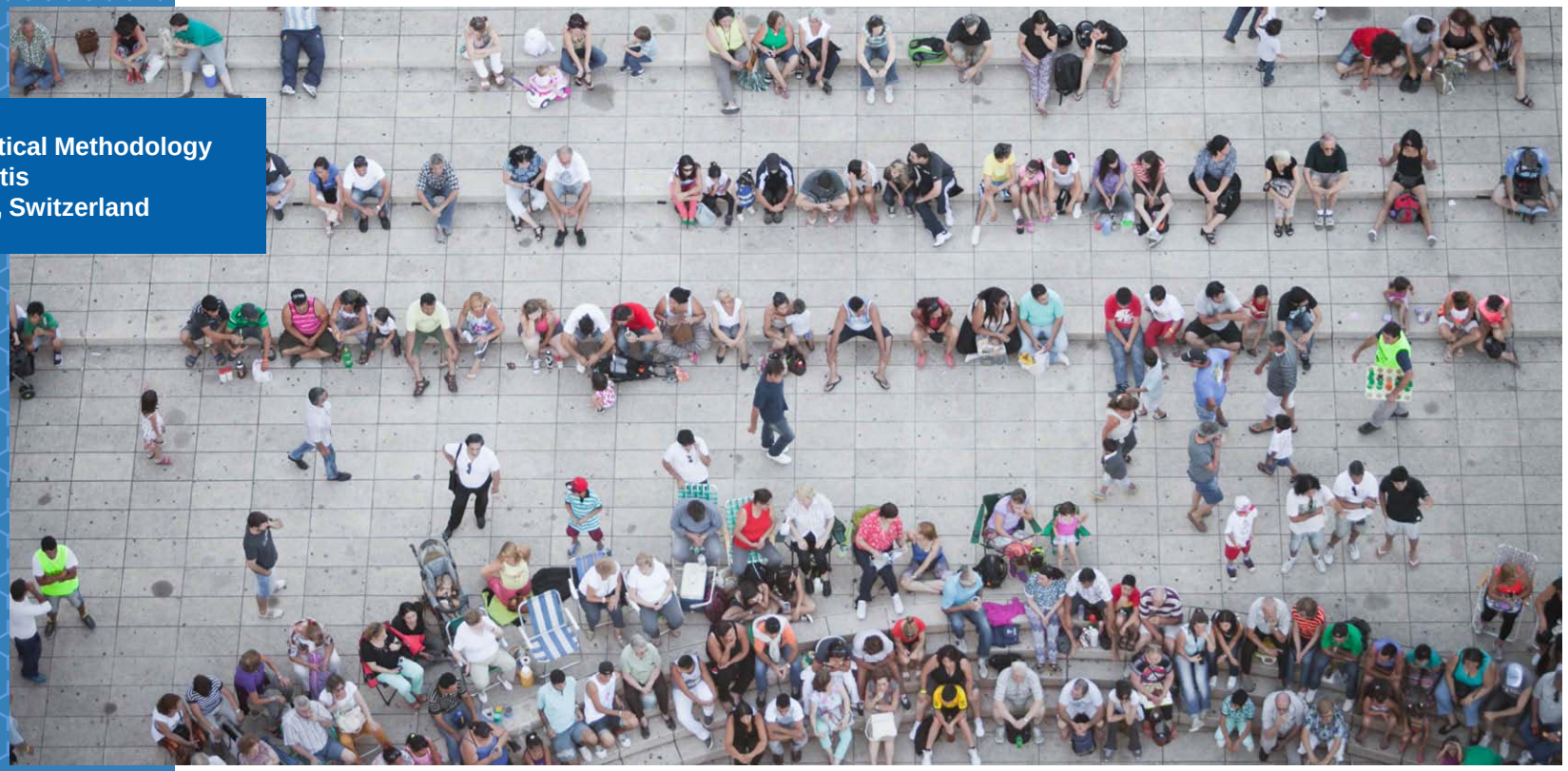
Statistical Methodology
Novartis
Basel, Switzerland



# Sample-size re-estimation in Multiple Sclerosis trials

**Heinz Schmidli**

***PSI Meeting on Sample Size Re-Estimation***
***London, November 2, 2016***

NOVARTIS

# Outline

- Multiple Sclerosis
- Sample size re-estimation
  - **Count endpoints**: number of lesions in the brain
  - **Recurrent event endpoints**: relapses
  - **Time-to-event endpoints**: disease progression
- Conclusions

