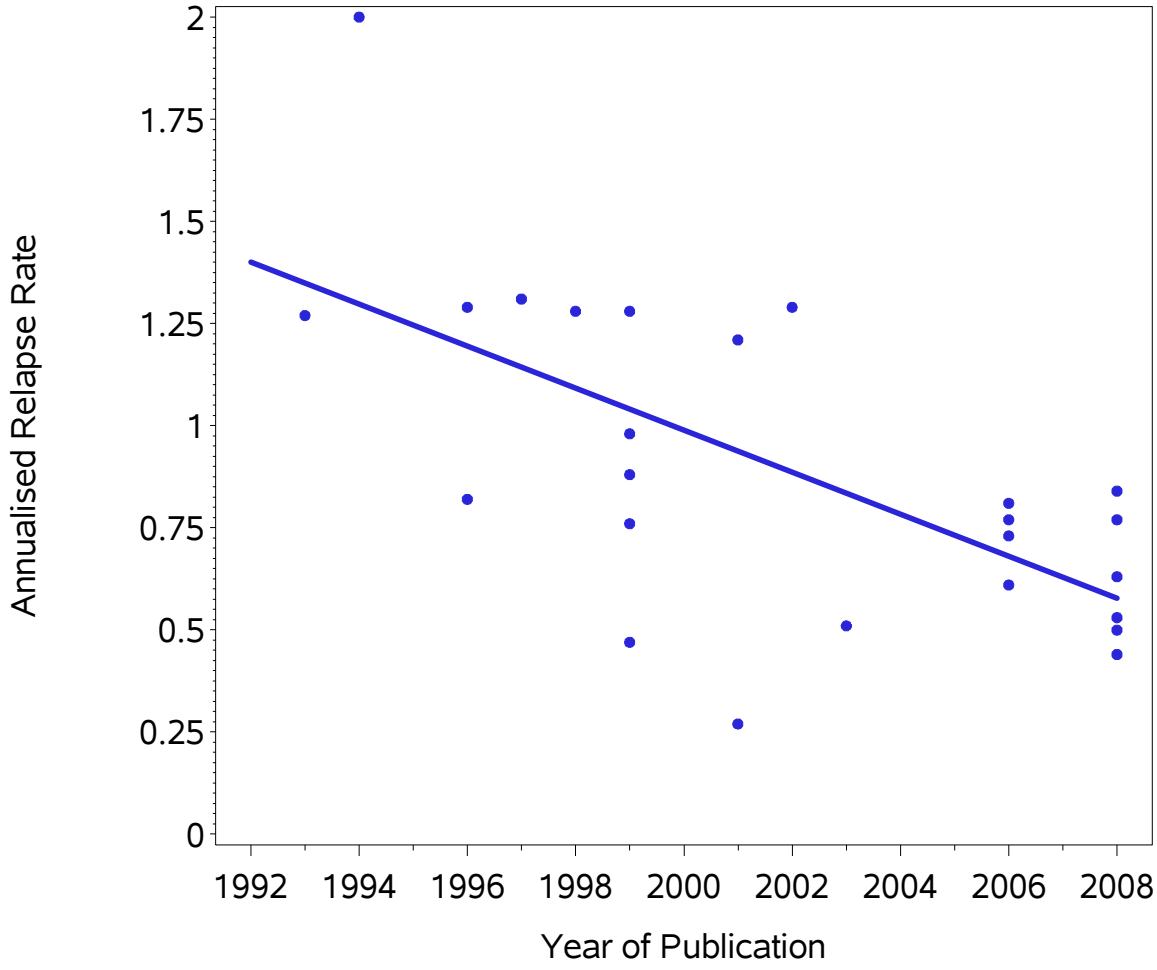
Blinded sample size reestimation with count data

Tim Friede[1] and Heinz Schmidli[2]

[1] Universtiy Medical Center Göttingen, Germany
[2] Novartis Pharma AG, Basel, Switzerland

Motivating Example: Relapse Counts in Multiple Sclerosis



- **placebo ARR in RRMS** (Sormani et al, 2009)
- **trend over time** and **substantial variation**
- recently published placebo-controlled fingolimod trial: placebo ARR 0.7 expected in sample size calculation, observed 0.4

