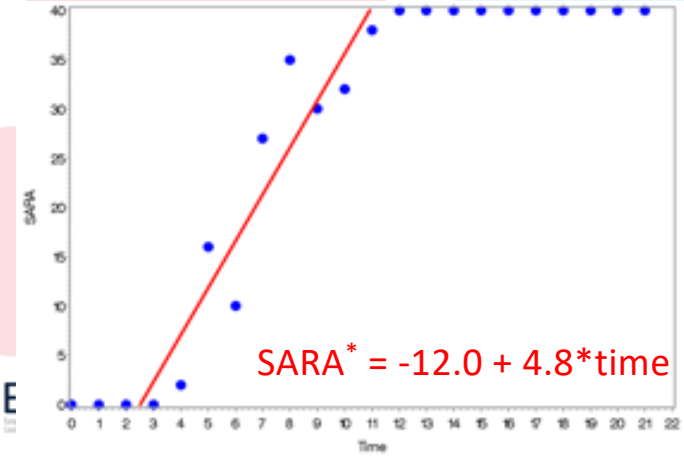
~

below/above limit of detection

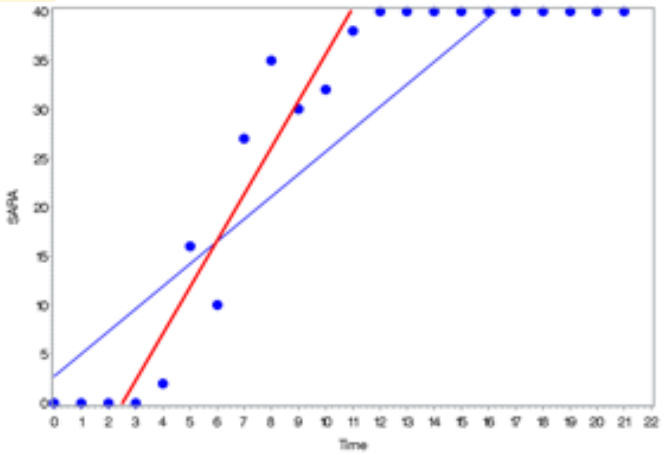




ceiled/floored observations



Ceiled/Floored Scores
Impact the estimation of linear trend



Tobit Model

Econometrica , Jan., 1958

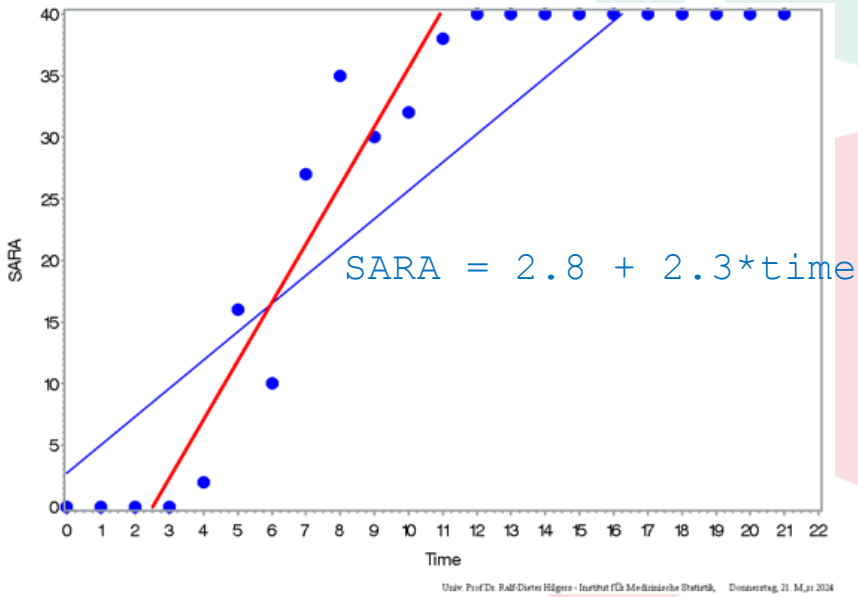
ESTIMATION OF RELATIONSHIPS FOR LIMITED DEPENDENT VARIABLES[1]

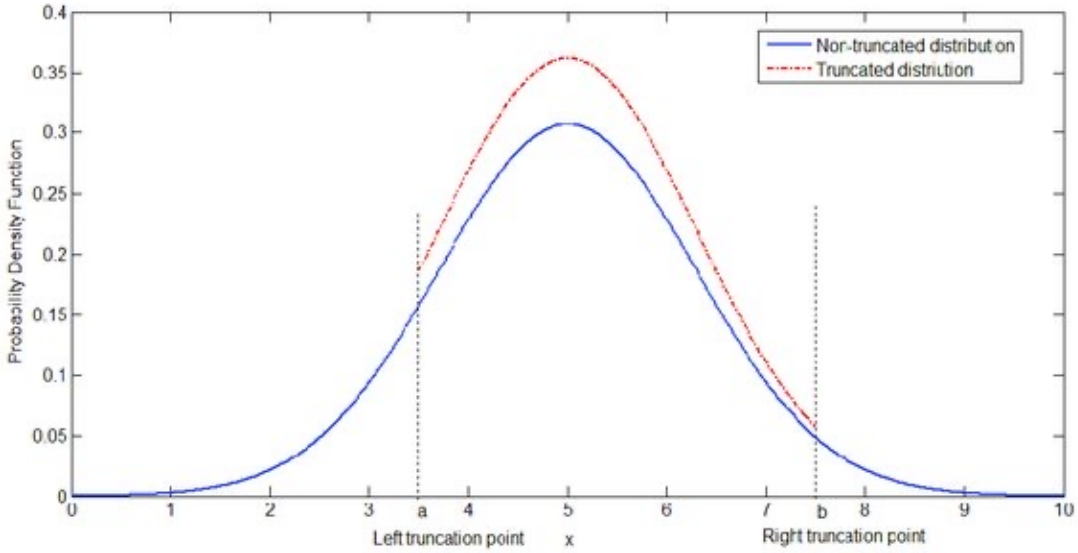
By James Tobin

"What do you mean, *less* than nothing?" replied Wilbur. "I don't think there is any such thing as *less* than nothing. Nothing is absolutely the limit of nothingness. It's the lowest you can go. It's the end of the line. How can something be less than nothing? If there were something that was less than nothing then nothing would not be nothing, it would be something—even though it's just a very little bit of something. But if nothing is *nothing*, then nothing has nothing that is less than *it* is."

