



## PERSPECTIVE

# Statistical analysis of exacerbation rates in COPD: TRISTAN and ISOLDE revisited

O.N. Keene\*, P.M.A. Calverley<sup>#</sup>, P.W. Jones<sup>†</sup>, J. Vestbo<sup>+,§</sup> and J.A. Anderson\*

**ABSTRACT:** Various statistical methods have been used to measure the impact of treatment on chronic obstructive pulmonary disease (COPD) exacerbations. Poisson regression has recently been recommended as the appropriate method but the model does not satisfactorily account for variability between patients. In contrast, use of a negative binomial model, which corresponds to assuming a separate Poisson parameter for each patient, offers a more appealing approach.

The present paper reviews analysis methods, with particular focus on the negative binomial model. To illustrate the differences that arise from using different analysis methods, we have reanalysed data from two large studies which, among other objectives, investigated the effectiveness of inhaled corticosteroids in reducing COPD exacerbation rates.

Using the negative binomial model to reanalyse data from the TRISTAN and ISOLDE studies, the overall estimates of exacerbation rates on each treatment arm are higher and the confidence intervals for comparisons between treatments are wider, but the overall conclusions of TRISTAN and ISOLDE regarding reduction of exacerbations remain unchanged.

The negative binomial approach appears to provide a better fit to the distribution of the data than earlier methods and is currently the method of choice. Research needs to continue on further methods to improve the analysis of exacerbation data.

**KEYWORDS:** Exacerbations, negative binomial model, Poisson, statistical analysis

Exacerbations of chronic obstructive pulmonary disease (COPD) result in a significant burden on the patient. Frequent exacerbations are associated with a more rapid decline in lung function [1, 2], a significant impact on patients' quality of life [3, 4] and increased mortality [5]. Therefore, decreasing the number of exacerbations is a key treatment goal for COPD management [6].

A number of different approaches to defining an exacerbation have been developed, ranging from daily recording of symptoms on a self-completed diary card [7, 8] and self-reported symptoms at a clinic visit [9], to episodes defined as an increase in symptoms that leads to a change in therapy [6, 10]. The latter approach has been used in many recent clinical trials, as it is relatively unambiguous and objective. However, there remains a need to improve the precision and reliability of this measurement.

Whatever definition is used, it is clear that exacerbations are discrete episodes with symptoms that peak and then resolve [7], although the

duration of episodes requiring medical treatment is not always clear. Moreover, some patients exacerbate relatively frequently, some do not report these episodes even over several years of observation and others may withdraw from a study before they have experienced an exacerbation [11, 12]. Therefore, providing an accurate summary of exacerbation frequency is much harder than reporting continuous variables like lung function or symptom scores, or discrete outcomes like death.

SUISSA [13] has discussed the important issue of the statistical analysis of exacerbation rates from COPD clinical trials and has proposed the use of a specific model, the Poisson. Many different methods have been used and statistical techniques continue to be developed. This is an important issue because different statistical methodologies can produce different estimates of the effectiveness of treatment. The present paper reviews the different approaches to the analysis of exacerbations, providing an explanation of the issues involved and a discussion of the merits of each method. The paper also presents a reanalysis

### AFFILIATIONS

\*GlaxoSmithKline Research and Development, Stockley Park, Uxbridge,

<sup>#</sup>University Hospital Aintree, Liverpool,

<sup>†</sup>St George's, University of London, London,

<sup>+</sup>Wythenshawe Hospital, Manchester, UK.

<sup>§</sup>Hvidovre Hospital, Hvidovre, Denmark.

### CORRESPONDENCE

O.N. Keene

GlaxoSmithKline Research and Development

1–3 Iron Bridge Road  
Uxbridge

Middlesex UB11 1BU  
UK

E-mail: oliver.n.keene@gsk.com

Received:

November 30 2007

Accepted after revision:

March 11 2008

### STATEMENT OF INTEREST

Statements of interest for all authors of this manuscript can be found at [www.erj.ersjournals.com/misc/statements.shtml](http://www.erj.ersjournals.com/misc/statements.shtml)

of two major trials using appropriate methods to account for patient variability, as requested by SUISSA [13], including details regarding the distribution of exacerbation rates in the trials and comparative data on average rates.

### SIMPLE ANALYSES

Analyses of exacerbation data are sometimes confined to simple comparisons of proportions of patients with an exacerbation or time to first exacerbation. These methods are familiar from analyses of other data and are relatively easy to understand.

