

Evaluate multiple adverse events in crossover design bioequivalence clinical trials

WANG Yong¹, LI Lin-Xian², WANG Zi-Can², WANG Yue-He³ (*Biogen Inc, Cambridge, MA, USA*; ²*Kunming Medical College, Kunming 650031, China*; ³*Lenox Hill Hospital, New York, NY, USA*)

KEY WORDS crossover clinical trial; adverse effects; repeated-measures analysis; statistical models

ABSTRACT

AIM: To establish a statistical model to appropriately evaluate the relationship between clinical adverse event and treatment. **METHODS:** By splitting the duration of each treatment period into several time intervals (use day as unit), a clinical adverse event was sampled in each time interval as presented or not presented. The number of presented cases was added for each time interval, and time of subjects spent in each time interval was cumulated as person-days. The onset of a clinical adverse event and its duration were represented as repeatedly measured count data. By using the generalized linear mixed model with fixed- and random-effects, the relative rate of clinical adverse events relative to different treatments was modeled by the generalized estimate equation (GEE) technique. **RESULTS:** Example shows that not only the onset of adverse events, but also its duration and total person-days subjects spent in study would influence the relative rate of clinic adverse events. **CONCLUSION:** Our proposed approach is a good alternative and supplemental method for evaluating clinical adverse events.

INTRODUCTION

In a crossover design clinical trial, a subject is treated for several treatment periods and received different treatments^(1,2). In the trial, a subject might experience a clinical adverse event started at one treatment period and stopped at another one. When both the onset and duration of a clinical adverse event need to be compared across different treatments, one event could be counted in more than one treatment group. These adverse event da-

ta in different treatment groups are highly correlated. Meanwhile, when a clinical adverse event recurs many times, its presence in different treatment groups would be highly correlated. Since correlated data would not have as much information as independent data, fail to consider correlation among repeated-measures data would lead to biased results⁽³⁻⁵⁾. In this article, we will discuss an approach for appropriately evaluate onset and duration of clinical adverse events occurring in a crossover trial.

Here we briefly discuss a clinical trial used in this paper as an example. We are not going to analyze the data of this study but only to demonstrate our method. The analyses results will be reported other where.

The study was a 2 by 2 crossover pharmacokinetic and pharmacodynamic clinical trial⁽⁶⁾. The study drug is a novel compound for non-narcotic analgesic use in moderate and severe pain. The compound is a selective inhibitor of cyclooxygenase with anti-inflammatory and analgesic activity. The study was designed to demonstrate the bioequivalence of using intramuscular (IM) and intravenous (IV) administration routes in drug's pharmacokinetic and pharmacodynamic activities (one type of study required by the Food and Drug Administration of the United States of American for New Drug Application).

In each treatment period, every subject was given a IM route and a IV route administration of the study drug and placebo (or two placebos) within three houses time period, so that both subjects and investigators were blinded from both real treatment drug and real administration route of the drug.

During the 17 of study period, eighteen subjects were treated with either the study drug or placebo, administered by both intramuscular and intravenous routes. Following the first 5 d of IM or IV administration of the study drug (treatment period 1), subjects underwent a 7-d washout period. During the second 5-d treatment period (treatment period 2), each subject received the study drug by the route of administration not received during the first treatment period.

¹ Correspondence to WANG Yong.

E-mail yong-wang@biogen.com

Received 1999-11-22

Accepted 2000-10-23

Subjects were randomized to one of three treatment sequences (AB, BA, or CC): A = IM 40 mg study drug plus IV placebo; B = IM placebo plus IV 40 mg study drug; C = IM placebo plus IV placebo. Thirteen subjects were randomized to each treatment sequence of AB and BA, respectively. Nine subjects randomized to the CC sequence received one IV placebo and one IM placebo during both treatment periods 1 and 2. One subject in the AB sequence withdrew before the end of the study.

Clinical adverse events occurred during the study periods were identified by subjects and investigators. All reported clinical adverse events were categorized by body systems and preferred terms. Clinical adverse events occurred during the washout period were assessed relative to the most recent treatment.

METHODS

Repeated measures in sequential time intervals

To consider effect of both onset of clinical adverse events and its duration, we divided the duration of each treatment period into several small time intervals. A clinical adverse event was counted in each time interval as presented or not presented. The number of presented cases was added for each time interval, and time of subjects spent in each time interval was summarized as person-days.