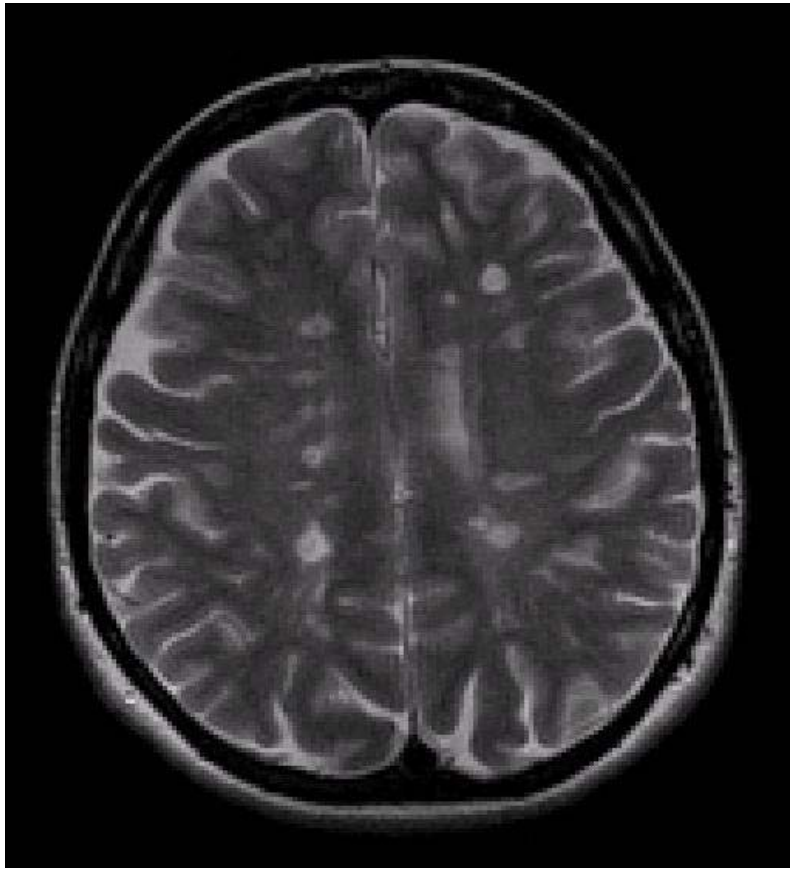
*Acknowledgements*

*Tim Friede, Simon Schneider*
*Medical Statistics, University Medical Center Göttingen, Germany*

*Sandro Gsteiger, Dieter Häring, David Ohlssen, Harald Pohlmann*

# Multiple Sclerosis (MS)

- Progressive, degenerative disease of central nervous system (CNS)
  - most common disorder of the CNS in adults (2.5 million worldwide)
  - affects young adults
- Circulating auto-aggressive lymphocytes cross the blood–brain barrier into the CNS, leading to inflammation and tissue damage
  - Lesions
  - Relapses
  - Disability progression



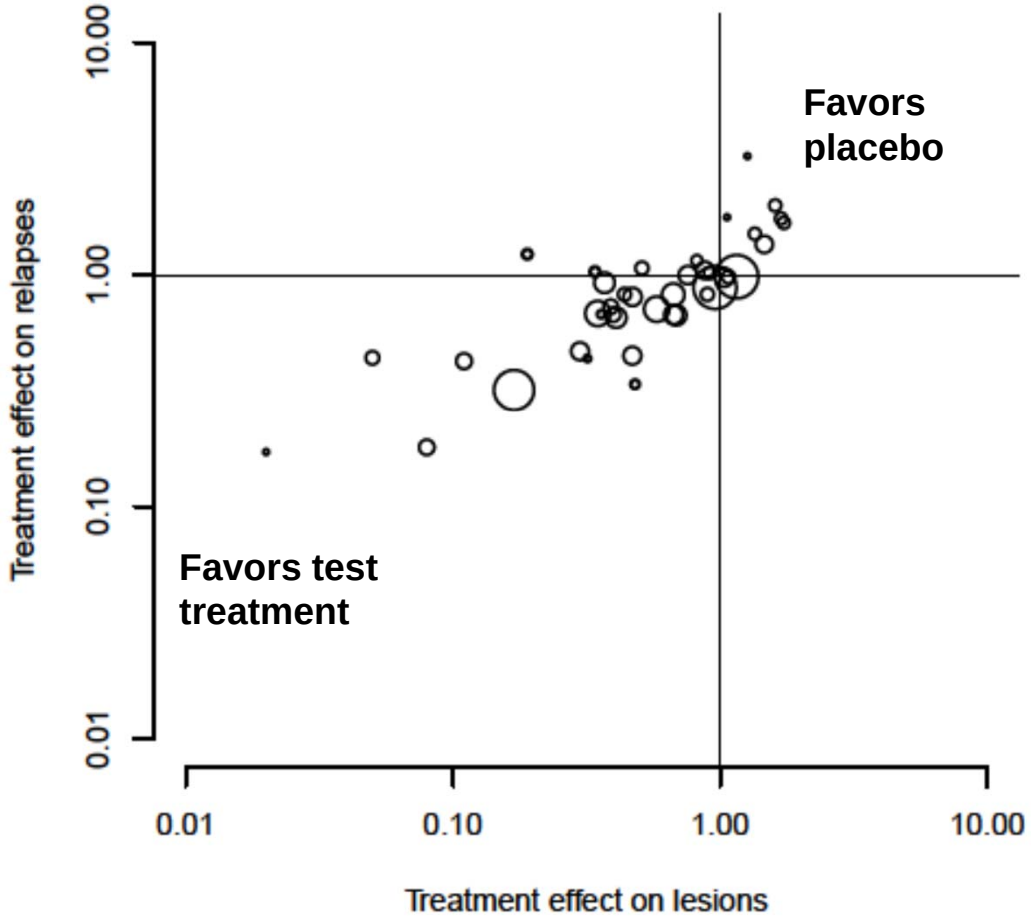
Lesions (MRI)

# Multiple Sclerosis

Treatment effects on

- Lesions
- Relapses

23 placebo-controlled studies (40 arms)




Sormani et al. (2009) *Annals Neurology*, Pozzi et al. (2016) *Pharmaceutical Statistics*

# Multiple Sclerosis

## *Clinical development*

- Phase II (proof-of-concept, dose-finding)
  - Primary endoint: Number of lesions (MRI) ***count endpoint***
  - 3-6 months, N ≈ 50 per group
- Phase III for relapsing forms of MS (RMS)
  - Primary endpoint: Relapses ***recurrent event endpoint***
  - 2 years, N ≈ 400 per group
- Phase III for secondary progressive MS (SPMS)
  - Primary endpoint: Disability progression ***time-to-event endpoint***
  - 2-4 years, N ≈ 800 per group

# Multiple Sclerosis

## *Clinical development*

- Phase II (proof-of-concept, dose-finding)
  - Primary endoint: Number of lesions (MRI) ***count endpoint***
  - 3-6 months, N ≈ 50 per group
- Phase III for relapsing forms of MS (RMS)
  - Primary endpoint: Relapses ***recurrent event endpoint***
  - 2 years, N ≈ 400 per group
- Phase III for secondary progressive MS (SPMS)
  - Primary endpoint: Disability progression ***time-to-event endpoint***
  - 2-4 years, N ≈ 800 per group