Comparison of proportions with the event does not account for the patients taking part in a study for variable lengths of time, which is a prominent feature of most long-term studies in COPD. Analysis of time to first exacerbation allows for incomplete follow-up of all patients but the estimation of a hazard ratio assumes a constant ratio between treatments across time. In studies where patients are discontinued from the study after experiencing exacerbations [14, 15], analysis of time to first event may provide the best approach.

However, in many studies, these methods omit a key feature of the data, namely whether a patient experiences an exacerbation on more than one occasion, as the data from each patient is reduced to a simple question of whether (and when) the patient experienced an exacerbation. From a clinical viewpoint, they ignore a key potential advantage of a treatment that may reduce multiple exacerbations in the same patient. From a statistical viewpoint they are also not efficient as they do not make use of all the data that have been collected.

### NONPARAMETRIC ANALYSES

In order to make use of the extra data from multiple exacerbations, one approach is to directly compare the number of exacerbations per subject [16], but this does not account for the varying lengths of treatment for each patient. It is preferable to derive an exacerbation rate separately for each patient and then compare these rates between treatment groups. Exacerbation rate per year can be calculated as the number of exacerbations divided by the time on treatment in years. The distribution of these rates is typically very skewed and it is not appropriate to compare rates between treatments using tests based on the normal distribution (such as t-tests). An alternative is to use a nonparametric method that does not assume any particular distribution for the data. One option for comparing exacerbation rates between treatment groups is to use a method based on the Wilcoxon rank-sum test. This method was used in the analysis of the ISOLDE (inhaled steroids in obstructive lung disease in Europe) trial [11] and is a planned analysis for the UPLIFT (understanding the potential long-term impacts on function with tiotropium) trial [17].

Typically, exacerbation rates vary with covariates such as baseline severity of disease (e.g. as described by baseline forced expiratory volume in one second (FEV<sub>1</sub>)). Therefore, greater power can be achieved by use of rank ANCOVA [18], at the expense of some increase in complexity.

For clinical trials in a regulatory setting, it is often advantageous to be able to support a primary model-based analysis with a nonparametric analysis that does not make sampling distribution assumptions. From a theoretical standpoint, such

analyses depend only on the fact that patients have been randomised into the trial and thus are free of controversy over whether modelling assumptions are satisfied.

The major limitation of applying nonparametric methods to exacerbation rates is that they do not provide appropriate methods for the estimation of treatment effects [19]. Therefore, a nonparametric method is best viewed as a supporting analysis rather than the primary analysis.

### POISSON REGRESSION

In order to directly analyse rates of exacerbations and to provide appropriate estimates of the size of a treatment effect, it is necessary to make extra assumptions and fit a statistical model to the data. A commonly used distribution of count data (such as number of events) is the Poisson distribution. In the Poisson analysis, each unit of time is weighted equally. The overall exacerbation rate is obtained by adding up all the exacerbations and dividing by the total time of all patients studied, e.g. a patient studied for 12 months is given 12 times the weight of one studied for 1 month. This approach will underestimate the true exacerbation rates, as patients with high exacerbation rates are more likely to be studied for less time than those with low underlying rates.

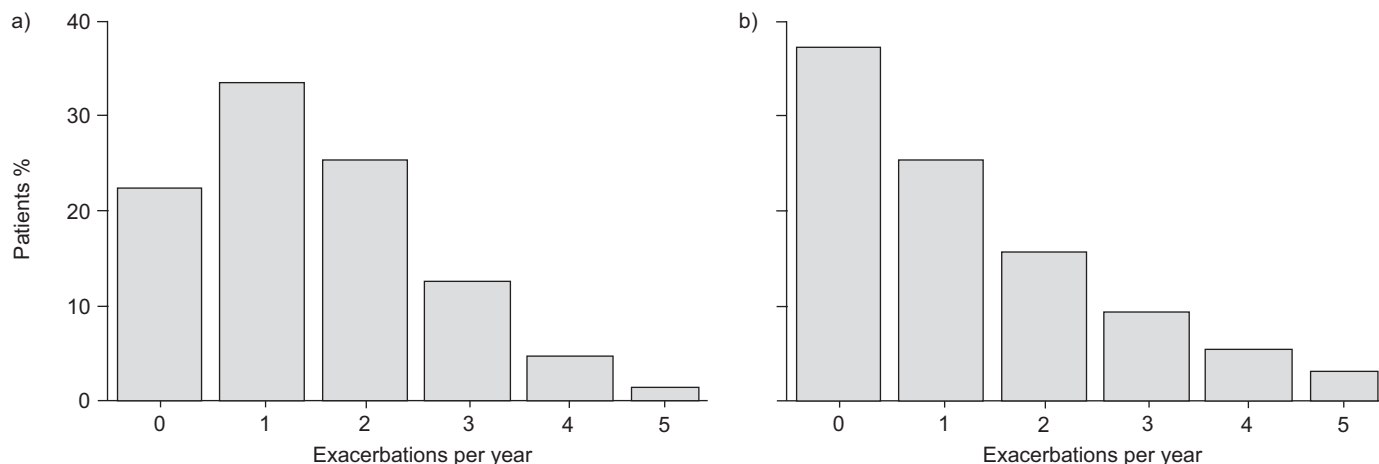
The methodology of generalised linear models allows this distribution to be fitted to exacerbation data. These models explicitly account for the length of time the patient has been on treatment and thus enable estimation of an average exacerbation rate for each treatment group, as well as estimation of the relative reduction in exacerbation rate when comparing treatments. This method was used in the analysis of the TRISTAN (trial of inhaled steroids and long-acting  $\beta_2$ -agonists) trial [20].

