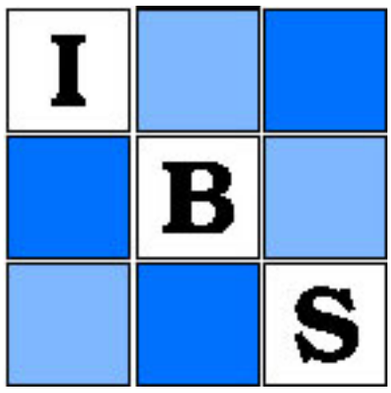

BIOMETRICS 48, 459–478
June 1992

On the Use of Historical Control Data to Estimate Dose Response Trends in Quantal Bioassay

Ross L. Prentice,[1] Robert T. Smythe,[2] Daniel Krewski,[3] and Mark Mason[1]

[1] Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, 1124 Columbia Street, MP-665, Seattle, Washington 98104, U.S.A.
[2] Department of Statistics, George Washington University, Washington, D.C. 20052, U.S.A.
[3] Health Protection Branch, Health and Welfare Canada, Ottawa, Ontario K1A 0L2, Canada

SUMMARY

New tests for trend in proportions, in the presence of historical control data, are proposed. One such test is a simple score statistic based on a binomial likelihood for the "current" study and beta-binomial likelihoods for each historical control series. A closely related trend statistic based on estimating equations is also proposed. Trend statistics that allow overdispersed proportions in the current study are also developed, including a version of Tarone's (1982, *Biometrics* **38,** 215–220) test that acknowledges sampling variation in the beta distribution parameters, and a trend statistic based on estimating equations. Each such trend test is evaluated with respect to size and power under both binomial and beta-binomial sampling conditions for the current study, and illustrations are provided.

1. Introduction

The need to test for trend in proportions arises in many areas of biomedical research. Studies of the carcinogenicity of chemicals or other substances in rodent bioassays constitute an area of particular importance. In a typical study, a control group of animals at dose level $d_0 = 0$, and one or more test groups, are followed to observe the number x_{1i} of animals developing tumors out of the n_{1i} at risk at dose level d_i, $i = 0, \ldots, k$. The Cochran–Armitage test is very often used to test for a trend in tumor rate with increasing dose level. It can be written

$$X^2_{CA} = \left(\sum_{i=0}^{k} x_{1i} d_i - \hat{p} \sum_{i=0}^{k} n_{1i} d_i \right)^2 \Bigg/ \left[\hat{p}(1 - \hat{p}) \left\{ \sum_{i=1}^{k} n_{1i} d_i^2 - (n_1 \bar{d})^2 n_1^{-1} \right\} \right], \tag{1}$$

where $x_1 = x_{10} + \cdots + x_{1k}$, $n_1 = n_{10} + \cdots + n_{1k}$, $\hat{p} = x_1 n_1^{-1}$, and $\bar{d} = (\sum n_{1i} d_i) n_1^{-1}$. This statistic quite generally has an asymptotic χ^2_1 distribution under the hypothesis of no trend in tumor rates assuming binomially distributed response counts. It can be derived as the efficient score test for $\beta = 0$ under a logistic dose response

$$p(d) = \alpha \exp(\beta d)\{(1 - \alpha) + \alpha e^{\beta d}\}^{-1}, \tag{2}$$

where α is the control group response (tumor) rate. In fact, Tarone and Gart (1980) show that the Cochran–Armitage statistic is locally optimal under a wide range of monotone dose response models.

Key words: Beta-binomial distribution; Binary regression; Carcinogenesis testing; Estimating equations; Historical controls; Logistic regression; Overdispersion; Trend tests for proportions.

There is a natural interest in methods to strengthen the trend test by using the control group experience from other studies involving similar animals and similar experimental conditions. While it may be too strong to assume that the control group response rate in the historical series is identical to that (α) in the current experiment, it may often be reasonable to assume that such historical response rates are independent and randomly distributed about α. Section 2 presents corresponding new historical control trend tests developed either under the assumption of a beta distribution for the historical control response rates, or under weaker moment assumptions. These tests make fundamental use of the historical control data, and corresponding dose–response estimation is straightforward.

The assumption that historical control response rates are random variables from a distribution with mean equal to the current control response probability α contrasts with the assumption underlying prior approaches to this problem. Specifically, earlier approaches, as reviewed by Krewski et al. (1988), assume that both historical and current control response rates are independent variates from a distribution with a common mean. For example, Tarone (1982) assumes that all control group response rates arise from a beta distribution, whereas Dempster, Selwyn, and Weeks (1983) assume a normal mixing distribution. Such assumptions replace the binomial model, routinely used for the current experiment in the absence of historical data, by a model with overdispersion. While allowance for overdispersion in the current experiment may or may not lead to a better-fitting model, the routine replacement of the standard assumptions about the current experiment simply because historical control data are being simultaneously analyzed, does not seem to be compelling. Furthermore, as elaborated in Section 4, the historical control data evidently play a rather minor role in trend testing under such overdispersed models. Specifically, such trend tests are generally asymptotically equivalent to the Cochran–Armitage statistic (1).

2. Testing and Estimation of Trend Under a Binomial Model for the Current Experiment

Suppose as above that there are x_{1i} responses among the n_{1i} individuals (animals) at dose level d_i ($i = 0, \ldots, k$; $d_0 = 0$), and that these responses arise as binomial observations with response probability $p(d_i)$ given by (2). Denote by x_j and n_j the sample size and number of responses in historical control group j, and suppose that x_j is binomial with response probability θ_j. Suppose also that θ_j ($j = 2, \ldots, s$) are independent random variates from a distribution with mean α, the current experiment control response rate. For example, a beta distribution with parameters a and b, where $a(a + b)^{-1} = \alpha$, leads to a beta–binomial distribution for x_j ($j = 2, \ldots, s$). The corresponding likelihood function can be written

$$L = L_1 L_2,$$

where

$$L_1 = \prod_{i=0}^{k} (p_i)^{x_{1i}} (q_i)^{n_{1i} - x_{1i}}$$

is the standard binomial likelihood from the current study, with $p_i = p(d_i)$ and $q_i = 1 - p_i$, and

$$L_2 = \prod_{j=2}^{s} \left\{ \prod_{0}^{x_j - 1} (\alpha + \gamma i) \prod_{0}^{n_j - x_j - 1} (1 - \alpha + \gamma i) \prod_{0}^{n_j - 1} (1 + \gamma i)^{-1} \right\}$$

is the product of beta–binomial likelihoods from the historical control series, where $\gamma = (a + b)^{-1} = \rho(1 - \rho)^{-1}$, and where ρ is the common correlation between binary responses

in the same historical series. Note that L_2 reduces to a familiar binomial likelihood product if $\gamma = \rho = 0$ and that the reparametrization from (a, b) to (α, γ) extends the model to include certain negative values of γ and ρ (Prentice, 1986). This parametrization also greatly facilitates likelihood maximization since L_2 expressed in terms of a and b tends to be highly skewed with the binomial special case ($\gamma = 0$), for example, occurring on the boundary $a + b = \infty$.

Standard asymptotic likelihood formulae apply to $l = \log L$. Appendix 1 provides the first and second derivatives of l with respect to $\tau = (\alpha, \beta, \gamma)$, and notes conditions for the asymptotic normality of the maximum likelihood estimate $\hat{\tau}$. A trend test may be obtained by evaluating $\partial l/\partial\beta$ at $\beta = 0$ and at $\hat{\alpha}_0$, $\hat{\gamma}_0$, the maximum likelihood estimates of α and γ at $\beta = 0$, giving

$$T = \sum x_{1i} d_i - \hat{\alpha}_0 \sum n_{1i} d_i.$$

The asymptotic variance of T, under $\beta = 0$, is consistently estimated, using the expressions of Appendix 1, by

$$V = \hat{\alpha}_0(1 - \hat{\alpha}_0) \sum n_{1i} d_i^2 - (\sum n_{1i} d_i)^2 \text{var } \hat{\alpha}_0,$$

where

$$\text{var } \hat{\alpha}_0 = -(\partial^2 l/\partial\gamma_0^2)\{(-\partial^2 l/\partial\alpha_0^2)(-\partial^2 l/\partial\gamma_0^2) - (-\partial^2 l/\partial\alpha_0\partial\gamma_0)^2\}^{-1}$$

estimates the asymptotic variance of $\hat{\alpha}_0$ under $\beta = 0$. Hence we propose

$$X_B^2 = T^2 V^{-1}$$

as an asymptotic χ_1^2 trend test that incorporates historical control data.