Let $T_{k,0}$ and $T_{k,E}$ be the start and stop time points of the k th treatment period, respectively. For simplifying the notation, washout period is included in its previous treatment period. Let the time duration between $T_{k,0}$ and $T_{k,E}$ be divided into S_k intervals as $(t_{k,s}, t_{k,s+1})$, where $T_{k,0} = t_{k,1} < t_{k,2} < \dots < t_{k,S_k} = T_{k,E}$, and $t_{k,s+1} - t_{k,s} = \Delta(s = 1, \dots, S_k)$. Based on this setting, we can count adverse events in each small time intervals. Let y_{ijkl} denote the total number of occurrences of adverse events to the i th subject in the j th time interval of the l th treatment sequence and the k th treatment period $(t_{k,l}, t_{k,l+1})$, where $j = 1, \dots, n_k, i = 1, \dots, q, l = 1, \dots, q$, and $k = 1, \dots, K$. Similarly, we would have m_{ijkl} be the person-days the subject spent in that time interval. We assume that y_{ijkl} is a random variable and its first and second moments exist. In this setting, the onset and duration of clinical adverse events would be presented by counts added in sequential time intervals. If more clinical adverse events occurred, or if the duration of an event was longer, more counts would be observed in each time interval. The onset and dura-

tion of clinical adverse events were reflected by counts repeatedly measured at different time intervals.

Generalized linear mixed model for counts of adverse events on crossover trials

The counts of adverse events discussed in previous section are highly correlated so that repeated-measures method is an appropriate approach to analyze these data. A mixed model with fixed- and random-effects is used, in which repeated measures in different treatment periods associated with a subject is evaluated by a random-effect, and other factors such as treatment effect, treatment period, treatment sequence, and covariates such as age, gender, etc, are described by fixed-effects⁽⁷⁻⁹⁾.

In the mixed-model, the occurrence of clinical adverse events to the j th subject in the s th time interval of the l th treatment sequence and the k th treatment period would be proportional to the total number of person-days subjects spent in that time interval. Factors that would influence model results are treatment group, treatment sequence, treatment period, and other within subject variations. The discrete counts were modeled as $y_{ijkl} = m_{ijkl} \exp(\mathbf{X}^T \boldsymbol{\beta})$, where \mathbf{X} is a design matrix, and $\boldsymbol{\beta}$ is a vector of interesting model parameters. Specifically, let h be a log-link function, and $\mu_{ijkl} = E(y_{ijkl})$, the generalized linear mixed model can be written as: $h(\mu_{ijkl}) = \tau_i + \lambda_k + \kappa_l + x_{ijkl}^T \boldsymbol{\beta}_c + \gamma_{jk(l)}$, with an error of the model ϵ_{ijkl} . The fixed-effects τ_i, λ_k and κ_l represent factors of treatment group, treatment sequence, and treatment period, respectively. The term $x_{ijkl}^T \boldsymbol{\beta}_c$ includes other interesting covariates and offset of person-days^(1,2,7,10). Random effect $\gamma_{jk(l)}$ reflects repeated-measures of the j th subject in the k th treatment period nested within the l th treatment sequence. The two random variables $\gamma_{jk(l)}$ and ϵ_{ijkl} are mutually independent, and both have their own means and variances. Without lose of generality, the correlation between different y_{ijklp} and y_{ijklq} is assumed to be the same. In case of log-link function, quantity $\exp(\tau_A - \tau_B)$ represents relative rate of clinical adverse events between treatment group A and B.

We suppose that the generalized linear mixed model has q model parameters as fixed-effects, and r model parameters reflecting random-effects. Let $\boldsymbol{\beta} = (\boldsymbol{\beta}_1, \dots, \boldsymbol{\beta}_q)^T$ and $\boldsymbol{b}_j = (\boldsymbol{b}_{1j}, \dots, \boldsymbol{b}_{rj})^T$ denote the fixed-effects for the model and the random-effects for the j th subject respectively, and μ_j be a vector of counts μ_{ijkl} for j th subject. The generalized linear mixed model could be presented as $h(\mu_j) = \mathbf{x}_j^T \boldsymbol{\beta} + \mathbf{z}_j^T \boldsymbol{b}_j$, where \mathbf{x}_j and \mathbf{z}_j are $q \times 1$ and $r \times 1$ vectors. The random-effects represent

the variations among repeated measurements within each subject^(2,7). This model could be fitted by the generalized estimation equation (GEE) method^(8,9,11).