### POISSON REGRESSION WITH OVERDISPERSION CORRECTION

Poisson regression, as described above, does not explicitly account for variability between patients, and exacerbation data is typically more dispersed than would be expected by a Poisson distribution because individuals vary intrinsically in their tendency to exacerbate.

Poisson regression can be improved by use of an overdispersion correction. Such a correction increases the standard errors of the treatment estimates to allow for a between-patient variability in rates that is not fully explained by the Poisson distribution. There are two commonly used corrections: deviance correction and Pearson correction [21]. An alternative correction is *via* use of generalised estimating equations [22]. One problem is that there is no specific statistical consensus on which to use. Poisson regression with overdispersion correction is the method recommended by SUISSA [13] and was used in the analysis of two budesonide/formoterol trials [23, 24].

An important issue is that this is a generic correction and variability between patients is not an explicit part of the model. As described above for simple Poisson regression, the model does not account appropriately for the situation in which subjects who withdraw early are more likely to have frequent exacerbations than those who complete a trial [25]. Failing to account properly for the contribution of these patients is likely to underestimate the true benefit of treatment.



**FIGURE 1.** Exacerbation rates from negative binomial distribution. a) Dispersion=0 (Poisson); b) dispersion=0.8.

### NEGATIVE BINOMIAL REGRESSION

More recently, statistical research has indicated that negative binomial regression [26, 27] provides an improved model for exacerbation rate data compared with Poisson regression. This model assumes that each individual has their own underlying rate of exacerbations. The number of exacerbations for each individual follows a Poisson distribution but the expected number is allowed to vary across patients according to a gamma distribution [26]. The gamma distribution is an alternative statistical distribution to the well-known normal distribution; in contrast to the normal distribution it does not allow negative values and has a longer positive tail.

The shape parameter from the gamma distribution explicitly represents variability between patients. The negative binomial model coincides with the Poisson model when this dispersion parameter is equal to zero. If there is no overdispersion, then nothing is lost by using this method. Figure 1 shows examples of the negative binomial distribution of exacerbation rates for an overall average rate of 1.5 per year. The larger the shape parameter, the more dispersed is the distribution. When the shape parameter is small, the rates cluster more closely around the mean value.

The model has several advantages over the Poisson model with overdispersion correction. First, it relies on a less simplistic assumption about variability from patient to patient than the Poisson model does: counts are assumed to vary about means that differ for each patient, with the means varying across the population, rather than varying about a common mean for the whole population. Secondly, estimates from a Poisson model assume a common rate for a population and weight each unit of time equally. For example, a patient who is followed for 12 months will be given 12 times the weight of a patient followed for 1 month. Since patients who exacerbate tend to withdraw earlier from a trial [25], this approach underestimates the true rate of exacerbations. In contrast, the negative binomial model explicitly includes modelling between-patient variability in the estimation of rates. If the number of exacerbations for each individual follows the distribution described above, then this model more effectively accounts for increased exacerbation events among patients

withdrawing early. Statistical details on application of the negative binomial model to analysis of COPD trials have been published elsewhere [19]. This method has been used for analysis of exacerbation rates from the TORCH (towards a revolution in COPD health) trial [12].

### MULTIPLE TIME-TO-EVENT METHODS

Exacerbations can also be considered as repeated events occurring across time. Such data can be analysed using multiple time-to-event methods, which extend the simple approach of analysing the time to first event. The most commonly used of these methods is the Andersen–Gill counting process model [28].