A Newton–Raphson iteration is convenient for the calculation of $(\hat{\alpha}_0, \hat{\gamma}_0)$. Trial values (α_0, γ_0) give rise to updated values (α_1, γ_1) via

$$\begin{pmatrix} \alpha_1 \\ \gamma_1 \end{pmatrix} = \begin{pmatrix} \alpha_0 \\ \gamma_0 \end{pmatrix} + \begin{pmatrix} -\partial^2 l/\partial\alpha_0^2 & -\partial^2 l/\partial\alpha_0\partial\gamma_0 \\ -\partial^2 l/\partial\alpha_0\partial\gamma_0 & -\partial^2 l/\partial\gamma_0^2 \end{pmatrix}^{-1} \begin{pmatrix} \partial l/\partial\alpha_0 \\ \partial l/\partial\gamma_0 \end{pmatrix} \tag{3}$$

with all quantities evaluated at $\beta = 0$. Suitable starting values are given by $\gamma_0 = 0$ and $\alpha_0 = \hat{\alpha}_0\,(\gamma = 0) = \hat{\alpha}_{00} = xn^{-1}$, where $x = x_1 + \cdots + x_s$, $n = n_1 + \cdots + n_s$. In this context $\gamma = 0$ represents a natural boundary condition under which the historical control data can merely be pooled with the current control data. Hence, even though the above iterative procedure may successfully identify values of $\hat{\gamma}_0 < 0$, we define $\hat{\gamma}_0 = 0$, $\hat{\alpha}_0 = \hat{\alpha}_{00}$, whenever $\partial l/\partial\gamma$ evaluated at $\gamma_0 = 0$, $\alpha_0 = \hat{\alpha}_{00}$, is equal to or less than zero, thereby avoiding the need for an iterative calculation in these circumstances. If $\hat{\gamma}_0 = 0$, $\hat{\alpha}_0 = \hat{\alpha}_{00}$, one obtains

$$T = T_0 = \sum x_{1i}\, d_i - \hat{\alpha}_{00} \sum n_{1i}\, d_i; \quad V = V_0 = \hat{\alpha}_{00}(1 - \hat{\alpha}_{00})\{\sum n_{1i}\, d_i^2 - (n_1 \bar{d})^2 n^{-1}\},$$

so that $X_B^2 = T_0^2 V_0^{-1}$ is then precisely the Cochran–Armitage statistic (1) following the pooling of current and historical controls. Note that, like X_{CA}^2, X_B^2 will have optimal local power properties against logistic alternatives, as well as against a wide range of other monotone dose response models.

Consider now an estimating equation approach to trend testing. The above binomial model for the current experiment specifies a mean vector $\boldsymbol{\mu}_1^T = (n_{10}p_0, \ldots, n_{1k}p_k)$ and covariance matrix $V_1 = \text{diag}(n_{10}p_0q_0, \ldots, n_{1k}p_kq_k)$ for the response vector $\mathbf{x}_1^T = (x_{10}, \ldots, x_{1k})$. As above, let us assume that x_j, the jth historical control response, has mean n_jp_0 and variance $V_j = n_jp_0q_0\{1 + (n_j - 1)\rho\}$, where ρ is the common pairwise correlation among the binary responses contributing to x_j, and that the response counts from each experiment are statistically independent. These assumptions alone, without the assumption of a specific parametric model for x_j ($j = 2, \ldots, s$), are enough to allow

estimation of α, β, and ρ. As elaborated in Appendix 1, the estimating equation approach of Liang and Zeger (1986) and Zeger and Liang (1986) now leads to a trend test statistic

$$T = \sum x_{1i}d_i - \hat{\alpha}_0 \sum n_{1i}d_i,$$

where

$$\hat{\alpha}_0 = \left[x_1 + \sum_2^s x_j\{1 + (n_j - 1)\hat{\rho}_0\}^{-1}\right] \Big/ \left[n_1 + \sum_2^s n_j\{1 + (n_j - 1)\hat{\rho}_0\}^{-1}\right]. \tag{4}$$

Under the assumed mean and variance structure, a generally consistent estimator of the variance of T is given by

$$V = \hat{\alpha}_0(1 - \hat{\alpha}_0)\left(\sum n_{1i}d_i^2 - (\sum n_{1i}d_i)^2\left[n_1 + \sum_2^s n_j\{1 + (n_j - 1)\hat{\rho}_0\}^{-1}\right]^{-1}\right). \tag{5}$$

The trend test development is completed by specifying $\hat{\rho}_0 = \hat{\rho}_0(\hat{\alpha}_0)$, an $s^{1/2}$-consistent estimator of ρ_0 in (4) and (5). Such a consistent estimator can be written

$$\hat{\rho}_0 = \sum_2^s \hat{w}_j\hat{S}_j \Big/ \sum_2^s \hat{w}_j, \tag{6}$$

where "$\hat{}$" denotes evaluation at $\hat{\alpha}_0$, and where

$$S_j = \{(x_j - n_jp_0)^2 - n_jp_0q_0\}\{n_j(n_j - 1)p_0q_0\}^{-1}$$

and

$$w_j = n_j(n_j - 1)^2p_0q_0\{1 + 2(n_j - 3)p_0q_0\}^{-1}.$$

A simple iteration between (4) and (6) allows $\hat{\alpha}_0$ and $\hat{\rho}_0$, and hence the asymptotic χ_1^2 estimating equation trend test statistic,

$$X_{\text{EQ}}^2 = T^2V^{-1},$$

to be calculated. Note that T reduces to T_0 and V to V_0 upon evaluation at $\hat{\rho}_0 = 0$ and $\hat{\alpha}_0 = \hat{\alpha}_{00}$. Therefore, in line with our handling of X_{B}^2 above, we define

$$X_{\text{EQ}}^2 = T_0^2V_0^{-1},$$

and set $\hat{\rho}_0 = 0$ and $\hat{\alpha}_0 = \hat{\alpha}_{00}$, whenever (6) evaluated at $p_0 = \hat{\alpha}_{00}$ is equal to or less than zero.

3. Illustrations and Simulation Results

Table 1 shows data from four rodent carcinogenesis experiments, each with a sizable number of historical control data sets. The first example involves alveolar–bronchiolar adenomas in mice. The Cochran–Armitage statistic takes value $X_{\text{CA}}^2 = 1.60$ with a corresponding (two-sided) significance level of $P = .21$. The new trend tests defined above have value $X_{\text{B}}^2 = 7.27$ ($P = .01$) and $X_{\text{EQ}}^2 = 7.25$ ($P = .01$). The inclusion of the historical data has led to much stronger evidence of trend in this example. In the calculation of X_{B}^2, the iterative procedure converged to within .0001 for both parameters in two iterations, giving $\hat{\alpha}_0 = .0979$, $\hat{\gamma}_0 = .0007$ ($\hat{a}_0 = 138.99$, $\hat{b}_0 = 1{,}279.81$). In the calculation of X_{EQ}^2, expression (6) evaluated at $\alpha = \hat{\alpha}_{00} = .0979$ was slightly negative so that we set $\hat{\alpha}_0 = \hat{\alpha}_{00}$ and $\hat{\gamma}_0 = 0$.

The second example involves mammary gland adenomas in male rats. Here the Cochran–Armitage statistic has the significant value of 7.16 ($P = .01$), while $X_{\text{B}}^2 = 5.37$ ($P = .02$), $X_{\text{EQ}}^2 = 5.39$ ($P = .02$). Compared to the previous example, there is a stronger internal

Table 1

Current study and historical control counts from the literature. Each entry gives the number of animals with tumor and the number of animals.