Modeling repeated measured counts by GEE procedure The GEE method developed by Zeger and Liang^(8,9,11) was used to fit the generalized linear mixed model discussed in the previous section. Following Zeger and Liang, we assume that the discrete counts have a variance function $g(\mu_i)$, and denote $\mathbf{A}_i = \text{diag}[g(\mu_{i1}), \dots, g(\mu_{iM})]$ as an diagonal matrix. Under the GEE model, the covariance matrix among a set of repeated measures is $\mathbf{V}_j = \Psi \mathbf{A}_j^{1/2} \mathbf{R} \mathbf{A}_j^{1/2}$, where \mathbf{R} is the working correlation matrix among the repeated measures, and $\Psi > 0$ is a scale parameter reflecting under ($0 < \Psi < 1$) or over ($\Psi > 1$) dispersion relative to the underlying stochastic model⁽¹¹⁾. The parameter Ψ is typically a nuisance parameter and is estimated from the data at the analysis stage. Based on the discussion for within-subject correlation in previous section, the compound symmetric or exchangeable covariance structure is used for \mathbf{R} in estimating covariance of repeatedly measured counts in GEE procedure.

The GEE method is available in a few statistical analysis softwares such as SAS and Splus^(12,13). We used SAS to implement the GEE estimation methods for mixed model. With its Repeated option of Genmod Procedure available in SAS version 6.12 or high, the number of adverse events was assumed to have a log-link function⁽¹¹⁾. Each treatment period was divided into several sequential time intervals. The length of a time interval was one day. In each time interval (day), the onsets of clinical adverse events were counted. For an example, suppose a subject had a clinical adverse event 3 d in the treatment period 1 (started at d 8) and 2 d in the treatment period 2. The length of both the treatment periods was 10 days. The counts calculated for this event was represented by vectors (0 0 0 0 0 0 1 1 1) and (1 1 0 0 0 0 0 0 0) for period 1 and 2, respectively. These repeatedly measured binary data, associated with various categories variables of treatment group, treatment sequence, treatment period, and demographic covariates such as gender and age, were added together for all subjects to form the data set for analysis. A category variable "repeat" was set to represent repeated-measures relative to individual subjects. Using Genmod Procedure of SAS, this data set was analyzed by following modeling options:

```
Proc Genmod data = data set name;  
class patient - id treatment sequence period repeat
```

sex age;

model counts = treatment period sex age/link = log
distribution = Poisson;

repeated subject = patient - id(sequence)/type = ex-
changeable within = repeat;

The interesting outputs were the estimated model parameters (β) and its covariance matrix [$\text{cov}(\beta)$]. With these estimated model parameters, we can express the specific hypothesis under consideration as $H_0: \mathbf{H}\beta = \mathbf{h}_0$ vs $H_1: \mathbf{H}\beta \neq \mathbf{h}_0$ where \mathbf{H} is an matrix of full row rank and \mathbf{h}_0 is a conformable vector of constant term. Under this setting, the Wald test statistics $Q_W = (\mathbf{H}\hat{\beta} - \mathbf{h}_0)^T \text{cov}(\hat{\beta})^{-1} (\mathbf{H}\hat{\beta} - \mathbf{h}_0)$ is asymptotically distributed as a $\chi^2_{(h)}$ distribution. The P -value on test H_0 is given by $\chi^2_{(h), 1-\alpha}$ where α is the type I error, and h is equal to the full row rank of \mathbf{H} , and is the degree of freedom for χ^2 distribution. The SAS program is available upon request.

RESULTS

The proposed method was used to estimate the relative rate of clinical adverse events of the crossover design clinical trail discussed in the section 1.2. The relative rate of clinical adverse events was defined as ratio of percentage of clinical adverse events in a treatment group versus that in another group. The clinical adverse events occurred in this study were tabulated as the incidence of adverse events in Tab 1.