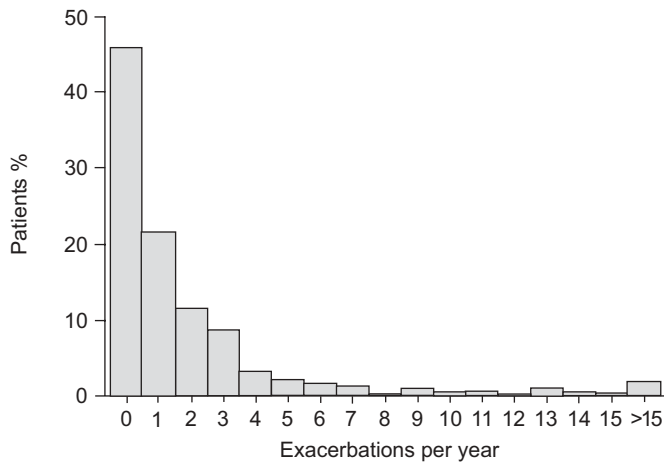
The advantage of these models over the direct analysis of exacerbation rates *via* Poisson or negative binomial regression is that they do not implicitly assume that the exacerbation rate for each patient remains constant over time.

In the respiratory area, the key clinical focus is on the relative rate of exacerbations. The direct approach leads to estimates of the relative rates of the events between treatment groups, a comparison which is of primary clinical interest. Time-to-event methods give estimates of hazard ratios that do not specifically compare rates and are more problematic to interpret. These methods also assume a constant hazard ratio between treatments over time. Although these methods have been a focus for statistical research on the analysis of recurrent events over time, they have not so far been widely used in the analysis of COPD exacerbations.

### THE IMPACT OF STATISTICAL METHODS ON CLINICAL TRIAL RESULTS

#### TRISTAN

TRISTAN [20] was a large (n=1,465), 1-yr, double-blind, randomised study in COPD comparing the effects of the salmeterol/fluticasone propionate combination product (SFC) to salmeterol alone (SAL), fluticasone propionate alone (FP) and placebo. Doses were 50/500 µg (SFC), 50 µg (SAL) and 500 µg (FP), all twice daily. Although the primary end-point was pre-bronchodilator FEV<sub>1</sub>, the number of exacerbations was an important secondary end-point. The occurrence of acute exacerbations was investigated at every clinic visit.



**FIGURE 2.** Exacerbation rates by patient: results from the TRISTAN trial (trial of inhaled steroids and long-acting  $\beta_2$ -agonists) [20].

Exacerbations were defined *a priori* as a worsening of COPD symptoms that required treatment with antibiotics, oral corticosteroids or both. Three patients with incomplete baseline covariate information were excluded from the exacerbation analysis.

Figure 2 presents the exacerbation rates by patient, clearly showing the skewed distribution of this data.

Table 1 shows the estimates of exacerbation rates by treatment. Covariates included in the models were treatment, age, sex, country, baseline severity and smoking status (current or former smoker). The original published estimates of exacerbation rates using the Poisson model are shown [20]. By default, statistical software weights the levels of the covariates equally, *e.g.* estimates rates as though there were equal numbers of males and females in the trial; this was the approach taken in the original publication [20]. However, in clinical trials it is better to reflect the sampled population in general and the estimates have been recalculated in this way, *e.g.* as an average of 73% males and 27% females. This does not affect comparisons between treatments.

This analysis shows that, as described above, the Poisson model underestimates the true exacerbation rate in the trial. Use of an overdispersion correction does not affect the exacerbation rate estimates.

TABLE 1	Summary of estimated exacerbation rates: results from the TRISTAN trial			
	Placebo	Salmeterol	Fluticasone	Combination
<b>Subjects n</b>	361	371	374	356
<b>Exacerbations per year</b>				
Original publication <sup>#</sup>	1.30	1.04	1.05	0.97
Poisson model <sup>†,‡</sup>	1.34	1.08	1.08	1.00
Negative binomial model	1.51	1.12	1.11	1.03

TRISTAN: trial of inhaled steroids and long-acting  $\beta_2$ -agonists. #: Poisson model with equal weights for each covariate level [20]; †: covariates weighted according to observed numbers of patients at each level; ‡: with and without overdispersion adjustment (does not affect the exacerbation rate estimates).

Table 2 shows the ratios of the exacerbation rates on each of the active treatments relative to placebo, with their associated 95% confidence intervals and p-values. A value of 1 would indicate the same rate as placebo; the lower the ratio the greater the reduction compared with placebo.