Example 1. Occurrence of alveolar–bronchiolar adenomas in mice following 102 weeks exposure to pivalolactone (Smythe et al., 1986)					
Current experiment:	Dose (mg/kg/day)	0	.75	1.50	
		2/20	6/49	10/49	
Historical controls:	0/20, 0/12, 1/19, 2/25, 1/10, 8/49, 0/20, 0/12, 1/19, 4/47, 6/54, 3/18, 0/19, 0/10, 1/17, 2/22, 3/20, 4/20, 0/17, 1/20, 1/15, 2/20, 3/20				
Example 2. Occurrence of mammary gland adenomas in male rats (Dempster et al., 1983)					
Current experiment:	Dose (ppm)	0	3	100	1,000
		3/55	3/57	5/60	10/55
Historical controls:	1/65, 1/47, 2/60, 2/47, 3/64, 3/55, 9/150, 7/107, 4/57, 6/64, 11/97, 12/100, 9/74, 5/41, 6/48, 7/50, 21/98, 13/60				
Example 3. Occurrence of follicular cell adenomas in selected studies from the NTP (Bickis and Krewski, 1989)					
Current experiment:	Dose (mg/kg/day)	0	.5	1.0	
		0/8	0/23	4/39	
Historical controls:	0/48, 0/42, 0/39, 0/23, 0/20, 0/20, 0/20, 0/20, 0/19, 0/18, 0/17, 0/17, 0/17, 0/17, 0/14, 0/13, 0/12, 0/11, 0/10, 0/9, 1/21, 1/14, 2/19				
Example 4. Occurrence of fibrosarcomas in selected studies from the NTP (Bickis and Krewski, 1989)					
Current experiment:	Dose (mg/kg/day)	0	.5	1.0	
		0/20	0/50	2/50	
Historical controls:	0/54, 0/33, 0/25, 0/25, 0/20, 0/20, 0/20, 0/20, 0/20, 0/20, 0/20, 0/20, 0/20, 0/20, 0/20, 0/15, 0/14, 0/14, 0/10, 0/9, 0/9, 0/8, 2/50, 2/20				

control in the current experiment, and greater overdispersion among the historical controls. As a result, the historical data have little influence on the trend test. The procedure used in the calculation of X^2_B gave $\hat{\alpha}_0 = .0936$, $\hat{\rho}_0 = \hat{\gamma}_0(1 + \hat{\gamma}_0)^{-1} = .0231$ in seven iterations, while that used in the calculation of X^2_{EQ} gave $\hat{\alpha}_0 = .0935$, $\hat{\rho}_0 = .0245$ in three iterations.

The next two examples involve extremely low tumor rates and therefore provide a test of the stability of the iterative procedures. These data were obtained from the National Cancer Institute/National Toxicology Program Carcinogenesis Bioassay Program (Bickis and Krewski, 1989). The Example 3 data (follicular cell adenomas) give $X^2_{CA} = 2.77$ ($P = .10$), $X^2_B = 13.77$ ($P < .001$), and $X^2_{EQ} = 17.66$ ($P < .001$). Hence, if a χ^2_1 approximation is appropriate with such extreme rates, the new tests again yield much stronger evidence of trend than does the Cochran–Armitage test. With these data the iterative procedure used in the calculation of X^2_B was slow to converge, taking ten iterations and giving $\hat{\alpha}_0 = .0325$ and $\hat{\rho}_0 = .2450$. In contrast, $\hat{\alpha}_0 = \hat{\alpha}_{00} = .0151$ and $\hat{\rho}_0 = 0$ was used in the calculation of X^2_{EQ} since (6) was slightly negative when evaluated at $\hat{\alpha}_{00}$. Evidently, estimates of α and ρ are highly correlated for this data set.

Example 4 (fibrosarcomas) gives $X^2_{CA} = 2.20$ ($P = .14$), $X^2_B = 3.54$ ($P = .06$), and $X^2_{EQ} = 3.41$ ($P = .06$). The new tests provide somewhat greater evidence of trend than does X^2_{CA}. The values of $(\hat{\alpha}_0, \hat{\rho}_0)$ used in calculating X^2_B and X^2_{EQ} are (.0105, .0805) and (.0096, .0183) with nine and two iterations required for convergence, respectively.

In order to study operating characteristics of these tests, simulation studies were carried out under the binomial beta–binomial model described above. Both the moderate tumor rate ($a = 6$, $b = 34$; $\alpha = 6/40 = .15$, $\gamma = (40)^{-1} = .025$) and the rare tumor rate ($a = 2$, $b = 100$; $\alpha = 2/102 = .0196$, $\gamma = (102)^{-1} = .0098$) cases used in the Tarone test simulations of Tamura and Young (1986) were considered. Calculations were carried out on a Compaq 386 Model 80 microcomputer. In each case, the current study had three dose levels, $d_i = 0, 1, 2$, with equal sample sizes at each dose. The binomial variates x_{1i} ($i = 0, \ldots, k$) and the beta-binomial variates x_j ($j = 2, \ldots, s$) were generated by comparing a uniform pseudorandom number to the cumulative distribution function for these distributions. Table 2 gives results of size simulations at selected values of the common current study sample size n_{1i}, the common historical control group sample size n_j, and the number of historical control series $s - 1$. Each entry of the table is based on 1,000 simulations.

The new tests X^2_B and X^2_{EQ} retain empirical sizes close to .05 even in the rare tumor situation, whereas X^2_{CA} is seriously conservative under rare tumor sampling conditions, as has been noted previously (Portier and Hoel, 1984). Corresponding trend test size estimation was also carried out for nominal 10%, 2%, and 1% level tests. Test size appeared to be well preserved by the new trend tests even at the 1% significance level in the moderate tumor case. Estimated sizes of the 2% and 1% tests were somewhat high in the rare tumor situation, in the vicinity of 3% for the nominal 2% test, and in the vicinity of 2% for the nominal 1% test, with X^2_{EQ} being slightly more liberal than X^2_B.

Calculation of X^2_B in the moderate tumor simulation led 16.2%, 5.9%, and .7% of the samples with 10, 20, and 40 historical control groups, respectively, to give $\partial l/\partial\gamma \leq 0$, when evaluated at $\alpha = \hat{\alpha}_{00}$, $\gamma = 0$. Otherwise, $\hat{\alpha}_0$ and $\hat{\gamma}_0$ were calculated iteratively, with convergence to within .0001 achieved in 15 or fewer iterations for all samples. Similarly, for X^2_{EQ}, 17.0%, 5.9%, and .8% of the samples having 10, 20, and 40 historical control

Table 2

Estimated trend test size under a binomial beta–binomial model. Each entry gives the fraction of 1,000 *simulated samples for which a trend statistic exceeds* 3.84, *the* 5% *critical value for a* χ^2_1 *distribution. Standard errors for estimated test sizes are approximately* .007.

Moderate tumor case ($\alpha = .15$, $\gamma = .025$)

	Historical control sample size = 40					
	Current sample size: 20 Number of historical control groups			Current sample size: 40 Number of historical control groups		
Trend statistic	10	20	40	10	20	40
X^2_{CA}	.057	.043	.063	.067	.043	.037
X^2_B	.047	.045	.039	.057	.039	.049
X^2_{EQ}	.047	.045	.039	.055	.038	.049

Rare tumor case ($\alpha = .0196$, $\gamma = .0098$)

	Historical control sample size = 40					
	Current sample size: 20 Number of historical control groups			Current sample size: 40 Number of historical control groups		
Trend statistic	10	20	40	10	20	40
X^2_{CA}	.010	.008	.002	.020	.019	.020
X^2_B	.057	.054	.050	.039	.049	.045
X^2_{EQ}	.060	.056	.051	.044	.050	.045

groups gave nonpositive values for (6) when evaluated at $\hat{\alpha}_{00}$. Otherwise, $\hat{\alpha}_0$ and $\hat{\rho}_0$ were computed iteratively, with convergence for all samples.