# Sample size re-estimation
# Count endpoint – number of lesions

- Number of lesions in the brain: overdispersed counts
- Negative binomial model $Y \sim NB(\mu,\kappa)$

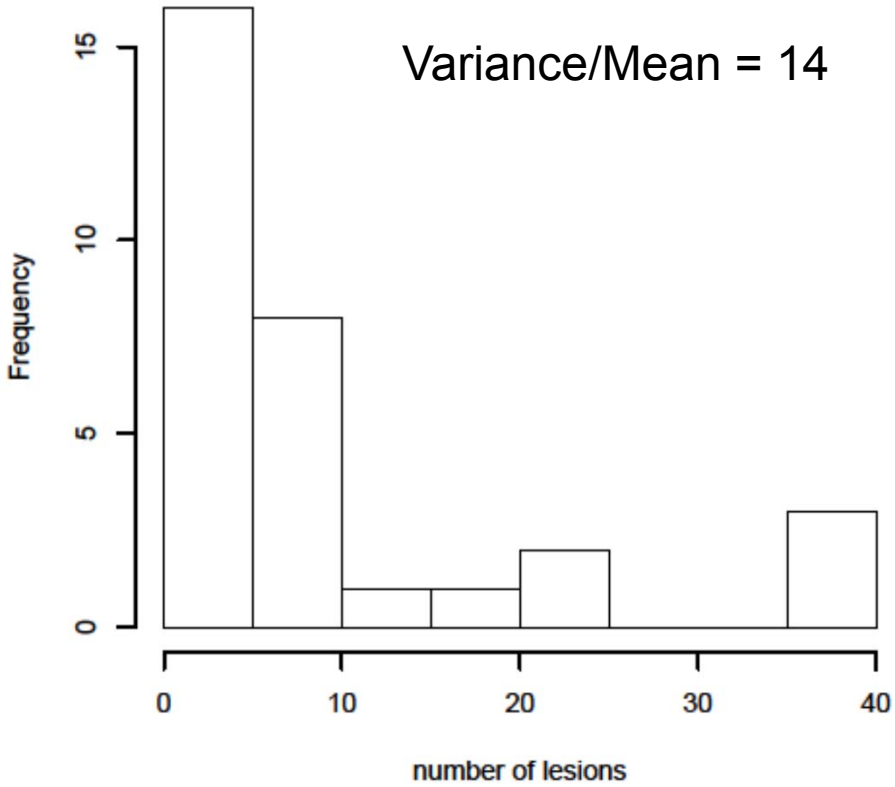
  Mean $\mu$, Dispersion $\kappa$

  $E(Y) = \mu \quad Var(Y) = \mu (1+ \kappa \mu)$

- Interpretation

  $Y \mid \lambda \sim Poisson(\lambda)$

  $\lambda \mid \mu,\kappa \sim Gamma(1/\kappa, 1/\mu\kappa)$




Tubridy et al. (1998) *J Neur Psych*

# Sample size re-estimation
# Count endpoint – number of lesions

- Design of a new study $*$

  Treatment (1) vs Control (0), 1:1 randomization

  Negative Binomial count data (different means, same dispersion)

- Sample size per group

$$N = \{ 2\,\kappa_* + (1+1/\Delta)/\mu_{0*} \} \quad \{ \phi^{-1}(\alpha) + \phi^{-1}(\beta) \}^2 / \log(\Delta)^2$$

  - Significance level (one-sided) , e.g. $\alpha=0.025$
  - Power $1-\beta$ , e.g. 0.8
  - Clinically relevant treatment effect $\Delta = \mu_{1*}/\mu_{0*}$ , e.g. 0.3
  - Dispersion parameter $\kappa_*$
  - Control group mean $\mu_{0*}$

  } Historical information on nuisance parameters

Keene (2007) *Pharm Stats*, Friede and Schmidli (2010) *Methods Inf Med*

# Sample size re-estimation
# Count endpoint – number of lesions

- Historical information
  - J historical studies with same endpoint, similar design
  - Lesion counts in control group from j-th study

    $Y_{ij} \mid \mu_{0j}, \kappa \sim \text{NegBin}(\mu_{0j}, \kappa) \quad j=1,\ldots,J \quad i=1,\ldots,n_j$

    Assuming same dispersion parameter in all studies
  - Control mean in new study $\mu_{0*}$
- Meta-Analytic-Predictive (MAP) approach

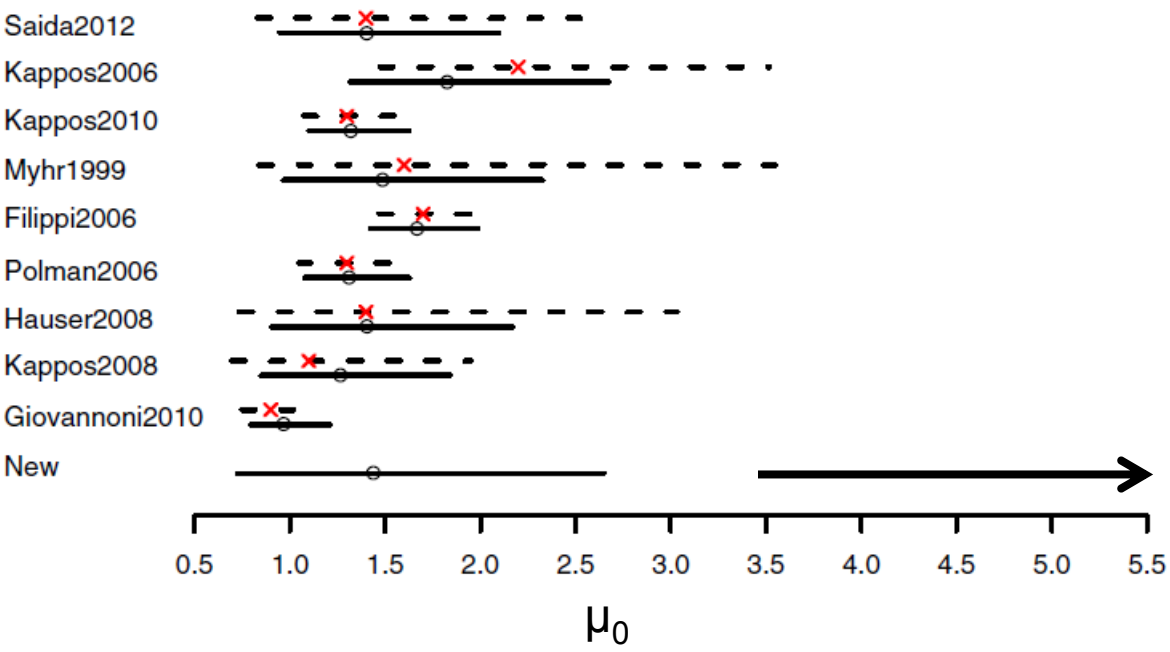
Hierarchical model for transformed parameters $\theta_j = \log(\mu_{0j})$