Poisson Counts, Rate Ratio and Wald Test

- **data**: D_T and D_C Poisson counts with . . .
  - event rates λ_T and λ_C
  - follow-up times y_T and y_C
- **hypotheses**: $H_0 : \theta = 1$ versus $H_1 : \theta < 1$ with $\theta = \lambda_T/\lambda_C$
- **hypothesis test**: reject H_0, iff $Z \geq z_{1-\alpha}$ whereby

$$Z = \frac{\log(\hat{\lambda}_T/\hat{\lambda}_C)}{\sqrt{1/D_T + 1/D_C}} \quad \text{with} \quad \hat{\lambda}_i = D_i/y_i$$

Sample Size Calculation for Poisson Counts

- **desired power** $1-\beta$ and **relevant effect size** $\theta^\star$
- assuming same follow-up for all patients and 1:1 randomisation
- **overall event rate** $\bar{\lambda} = (\lambda_C + \lambda_T)/2$
- approximate **sample size** per group (see e.g. Ng & Tang, 2005)

$$n = \frac{1}{\bar{\lambda}} \, \frac{(1+\theta^\star)^2}{2\ \theta^\star} \, \frac{(z_\alpha + z_\beta)^2}{(\log \theta^\star)^2}$$

Internal Pilot Study Design (Wittes & Brittain, 1990)

- **initial sample size estimation** $n_0 = n(\alpha, 1-\beta, \Delta^\star, \widehat{\sigma}_0^2)$
  - significance level α, desired power $1-\beta$, clinically relevant effect $\Delta^\star$
  - initial estimate $\widehat{\sigma}_0^2$ of the nuisance parameter σ^2 (from other studies)
- **sample size review**:
  - after recruitment of $n_1 = \pi n_0$ patients (e.g., $\pi = 1/2$)
  - estimation of nuisance parameter $\rightarrow \widehat{\sigma}^2$
  - sample size *re-estimation* $\widehat{N} = n(\alpha, 1-\beta, \Delta^\star, \widehat{\sigma}^2)$
    - "restricted": $n_2 = \max(n_0, \widehat{N}) - n_1$
    - "unrestricted": $n_2 = \max(n_1, \widehat{N}) - n_1$ (Birkett & Day, 1994)
- **final analysis**
  - estimation of treatment effect and hypothesis test
  - with all $n_1 + n_2$ patients

Sample Size Re-estimation and International Guidelines

- **ICH Guideline E9 (1998)**

  The steps taken to preserve blindness and consequences, if any, for the type I error [. . .] should be explained.

- **CHMP Reflection Paper on Adaptive Designs (2007)**

  Whenever possible, methods for blinded sample size re-assessment [. . .] that properly control the type I error should be used.

- **draft FDA guidance on adaptive designs (2010)**, Sec. V.B

  Sample size adjustment using blinded methods to maintain desired study power should generally be considered for most studies.

Blinded Sample Size Recalculation for Poisson Counts

- overall event rate $\bar{\lambda}$ estimated in blinded fashion from data of IPS

$$\hat{\bar{\lambda}}_{1\cdot} = \frac{D_{1T} + D_{1C}}{y_{1T} + y_{1C}} = \frac{D_{1\cdot}}{y_{1\cdot}}$$

  similar to Gould's approach for binomial data, see Gould (1992), Friede & Kieser (2004)

- assuming constant event rates within treatment groups over the course of the study, all data available can be used

- plugging in the observed overall event rate $\hat{\bar{\lambda}}_{1\cdot}$ for $\bar{\lambda}$ in sample size formula leads to a new sample size estimate

Extension to Overdispersed Poisson Counts

- **overdispersion**:
  - so far $Var(D_i) = \lambda_i y_i$
  - now $Var(D_i) = \sigma^2 \lambda_i y_i$ with overdispersion parameter $\sigma^2 \geq 1$
  - quasi-likelihood approach (McCullagh & Nelder, 1989)
- **test statistic** allowing for overdispersion: $Z^{(O)} = Z/\tilde{\sigma}$
- **sample size**: $n^{(O)} = \sigma^2 n$

Extension to Overdispersed Poisson Counts (cont.)