Tab 1 shows that the number of subjects had clinical adverse events were reported by 4 (44 %) subjects in the protocol group, 7 (58 %) subjects in the IM-route group, and 6 (46 %) in IV-route group, respectively. The most frequently occurring adverse event was headache [2 (22 %), 3 (25 %), and 4 (31 %) subjects in each protocol, IM- and IV-route group, respectively]. The most commonly reported adverse event (at least twice) in any of the protocol, IM- and IV-route groups were headache (2:3:4 subjects) and fever (0:3:2 subjects). In each group, subjects might have multiple adverse events.

From Tab 1 one could calculate that the relative incidence of adverse events of IM and IV-route group is 1.32 and 1.05 times higher than that in the protocol group. The ratio of incidence of adverse events between IM- and IV-administration groups was 1.26.

The method used for Tab 1 excluded multiple and re-occurred adverse events. This method could not re-

flect global effect of treatments on multiple adverse events occurred to individual subjects. Further, the method could not provide quantitative statistical criteria to judge difference among multiple adverse events associated with different treatments.

Tab 1. Incidence of adverse event¹⁾. Subjects with at least one dose of study drug.

	Placebo n (%)	IM-Route drug ²⁾ n (%)	IV-Route drug n (%)
Number of subject dosed	9 (100)	12 (100)	13 (100)
Number with any event	4 (44)	7 (58)	6 (46)
Fever	0 (0)	3 (25)	2 (15)
Headache	2 (22)	3 (25)	4 (31)
Peripheral pain	0 (0)	1 (8)	0 (0)
Edema generalized	0 (0)	1 (8)	0 (0)
Rhinitis	0 (0)	1 (8)	0 (0)
Vein pain	0 (0)	1 (8)	0 (0)
Flushing	0 (0)	1 (8)	1 (8)
Hypouricemia	0 (0)	1 (8)	0 (0)
Sgpt increased	0 (0)	1 (8)	0 (0)
Sgot increased	0 (0)	1 (8)	0 (0)
Diarrhea	0 (0)	1 (8)	1 (8)
cramps legs	0 (0)	1 (8)	0 (0)
Dizziness	0 (0)	1 (8)	1 (8)
Arthritis	0 (0)	1 (8)	0 (0)
Arthralgia	0 (0)	1 (8)	1 (8)
Rash maculo-papular	0 (0)	1 (8)	1 (8)
Injection site	0 (0)	0 (0)	1 (8)
Pain	0 (0)	0 (0)	1 (8)
Back pain	0 (0)	0 (0)	1 (8)
Edema peripheral	0 (0)	0 (0)	2 (15)
Urinary incontinence	1 (11)	0 (0)	0 (0)
Upper resp tract infection	0 (0)	0 (0)	1 (8)
Pharyngitis	0 (0)	0 (0)	2 (15)
Tachycardia	0 (0)	0 (0)	1 (8)
Hypertalemia	0 (0)	0 (0)	1 (8)
Bun increased	0 (0)	0 (0)	1 (8)
Abdominal pain	0 (0)	0 (0)	1 (8)
Constipation	1 (11)	0 (0)	0 (0)
Somnolence	1 (11)	0 (0)	0 (0)
Insomnia	1 (11)	0 (0)	0 (0)
Confusion	1 (11)	0 (0)	0 (0)
Earache	0 (0)	0 (0)	1 (8)
Myalgia	0 (0)	0 (0)	2 (15)
Rash	0 (0)	0 (0)	1 (8)
Pruritus	1 (11)	0 (0)	1 (8)

1) If a subject had more than one of the same type of adverse event, that subject was counted once under each adverse event.

2) Adverse events were sorted by descending incidence of this treatment column.

Our proposed method was used to model the relative rate of clinical adverse events for these data. The results were summarized in Tab 2. In Tab 2, the number of subjects having adverse events, the number of multiple

adverse events, the duration of adverse events, and person-days subjects spent in the study were tabulated by treatment groups. The relative rate of adverse events, denoted as $\exp(\tau_A - \tau_B)$ where τ_A and τ_B were estimated model parameters representing treatment effects for group A and B, was estimated by our proposed method discussed in section 2. These model parameters were estimated through a generalized mixed model by the GEE method as discussed in section 2.