E. B. White, *Charlotte's Web*
(New York: Harper, 1952) p. 28.

$SARA^{*} = -12.0 + 4.8*time$

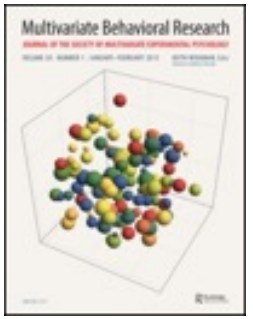




EUROPEAN JOINT PROGRAMME RARE DISEASES

ERICA

Linear Mixed Effect Model versus Tobit Regression



From our simulations, even 18% of the participants reaching ceiling at one occasion (ceiling threshold = 15) could lead to some problems in longitudinal data analysis. Therefore, it is important to detect potential ceiling data before doing longitudinal data analysis. To detect ceiling data, a longitudinal plot of the data could help us to visually check if there is a substantial proportion of participants who obtained maximum scores. Frequency table of the maximum scores across occasions is quantitatively helpful. For growth curve modeling, researchers could also try to use both the regular growth curve method and the Tobit growth curve model to analyze the data. If the percentage of the participants reaching ceiling at one occasion is larger than 20% or there are some important discrepancies between the parameter estimates from two methods, researchers should be cautious about the influences of ceiling effects in the data.

Wang L, Zhang Z, McArdle JJ, Salthouse TA. Investigating Ceiling Effects in Longitudinal Data Analysis. Multivariate Behav Res. 2009 Jul 1;43(3):476-496. doi: 10.1080/00273170802285941.

EUROPEAN JOINT PROGRAMME RARE DISEASES ERICA

ELSEVIER

Journal of Clinical Epidemiology 62 (2009) 953–958

Journal of Clinical Epidemiology

Longitudinal tobit regression: A new approach to analyze outcome variables with floor or ceiling effects

Jos Twisk[a,b,*], Frank Rijmen[a]

[a]Department of Clinical Epidemiology and Biostatistics, VU Medical Centre, Amsterdam, The Netherlands
[b]Department of Methodology and Applied Biostatistics, Institute of Health Sciences, Vrije Universiteit, Amsterdam, The Netherlands

Accepted 1 October 2008

Abstract

Background: In many epidemiologic longitudinal studies, the outcome variable has floor or ceiling effects. Although it is not correct, these variables are often treated as normally distributed continuous variables.

Objectives: In this article, the performance of a relatively new statistical technique, longitudinal tobit analysis, is compared with a classical longitudinal data analysis technique (i.e., linear mixed models).

Study Design and Setting: The analyses are performed on an example data set from rehabilitation research in which the outcome variable of interest (the Barthel index measured at on average 16.3 times) has typical floor and ceiling effects. For both the longitudinal tobit analysis and the linear mixed models an analysis with both a random intercept and a random slope were performed.

Results: Based on model fit parameters, plots of the residuals and the mean of the squared residuals, the longitudinal tobit analysis with both a random intercept and a random slope performed best. In the tobit models, the estimation of the development over time revealed a steeper development compared with the linear mixed models.

Conclusion: Although there are some computational difficulties, longitudinal tobit analysis provides a very nice solution for the longitudinal analysis of outcome variables with floor or ceiling effects.

Keywords: Longitudinal studies; tobit analysis; Linear mixed models; Statistical methods; Floor effects; Ceiling effects

models compared with the linear mixed models. However, it should be borne in mind that the parameters of the longitudinal tobit analysis are difficult to estimate, especially when there are more random coefficients and/or the model becomes more extensive. So, because of the computational difficulty and instability of the models, one should be very careful with the use of complicated longitudinal tobit models. Furthermore, the smaller the number of censored

> Compare Parameter Estimates under both models

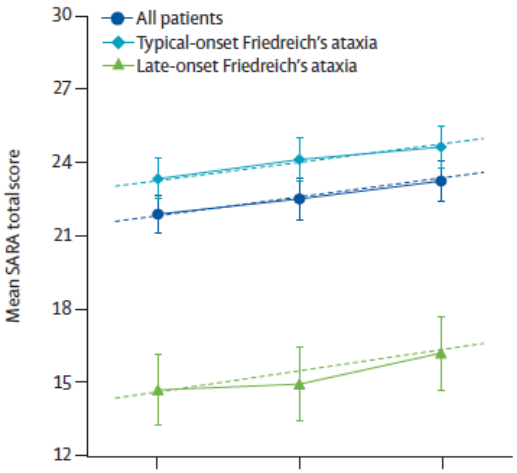
EUROPEAN JOINT PROGRAMME RARE DISEASES ERICA

Application to Friedreich Ataxia

SARA in Friedreich Ataxia

Assumption 8: linear (monotone) trend

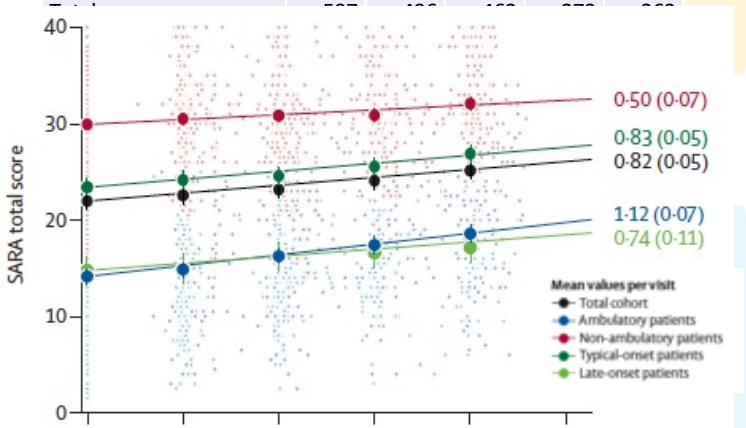
Assumption 9: distributional assumption



Fairly linear

Reetz 2016

SARA	Visit				
Frequency	0	1	2	3	4
within (1-39)	591	489	457	368	357
	98.99	98.59	98.92	98.66	98.35
high (score 40)	6	7	5	5	6
	1.01	1.41	1.08	1.34	1.65



Fairly linear

Reetz 2021

Among others

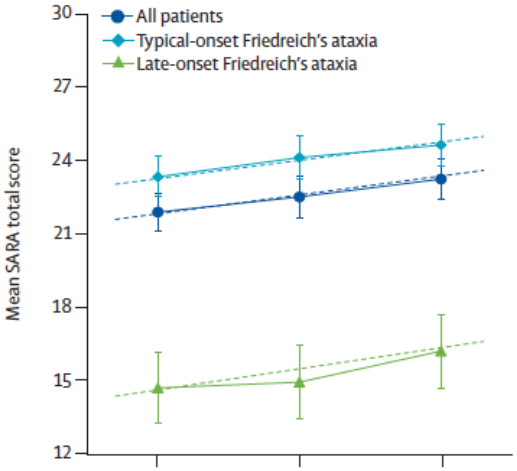
But risk for truncated distribution

Data Form EFACTS Register (I have to ask for Permisson)