- various estimators for σ^2 proposed, e.g. method of moments estimator

$$\tilde{\sigma}^2 = \left(\sum_{i=T,C} \sum_{j=1}^{n_i} \frac{(D_{ij} - \hat{\lambda}_i)^2}{\hat{\lambda}_i} \right) \Big/ (n_T + n_C - 2)$$

- for the purpose of a blinded review the dispersion parameter σ^2 estimated in blinded fashion by

$$\tilde{\sigma}^2_{1.} = \left(\sum_{i,j} \frac{(D_{1ij} - \hat{\bar{\lambda}}_{1.y_{1.}})^2}{\hat{\bar{\lambda}}_{1.y_{1.}}} \right) \Big/ (n_T + n_C - 1)$$

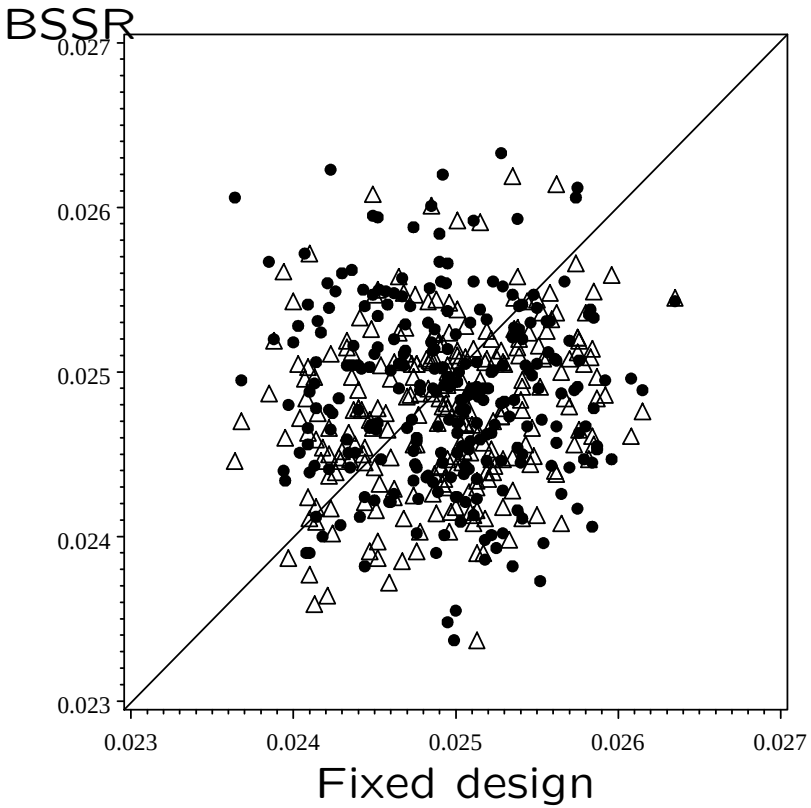
Simulation Study

- **motivated by trials in relapsing MS**
  - phase III: relapse counts (no or small overdispersion)
  - phase II: MRI lesion counts (large overdispersion)
- **number of simulated trials per scenario**: 100,000 for power (type I error rate) and 10,000 for sample size distribution
- **distributional assumptions**: Poisson and Negative Binomial (as an example for overdispersed counts)