The 95 % confidence intervals of the estimated relative incidence rate of adverse events were calculated by $\exp\{\tau_A - \tau_B \pm 1.96 [\text{var}(\tau_A) - 2 \text{cov}(\tau_A, \tau_B) + \text{var}(\tau_B)]^{1/2}\}$, where $\text{var}(\tau_A)$, $\text{var}(\tau_B)$, and $\text{cov}(\tau_A, \tau_B)$ were estimated variances and covariance of τ_A and τ_B . The quantity $(\tau_A - \tau_B)2/[\text{var}(\tau_A) - 2 \text{cov}(\tau_A, \tau_B) + \text{var}(\tau_B)]$ had a χ^2 distribution with one degree of freedom. The estimated relative incidence rate of adverse events, its 95 % confidence interval, and *P*-value of χ^2 test were listed in the Tab 2.

Tab 2 shows that 9 subjects took placebo plus placebo treatment sequence, and each 13 subjects took IM-route plus IV-route and IV-route plus IM-route treatment sequences. One subject dropped from IV-route treatment in treatment period 2. Totally 13, 12, and 9 subjects were treated with IM- and IV-route administrated study drug, and placebo, respectively. The person-days these subjects spent in study were 39.5, 38.9, and 28.3 d in IM- and IV-route, and placebo groups, respectively. The number of multiple clinical adverse events and their re-occurrences were 23, 29, and 6 cases in these three treatment groups, and the duration of adverse events were 50, 87, and 5 d, respectively. The number and duration of adverse events in IM- and IV-route groups were similar, but were bigger and longer than those in placebo group.

The rate of multiple clinical adverse events in IM-route group was estimated as 3.361 times higher than that in placebo group with 95 % confidence interval (CI) as (1.237, 9.135). The rate of clinical adverse events in IM-route group was statistically significantly higher than that in the placebo group (*P* = 0.021). Similarly, the rate of multiple clinical adverse events in IV-route was estimated as 3.564 times higher than that in placebo group with 95 % CI as (1.322, 9.606). Again, the rate of clinical adverse events in IV-route group was statistically significantly higher than that in the placebo group (*P* = 0.015). On the other hand, the rate of multiple clinical adverse events in IM-route group was slightly lower than

Tab 2. Relative incidence rate of adverse events between two treatment groups subjects with at least one dose of study drug.

Estimates ¹⁾	IM-Route drug vs IV-route drug	IM-Route drug vs placebo	IV-Route drug vs placebo
Subjects	13/12	13/9	12/9
Person-days spent in treatment	221/204	221/153	204/153
Adverse events	23/29	23/6	29/6
Adverse events duration (day)	50.0/87.0	50.0/5.0	87.0/5.0
Relative incidences rate ²⁾	0.9	3.361	3.564
95 % CI of relative incidences	(0.625, 1.423)	(1.237, 9.135)	(1.322, 9.606)
P-value of chi-square test ³⁾	0.787	0.021	0.015

1) Analysis incorporated information of single and recurrent adverse events, and the duration of each event. Table shows Population-Average Estimates based on the Generalized Estimation Equation (GEE) method. The treatment, sequence and period factors were fixed-effects, and repeated-measures per subject were random-effects of a generalized linear mixed model with log-link function.

2) Relative incidence rate presents the ratio of the incidence of one group to the incidence of the other group.

3) Chi-square test on difference between two treatment groups.

that in the IV-route group; the ratio of the two incidences was estimated as 0.9 with 95 % CI as (0.625, 1.423). This small difference is not statistically significant ($P = 0.787$).

From Tab 2 one can see that both the onset and the duration of clinical adverse events would influence the relative rate. Statistical criteria were provided to judge these effects across difference between treatment groups.

DISCUSSION

Crossover design clinical trial is usually used for study such as pharmacokinetic and pharmacodynamic study of new compound in which variation in efficacy measures associated with individual subjects is significantly large. However, evaluating the onset and the duration of clinical adverse events in crossover design clinical trial arises a challenge for statistician because an adverse event might undergo more than one treatment periods so that this event would be related to more than one treatment groups. In this paper, we discussed an approach that is suitable for analyzing both onset and duration of multiple clinical adverse events in a crossover design clinical trial. We combined information of the onset and the duration of adverse events relative to individual subjects into repeatedly measured counts for the subjects. The correlation among repeatedly measured counts was appropriated handled by the GEE method so that elegant statistical criteria can be established to evaluate difference in relative rate of clinical adverse events across treatments.