$$\theta_*, \theta_1, \ldots, \theta_J \mid \mu_0, \tau \sim \text{Normal}(\mu_0, \tau^2)$$

Mean $\mu$ , between-trial standard deviation $\tau$

Gsteiger et al. (2013) *Stats in Medicine*, Schmidli et al. (2014) *Biometrics*

# Sample size re-estimation
# Count endpoint – number of lesions




MAP prior
Median (95% prob interval)
1.38 (0.73, 2.65)

Multiple Sclerosis (relapsing-remitting):
Number of lesions (gadolinium-enhanced), based on MRI scans at 6mo

Dispersion parameter κ: Median (95% prob interval) 3.48 (2.85, 4.26)

# Sample size re-estimation
# Count endpoint – number of lesions

Sample size per group

$$N = \{ 2\,\kappa_* + (1+1/\Delta)/\mu_{0*} \} \quad \{ \phi^{-1}(\alpha) + \phi^{-1}(\beta) \}^2 / \log(\Delta)^2$$

- Significance level (one-sided) , e.g. $\alpha=0.025$
- Power $1-\beta$ , e.g. 0.8
- Clinically relevant ratio $\Delta = \mu_{1*}/\mu_{0*}$ , e.g. 0.3
- Dispersion parameter $\kappa^*$ 3.48 (2.85, 4.26)
- Control group mean $\mu_{0*}$ 1.38 (0.73, 2.65)

- Using point estimates for $\kappa_*=3.48$, $\mu_{0*}=1.38$ : N=55
- Uncertainty, e.g.

  $\kappa_*=3.48$, $\mu_{0*}=0.73$ N=70 $\kappa_*=3.48$, $\mu_{0*}=2.65$ N=47

  $\kappa_*=2.85$, $\mu_{0*}=1.38$ N=48 $\kappa_*=4.26$ $\mu_{0*}=1.38$ N=63

# Sample size re-estimation
# Count endpoint – number of lesions

- ***Blinded sample size re-estimation***
  - Choose initial sample size (typically optimistic), e.g. N=50
  - Blinded interim review (typically towards the end of recruitment)
  - Decide on whether to adapt sample size
- Blinded review
  - Fit negative binomial model to lumped data from both treatment groups
  - Derive means for treatment groups, based on assumed treatment effect
  - Plug-in point estimates in sample size formula
- Controls Type I error
- Typically maintains power at desired level

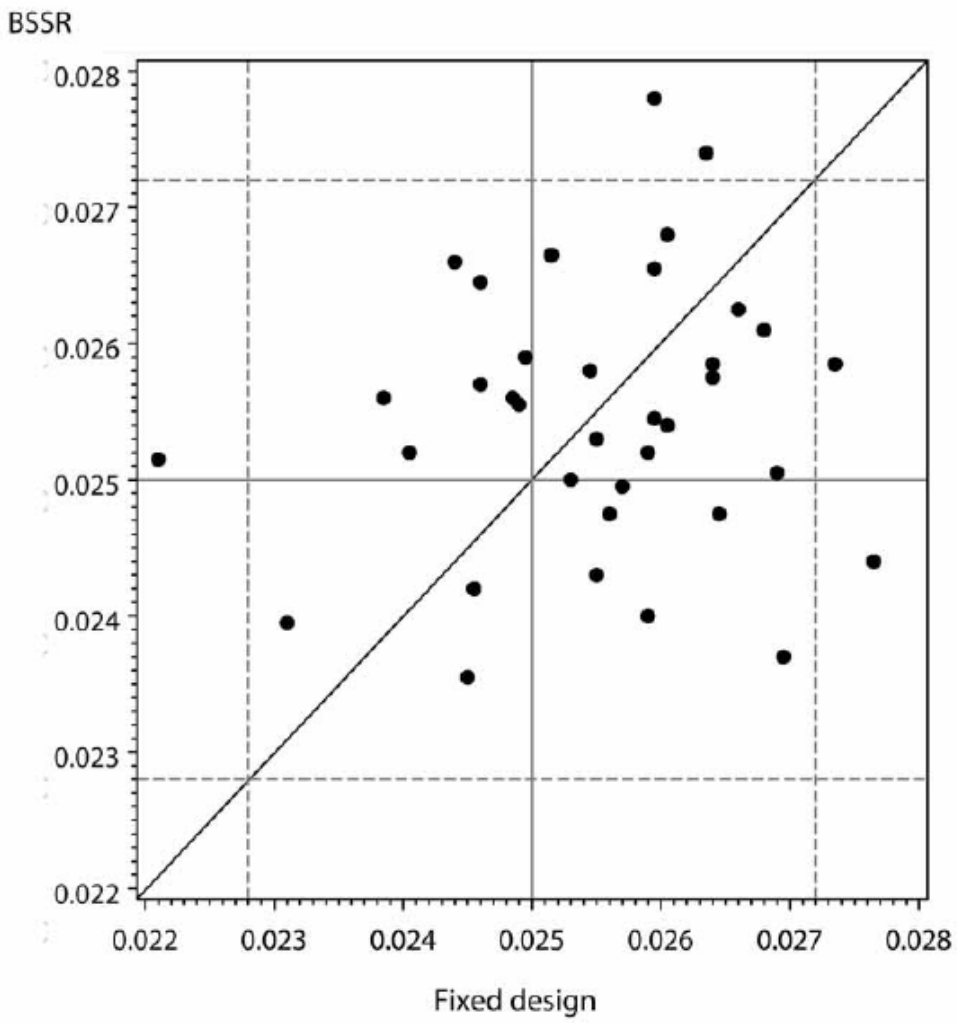
Friede and Schmidli (2010) *Stats in Medicine* Friede and Schmidli (2010) *MIM*
Schneider et al. (2013) *Stats in Medicine* Schneider et al. (2013*) Biometrical J*