Nonconvergence was also not a serious issue in the rare tumor analyses of Table 2. Five of the 2,000 samples with 10 historical control series failed to give convergent results in the beta–binomial parameter iteration, and 8 of these 2,000 samples failed to yield convergent estimates in the estimating equation iteration. There was no nonconvergence among samples having 20 or 40 historical control groups. Note that a rather large fraction of samples give $\hat{\alpha}_0 = \hat{\alpha}_{00}$, $\hat{\gamma}_0 = \hat{\rho}_0 = 0$, in this rare tumor situation. For example, among the simulations having 10 historical control groups, 47% gave $\hat{\gamma}_0 = 0$ in the beta–binomial simulation, and 49% gave $\hat{\rho}_0 = 0$ in the estimating equation simulation. Hence, a beta–binomial iteration on the beta distribution parameters (a, b) rather than (α, γ) would be expected to give rise to very frequent nonconvergence under these sampling conditions.

The sampling distribution of $\hat{\alpha}_0$ from the beta–binomial iteration was very nearly symmetric about the true α in each configuration of Table 2. For example, the sample mean of $\hat{\alpha}_0$ in the moderate tumor situation was equal to its theoretical value of .150 in each of the six sampling configurations, while it equaled .020 or .021, as compared to the theoretical value of .020, in each of the rare tumor sampling configurations. The corresponding estimators of $\hat{\gamma}_0$ were nearly median unbiased in all sampling configurations in both the moderate and rare tumor situations, with only moderate skewness to the right even in the rare tumor setting. These desirable properties contrast sharply with those for $\hat{a}_0$ and $\hat{b}_0$ as discussed by Tamura and Young (1986, pp. 345–346). Properties of $\hat{\alpha}_0$ and $\hat{\rho}_0$ from the estimating equation simulation were similar to those from the beta–binomial maximum likelihood simulation.

Table 3 gives corresponding power calculations under the dose response model (2) with β values selected to give power in the vicinity of 45% or 75% for the new tests under each of the sample size configurations of Table 2. Five hundred simulations were carried out for each table entry with very few convergence problems. Each of the trend statistics X^2_{B} and X^2_{EQ} has considerably greater power under binomial beta–binomial sampling than does the Cochran–Armitage test. Estimated power may be very slightly greater for the estimating-equation-based statistic X^2_{EQ} than for the likelihood score test X^2_{B}, in spite of the local optimality of X^2_{B} under these sampling conditions.

4. Trend Testing with Overdispersion in the Current Study

The previous results apply to a repeated sampling situation in which the current study has a constant control group response rate of α, while historical control group response rates vary independently and randomly about α. In some contexts it may be natural, instead, to view the current and historical response rates as exchangeable, in which case one may assume current and historical response rates to be independent and random from some distribution. The current study response counts will then generally be overdispersed relative to binomial counts and the methods discussed above for incorporating historical control information may not apply. Because an assumption of control group exchangeability underlies the previous proposals for this historical control problem, we will go into some detail concerning possible trend tests in this situation.

In contrast to the assumptions of Section 2, suppose that the current study control response probability, θ_1, like $\theta_2, \ldots, \theta_s$ from the historical control series, arises as an independent variate from a distribution with mean α. For example, a beta distribution with parameters a and b, where $a(a + b)^{-1} = \alpha$ and $(a + b)^{-1} = \gamma$, leads to a likelihood function

$$L = L_1 L_2,$$

Table 3

Estimated trend test power under a binomial beta–binomial model. Each entry gives the fraction of 500 *simulated samples for which a trend statistic exceeds* 3.84, *the* 5% *critical value for a* χ^2_1 *distribution.* β *values were selected to give power in the range of* 45% *on the left side of the table and in the vicinity of* 75% *on the right side for new trend tests. Corresponding standard errors are approximately* .022 *and* .018, *respectively. Historical control groups are of the size* 40.

Moderate tumor case ($\alpha = .15$, $\gamma = .025$)

	$\beta = .4503$ Current sample size = 20 Hist. control groups			$\beta = .3303$ Current sample size = 40 Hist. control groups			$\beta = .6113$ Current sample size = 20 Hist. control groups			$\beta = .4536$ Current sample size = 40 Hist. control groups		
Trend statistic	10	20	40	10	20	40	10	20	40	10	20	40
X^2_{CA}	.21	.21	.20	.22	.21	.17	.36	.37	.40	.39	.39	.39
X^2_B	.44	.46	.49	.43	.45	.47	.72	.77	.80	.70	.73	.77
X^2_{EQ}	.45	.46	.49	.44	.45	.48	.73	.77	.80	.70	.73	.77

Rare tumor case ($\alpha = .0196$, $\gamma = .0098$)

	$\beta = .8400$ Current sample size = 20 Hist. control groups			$\beta = .6662$ Current sample size = 40 Hist. control groups			$\beta = 1.0440$ Current sample size = 20 Hist. control groups			$\beta = .8457$ Current sample size = 40 Hist. control groups		
Trend statistic	10	20	40	10	20	40	10	20	40	10	20	40
X^2_{CA}	.14	.14	.14	.21	.18	.20	.28	.29	.27	.28	.32	.34
X^2_B	.50	.60	.54	.45	.50	.53	.69	.77	.77	.66	.73	.78
X^2_{EQ}	.54	.60	.54	.49	.52	.54	.73	.77	.78	.70	.75	.78

where L_2 is as above (Section 2), whereas the likelihood L_1 from the current study can be written

$$L_1 = \int_0^1 (\theta_1)^{x_1+a-1}(1-\theta_1)^{n_1-x_1+b-1}\exp\left\{\beta \sum_0^k n_{1i}d_i\right\} \prod_0^k \{(1-\theta_1)+\theta_1 e^{\beta d_i}\}^{-n_{1i}}\beta(a,b)^{-1}\, d\theta_1,$$

where $\beta(a, b) = \int_0^1 \theta^{a-1}(1-\theta)^{b-1}\, d\theta$. Let $l = \log L$. A score test $T = \partial l/\partial\beta$, evaluated at $\beta = 0$ and $(\hat{a}_0, \hat{b}_0)$, or equivalently at $(\hat{\alpha}_0, \hat{\gamma}_0)$, the maximum likelihood estimators under $\beta = 0$, reduces to

$$T = \sum x_{1i}d_i - \tilde{p} \sum n_{1i}d_i,$$

where $\tilde{p} = (x_1 + \hat{a}_0)(n_1 + \hat{a}_0 + \hat{b}_0)^{-1}$, as was shown by Tarone (1982). This development assumes $\hat{a}_0 + \hat{b}_0$ to be finite. A natural extension of Tarone's statistic removes this requirement by reparametrization to (α, γ), giving

$$\tilde{p} = (x_1\hat{\gamma}_0 + \hat{\alpha}_0)(n_1\hat{\gamma}_0 + 1)^{-1},$$

and $\tilde{p} = \hat{\alpha}_0 = \hat{\alpha}_{00}$ if $\hat{\gamma}_0 = 0$. Hence we define $\hat{\gamma}_0 = 0$, $\tilde{p} = \hat{\alpha}_{00}$, and $T = T_0$ whenever $\partial l/\partial\gamma_0 \leqslant 0$ when evaluated at $\alpha_0 = \hat{\alpha}_{00}$, $\gamma_0 = 0$. Iterative calculation of $(\hat{\alpha}_0, \hat{\gamma}_0)$ can be carried out using the same type of Newton–Raphson iteration (3), with the same starting values, as was described in Section 2. A variance estimator for T is given by

$$V = -\partial^2 l/\partial\beta_0^2 - (n_1\bar{d})^2(n_1\hat{\gamma}_0 + 1)^{-2}\text{var}\ \hat{\alpha}_0,$$

giving rise to an overdispersed beta–binomial trend test statistic $X^2_{\text{OB}} = T^2V^{-1}$. The details for the calculation of X^2_{OB} are given in Appendix 2. In particular, if $\hat{\gamma}_0 = 0$ then V reduces to V_0 given above, so that we define $X^2_{\text{OB}} = T_0^2V_0^{-1}$ whenever $\partial l/\partial\gamma_0 \leqslant 0$, when evaluated at $\alpha_0 = \hat{\alpha}_{00}$, $\gamma_0 = 0$.

