**ONLINE DEPOSITORY**

# OBJECTIVE MEASUREMENT OF COUGH IN OTHERWISE HEALTHY VOLUNTEERS WITH ACUTE COUGH

Kanchan Sunger[1], William Powley[2], Angela Kelsall[3], Helen Sumner[1], Robert Murdoch[2], Jaclyn A. Smith[1]

[1]Respiratory Research Group, The University of Manchester, Manchester, UK

[2]GlaxoSmithKline, Stevenage, UK

## METHODS

### Calculations of Sample Sizes and Power for Future Study Design

The data were used to estimate variance parameters corresponding to parallel group (PG) and cross over designs (XO). Two Generalised Linear Mixed Models were fitted to the total daytime cough count using PROC GLIMMIX (PC SAS v9.2). Each model had a fixed effect for recording session, used length of daytime recording as a log offset term and each model assumed the responses followed a Negative Binomial distribution (log link function). The cross over design model had an additional random subject effect term fitted on the linear predictor. The corresponding variance parameter estimates were used to obtain the standard error of a treatment effect under simple future PG and XO study designs. This standard error was used to compute the power of detecting a 50% reduction in cough rate on active relative to the cough rate on placebo (2-sided test, 5% alpha). Since the standard error of the treatment effect (and hence power) also depends upon the placebo response, several power curves are shown that cover a range of plausible future placebo response rates.

**Table E1** Variance parameter estimates appropriate to future Parallel Group and Cross Over study designs

| **Future study type** | **Subject Allocation** | **Estimate for Variance Parameter** | |
|---|---|---|---|
| | | **k (over-dispersion)** | **Between Subject Variability** $\sigma^2_{\text{between}}$ |
| **Parallel Group (2 Arm)** | Equal allocation to Active (A) and Placebo (P) arms | 0.6809 | n/a |
| **Cross Over (2x2)** | Equal allocation to sequences<br>Seq. #1: *A P*<br>Seq. #2: *P A* | 0.1147 | 0.6993 |

Let...

$\sigma^2_{\text{between}}$ = Pure between subject variance component

$k$ = Overdispersion parameter for Negative Binomial distribution (SAS notation)

$N$ be the number of observations in the dataset from the future study

*Nsub* be the number of subjects in the future study

**X** be a $N$ x 2 design matrix for fixed effects

**β** be a 2 x 1 vector of model parameters ($\beta_0$ = Intercept (Placebo response), $\beta_1$ = Treatment effect)

Under $H_0$ : $\boldsymbol{\beta} = \begin{bmatrix} \ln(\text{Mean Pbo cough rate}) \\ 0 \end{bmatrix}$ , Under $H_1$ : $\boldsymbol{\beta} = \begin{bmatrix} \ln(\text{Mean Pbo cough rate}) \\ \ln((100-\omega)/100) \end{bmatrix}$

$\omega$ = Treatment effect under $H_1$ (expressed as a % reduction relative to Pbo response, e.g $\omega = 50\%$)

**Z** be a $N$ x *NSub* design matrix for random subject effects

**γ** be a *Nsub* x 1 vector of subject effects $\boldsymbol{\gamma} \sim N(\mathbf{0}, \mathbf{G})$

**R** be the $N$ x $N$ Identity matrix

**G** = *Nsub* x *Nsub* matrix with diagonal elements equal to $\sigma^2_{\text{between}}$

$\boldsymbol{\eta} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\boldsymbol{\gamma}$ = The linear predictor

$\boldsymbol{\mu} = \exp(\boldsymbol{\eta})$

$\mathbf{A} = \text{diag}(\mu + k\mu^2)$

$\boldsymbol{\Delta} = \text{diag}(\mu)$

$L$ = [0,1] (the 1 x 2 vector that computes the treatment effect from the estimated model parameters)

$\alpha = 0.05$ (Type I error rate : 5%), $\Phi^{-1}$ = Quantile function of a standard N(0,1) distribution

Then for the Cross over design (Note assumes no period effect in model)...