SARA in Friedreich Ataxia

Assumption 8: linear (monotone) trend

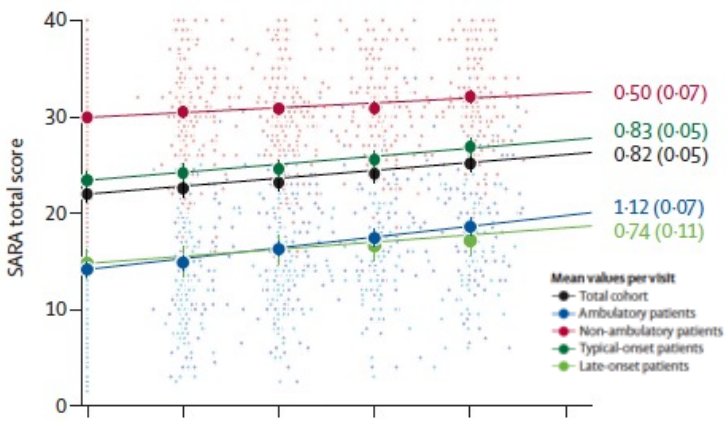
Assumption 9: distributional assumption



Fairly linear

Reetz 2016

SARA	Visit				
Frequency	0	1	2	3	4
within (1-39)	591	489	457	368	357
	98.99	98.59	98.92	98.66	98.35
high (score 40)	6	7	5	5	6
	1.01	1.41	1.08	1.34	1.65



Fairly linear

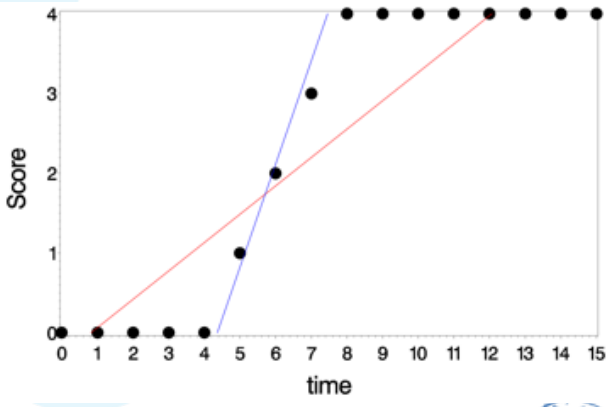
Data Form EFACTS Register (I have to ask for Permisson)

Among others

Reetz 2021

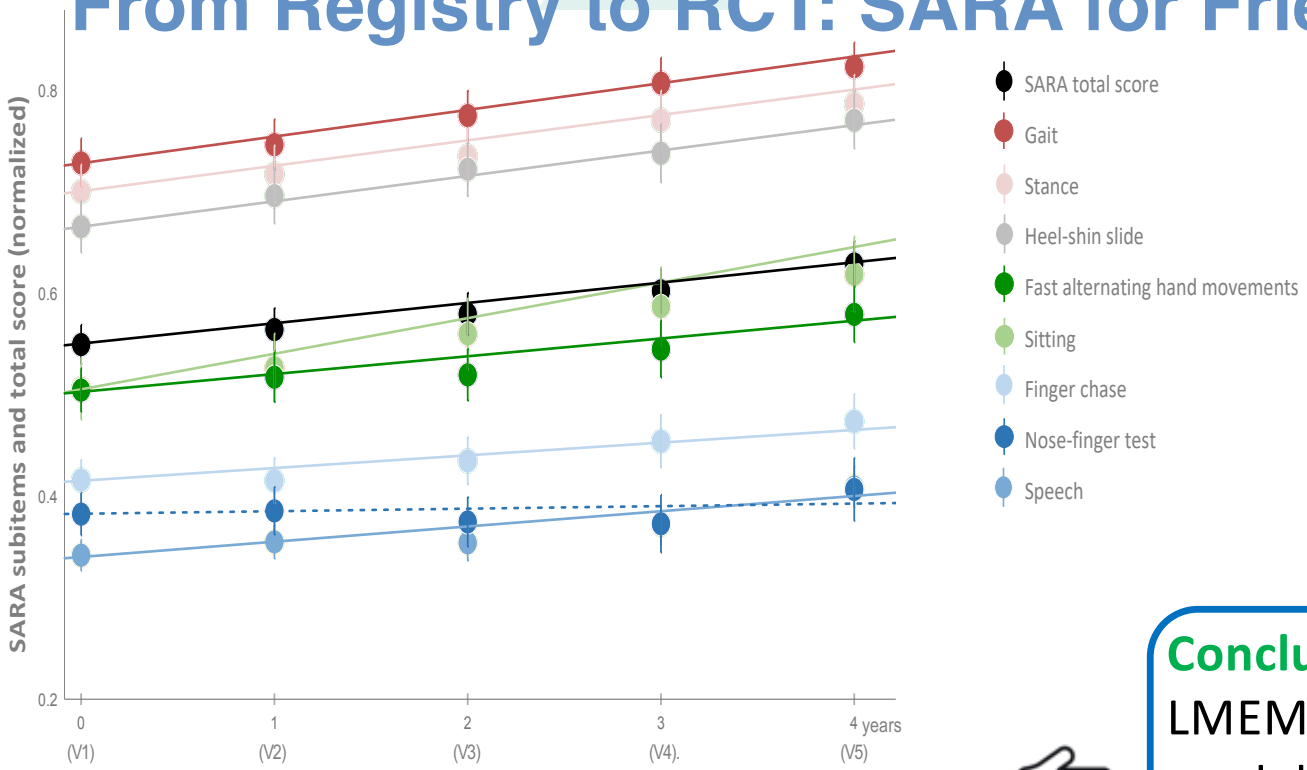
But risk for truncated distribution

Adjusted NLME model necessary if 20%+ ceiled



Tobin 1958

From Registry to RCT: SARA for Friedreich Ataxia



Reetz, K et al. EFACTS 2016, 2021
Reetz, K et al. Nicofa Trial2019

- **SARA as composite**
- important components ?
- Linear Mixed Effects Model (LMEM)