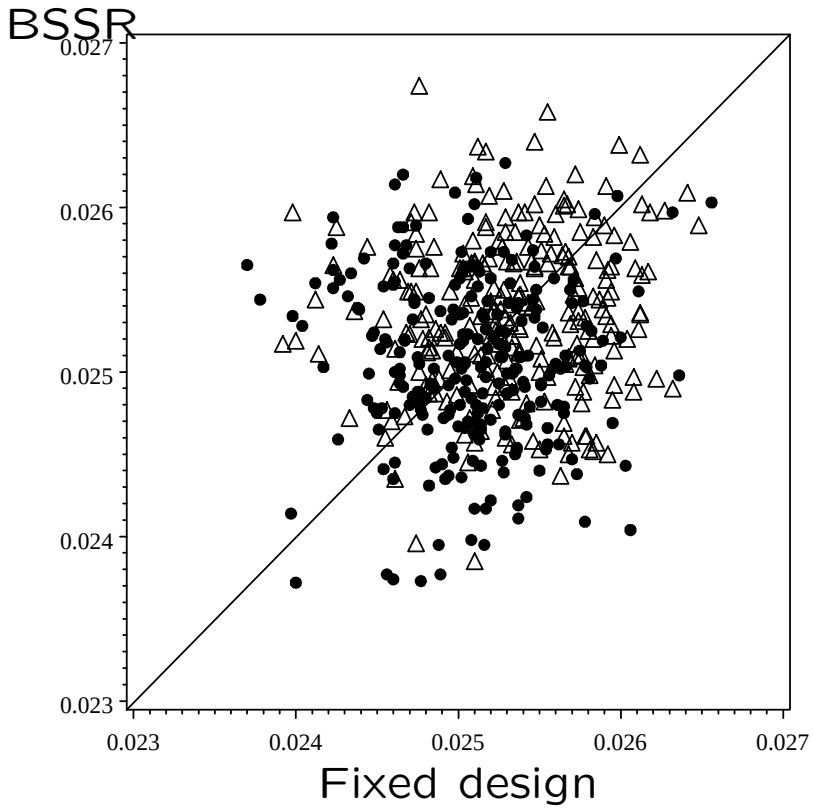
Simulation Study Setup

	Relapses	MRI lesions
Proportion π in IPS	0.2, 0.5	0.4
Overall event rate $\bar{\lambda}$	0.5, 0.51, . . . , 1	5, 6, . . . , 15
Dispersion parameter σ^2	1, 1.1, . . . , 2	20, 25, . . . , 40
Under the null hypothesis $H_0 : \theta = 0$		
Required sample size N	100, 200, . . . , 500	50, 75, . . . , 150
Under the alternative $H_1 : \theta = \theta_a$		
Assumed overall event rate $\bar{\lambda}_a$	0.75	10
Assumed dispersion σ_a^2	1	30
Assumed rate ratio θ_a	0.6, 0.75	0.5
Target power $1 - \beta$	0.90	0.80