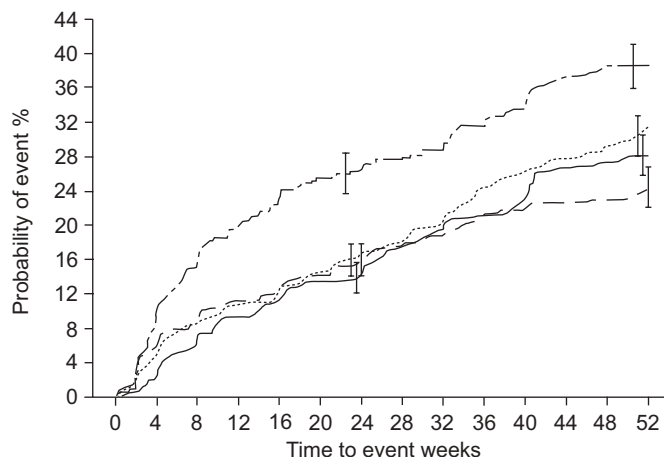
Applying an overdispersion correction to the Poisson model increases the size of the confidence intervals and reduces the size of the p-values for the comparisons between treatments, compared with Poisson regression without overdispersion. Nevertheless, these all remain significantly different from placebo and the sizes of the estimated effects are unchanged. However, investigation of how well the models fit the data through examination of residuals indicates a better fit for the negative binomial model compared with the Poisson model with overdispersion [12].

The negative binomial model provides somewhat larger estimates of treatment effect and it is relevant to understand why this occurs. Figure 3 shows a plot of time to withdrawal by treatment. This clearly indicates a differential withdrawal pattern across treatment groups.

Withdrawal rates were high, particularly on placebo, where 39% of patients withdrew before the end of the trial. Table 3 shows overall exacerbation rates by time on treatment for exacerbations. Exacerbation rates were calculated as number of exacerbations divided by the length of time on treatment in years. As can be seen, patients withdrawing early had much

TABLE 2	Summary of treatment comparisons: results from the TRISTAN trial								
	Salmeterol versus placebo			Fluticasone propionate versus placebo			Salmeterol and fluticasone versus placebo		
	RR	95% CI	p-value	RR	95% CI	p-value	RR	95% CI	p-value
<b>Original publication<sup>#</sup></b>	0.80	0.69–0.93	0.003	0.81	0.70–0.93	0.003	0.75	0.64–0.87	<0.001
<b>Poisson model<sup>†</sup></b>	0.80	0.67–0.96	0.015	0.81	0.68–0.96	0.017	0.75	0.62–0.90	0.002
<b>Negative binomial model</b>	0.74	0.62–0.89	0.001	0.74	0.61–0.89	0.001	0.68	0.57–0.83	<0.001

TRISTAN: trial of inhaled steroids and long-acting  $\beta_2$ -agonists; RR: estimated rate ratio; CI: confidence interval. #: Poisson model with no adjustment [20]; †: deviance overdispersion adjustment.



**FIGURE 3.** Time to study drug discontinuation: results from the TRISTAN trial (trial of inhaled steroids and long-acting  $\beta_2$ -agonists) [20]. Cumulative incidence curve. - · - · - : placebo; · · · · · : salbutamol (50 µg); —: fluticasone propionate (500 µg); - - - - - : salbutamol/fluticasone propionate (50/500 µg) combination. The numbers at risk at 0, 8, 16, 24, 32, 40 and 52 weeks were as follows. For the placebo group: 361, 302, 277, 265, 255, 238 and 140; for the salbutamol (50 µg) group: 372, 340, 325, 309, 297, 275 and 159; for the fluticasone propionate (500 µg) group: 374, 348, 331, 321, 301, 286 and 156; for the salbutamol/fluticasone propionate (50/500 µg) combination group: 358, 325, 312, 301, 290, 278 and 166.

higher exacerbation rates than those who completed the study. The patients who complete the trial tend therefore to be the ones with lower rates of exacerbation.

As stated above, in the Poisson analysis (with or without overdispersion) each unit of time is weighted equally. In contrast, the negative binomial model, which allows separate underlying exacerbation rates for each patient, gives more representative estimates of exacerbation rates. Similarly, because placebo patients tend to withdraw earlier than those on the active treatment arms, the treatment comparison from the Poisson models will underestimate the size of the treatment effect compared with the more representative estimate from the negative binomial model.

**ISOLDE**

ISOLDE [11] randomised 751 subjects to receive FP and placebo twice daily for 3 yrs. Although the primary end-point was decline in FEV<sub>1</sub>, the frequency of exacerbations was an important secondary end-point. Exacerbations were defined *a priori* as a worsening of respiratory symptoms that required treatment with antibiotics, oral corticosteroids or both. Figure 4 shows the exacerbation rate by patient. These rates are somewhat less dispersed than those observed for the TRISTAN trial.

Table 4 shows the estimates of exacerbation rates by treatment using the original nonparametric analysis and from Poisson regression and negative binomial analyses. The rates reported for the nonparametric analysis were the medians of the observed rates for each patient.