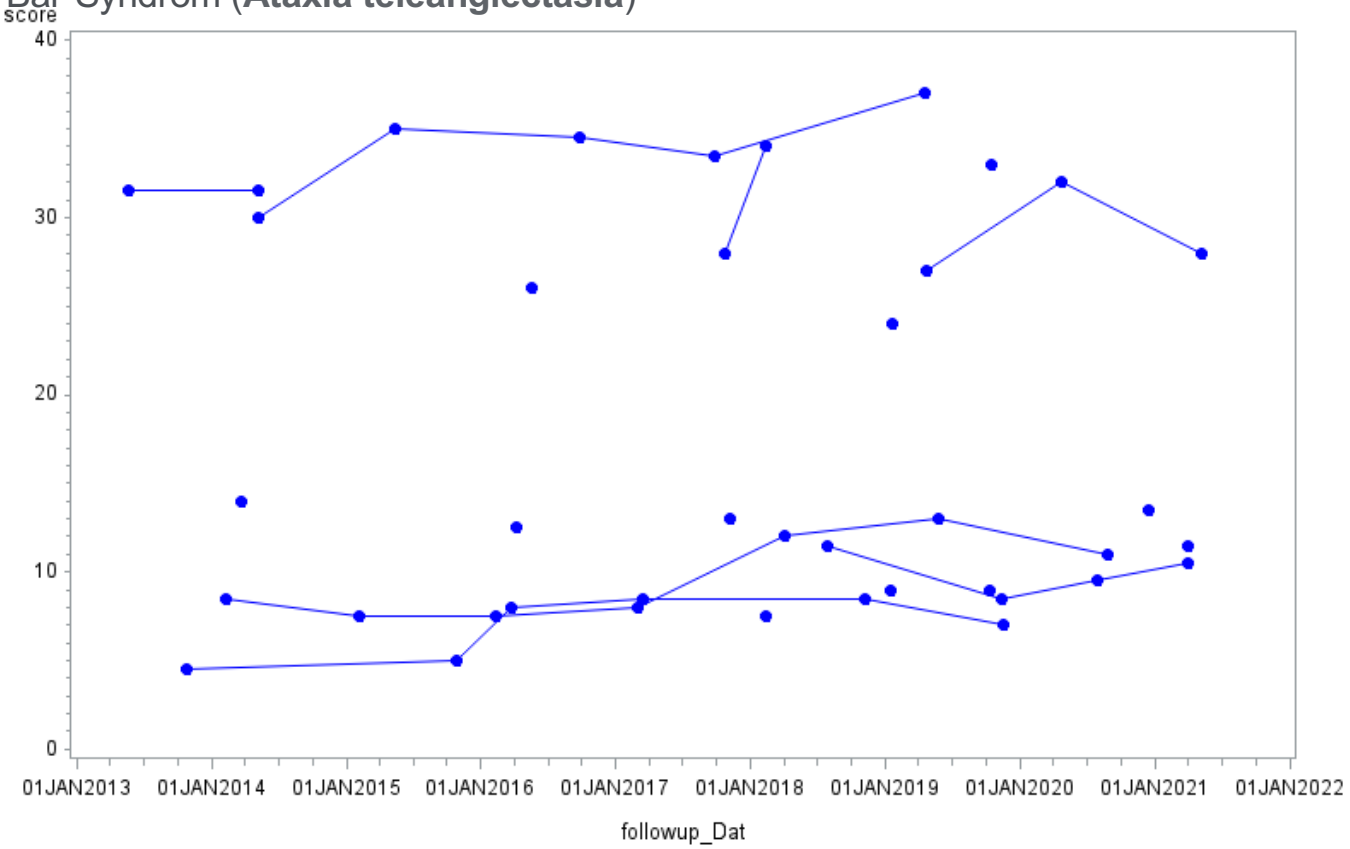
Conclusion - Analysis

LMEM provide a sufficient solution to model longitudinal data where ceiled and floored observations are ignorable.

Application to ATM Data with SAS Implementation

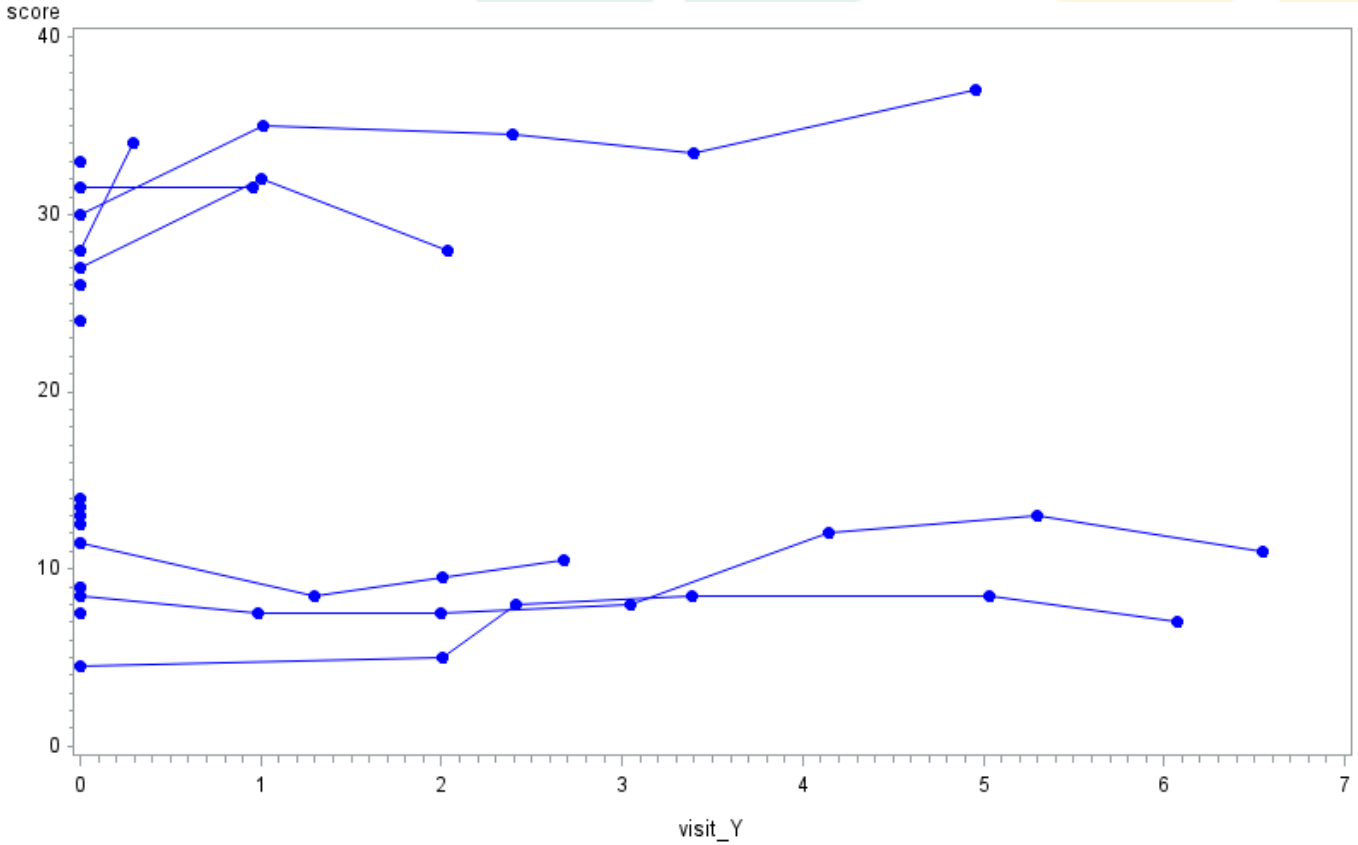
Data of selected 18 ATM cases

18 patients with Louis-Bar-Syndrom (**Ataxia teleangiectasia**)