# Sample size re-estimation
# Count endpoint – number of lesions

**Type I error**

Various scenarios

blinded sample size re-estimation (BSSR)
vs
Fixed design



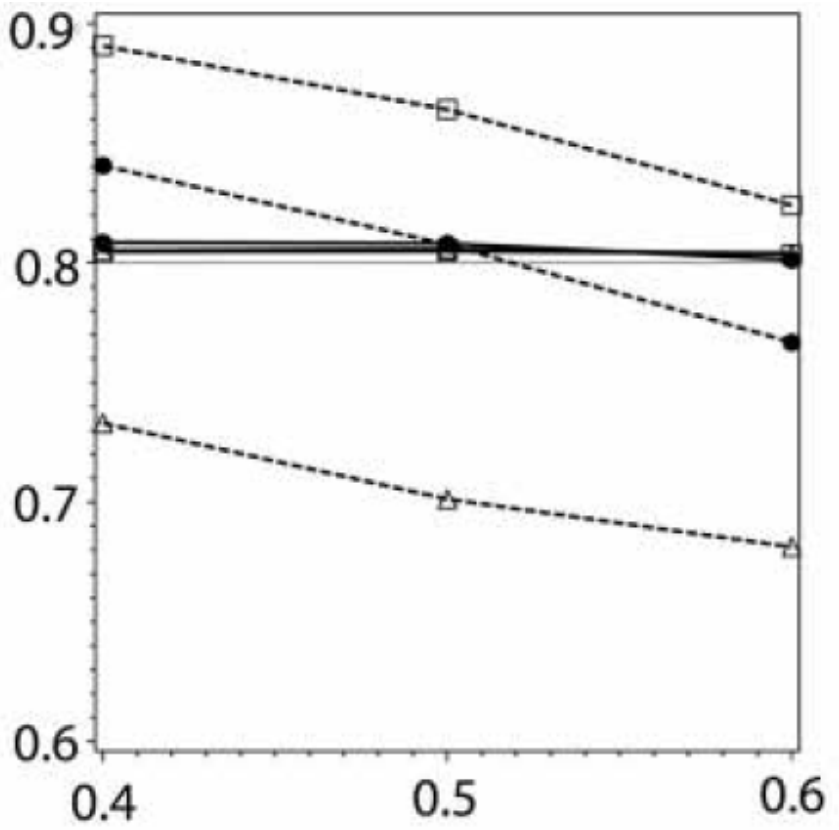
# Sample size re-estimation
# Count endpoint – number of lesions

**Power**

Solid lines:
blinded sample size re-estimation (BSSR)

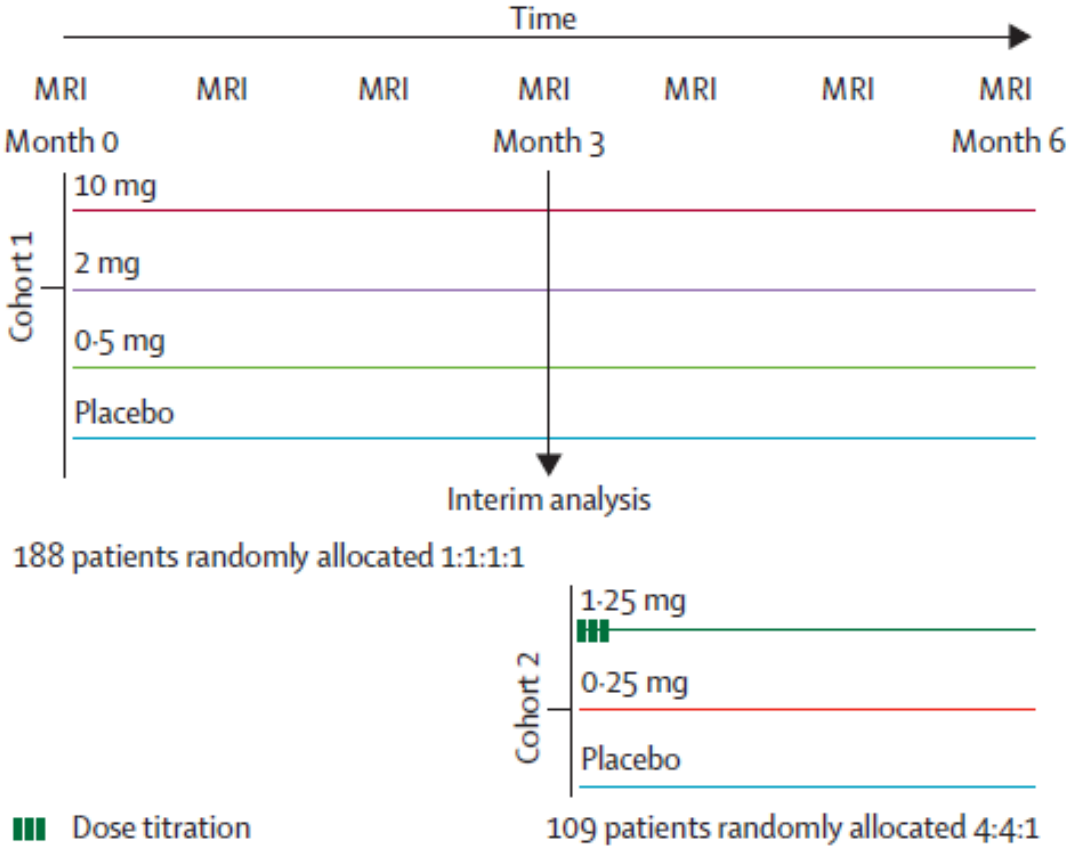
Dashed lines:
Fixed design

Scenarios consist of 3 different true values of:
- control mean
- dispersion parameter