For comparison we define the Tarone trend test as $X^2_{\text{T}} = T^2/\{-\partial^2 l/\partial\beta_0^2\}$, since the denominator of this expression is the variance estimator for T proposed by Tarone (1982). The above development indicates that this trend statistic may be systematically too small because it ignores the variance reduction that derives from the estimation of the beta distribution parameters. The practical implications of this variance adjustment will be discussed further below. The denominator in X^2_{T} was defined to be $\hat{\alpha}_{00}(1-\hat{\alpha}_{00})\sum n_{1i}d_i^2$, the value of $-\partial^2 l/\partial\beta_0^2$ at $\hat{\gamma}_0 = 0$, whenever $\partial l/\partial\gamma_0 \leqslant 0$ at $\alpha_0 = \hat{\alpha}_{00}$, $\gamma_0 = 0$. Note that X^2_{T} differs from Tarone's proposal in using the current study, in addition to the historical studies, to estimate $(\hat{\alpha}_0, \hat{\gamma}_0)$. See also Margolin and Risko (1984) for an alternate approach to acknowledging sampling variation in the beta distribution parameters.

An estimating equation approach may also be considered with overdispersed response counts in the current study. For this purpose it is convenient to replace the rather complicated mean and variance structure implied by L_1 by $\boldsymbol{\mu}_1^{\text{T}} = (n_{10}p_0, \ldots, n_{1k}p_k)$, with $p_i = p(d_i)$ given by (2) and

$$V_1 = (1-\rho)\text{diag}(n_{10}p_0q_0, \ldots, n_{1k}p_kq_k) + \rho\boldsymbol{l}\boldsymbol{l}^{\text{T}},$$

where $\boldsymbol{l}^{\text{T}} = \{n_{10}(p_0q_0)^{1/2}, \ldots, n_{1k}(p_kq_k)^{1/2}\}$. This variance matrix arises from a common pairwise correlation among all the binary variates in the current study. As detailed in Appendix 2, the mean and variance structure just given for $\mathbf{x}_1$, along with that given in Section 2 for $x_2, \ldots, x_s$, gives a trend statistic

$$T = \sum x_{1i}d_i - \tilde{p} \sum n_{1i}d_i,$$

where

$$\tilde{p} = x_1n_1^{-1}[n_1\hat{\rho}_0\{1 + (n_1-1)\hat{\rho}_0\}^{-1}] + \hat{p}_0[(1-\hat{\rho}_0)\{1 + (n_1-1)\hat{\rho}_0\}^{-1}]$$

is a weighted linear combination of the observed response rate $x_1 n_1^{-1}$ in the current study, and the overall estimated control group response rate $\hat{p}_0 = \hat{\alpha}_0$. Note that $\tilde{p} = \hat{\alpha}_0 = \hat{\alpha}_{00}$ at $\hat{\rho}_0 = 0$ and $\tilde{p} \to x_1 n_1^{-1}$ as $\hat{\rho}_0 \to 1$. Note also that

$$\hat{\alpha}_0 = \hat{\alpha}_0(\hat{\rho}_0) = \frac{\sum_1^s x_j\{1 + (n_j - 1)\hat{\rho}_0\}^{-1}}{\sum_1^s n_j\{1 + (n_j - 1)\hat{\rho}_0\}^{-1}} \tag{7}$$

so that $\hat{\alpha}_0 = \hat{\alpha}_{00} = xn^{-1}$ if $\hat{\rho}_0 = 0$ and $\hat{\alpha}_0 \to s^{-1} \sum x_j n_j^{-1}$ as $\hat{\rho}_0 \to 1$. A simple iteration between (7) and (6), with j ranging from 1 to s and with x_1 and n_1 replaced by x_{10} and n_{10} in (6), can be used to calculate $\hat{\rho}_0$ and $\hat{\alpha}_0$. A variance estimator for T is given by

$$V = \hat{\alpha}_0(1 - \hat{\alpha}_0)\left\{(1 - \hat{\rho}_0)\sum_0^k n_{1i}d_i^2 - \{1 + (n_1 - 1)\hat{\rho}_0\}^{-1}(n_1\bar{d})^2 \times \left(\hat{\rho}_0(1 - \hat{\rho}_0) + (1 - \hat{\rho}_0)^2\{1 + (n_1 - 1)\hat{\rho}_0\}^{-1}\left[\sum_1^s n_j\{1 + (n_j - 1)\hat{\rho}_0\}^{-1}\right]^{-1}\right)\right\}.$$

Hence we define an overdispersed estimating equation trend test by

$$X_{\mathrm{OE}}^2 = T^2 V^{-1}.$$

Since T reduces to T_0 and V reduces to V_0 at $\hat{\rho}_0 = 0$, we again define $X_{\mathrm{OE}}^2 = T_0^2 V_0^{-1}$ if (6), with j ranging from 1 to s, evaluated at $\alpha_0 = \hat{\alpha}_{00}$ and $\gamma_0 = 0$ is equal to or less than zero. Under $\beta = 0$ the beta–binomial model of this section yields the mean and variance structure assumed in this estimating equation approach. Hence one expects X_{OE}^2 to yield a valid trend test under the beta–binomial model.

Consider now the asymptotic distribution of the trend test statistics of this section while assuming that current and historical response rates $\theta_1, \ldots, \theta_s$ are independent observations from a mixing density f. Conditional on θ_1, the square root, X_{CA}, of the Cochran–Armitage statistic quite generally has a standard normal distribution as $n_1 \to \infty$. It follows that the unconditional asymptotic distribution is also standard normal, since

$$\Pr(X_{\mathrm{CA}} \leq x) \approx \int_0^1 \Phi(x) f(\theta)\, d\theta = \Phi(x), \tag{8}$$

where $\Phi(x)$ is the standard normal distribution function. Hence X_{CA}^2 retains an asymptotic χ_1^2 distribution under this overdispersed model for the current study. As shown in Appendix 2, the numerators of the trend statistics X_{OB}^2, X_{T}^2, and X_{OE}^2 are each asymptotically equivalent to that for X_{CA}^2 assuming sn_1^{-1} approaches a nonzero constant, so that the historical control data are playing a comparatively minor role in this overdispersed model, a role that becomes negligible as $n_1 \to \infty$. Since X_{OB}^2 and X_{T}^2 are standardized by an "observed" information variance pertinent to the current study control response rate, they too generally have asymptotic χ_1^2 distributions. In contrast, since X_{OE} is standardized by an "expected" variance formula, it has a more complicated asymptotic distribution which, under a beta–binomial mixing distribution, can be written

$$\Pr(X_{\mathrm{OE}} \leq x) \approx \beta(a, b)^{-1} \int_0^1 \Phi(x/\sigma_{\theta_1})\theta_1^{a-1}(1 - \theta_1)^{b-1}\, d\theta_1, \tag{9}$$

where $\sigma_\theta^2 = \theta(1 - \theta)\{a(1 - \alpha)(1 + \gamma)^{-1}\}$. Hence X_{OE}^2 does not have an asymptotic χ_1^2 distribution for $\gamma \neq 0$.