The estimates from the Poisson and negative binomial models are higher than those from the nonparametric analysis reflecting, the skewed distribution of the data. The p-values from the Poisson and negative binomial analyses are smaller than those from the original analysis, reflecting the greater power of a model-based analysis compared with a nonparametric analysis. There is less difference in the ISOLDE trial than in TRISTAN between the results from the negative binomial and Poisson analyses. In ISOLDE, in contrast to TRISTAN, patients were withdrawn if the number of exacerbations that required corticosteroids exceeded two in any 3-month period; these withdrawals may have reduced the dispersion of the data.

**GENERAL CONCLUSIONS**

All of the methods described above have advantages and disadvantages. Simple statistical methods that only use data from the first exacerbation can show the direction of a difference but do not appropriately quantify the difference. Use of exacerbation rates is more informative about the signal size for the population in the trial.

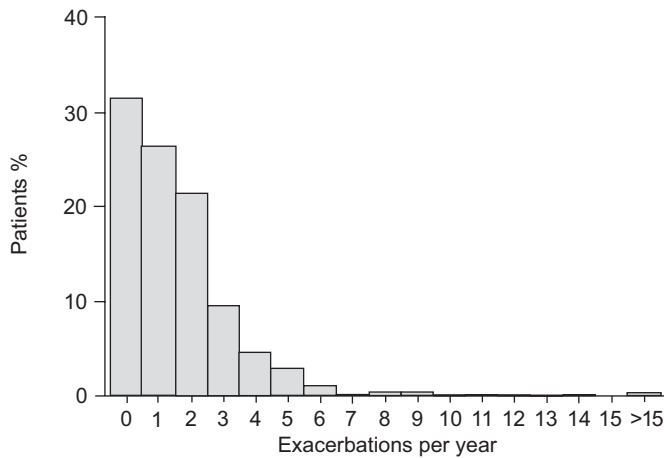
Currently, negative binomial regression appears to be a particularly valuable method. This model directly analyses rates of exacerbations and provides estimates of the relative

**TABLE 3** Summary of mean exacerbation rates by length of time on treatment: results from the TRISTAN trial

	Placebo		Salmeterol		Fluticasone propionate		Combination	
	Subjects n	Exacerbations per year	Subjects n	Exacerbations per year	Subjects n	Exacerbations per year	Subjects n	Exacerbations per year
<b>Length of time on treatment months</b>								
<1	39	8.6	22	4.8	13	5.3	24	2.9
1-<3	35	9.4	18	5.5	22	1.2	15	4.4
3-<6	26	3.9	23	3.5	25	3.4	22	3.7
6-<9	20	3.8	31	2.5	21	2.1	16	2.4
≥9	241	1.0	277	1.0	293	1.1	279	0.9
<b>Total</b>	361		371		374		356	

TRISTAN: trial of inhaled steroids and long-acting  $\beta_2$ -agonists. Reproduced from [19] with permission from the publishers.





**FIGURE 4.** Exacerbation rates by patient: results from the ISOLDE trial (inhaled steroids in obstructive lung disease in Europe) [11].

rate of exacerbations between treatment groups, in contrast to time-to-event methods. The negative binomial model is based on the reasonable assumption that each individual has their own rate of exacerbations. It addresses the concerns raised by SUISSA [13] over adjustment for between-subject variability. The model appears more plausible and scientifically reasonable than the Poisson approach, which assumes one single rate for all patients and then introduces an arbitrary correction for overdispersion. If it is true that the number of exacerbations for each individual follows a Poisson distribution and the expected number varies across patients according to a gamma distribution, then the negative binomial model will provide unbiased estimates of exacerbation rates [27].

The evaluation of whether one particular model fits a set of data better than another cannot be proved mathematically. The standard approach to the assessment of model fit is through residual plots, and these indicate a better fit for the negative binomial model compared with the Poisson model [19]. A model that better describes the data will have reduced noise compared with a less precise one. Reduction in noise associated with a signal leads to more accurate estimates. It is reassuring that the reanalysis of TRISTAN and ISOLDE supports the conclusion that inhaled corticosteroids reduce the rate of exacerbations.

Patients withdrawing early reveal important information on treatment efficacy. The time-weighted estimates from the Poisson model, based on number of exacerbations and person-years, do not account for the overall higher exacerbation rate among patients who withdrew and, therefore, substantially underestimate the true exacerbation rates. In contrast, the negative binomial model provides more appropriate estimates of exacerbation rates.