$$\text{var}(\theta) = \left(\mathbf{ZGZ^T}\right) + \left(\boldsymbol{\Delta}^{-1}\mathbf{A}^{1/2}\mathbf{R}\mathbf{A}^{1/2}\boldsymbol{\Delta}^{-1}\right)$$

$$s.e.(\text{Treatment effect}) = s.e.(\beta_1) = \sqrt{\mathbf{L}\mathbf{X}^{\mathbf{T}}\,\text{var}(\theta)\mathbf{X}\mathbf{L}^{\mathbf{T}}}$$

Note : Compute the value of $s.e.(\beta_1)$ under $H_0$ and $H_1$

$$\text{Power}\,(\%) = 100 * \Phi^{-1}\left(\frac{\left[-\,probit(1-\alpha/2) * s.e.(\beta_1)_{H_0}\right] - \ln((100-\omega)/100)}{s.e.(\beta_1)_{H_1}}\right)$$

For the Parallel group design the same formulae are used except

$\boldsymbol{\eta} = \mathbf{X}\boldsymbol{\beta}$ and $\text{var}(\theta) = \left(\boldsymbol{\Delta}^{-1}\mathbf{A}^{1/2}\mathbf{R}\mathbf{A}^{1/2}\boldsymbol{\Delta}^{-1}\right)$

**Discussion about Statistical Analysis Models for cough count data**

By their nature coughs are discrete counts, not continuous responses (e.g. one would not observe 2.5 coughs in an individual). Historically, a typical statistical analyses of cough count data involved log transformation of the observed cough counts (or cough rates); with subsequent statistical analyses assuming that these transformed responses are continuous with an approximately normal distribution. A general linear model would then be used to model the data. Advantages of this approach are its simplicity (less statistical expertise required to implement it), and the widespread availability of statistical analysis packages and routines to implement general linear models. Disadvantages of this approach occur with zero cough counts, because the log transformation is undefined, and the choice of constant to add onto each value prior to transformation can have a strong influence on the analysis results.

Today however, many commercially available statistical analysis packages contain routines to perform generalised linear modelling. This allows a more appropriate distribution to be used to model the cough count data (e.g. the Poisson, or Negative Binomial distributions). The advantage of generalised linear modelling is that it reflects the nature of the response variable and can model zero responses. Given the dearth of new chemical entities in the cough arena, it would be good to see the establishment of generalised linear models as the de-facto statistical analysis technique for modelling cough count data.

A large obstacle to the wider use of generalised linear modelling is the additional level of statistical expertise required, since basic statistical training courses may not cover them in great detail. A certain amount of pragmatism is also required, because when cough counts are large, general linear modelling and generalised linear modelling should give similar conclusions. For example, if developing a new chemical entity with a long duration of action the epoch of interest may be a 24 hour period and either modelling approach should suffice. However, if the duration of action of a new chemical entity is expected to be short the epoch

of interest may only be one hour, in which case there is a large possibility of observing zero coughs, so the use of generalised linear models would be preferable.

In this paper, the power calculations have been performed assuming generalised linear modelling will be used in the subsequent studies.

**RESULTS**

**Table E2:** Correlations between objective and subjective measures of cough at each 24HR study session, Spearman's correlation coefficients

| **SESSION 1** | | **SESSION 2** | |
|---|---|---|---|
| | Day 1 Objective Cough Frequency | | Day 2 Objective Cough Frequency |
| Day 1 VAS frequency | r=0.46 p<0.001 | Day 2 VAS frequency | r=0.59 p<0.001 |
| Day 1 VAS severity | r=0.28 p=0.03 | Day 2 VAS severity | r=0.42 p=0.001 |
| | | | |
| | Night 1 Objective Cough Frequency | | Night 2 Objective Cough Frequency |
| Night 1 VAS frequency | r=0.43 p=0.001 | Night 2 VAS frequency | r=0.47 p<0.001 |
| Night 1 VAS severity | r=0.50 p<0.001 | Night 2 VAS severity | r=0.43 p=0.001 |
| Night 1 VAS sleep latency | r=0.29 p=0.03 | Night 2 VAS sleep latency | r=0.49 p<0.001 |

**Figure E1:** Bland Altman plot of 24hr cough rates. Solid line shows median difference in cough rates between sessions 1 and 2, with the interquartile range represented by the dotted lines.