5. Illustrations and Simulation Results

Consider the application of the trend tests of Section 4 to the four examples of Table 1. All iterative routines converged for each example. The first example (alveolar–bronchiolar

adenoma), gives $X^2_{OB} = 4.02$ ($P = .04$), $X^2_{T} = 3.78$ ($P = .05$), and $X^2_{OE} = 7.25$ ($P = .01$), as compared to $X^2_{CA} = 1.60$ ($P = .21$), with all significance levels based on a χ^2_1 assumption. These values for X^2_{OB} and X^2_{T} are intermediate to X^2_{CA} and X^2_{B} of Section 2, while X^2_{OE} is identical to X^2_{EQ} of Section 2 owing to $\hat{\rho}_0 = 0$. Estimates of $(\hat{\alpha}_0, \hat{\rho}_0)$ used in the calculation of X^2_{OB} (and X^2_{T}) and X^2_{OE} are, respectively, (.9068, .0100) and (.0979, 0). The second example (mammary adenomas) gives $X^2_{OB} = 6.67$ ($P = .01$), $X^2_{T} = 6.63$ ($P = .01$), and $X^2_{OE} = 6.79$ ($P = .01$), all rather close to $X^2_{CA} = 7.16$ ($P = .01$). Estimates of $\hat{\alpha}_0$ and $\hat{\rho}_0$ used in calculating X^2_{OB} and X^2_{OE} are (.0940, .0208) and (.0939, .0236), respectively.

The rare tumor third example (follicular cell adenoma) gives $X^2_{OB} = 3.87$ ($P = .05$), $X^2_{T} = 3.83$ ($P = .05$), and $X^2_{OE} = 17.66$ ($P < .001$) as compared to $X^2_{CA} = 2.77$ ($P = .10$). The overdispersed tests X^2_{OB} and X^2_{T} indicate somewhat greater evidence of trend than does X^2_{CA}, but considerably less so than does X^2_{B} of Section 2, while X^2_{OE} is identical to X^2_{EQ}. Estimates of α_0 and ρ_0 used in X^2_{OB} and X^2_{OE} are, respectively, (.0128, .0367) and (.0151, 0). Similarly, the rare tumor fourth example (fibrosarcoma) gives $X^2_{OB} = 2.24$ ($P = .14$), $X^2_{T} = 2.18$ ($P = .14$), and $X^2_{OE} = 4.35$ ($P = .04$), as compared to $X^2_{CA} = 2.20$ ($P = .14$). The values of $(\hat{\alpha}_0, \hat{\rho}_0)$ used in X^2_{OB} and X^2_{OE} are, respectively, (.0081, .0177) and (.0077, .0131).

In each of the four examples X^2_{T} was only slightly smaller than X^2_{OB}, suggesting that acknowledging the sampling variation in $(\hat{\alpha}_0, \hat{\rho}_0)$ may be unimportant, at least if the number of historical control groups $(s - 1)$ is fairly large. The test X^2_{OE} exceeded X^2_{OB} in each example. Inappropriately large values of X^2_{OE} could reflect departures from a χ^2_1 distribution under the null hypothesis.

Consider now the performance of the above trend statistics under the overdispersed beta–binomial sampling model of Section 4. Table 4 presents estimates of test size for the same sample configurations as in Table 2. The response counts $\mathbf{x}_1$ for the current experiment were generated in this null hypothesis situation by first generating a beta–binomial variate $x_1 = x_{10} + \cdots + x_{1k}$ as above and then generating each element of $\mathbf{x}_1$, except the last, using the marginals from the hypergeometric distribution for $\mathbf{x}_1$ given x_1. Once again the iterative procedures described above were successful in virtually all simulated samples. The first four trend statistics listed in the upper part (moderate tumor) of Table 4 are intended to apply to the overdispersed model from which the data were generated. X^2_{OB} and X^2_{T} give empirical sizes close to the nominal .05 level in the moderate tumor situation, while X^2_{OE} and X^2_{CA} appear to be somewhat liberal, with estimated sizes mostly in the 6%–9% range, under these sampling conditions. Numerical integration using (9) gives an actual asymptotic probability of .0525 that X^2_{OE} exceeds the χ^2_1 critical value of 3.84, so that such liberality is not explained by departure of the asymptotic distribution from χ^2_1. The lack of adjustment of the variance of X^2_{T} for sampling variation in $(\hat{\alpha}_0, \hat{\rho}_0)$ appears to be unimportant, particularly if there are 20 or more historical control groups. Note that the size estimates given here for our extension of Tarone's test do not agree with the Tarone statistic size simulations of Tamura and Young (1986, Table 2b), who reported size estimates in the 6%–10% range for these table entries. Presumably the inability of Tamura and Young's iterative procedures to converge whenever $\hat{\gamma}_0 = 0$ $(\hat{a}_0 + \hat{b}_0 = \infty)$ has adversely affected their simulation results. The final two trend statistics are designed for binomial responses in the current experiment. As expected, these tests do not adequately preserve test size under the simulated overdispersion.

The lower part of Table 4 gives similar trend test size calculations in the rare tumor situation ($\alpha = .0196$, $\gamma = .0098$). Nominal χ^2_1 critical values were exceeded far too often for each of X^2_{OB}, X^2_{T}, and X^2_{OE}. The liberality of X^2_{OE} is not explained by departure of the asymptotic distributions from χ^2_1 in that the true asymptotic significance level, using (9), is .0586 in this rare tumor situation. The Cochran–Armitage statistic, on the other hand, is substantially conservative at the smaller sample size ($n_{1i} = 20$). These simulation results are quite different from those of Tamura and Young (1986, Table 2a), who, using the

Table 4

Estimated trend test size under certain overdispersed beta-binomial models for the current study. Each entry gives the fraction of 1,000 *simulated samples for which a trend statistic exceeds* 3.84, *the* 5% *critical value for a* χ^2_1 *distribution. Standard errors for the estimated test sizes are approximately* .007.

Moderate tumor case ($\alpha = .15$, $\gamma = .025$)

	Historical control sample size: 40					
	Current sample size: 20 Number of historical control groups			Current sample size: 40 Number of historical control groups		
Trend statistic	10	20	40	10	20	40
X^2_{OB}	.043	.054	.068	.040	.056	.042
X^2_{T}	.033	.051	.068	.037	.056	.042
X^2_{OE}	.070	.072	.081	.075	.088	.059
X^2_{CA}	.086	.071	.086	.049	.068	.063
X^2_{B}	.110	.121	.135	.135	.183	.209
X^2_{EQ}	.107	.120	.136	.112	.178	.210

Rare tumor case ($\alpha = .0196$, $\gamma = .0098$)

	Historical control sample size: 40					
	Current sample size: 20 Number of historical control groups			Current sample size: 40 Number of historical control groups		
Trend statistic	10	20	40	10	20	40
X^2_{OB}	.119	.119	.112	.168	.139	.119
X^2_{T}	.103	.114	.107	.134	.129	.113
X^2_{OE}	.123	.118	.125	.141	.150	.107
X^2_{CA}	.017	.015	.012	.040	.053	.041
X^2_{B}	.116	.117	.135	.140	.154	.141
X^2_{EQ}	.122	.120	.136	.144	.160	.144

Tarone statistic, did not see evidence of liberality in this rare tumor situation. Once again this discrepancy presumably relates to their manner of handling samples in which $\hat{\gamma}_0 = 0$ which, for example, amounted to about 45% of samples with 10 historical control data sets in this rare tumor situation. The trend tests based on binomially distributed response counts in the current study (X^2_B and X^2_{EQ}) are again substantially liberal.

Table 5 gives comparative power estimates for tests having close to the desired size under the overdispersed model of Section 4, using the same β values as in Table 3. Here $\mathbf{x}_1$ was generated by selecting a pseudorandom control rate θ_1 from a beta distribution with parameters a and b, followed by the generation of x_{1i} as a binomial variate with mean p_i given by (2) with $\alpha = \theta_1$. Each entry in Table 5 is based on 500 simulated samples. In view of the size estimates of Table 4, only the moderate tumor situation is displayed, and only the trend tests designed to accommodate this overdispersion are listed. X^2_{OB} and X^2_T exhibit modest increases in power relative to X^2_{CA}, but there appears to be little reason to include more than 10 historical control groups under these sampling conditions. X^2_{OE} exhibits somewhat larger empirical power than does the locally optimal test X^2_{OB}, but this may reflect a possible lesser ability to adequately preserve test size, as was shown in Table 4.