The negative binomial analysis of rate of exacerbations implicitly assumes that the treatment effect is constant over time after randomisation. In order to further explore this inherent assumption of no interaction between treatment effects and time, one approach is to partition time into a series of intervals and to count the number of exacerbations experienced by the patient in each interval. This analysis requires use of more sophisticated statistical techniques. The model could use a similar negative binomial model within each period with separate rates for each period. However, complicated adjustment is needed to account for the repeated measures nature of the data (sequence of observations within each patient), either *via* a random effects model or by using generalised estimating equations methodology [29]. One problem is that the methods make strong assumptions on the nature of the missing data over time and thus it is likely to be necessary to impute data for patients who withdraw early. These methods deserve further research on their application to the analysis of exacerbation rates but cannot be recommended as the primary analysis at present.

It is important to investigate the sensitivity of trial conclusions to assumptions made in the statistical analysis. Therefore, particularly in major trials, primary analysis using the negative binomial model could be supplemented by secondary multiple time-to-event analysis, which makes different assumptions about missing data, and by nonparametric analysis, which makes no distributional assumptions. Indeed, one single analysis is unlikely to completely summarise the entire complexity of distribution of exacerbations in any COPD study.

As in any clinical trial analysis, the estimates of treatment effects apply to the observed group as a whole. The reductions in exacerbations seen reflects an average effect on a population, which includes patients who will exacerbate frequently without active treatment as well as patients who will exacerbate less frequently, if at all.

TABLE 4	Summary of estimated exacerbation rates: results from the ISOLDE trial				
	Placebo group* exacerbations per year	Fluticasone propionate group† exacerbations per year	Fluticasone propionate/placebo		
			Exacerbations per year	95% CI	p-value
Nonparametric analysis‡	1.32	0.99	0.75	NA <sup>f</sup>	0.026
Poisson model§	1.52	1.29	0.85	0.75–0.96	0.010
Negative binomial model	1.58	1.32	0.83	0.73–0.95	0.006

ISOLDE: inhaled steroids in obstructive lung disease in Europe; CI: confidence interval; NA: not available. #: n=375; †: n=376; ‡: original analysis; §: with overdispersion adjustment; <sup>f</sup>: no CI given for ratio in original paper [11].

Many factors influence the likelihood of a patient experiencing an exacerbation, including the definition of the event itself, the occurrence of previous exacerbations, baseline lung function and previous therapy. Differences in these variables help explain some of the apparently puzzling differences in reported event rates in clinical trials. The subtler, but equally important, differences owing to the statistical method adopted have an effect which is at least as great as that produced by these more obvious factors. The critical reader trying to assess this field should be aware of this and be cautious when pooling data analysed by different methods in any statistical summary of the effectiveness of treatment. Hopefully, future trials will adopt more standardised and robust approaches, like those described herein, to decrease confusion about how frequently exacerbations occur when COPD patients are studied in large clinical trials.

The statistical analysis of exacerbations in clinical trials is challenging. With each clinical trial carried out in chronic obstructive pulmonary disease, more understanding is gained regarding how exacerbation rates are distributed within a population and how they may be affected by patient withdrawal. The statistical methodology for the analysis of such data has evolved over time and there continues to be active scientific research in this area. It is important for clinicians and those who evaluate clinical evidence to be aware of developments in this changing field.

#### ACKNOWLEDGEMENTS

The authors thank L. Willits (GlaxoSmithKline Research and Development, Uxbridge, UK) for programming the statistical analyses.