# Sample size re-estimation
# Count endpoint – number of lesions

- ***Unlinded sample size re-estimation***
  - Less control of type I error
  - Discouraged by regulators in phase III
  - Option in settings with complex adaptations




Selmaj et al. (2013)
*Lancet Neurology*

Mercier et al. (2015)
*Pharm Stat*

# Multiple Sclerosis

## *Clinical development*

- Phase II (proof-of-concept, dose-finding)
  - Primary endoint: Number of lesions (MRI) ***count endpoint***
  - 3-6 months, N ≈ 50 per group
- Phase III for relapsing forms of MS (RMS)
  - Primary endpoint: Relapses ***recurrent event endpoint***
  - 2 years, N ≈ 400 per group
- Phase III for secondary progressive MS (SPMS)
  - Primary endpoint: Disability progression ***time-to-event endpoint***
  - 2-4 years, N ≈ 800 per group

# Sample size re-estimation
# Recurrent event endpoint – relapses

- Relapses as over-dispersed recurrent events

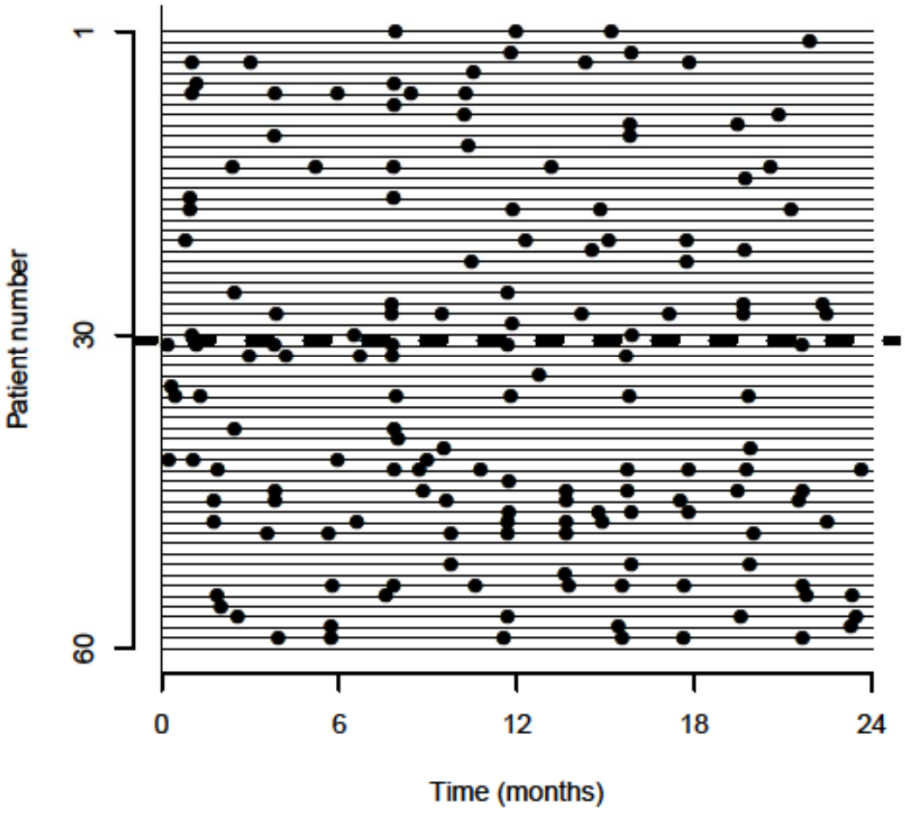
- Negative binomial model
  - Number of relapses within follow-up time T (in years)
  - $Y \sim NB(\mu T, \kappa)$
  - Annualized Relapse Rate $\mu$

- Interpretation

  $Y \mid \lambda \sim Poisson(\lambda T)$

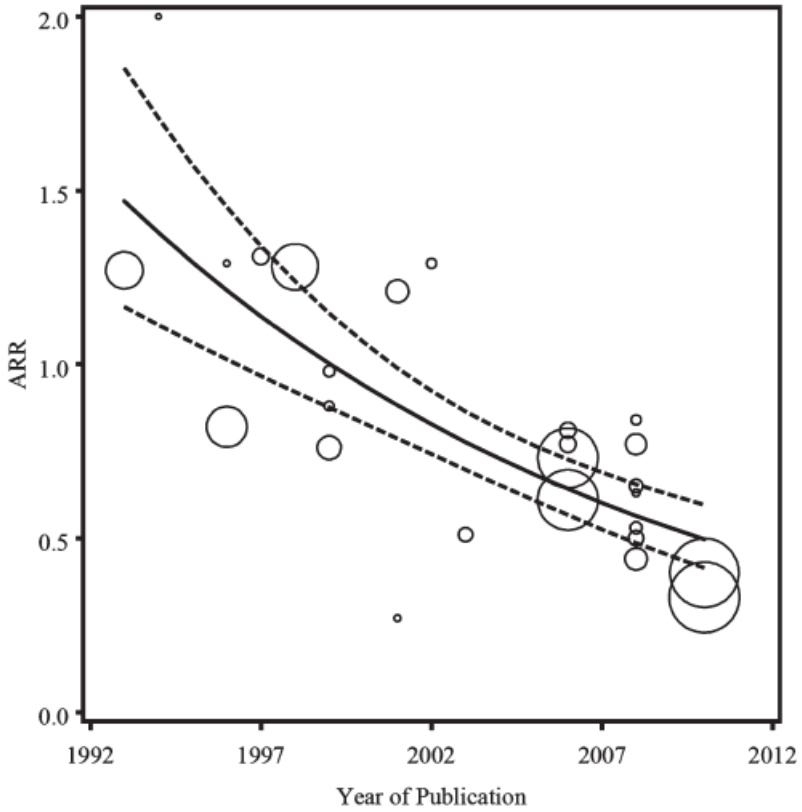
  $\lambda \mid \mu, \kappa \sim Gamma(1/\kappa, 1/\mu\kappa)$