For completeness it can be noted that the trend tests of Table 5 were also studies under the binomial sampling conditions of Section 2. Each test was noticeably conservative with estimated sizes mostly in the 2%–4% range, with greater conservatism in the moderate

Table 5

Estimated trend test power under certain overdispersed beta–binomial models for the current study. Each entry gives the fraction of 500 simulated samples for which a trend statistic exceeds 3.84, the 5% critical value for a χ^2_1 distribution. β values are as in Table 3. Standard errors are in the vicinity of .02. Historical control groups are of size 40.

	Moderate tumor case ($\alpha = .15$, $\gamma = .025$)											
	$\beta = .4503$ Current sample size = 20 Hist. control groups			$\beta = .3303$ Current sample size = 40 Hist. control groups			$\beta = .6113$ Current sample size = 20 Hist. control group			$\beta = .4536$ Current sample size = 40 Hist. control groups		
Trend statistic	10	20	40	10	20	40	10	20	40	10	20	40
X^2_{OB}	.30	.31	.33	.25	.28	.25	.47	.49	.54	.46	.41	.46
X^2_T	.28	.30	.32	.24	.27	.25	.46	.48	.54	.45	.40	.46
X^2_{OE}	.40	.42	.40	.37	.37	.33	.58	.58	.64	.57	.52	.55
X^2_{CA}	.20	.19	.21	.19	.24	.20	.38	.36	.35	.39	.35	.38

tumor situation. As an exception, the estimating equation statistic X^2_{OE} was alternately conservative and liberal in the moderate and rare tumor situations. Correspondingly, power estimates for the overdispersed trend tests were lower than those shown in Table 3 for X^2_B and X^2_{EQ}. Typically if the power of these tests was in the vicinity of 45% or 75% the corresponding estimated power for each of X^2_{OB}, X^2_T, and X^2_{OE} was in the vicinity of 30% or 60%, respectively.

6. Discussion

Two new trend tests have been proposed that incorporate historical control data into the binomially-based model that underlies the standard Cochran–Armitage statistic. One statistic (X^2_B) assumes that the historical control data arise as independent observations from a beta–binomial distribution with mean response equal to the control group response in the current study. The second (X^2_{EQ}) makes the weaker assumption that the binary variates contributing to a historical control response count have a common pairwise correlation. Both trend tests have asymptotic χ^2_1 distributions under a binomial beta–binomial sampling model and both exhibit attractive size and power characteristics under simulation conditions (Tables 2 and 3) in both moderate and rare tumor situations. Hence we think that these statistics can quite generally be advocated for testing trend in proportions. The estimating equation statistic may possess a greater robustness to departure of the historical control counts from a beta–binomial distribution, but the two statistics are likely virtually interchangeable for practical purposes. Note also that either the binomial beta–binomial model that underlies X^2_B or the corresponding mean and variance model that underlies X^2_{EQ} gives rise to convenient estimation as well as testing of the dose response parameter. Additional covariates may also be added to (2), for example to control confounding, and the pairwise correlation (ρ) among historical control responses could be allowed to vary among groups of historical studies, or otherwise to depend on covariates (e.g., Prentice, 1988; Zhao and Prentice, 1990).

In the event that it is thought necessary to allow the response counts in the current study to be overdispersed, the trend testing and estimation problem is more complex. Assuming current and historical control rates are exchangeable, various trend test statistics can be developed. One of these, X^2_{OB}, is a score test under a beta–binomial model for the current experiment. X^2_{OB} is a refinement of Tarone's test X^2_T that is intended to remove test conservatism that arises if the number of historical control series is moderate. This conservatism disappears asymptotically and appeared to be unimportant in our simulation studies. Another test, X^2_{OE}, was based on estimating equations with a somewhat different current study mean response vector than arises from the beta–binomial model. X^2_{OE} has a rather complicated asymptotic distribution that differs somewhat from χ^2_1. These tests did not preserve test size relative to χ^2_1 critical values in our rare tumor simulation model (Table 4), while X^2_{OE} may also be liberal in the moderate tumor case. Hence these tests should be cautiously applied, pending further development.

The statistics X^2_{OB}, X^2_T, and X^2_{OE} use overdispersion parameter estimates based on an analysis of historical control and the current study data, assuming $\beta = 0$. A parallel set of estimators is readily specified by restricting the overdispersion parameter estimation to the historical control data. Such restriction may yield slightly improved power properties, particularly for distant alternatives, since departure from $\beta = 0$ in the current experiment may lead to overestimates of dispersion. For example, for the first data set of Table 1, the value of 3.78 for the Tarone statistic is increased to 6.45 upon excluding the current experiment from the estimation of $(\hat{\alpha}_0, \hat{\gamma}_0)$.

Note also that other classes of estimation procedures, both under binomial and overdispersed models for the current study, can readily be developed by replacing the logistic model (2) by other dose response functions—for example, the exponential function considered by Hoel (1983).

The estimating equation approach to this type of problem, with or without allowance for current study overdispersion, appears to merit emphasis and further development. In spite of a rather simple approach to pairwise correlation (ρ) estimation, specifically using a variance-weighted (at $\rho = 0$) linear combination of correlation estimates from individual control groups, the iterative procedure for calculating $(\hat{\alpha}_0, \hat{\rho}_0)$ was highly stable and the corresponding trend tests appear to have attractive power properties. More sophisticated approaches to correlation parameter estimation may yield further improvements.

The principal assumption linking the current to historical studies in Section 2 is that the current study control response rate is equal to the expected control response rate for the historical control series. One could consider testing this assumption by relaxing the current study control response rate to $\alpha + \Delta$ and testing $\Delta = 0$. For example, an asymptotic χ_1^2 score test for $\Delta = 0$, using only the control data from the current experiment in conjunction with the historical control data, can be written

$$(x_{10} - n_{10}\hat{\alpha}_0)^2/\{n_{10}\hat{\alpha}_0(1 - \hat{\alpha}_0)C\},$$

where $C = 1 - n_{10}\text{var}\,\hat{\alpha}_0/\{\hat{\alpha}_0(1 - \hat{\alpha}_0)\}$. The estimating equation approach of Section 2 gives an asymptotic χ_1^2 statistic of the same form with

$$C = \left[n_{10} + \sum_2^s n_j\{1 + (n_j - 1)\hat{\rho}_0\}^{-1}\right] \Big/ \sum_2^s n_j\{1 + (n_j - 1)\hat{\rho}_0\}^{-1},$$

where $\hat{\alpha}_0$ and $\hat{\rho}_0$ are calculated as in Section 2, using only current and historical control data. However, these tests are unlikely to be very powerful in situations in which one would consider incorporating historical control information—that is, when x_{10} and n_{10} are small. For example, for the four data sets of Table 1, the maximum likelihood and estimating-equation-based statistics take values (.063, .064), (.941, .931), (.076, .073), and (.134, .134), respectively. Instead, thorough knowledge of the origins of the current study data and of the degree of comparability of the current and historical experiments will be required to justify the use of historical control data in the manner described in Section 2. Appropriate historical control groups are those that are expected to have the same control group response rate as the current experiment. Known differences in study conditions that may be pertinent to control response rates should lead to the exclusion of the historical series or to specific accommodation in the analysis.

The circumstances in which one would use the overdispersed model of Section 4 are less clear. This model replaces the binomial assumption of Section 2 by requiring a common degree of overdispersion in the current and historical control series. This assumption could be examined empirically, for example, by testing the need to relax the marginal mean α for the current study control group to $\alpha + \Delta$, but such examination would presumably be of limited power in situations of practical interest. If knowledge of the current and historical experiments suggests exchangeability rather than commonality of control response rates, then the assumption of Section 4 may be more natural than those of Section 2. However, such an assumption implies a much reduced contribution of the historical data to trend testing and estimation.

ACKNOWLEDGEMENTS

This work was supported by National Institutes of Health Grants GM24472/CA53996 and CA-47658 (R. Prentice) and by Grant A-8664 from the National Sciences and Engineering Research Council (D. Krewski).