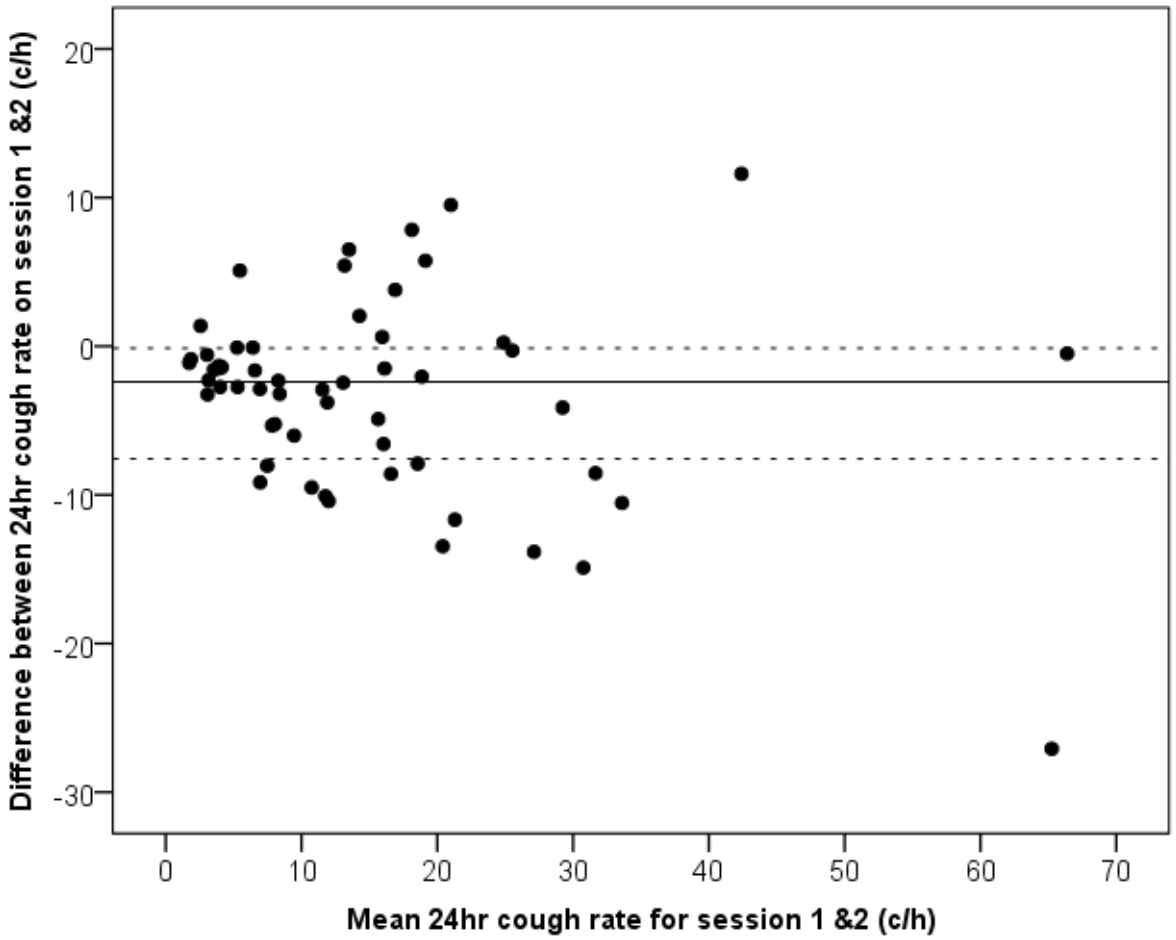

**Figure E2:** Bland Altman plot of daytime VAS measures of cough frequency. Solid line shows median difference in cough rates between sessions 1 and 2, with the interquartile range represented by the dotted lines.