In the example, day was used as time unite since the whole study period was 17 d. In crossover trails when time period for each treatment is longer (such as a few

months), potentially more subjects would drop out study than that in a short time period one. Using day as time unite would help to measure time subjects spent in study more correctly. The recommend approach is to select a time unit based on how often adverse events would occur in a trail.

The method provides an optional approach to evaluate treatment difference relative to both the onset and the duration of clinical adverse events, and the results are easily interpreted to medical investigators. Besides, the method can be easily implemented into most commercial statistical software with little programming. Therefore, our proposed approach is a good alternative and supplemental method for evaluating clinical adverse events in crossover designed clinical trials, which would allow investigators to explore information hidden in these complicated data, and to address less biased and more accurate conclusions.

REFERENCES

- 1 Chow SC, Liu JP. Design and analysis of bioavailability and bioequivalence studies. Marcel Dekker, Inc; 1992. p 23 - 70.
- 2 Vonesh EF, Chinchilli VM. Linear and nonlinear model for the analysis of repeated measurements. Marcel Dekker, Inc; 1997. p 45 - 202.
- 3 Feng Z, Grizzle JE. Correlated binomial variates, properties of estimator of intraclass correlation and its effects on sample size calculation. Statistics Med 1999; 11: 1607 - 14.
- 4 Burton P, Gurrin L, Sly P. Extending the simple linear regression model to account for correlated response: An introduction to generalized estimating equations and multi-level mixed modeling. Statistics Med 1998; 17: 1261 - 91.
- 5 Steiner JL, Mallet A, Mentre F. Estimating interindividual

pharmacokinetic variability. In: Rowland M, Sheiner LB, Steiner JL, editors. Variability in drug therapy. New York: Raven press; 1985. p 95 - 101.

6 Buckley MM, Brogden RN, Ketorolac J. A Review of pharmacodynamic and pharmacokinetic properties, and therapeutic potential. *Drugs* 1990; 39: 86 - 109.

7 Breslow NE, Clayton DG. Approximate inference in generalized linear mixed model. *J Am Statistic Assoc* 1993; 88: 9 - 25.

8 Zeger S, Liang KY, Albert PS. Generalized linear models for longitudinal data: A general estimating equation approach. *Biometrics* 1988; 44: 1049 - 60.

9 Zeger S, Liang KY. Longitudinal data analysis for discrete and continuous outcomes. *Biometrics* 1988; 42: 121 - 30.

10 McCullagh P, Nelder JA. Generalized linear model. 2nd ed. Oxford: Chapman and Hall; 1989. p 21 - 47, 404 - 409, 471.

11 Zeger S, Karim R. Generalized linear models with random effects. *J Am Statistic Assoc* 1991; 86: 79 - 86.

12 SAS Institute Inc. UNIX release 6.12. Cary: SAS Institute; 1998. p 604 - 49.

13 Carey VJ. Implementation of gee for spls. Boston: Harvard University; 1992. p 92 - 178.

王壖¹, 李麟仙², 王子灿², 王跃和³ (*Biogen Inc, Cambridge, MA, USA*; ²昆明医学院, 昆明 650031, 中国; ³*Lenox Hill Hospital, New York, NY, USA*)

关键词 交叉临床试验; 副作用; 重复测量分析; 统计学模型

目的: 建立一种统计学模型来评价临床不良事件与治疗间的关系. **方法:** 将交叉试验的治疗期与间歇期分开, 记录每一个阶段不良事件发生的次数和持续天数, 累积各受试者每一阶段发生不良事件的总时间, 用日/人表示. 以泊松分布法计数发生事件反应的次数及其持续时间; 以普通线性混合模型估计各不同处理间不良事件的发生率; 以广义估计方程估计各交叉试验组间的相关性. **结果:** 不良事件的发生率不仅受发生次数的影响, 也受其持续时间和受试者耗于不良事件(日/人)总时间的影响. **结论:** 本研究提供了一个实用的评价各种不良事件发生率的方法, 此方法比单纯考虑不良事件的发生次数具有更少的偏性.

在交叉设计的生物等效临床试验中评价多种不良事件

(责任编辑 朱倩蓉)