Calender Time

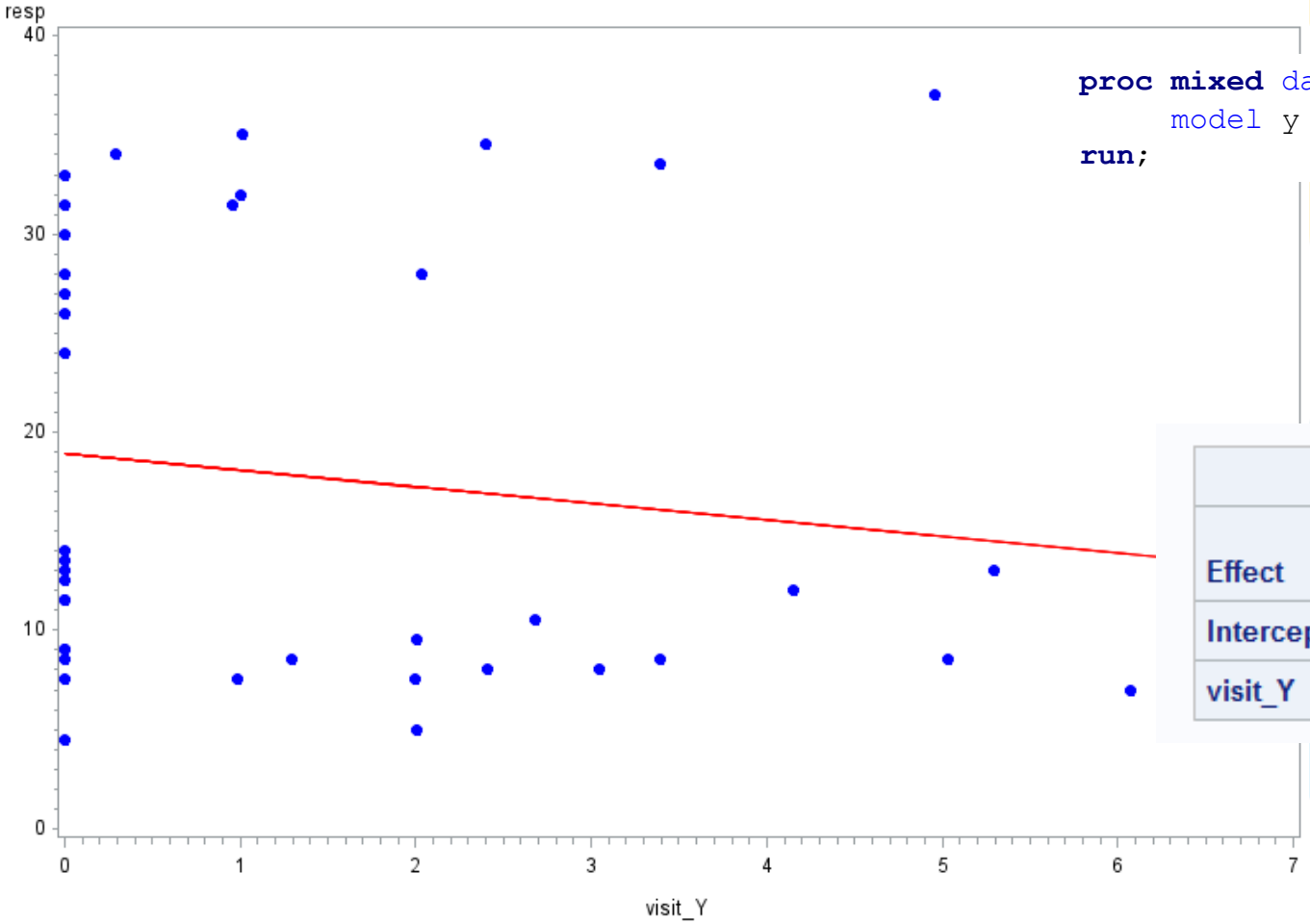
Data of selected ATM cases



Time from Baseline Visit

(Assumption: Comparability of courses from Baseline Visit on)

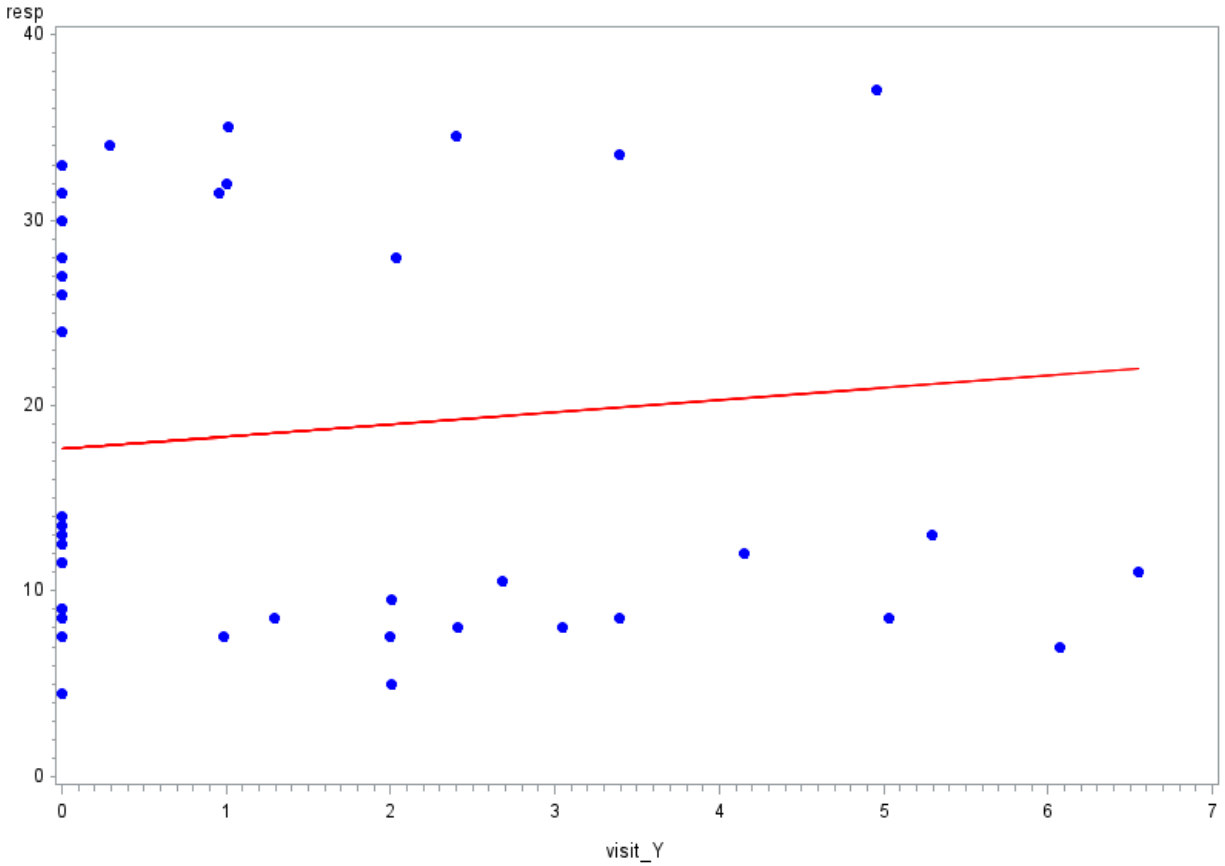
Joint Regression model ignoring longitudinal Character




```
proc mixed data=xxx method=reml ;
     model y = time;
run;
```