#### REFERENCES

- 1 Donaldson GC, Seemungal TAR, Bhowmik A, Wedzicha JA. Relationship between exacerbation frequency and lung function decline in chronic obstructive pulmonary disease. *Thorax* 2002; 57: 847–852.
- 2 Kanner RE, Anthonisen NR, Connett JE. Lower respiratory illnesses promote FEV1 decline in current smokers but not ex-smokers with mild chronic obstructive pulmonary disease: results from lung health study. *Am J Respir Crit Care Med* 2001; 164: 358–364.
- 3 Seemungal TAR, Donaldson GC, Paul EA, Bestall JC, Jeffries DJ, Wedzicha JA. Effect of exacerbation on quality of life in patients with chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 1998; 157: 1418–1422.
- 4 Spencer S, Calverley PMA, Burge PS. Impact of preventing exacerbations on deterioration of health status in COPD. *Eur Respir J* 2004; 23: 698–702.
- 5 Soler-Cataluña JJ, Martínez-García MÁ, Román Sánchez P, Salcedo E, Navarro M, Ochando R. Severe acute exacerbations and mortality in patients with chronic obstructive pulmonary disease. *Thorax* 2005; 60: 925–931.
- 6 Rabe KF, Hurd S, Anzueto A, et al. Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease: GOLD Executive Summary. *Am J Respir Crit Care Med* 2007; 176: 532–555.
- 7 Seemungal TA, Donaldson GC, Bhowmik A, Jeffries DJ, Wedzicha JA. Time course and recovery of exacerbations in patients with chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2000; 161: 1608–1613.
- 8 Calverley P, Pauwels R, Löfdahl CG, et al. Relationship between respiratory symptoms and medical treatment in exacerbations of COPD. *Eur Respir J* 2005; 26: 406–413.
- 9 Anthonisen NR, Manfreda J, Warren CP, Hershfield ES, Harding GK, Nelson NA. Antibiotic therapy in exacerbations of chronic obstructive pulmonary disease. *Ann Intern Med* 1987; 106: 196–204.
- 10 Rodriguez-Roisin R. Toward a consensus definition for COPD exacerbations. *Chest* 2000; 117: Suppl. 2, 398S–401S.
- 11 Burge PS, Calverley PMA, Jones PW, Spencer S, Anderson JA, Maslen TK. Randomised, double blind, placebo controlled study of fluticasone propionate in patients with moderate to severe chronic obstructive pulmonary disease: the ISOLDE trial. *BMJ* 2000; 320: 1297–1303.
- 12 Calverley PMA, Anderson JA, Celli B, et al. Salmeterol and fluticasone propionate and survival in chronic obstructive pulmonary disease. *N Engl J Med* 2007; 356: 775–789.
- 13 Suissa S. Statistical treatment of exacerbations in therapeutic trials of chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2006; 173: 842–846.
- 14 Hanaia NA, Darken P, Horstman D, et al. The efficacy and safety of fluticasone propionate (250 µg)/salmeterol (50 µg) combined in the Diskus inhaler for the treatment of COPD. *Chest* 2003; 124: 834–843.
- 15 Mahler DA, Wire P, Horstman D, et al. Effectiveness of fluticasone propionate and salmeterol combination delivered via the Diskus device in the treatment of chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2002; 166: 1084–1091.
- 16 Paggiaro PL, Dahle R, Bakran I, Frith L, Hollingworth K, Efthimiou J. Multicentre randomised placebo-controlled trial of inhaled fluticasone propionate in patients with chronic obstructive pulmonary disease. *Lancet* 1998; 351: 773–780.
- 17 Decramer M, Celli B, Tashkin DP, et al. Clinical trial design considerations in assessing long-term functional impacts of tiotropium in COPD: the UPLIFT trial. *COPD* 2004; 1: 303–312.
- 18 Quade D. Rank analysis of covariance. *J Am Stat Assoc* 1967; 62: 1187–1200.
- 19 Keene ON, Jones MR, Lane PW, Anderson J. Analysis of exacerbation rates in asthma and chronic obstructive pulmonary disease: example from the TRISTAN study. *Pharm Stat* 2007; 6: 89–97.
- 20 Calverley P, Pauwels R, Vestbo J, et al. Combined salmeterol and fluticasone in the treatment of chronic obstructive pulmonary disease: a randomised controlled trial. *Lancet* 2003; 361: 449–456.
- 21 McCullagh P, Nelder JA. Generalized Linear Models. 2nd Edn. London, Chapman and Hall, 1989.
- 22 Stokes ME, Davis CS, Koch GG. Categorical Data Analysis Using the SAS® System. 2nd Edn. Cary, SAS publishing, 2000; pp. 542–549.
- 23 Szafranski W, Cukier A, Ramirez A, et al. Efficacy and safety of budesonide/formoterol in the management of chronic obstructive pulmonary disease. *Eur Respir J* 2003; 21: 74–81.
- 24 Calverley PM, Boonsawat W, Cseke Z, Zhong N, Peterson S, Olsson H. Maintenance therapy with budesonide

- and formoterol in chronic obstructive pulmonary disease. *Eur Respir J* 2003; 22: 912–919.
- 25** Calverley PMA, Spencer S, Willits LR, Burge PS, Jones PW, ISOLDE Study Group. Withdrawal from treatment as an outcome in the ISOLDE study of COPD. *Chest* 2003; 124: 1350–1356.
- 26** Agresti A. *Categorical Data Analysis*. 2nd Edn. New Jersey, Wiley, 2002; pp. 559–563.
- 27** Metcalfe C, Thompson SG. The importance of varying the event generation process in simulation studies of statistical methods for recurrent events. *Stat Med* 2006; 25: 165–179.
- 28** Andersen PK, Gill, RD. Cox's regression model for counting processes: a large sample study. *Ann Stat* 1982; 10: 1100–1120.
- 29** Liang K-Y, Zeger SL. Longitudinal data analysis using generalized linear models. *Biometrika* 1986; 73: 13–22.