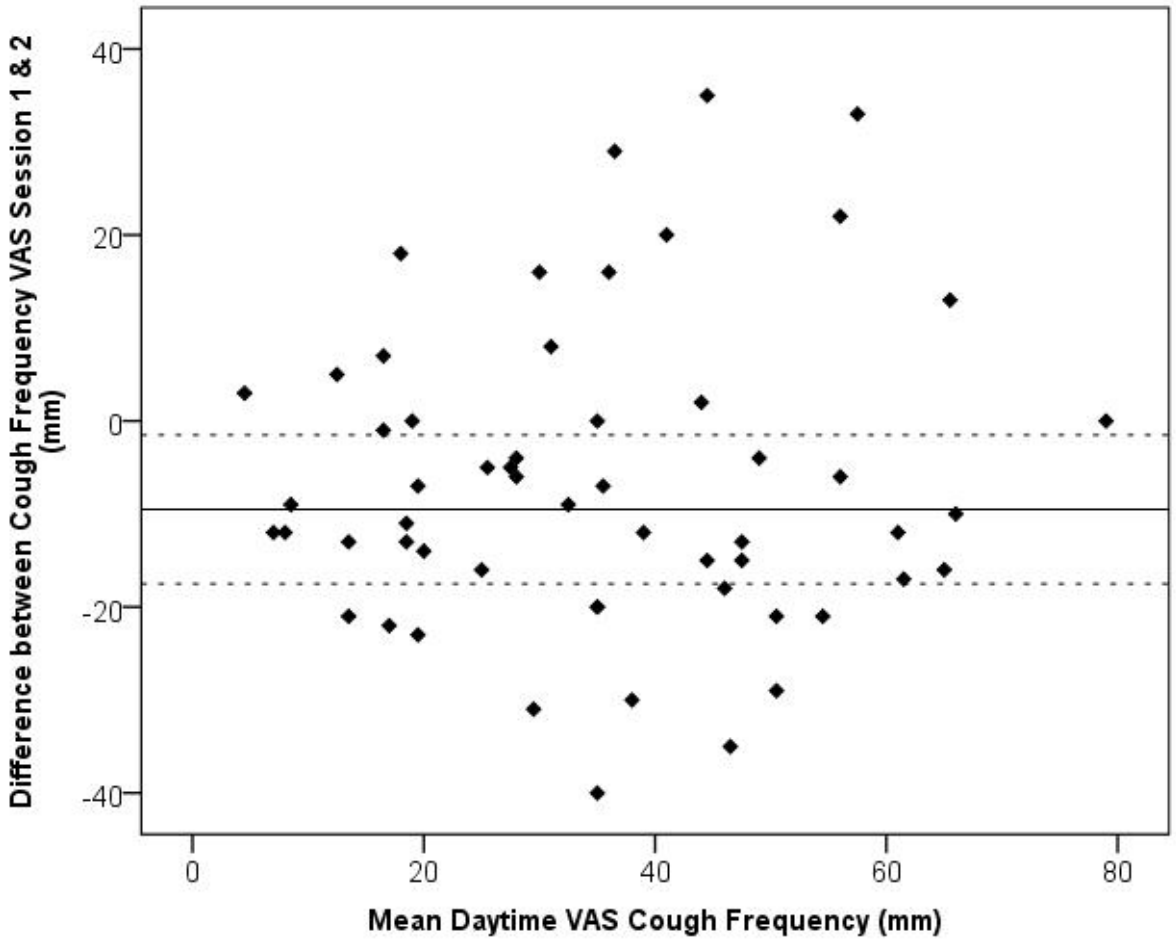

**Figure E3:** Bland Altman plot of night VAS measures of cough frequency. Solid line shows median difference in cough rates between sessions 1 and 2, with the interquartile range represented by the dotted lines.