Solution for Fixed Effects					
Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	18.9184	2.2356	38	8.46	<.0001
visit_Y	-0.8377	0.8983	38	-0.93	0.3569

Joint Regression model with random intercept



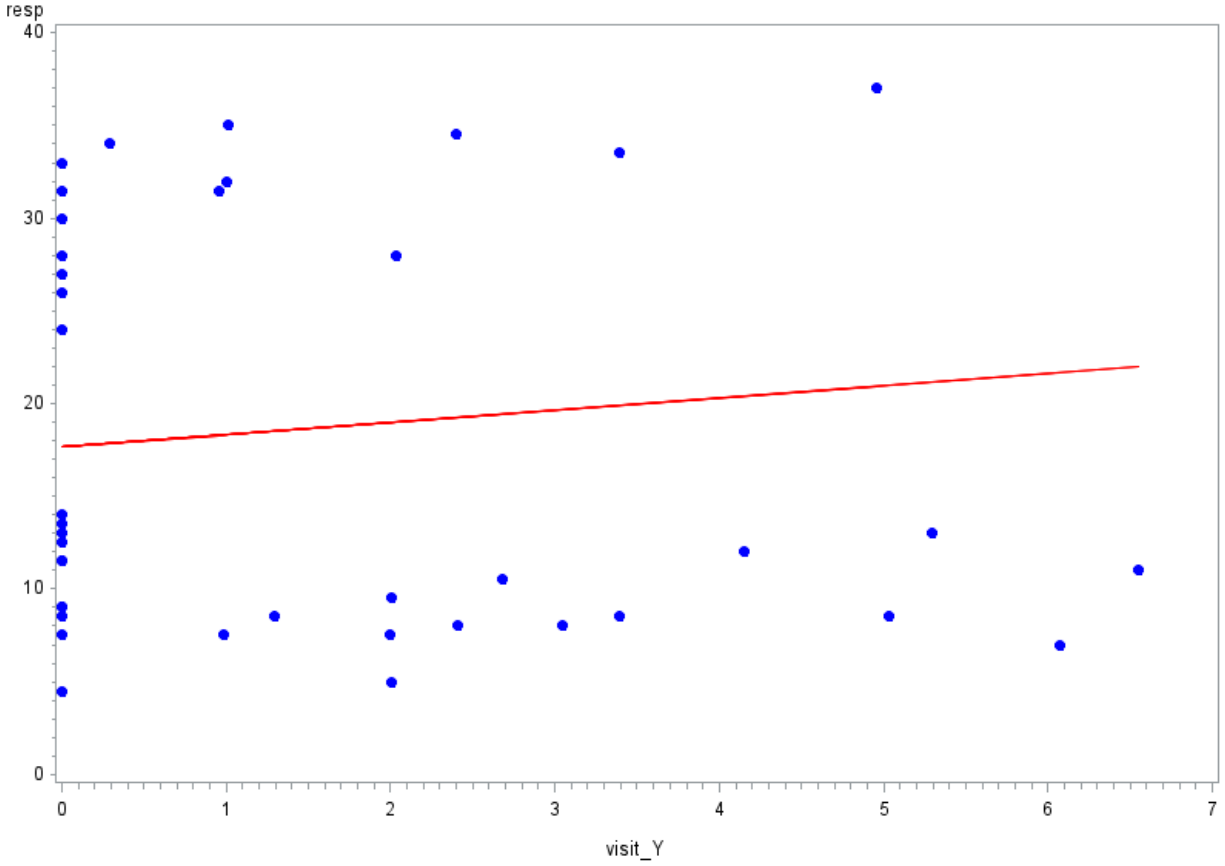

```
proc mixed data=xxx method=reml ;
     class Patno ;
     model y = time  /  ddfm=kr;
     random intercept / subject=Patno;
run;
```

Solution for Fixed Effects					
Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	17.6373	2.4056	17	7.33	<.0001
visit_Y	0.6630	0.2115	21	3.13	0.0050

EUROPEAN JOINT PROGRAMME RARE DISEASES

ERICA

Regression model with random effects



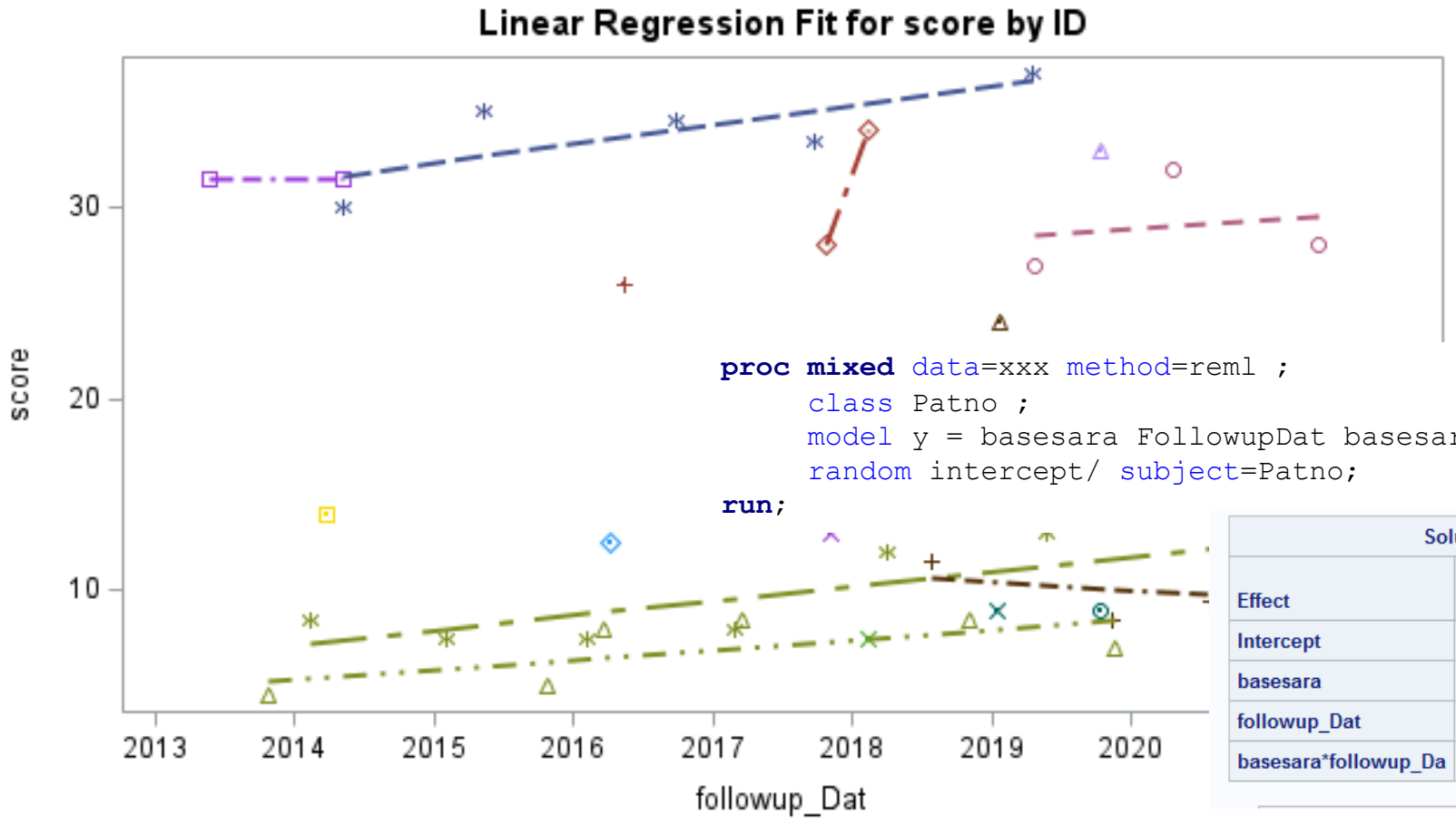

```
proc mixed data=xxx method=reml ;
     class Patno ;
     model y = time  /  ddfm=kr;
     random intercept time / subject=Patno;
run;
```