Relapses: Type I Error Rate

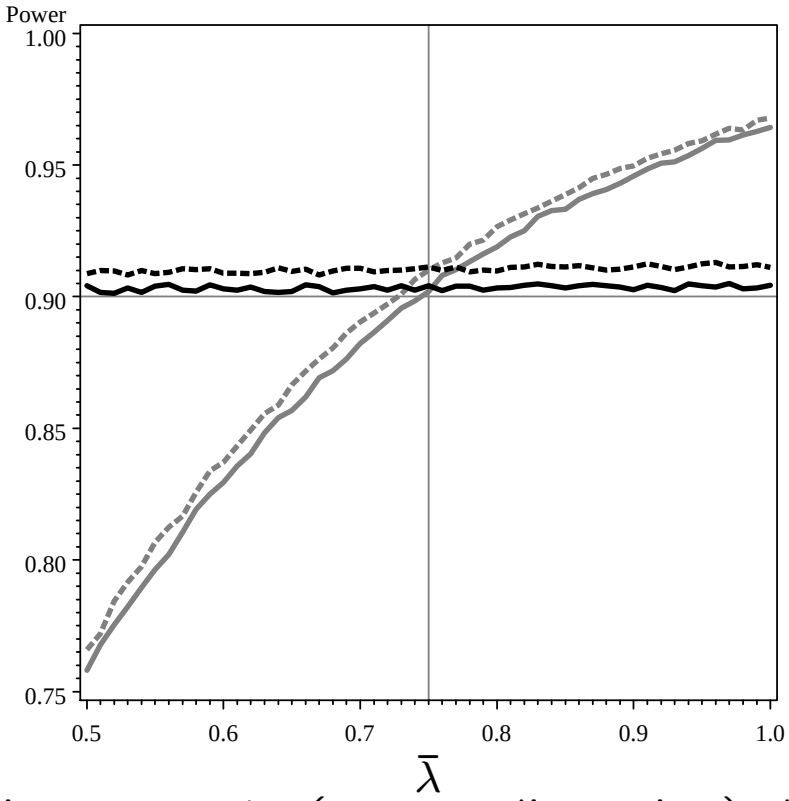


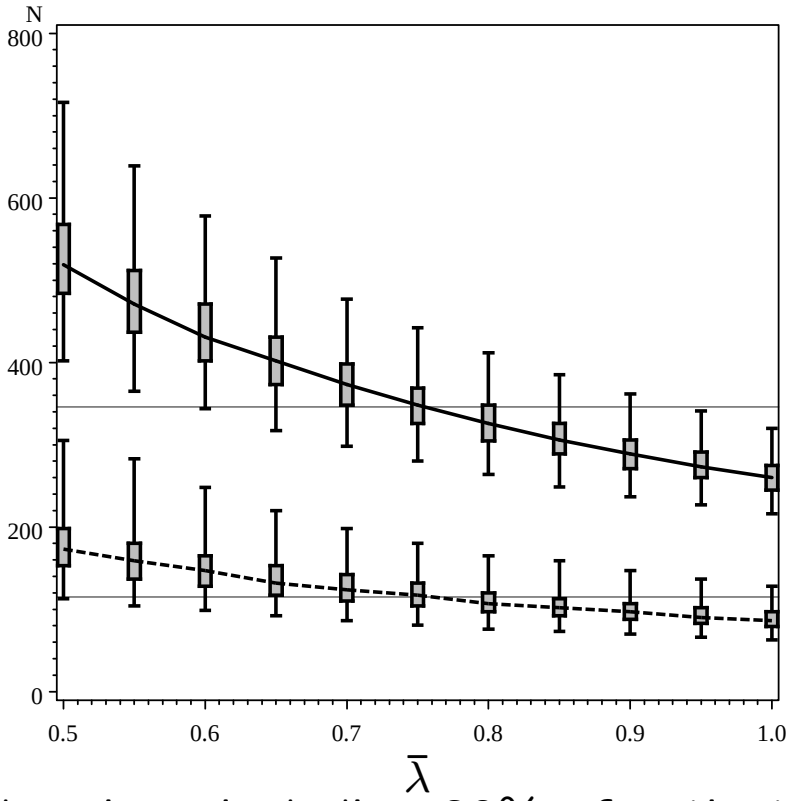
Poisson data with reviews including 20% (•) and 50% (△) of the patients



Reviews with 50% of the patients; $\sigma^2 = 1$ (•) and $\sigma^2 = 2$ (△)