Lycke et al. (1996) *J Neur*
Chen et al. (2013) *Biometrika*



# Sample size re-estimation
# Recurrent event endpoint – relapses

- Same sample size formula as for count data
- Historical information on placebo relapse rates




Annualized relapse rate (ARR)
vs
Year of publication

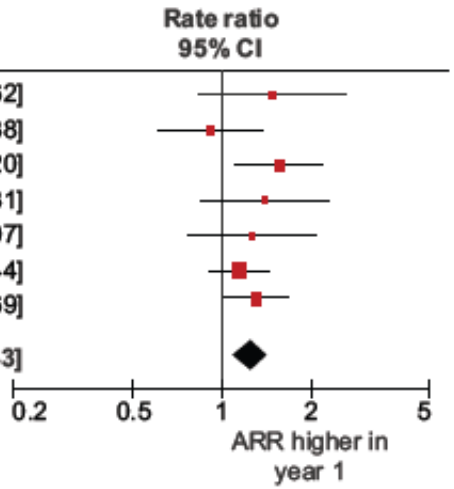
Nicholas et al. (2011)
*Multiple Sclerosis Journal*

# Sample size re-estimation
# Recurrent event endpoint – relapses

- Blinded sample size re-estimation as for count data
- Possible time trends of relapse rates within study (can be taken into account Schneider et al., 2013 *Stats in Medicine* )

| Study or Subgroup | log[Rate ratio] | SE | Year 1 Total | Year 2 Total | Weight | Rate ratio 95% CI |
|---|---|---|---|---|---|---|
| 1991 Goodkin [11] | 0.39272 | 0.29124 | 25 | 25 | 5.0% | 1.48 [0.84, 2.62] |
| 1996 Lycke [12] | -0.08311 | 0.20647 | 30 | 30 | 10.0% | 0.92 [0.61, 1.38] |
| 1997 Fazekas [13] | 0.44869 | 0.17222 | 73 | 63 | 14.4% | 1.57 [1.12, 2.20] |
| 1997 Millefiorini [14] | 0.33647 | 0.25591 | 24 | 24 | 6.5% | 1.40 [0.85, 2.31] |
| 1998 Achiron [15] | 0.23712 | 0.25098 | 20 | 20 | 6.8% | 1.27 [0.78, 2.07] |
| 2006 Polman [19] | 0.1372 | 0.11545 | 315 | 281 | 32.1% | 1.15 [0.91, 1.44] |
| 2010 Giovannoni [21] | 0.27029 | 0.13037 | 437 | 380 | 25.1% | 1.31 [1.01, 1.69] |
| Total (95% CI) | | | 924 | 823 | 100.0% | 1.25 [1.10, 1.43] |

Heterogeneity: Tau² = 0.00; Chi² = 5.14, df = 6 ($p$ = 0.53); $I^2$ = 0%
Test for overall effect: Z = 3.46 ($p$ = 0.0005)




Placebo annualized relapse rate (ARR): Year 2 vs Year 1
Nicholas et al. (2012) *Multiple Sclerosis Journal*

# Multiple Sclerosis

## *Clinical development*

- Phase II (proof-of-concept, dose-finding)
  - Primary endoint: Number of lesions (MRI) ***count endpoint***
  - 3-6 months, N ≈ 50 per group
- Phase III for relapsing forms of MS (RMS)
  - Primary endpoint: Relapses ***recurrent event endpoint***
  - 2 years, N ≈ 400 per group
- Phase III for secondary progressive MS (SPMS)
  - Primary endpoint: Disability progression ***time-to-event endpoint***
  - 2-4 years, N ≈ 800 per group

# Sample size re-estimation
# Time-to-event – disability progression

- Secondary Progressive MS (SPMS)
- Expanded Disability Status Scale (EDSS)
  - Ordinal scale for assessing neurologic impairment
  - 0 (normal) to 10 (death)
- Disability progression

  Increase from baseline of 1 point (in patients with baseline EDSS of 3.0 to 5.0) or 0.5 point (in patients with baseline EDSS of 5.5 to 6.5)
- Time to 3-month confirmed disability progression

  Same or higher EDSS score in the 3 months following a progression
- Cox proportional hazard model for analysis (or log-rank test)

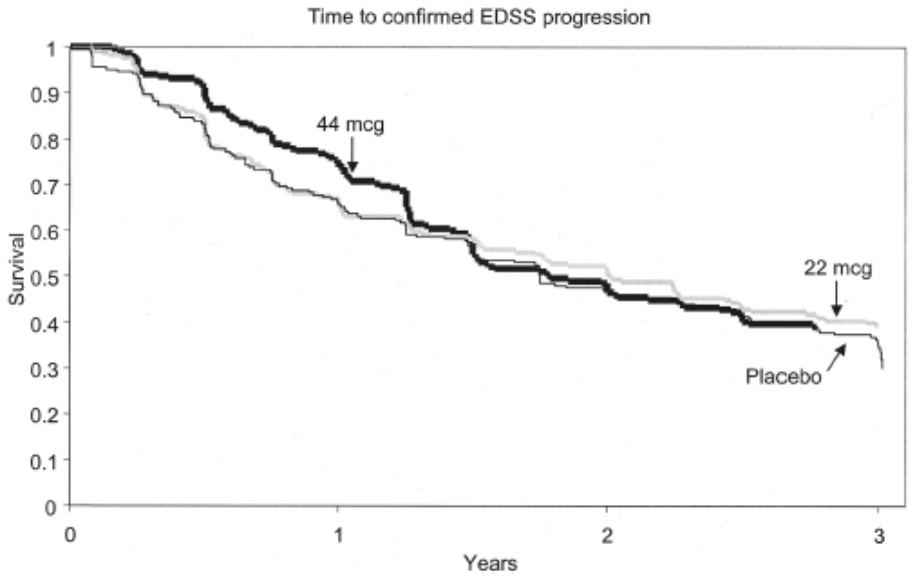
# Sample size re-estimation
# Time-to-event – disability progression