Solution for Fixed Effects					
Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	17.6373	2.4056	17	7.33	<.0001
visit_Y	0.6630	0.2115	6	3.13	0.0202

EUROPEAN JOINT PROGRAMME RARE DISEASES

ERICA

Evolution of the Model




```
proc mixed data=xxx method=reml ;
     class Patno ;
     model y = basesara FollowupDat basesara *FollowupDat /ddfm=kr;
     random intercept/ subject=Patno;
run;
```

Solution for Fixed Effects					
Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	-18.5490	16.8276	16	-1.10	0.2866
basesara	0.9525	0.9025	20	1.06	0.3038
followup_Dat	0.000865	0.000794	20	1.09	0.2886
basesara*followup_Da	5.122E-6	0.000043	20	0.12	0.9061

Sound Model Fit

- Consider Residual Plot for Distributional Assumptions
- Conduct Influence Diagnostic to detect observations with impact on estimates
- Consider model fit improvement
- Consider formulation of appropriate covariance structure

Inform the design of an RCT

Clinical Trials - Different Story

Analysis Objective: **Estimate the treatment effect**

In Randomised Clinical Trials we typically have:

	Feature	Problem	Solutions
1.	Scheduled Visits	→ Missing Data	→ LMEM & MI incl. MMRM
2.	Similarity	→ Stratification	→ Sensitivity Ana [covariables]
3.	Repeated Observations	→ Multiple Testing	→ LMEM incl. MMRM
4.	MultiComponent Scale	→ Indiv Scales	→ Item Response Models, LMEM

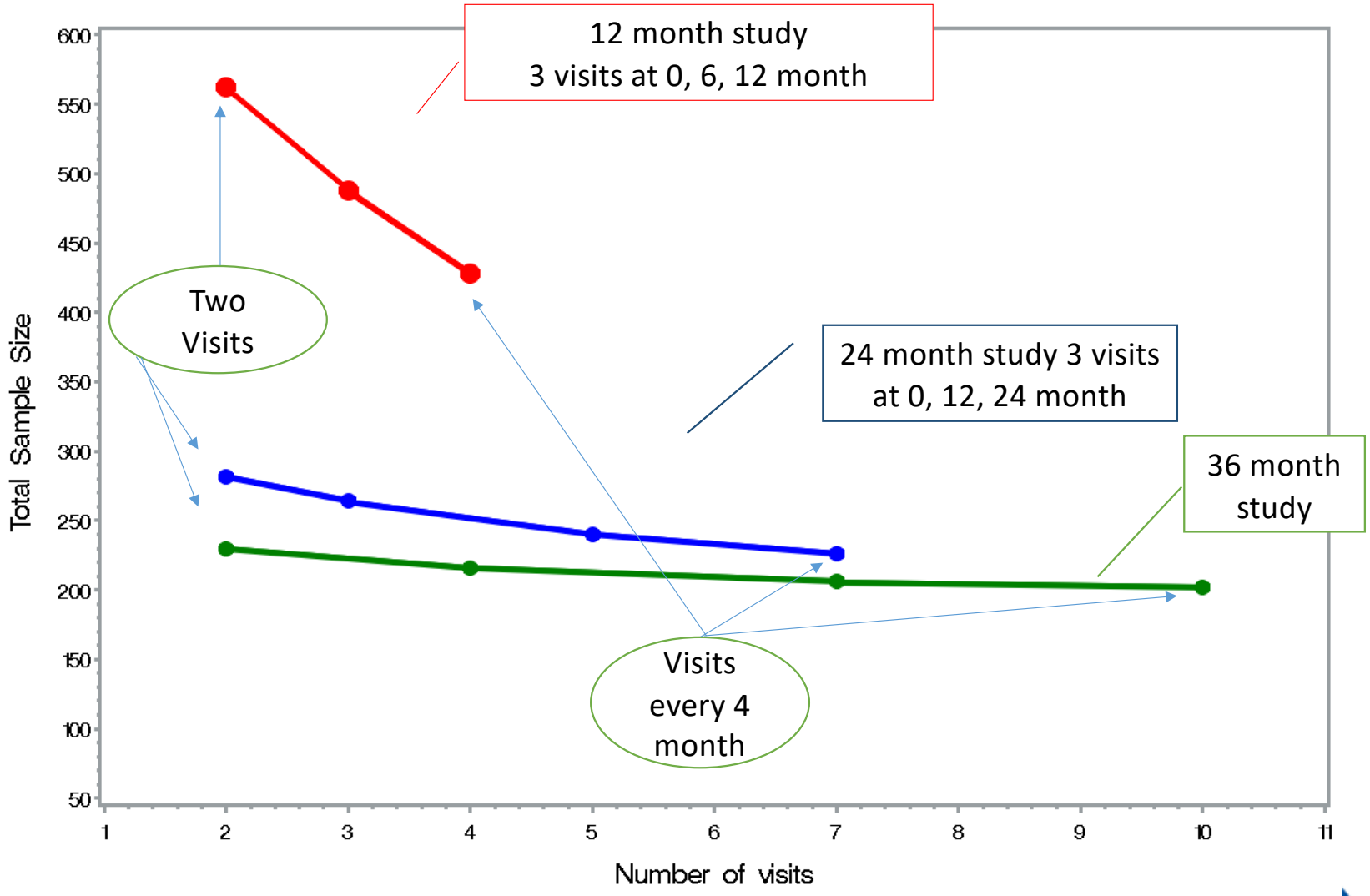
Support of comparative clinical trials

Support **sample size justification** for two arm parallel group trial showing difference in linear progress of SARA score