Relapses: Misspecification of the Overall Event Rate

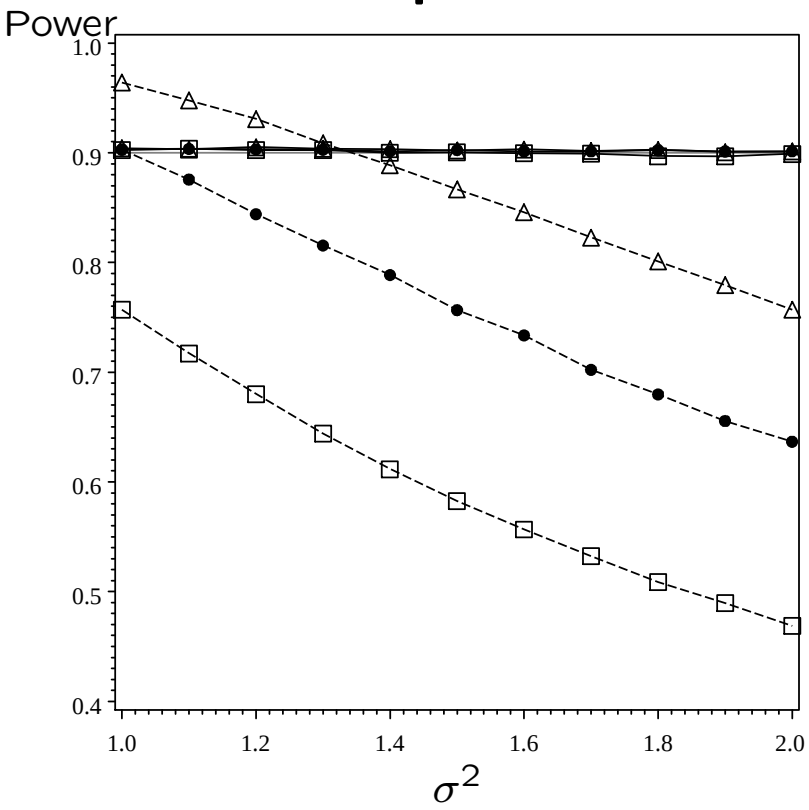




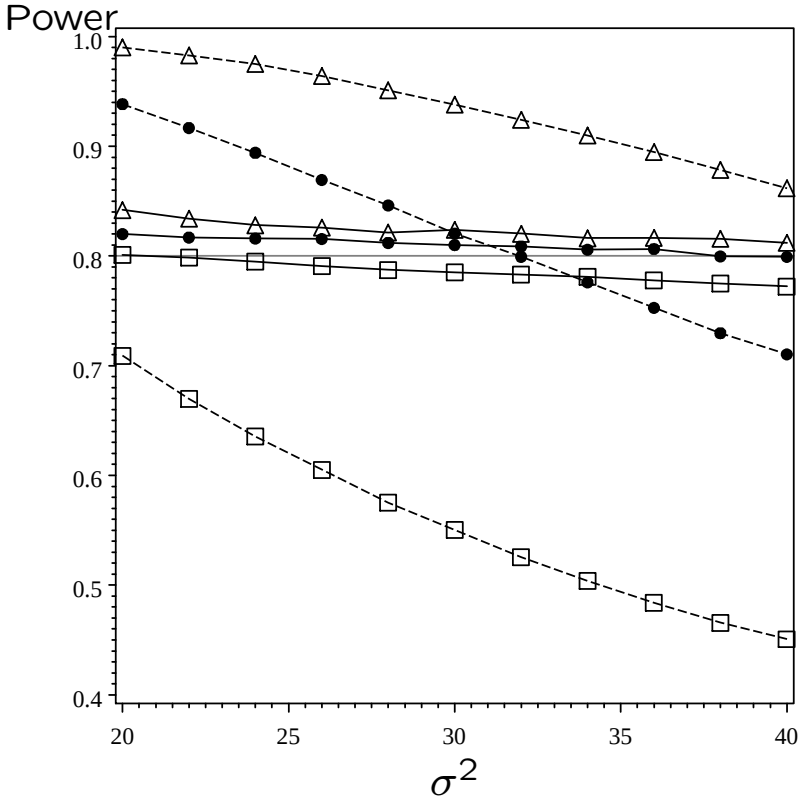
Poisson counts (no overdispersion); blinded reviews including 20% of patients in the internal pilot; rate ratios $\theta = 0.6$ (dashed) and $\theta = 0.75$ (solid); fixed design (grey) for comparison.

Misspecification of Overall Event Rate and Overdispersion

Relapses



MRI Lesion Counts



Likelihood Approach: Negative Binomial Distribution

- **likelihood-based inference** assuming a particular mixture distribution, e.g. negative binomial (Aban et al, 2009)

- Cook et al (2009) proposed a blinded procedure based on an **EM algorithm**
  - computationally more demanding than our approach outlined above
  - on the other hand, our approach does not utilize all information under sampling from negative binomial distributions

- **new approach**: assuming negative binomial distributions ML estimates of the overall event rate and of the shape parameter are derived in blinded review ignoring treatment groups

Non-inferiority Trials

- other indications: for instance exacerbation counts in asthma and COPD (Keene et al 2007, 2008)

- in asthma / COPD standard treament exist and placebo therefore (at least long-term) unethical

- active controlled non-inferiority trials

Non-inferiority: Type I Error Rates and Power

Simulation study motivated by trials in COPD with average study sizes ranging from 700 to 1300 patients

Scenario		Type I	Power
ϕ	λ	error rate	
0.4	1	0.0253	0.7985
	1.5	0.0255	0.7984
	2	0.0269	0.8054
0.5	1	0.0227	0.8073
	1.5	0.0241	0.7984
	2	0.0243	0.8010
0.6	1	0.0235	0.8015
	1.5	0.0259	0.7947
	2	0.0259	0.7991

Conclusions

- type I error rate: similar to fixed design tests
- power robust against misspecifations of both overall event rate and overdispersion parameter
- blinded reviews fulfill regulatory requirements

References

- Aban IB, Cutter GR, Mavinga N. Inferences and power analysis concerning two negative binomial distributions with an application to MRI lesion counts data. *Computational Statistics & Data Analysis* 2009; **53**: 820–833.

- Cook RJ, Bergeron PJ, Boher JM, Lie Y. Two-stage design of clinical trials involving recurrent events. Statistics in Medicine 2009; 28; 2617-2638.

- Friede T, Kieser M. Sample size recalculation for binary data in internal pilot study designs. *Pharmaceutical Statistics* 2004; **3**: 269–279.

- Gould AL. Interim analyses for monitoring clinical trials that do not materially affect the type I error rate. *Statistics in Medicine* 1992; **11**: 55–66.

- Keene ON, Calvery PMA, Jones PW, Vestbo J, Anderson JA. Statistical analysis of exacerbation rates in COPD: TRISTAN and ISOLDE revisited. European Respiratory Journal 2008; 32: 17-24.

- Keene ON, Jones MRK, Lane PW, Anderson J. Analysis of exacerbation rates in asthma and chronic obstructive pulmonary disease: example from the TRISTAN study. Pharmaceutical Statistics 2007; 6: 89-97.

- McCullagh P, Nelder JA. *Generalized Linear Models* (2nd edn). Chapman & Hall, 1989.

- Ng HKT, Tang M-L. Testing the equality of two Poisson means using the rate ratio. *Statistics in Medicine* 2005; **24**: 955–965.

- Wittes J, Brittain E. The role of internal pilot studies in increasing the efficacy of clinical trials. Statistics in Medicine 1990; **9**: 65–72.