- Phase III EXPAND trial NCT01665144
  - Patients with Secondary Progressive MS (SPMS)
  - Siponimod (BAF312) vs Placebo (2:1)
  - 1651 patients
  - Completed Sep 2016
  - 3-month confirmed progression: 21% risk reduction (p=0.013)
- Design stage: limit study duration to 42 months
  - Uncertainty on placebo event rate
  - Uncertainty on recruitment rate and dropout rate
- Blinded sample size re-estimation prospectively implemented

# Sample size re-estimation
# Time-to-event – disability progression

- Initial design
  - Determine required number of events, based on significance level, power, and HR (Schoenfeld, 1983)
  - Determine sample size such that expected number of events at 42 months is equal to required number of events

    Based on model for time-to-event, recruitment, dropout




SPECTRIMS study group (2001) N*eurology*

# Sample size re-estimation
# Time-to-event – disability progression

- Blinded sample size re-estimation at interim review
  - update model for time-to-event, recruitment, dropout
  - reevaluation of sample size
- Main challenge: extrapolation for time-to-event model
  - Survivor function needed for 42 months (3.5 years)
  - Maximal follow-up at interim review much shorter (< 2 years)
- Options for extrapolation
  - Shift originally assumed survivor function on complementary log-log scale (several possibilities to estimate shift)
  - Parametric models (exponential, Weibull, piece-wise exponential,...)

Whitehead et al. (2001) *Stats in Medicine* Whitehead (2001) *Drug Inf J*
Hade et al. (2010) *Clinical Trials* Todd et al. (2012) *Pharm Stat*

# Conclusions

- Blinded sample size re-estimation
  - Corrects for (some) wrong assumptions at design stage
  - Controls type I error
  - Generally accepted by regulators

    *CHMP Reflection Paper on Adaptive Designs (2007)*

    *Draft FDA guidance on adaptive designs (2010)*
- Methodology developed for various endpoints
  - counts, recurrent events, time-to-event, ...
- Implementation in clinical trials
  - Much easier than for other adaptive designs (e.g. no DMC required)
  - Still needs care

# References


- Chen Q, Zeng D, Ibrahim JG, Akacha M, Schmidli H (2013) Estimating time-varying effects for overdispersed recurrent events data with treatment switching. *Biometrika* 100(2):339-354.
- Friede T, Schmidli H (2010) Blinded sample size reestimation with negative binomial counts in superiority and non-inferiority trials. *Methods Inf Med*. 49(6):618-24.
- Friede T, Schmidli H (2010) Blinded sample size reestimation with count data: methods and applications in multiple sclerosis. *Stat Med.* 29(10):1145-56.
- Gsteiger S, Neuenschwander B, Mercier F, Schmidli H (2013) Using historical control information for the design and analysis of clinical trials with overdispersed count data. *Stat Med.* 2013 Sep 20;32(21):3609-22.
- Keene ON, Jones MR, Lane PW, Anderson J (2007) Analysis of exacerbation rates in asthma and chronic obstructive pulmonary disease: example from the TRISTAN study. *Pharm Stat.* 6(2):89-97.
- Mercier F, Bornkamp B, Ohlssen D, Wallstroem E (2015) Characterization of dose-response for count data using a generalized MCP-Mod approach in an adaptive dose-ranging trial. *Pharm Stat.* 14(4):359-67.
- Nicholas R, Straube S, Schmidli H, Pfeiffer S, Friede T (2012) Time-patterns of annualized relapse rates in randomized placebo-controlled clinical trials in relapsing multiple sclerosis: a systematic review and meta-analysis. *Mult Scler.*18(9):1290-6.
- Nicholas R, Straube S, Schmidli H, Schneider S, Friede T (2011) Trends in annualized relapse rates in relapsing-remitting multiple sclerosis and consequences for clinical trial design. *Mult Scler.* 17(10):1211-7.

# References


- Pozzi L, Schmidli H, Ohlssen DI (2016) A Bayesian hierarchical surrogate outcome model for multiple sclerosis. *Pharm Stat* 15(4):341-8.
- Schmidli H, Gsteiger S, Roychoudhury S, O'Hagan A, Spiegelhalter D, Neuenschwander B (2014) Robust meta-analytic-predictive priors in clinical trials with historical control information. *Biometrics* 70(4):1023-32.
- Schneider S, Schmidli H, Friede T (2013) Blinded sample size re-estimation for recurrent event data with time trends. *Stat Med*. 32(30):5448-57.
- Schneider S, Schmidli H, Friede T (2013) Blinded and unblinded internal pilot study designs for clinical trials with count data. *Biom J*. 55(4):617-33.
- Schneider S, Schmidli H, Friede T (2013) Robustness of methods for blinded sample size re-estimation with overdispersed count data. *Stat Med*. 32(21):3623-35.
- Todd S, Valdés-Márquez E, West J (2012) A practical comparison of blinded methods for sample size reviews in survival data clinical trials. *Pharm Stat.* 11(2):141-8.
- Whitehead J (2001) Predicting the duration of sequential survival studies. *Drug Information Journal* 35: 1387-1400.
- Whitehead J, Whitehead A, Todd S, Bolland K, Sooriyarachchi MR (2001) Mid-trial design reviews for sequential clinical trials. *Stat Med*. 20(2):165-76.