Sample size

Example:

- two arm parallel group study
- Power 80%
- 50% Reduction
- Slope 0.82
- Random slope model
- two sided
- 1:1 allocation ratio
- **Various number of visits**
- **12/24/36 month**



Sensitivity Analyse

Example:
12 month
two arm parallel group study
Power 80%
two sided
Random slope model
4 month visits
1:1 allocation ratio

Variance Slope	Residual Variance	Reduction of slope from 0.82	Total sample size
0.9930	2.0981	50%	444
1	2	50%	434
1	4	50%	676
0.5	2	50%	340

Support of comparative clinical trials

Support a two arm parallel group trial showing difference in linear progress of SARA score by borrow information of controls

Use of Historical Controls with Randomized Trials

E10, a hybrid approach of using external control data to add to a concurrent randomized control arm in a clinical trial may sometimes be useful.

Rare Diseases: Natural History Studies for Drug Development
Guidance for Industry

FDA, 2019

Pocock's (1976) Criteria

1. *The* **standard treatment** *has to be* **precisely defined** *and must be the same treatment for randomized controls.*
2. *The* **historical control group must have been part of a clinical study** *with the same requirements for patient eligibility.*
3. *The* **methods of treatment evaluation** *have to be the same.*
4. **Patient characteristics** *have to be comparable.*
5. *The studymust have been performed in the same* **organization** *with the same investigators.*
6. *There should be* **no indications** *leading one to* **expect a difference**.

Methoden:

- Bennett (2021) Bayesianische Methoden (Power Prior),
- Schmidli (2014) Meta-Analytic-Predictive Priors
- Viele (2014): Pool than Test Approach

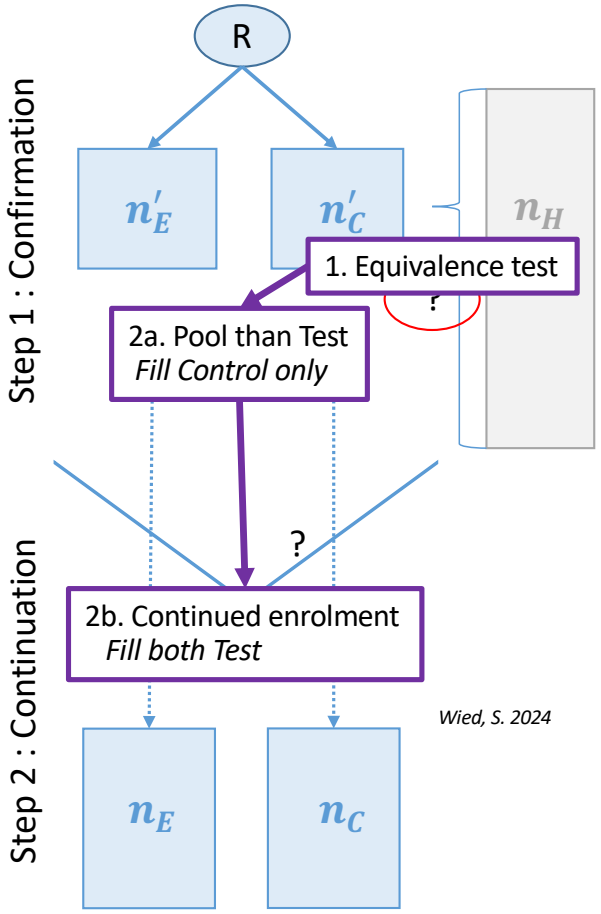
Historical Control– Fill-it-Up Design

Motivation

- Superiority of E (experimental) versus C (control)
- Reduce the necessary RCT sample size, by using a large number of Historical controls (resulting in an unbalanced Design to fulfil the Powerrequirement).

Aim

- Justify the use of historical controls by applying an equivalence proof of the historical and randomized controls
- Combine the historical and randomized controls
- Proof of superiority



EUROPEAN JOINT PROGRAMME RARE DISEASES

ERICA

Use of Historical Controls

- **Hybride Approach** in line with the FDA Guideline, but under consideration of the implied multiple testing problem.
  - last test proves **Non-Inferiority Hypothese**!
- At least 50% of the initial total sample size must be in step 1
- **Power - Gain ?**
- *In situations, where a few patients can be enrolled, the procedure can have some power gain, if good and suitable controls can be identified.*

Powergain (Fill-it-up versus randomized)

One sided test

Table 1: Overall power and sample sizes for different scenarios of the Fill-it-up-design depending on the choice of the equivalence margin Δ considering a medium effect size $\delta = 0.5$ when $n_H = 500$ are available and $\alpha_{S_1} = \alpha_{S_2} = 0.05$.

α_{Ept}	Δ	N_{FIU}	γN_{FIU}	AVN	$1-\beta_{S_1} = 1-\beta_{S_2}$	$1-\beta_{Ept,S_1}$	$1-\beta_{Ept^c,S_2}$	$1-\beta_{FIU}$
0.01	0.4596	100	54	100	0.80	0.9530	0.8471	0.8001
0.01	0.4798	100	54	100	0.80	0.9871	0.8130	0.8001
0.01	0.5000	116	62	116	0.85	0.9999	0.8001	0.8000
0.05	0.3250	102	54	100	0.81	0.9596	0.8403	0.8000
0.05	0.3901	138	74	136	0.90	0.9608	0.8392	0.8000
0.05	0.5000	112	60	110	0.84	0.8175	0.9825	0.8000
0.10	0.2303	124	66	120	0.87	0.9637	0.8368	0.8005
0.10	0.3652	124	66	120	0.87	0.8442	0.9558	0.8000
0.10	0.5000	100	54	96	0.80	0.9889	0.8113	0.8002
0.20	0.1663	102	54	94	0.81	0.9722	0.8279	0.8001
0.20	0.3331	102	54	94	0.81	0.9403	0.8599	0.8002
0.20	0.5000	132	72	120	0.89	0.8397	0.9607	0.8003

Total Sample size for the design without historical controls

Sample size at stage of equivalence proof.

Conclusion

- Do not squeeze the data by time points
- Linear progression model may suffice to show trend
- Linear mixed effects models allow formulation of a progression model
- They can be informative to plan a clinical trial
  - With respect to sample size justification
  - With respect to borrow information


Done! Thanks!