RÉSUMÉ

Pour les cas où on dispose de données historiques de contrôle, on propose de nouveaux tests de tendance appliqués à des proportions. Un de ces tests est un simple test du score utilisant une vraisemblance binomiale pour les données en cours d'étude et une vraisemblance béta–binomiale pour les données historiques. Un test très voisin, basé sur des équations d'estimation, est aussi proposé. On développe également des statistiques de tendance qui tiennent compte d'une surdispersion des proportions observées dans l'étude en cours. Elles incluent un test de Tarone (1982, *Biometrics* **38,** 215–220) qui prend en compte une fluctuation d'échantillonnage des paramètres des lois de distribution béta. Un test de tendance, basé sur des équations d'estimation, est aussi présenté. Pour des conditions d'échantillonnage de l'étude en cours aussi bien binomiales que béta–binomiales, on fait un examen en termes de taille et de puissance de chacun de ces tests. Des cas d'illustration sont présentés.

REFERENCES

Bickis, M. and Krewski, D. (1989). Statistical issues in the analysis of long-term carcinogenicity bioassay in small rodents: An empirical evaluation of statistical decision rules. *Fundamental and Applied Toxicology* **12,** 202–221.

Dempster, A. P., Selwyn, M. D., and Weeks, B. J. (1983). Combining historical and randomized controls for assessing trends in proportion. *Journal of the American Statistical Association* **78,** 221–227.

Hoel, D. G. (1983). Conditional two-sample tests with historical controls. In *Contributions to Statistics: Essays in Honour of Norman L. Johnson and P. K. Sen*, 229–236. Amsterdam: North-Holland.

Hoel, D. G. and Yanagawa, T. (1986). Incorporating historical controls in testing for a trend in proportions. *Journal of the American Statistical Association* **396,** 1095–1099.

Krewski, D., Smythe, R. T., Dewanji, A., and Colin, D. (1988). Statistical tests with historical controls. In *Carcinogenicity: The Design, Analysis and Interpretation of Long-Term Animal Studies*, H. C. Price and J. L. Ciminera (eds), 23–28. New York: Springer-Verlag.

Liang, K. Y. and Zeger, S. L. (1986). Longitudinal data analysis using generalized linear models. *Biometrika* **73,** 13–22.

Margolin, B. H. and Risko, K. J. (1984). The use of historical controls in the analysis of laboratory studies. In *Proceedings of 12th International Biometrics Conference*, 21–30. Alexandria, Virginia: The Biometric Society.

Portier, C. and Hoel, D. (1984). Type I error rate of trend tests in proportions and the design of cancer screens. *Communications in Statistics, Series A* **13,** 1–14.

Prentice, R. L. (1986). Binary regression using an extended beta–binomial distribution, with discussion of correlation induced by covariate measurement errors. *Journal of the American Statistical Association* **394,** 321–327.

Prentice, R. L. (1988). Correlated binary regression with covariates specific to each binary observation. *Biometrics* **44,** 1033–1048.

Smythe, R. T., Krewski, D., and Murdoch, D. (1986). The use of historical controls in modelling dose–response relationships in carcinogenesis. *Statistics and Probability Letters* **4,** 87–93.

Tamura, R. N. and Young, S. S. (1986). The incorporation of historical control information in tests of proportions: Simulation of Tarone's procedure. *Biometrics* **42,** 343–350.

Tarone, R. E. (1982). The use of historical control information in testing for a trend in proportions. *Biometrics* **38,** 215–220.

Tarone, R. E. and Gart, J. J. (1980). On the robustness of combined tests for trend in proportions. *Journal of the American Statistical Association* **75,** 110–116.

Zeger, S. L. and Liang, K. Y. (1986). Longitudinal data analysis for discrete and continuous outcomes. *Biometrics* **42,** 121–130.

Zhao, L. P. and Prentice, R. L. (1990). Correlated binary regression using a quadratic exponential model. *Biometrika* **77**, 642–648.

Received June 1989; *revised April* 1990; *accepted January* 1991.

Appendix 1

Likelihood and Estimating Equation Inference with a Binomial Model for the Current Experiment

Testing and estimation on $\tau = (\alpha, \beta, \gamma)$ are readily carried out using the log-likelihood function $l = \log L = \log L_1 + \log L_2$ of Section 2. In particular, maximum likelihood estimates of $\hat{\tau} = (\hat{\alpha}, \hat{\beta}, \hat{\gamma})$ are solutions to

$$\frac{\partial l}{\partial \alpha} = (x_1 - \sum n_{1i} p_i)\alpha^{-1}(1-\alpha)^{-1} + \sum_2^s \left\{ \sum_0^{x_j-1} (\alpha + \gamma i)^{-1} - \sum_0^{n_j-1} (1 - \alpha + \gamma i)^{-1} \right\} = 0,$$

$$\frac{\partial l}{\partial \beta} = \sum x_{1i} d_i - \sum n_{1i} p_i d_i = 0,$$

$$\frac{\partial l}{\partial \gamma} = \sum_2^s \left\{ \sum_0^{x_j-1} (\alpha + \gamma i)^{-1} i + \sum_0^{n_j-x_j-1} (1 - \alpha + \gamma i)^{-1} i - \sum_0^{n_j-1} (1 + \gamma i)^{-1} i \right\} = 0.$$

Note that $n^{1/2}(\hat{\tau} - \tau)$ quite generally has an asymptotic normal distribution with mean zero and with variance consistently estimated by $n_1 I^{-1}(\hat{\tau})$, where $I(\tau) = -\partial^2 l / \partial\tau \partial\tau^{\mathrm{T}}$ has entries

$$-\frac{\partial^2 l}{\partial \alpha^2} = \{\sum n_{1i} p_i q_i + (x_1 - \sum n_{1i} p_i)(1 - 2\alpha)\}\alpha^{-2}(1-\alpha)^{-2} + \sum_2^s \left\{ \sum_0^{x_j-1} (\alpha + \gamma i)^{-2} + \sum_0^{n_j-x_j-1} (1 - \alpha + \gamma i)^{-2} \right\},$$

$$-\frac{\partial^2 l}{\partial \alpha \partial \beta} = \sum n_{1i} d_i p_i q_i \alpha^{-1}(1-\alpha)^{-1},$$

$$-\frac{\partial^2 l}{\partial \alpha \partial \gamma} = \sum_2^s \left\{ \sum_0^{x_j-1} (\alpha + \gamma i)^{-2} i - \sum_0^{n_j-x_j-1} (1 - \alpha + \gamma i)^{-2} i \right\},$$

$$-\frac{\partial^2 l}{\partial \beta^2} = \sum n_{1i} d_i^2 p_i q_i, \quad -\frac{\partial^2 l}{\partial \beta \partial \gamma} = 0,$$

$$-\frac{\partial^2 l}{\partial \gamma^2} = \sum_2^s \left\{ \sum_0^{x_j-1} (\alpha + \gamma i)^{-2} i^2 + \sum_0^{n_j-x_j-1} (1 - \alpha + \gamma i)^{-2} i^2 - \sum_0^{n_j-1} (1 + \gamma i)^{-2} i^2 \right\}.$$

Standard conditions $n_1 \to \infty$, $\sum n_{1i}(d_i - \bar{d})^2 \to \infty$, where $\bar{d} = (\sum n_{1i} d_i) n_1^{-1}$, will ensure the usual asymptotic results for (α, β) estimation from L_1, while requiring $s n_1^{-1}$ to approach a constant will ensure such results for (α, γ) estimation from L_2. In fact, it would be technically possible to carry out standard likelihood inference even if $\sum n_{1i}(d_i - \bar{d})^2$ is bounded, provided $\bar{d} \neq 0$, as would occur, for example, if the current experiment had a single dose level and no control. Such inference would, however, depend very strongly on the suitability of the historical control modelling assumptions. An asymptotic χ_1^2 distribution evidently also holds for X_{B}^2 as $n_1 \to \infty$ even if the number of control groups $(s-1)$ is bounded provided the control group sizes (n_j) are also bounded. In this case the contribution of the historical series to X_{B}^2 is asymptotically negligible.

Consider now inference using estimating equations under the model of Section 2. The results of Liang and Zeger (1986) and the expressions for $(\boldsymbol{\mu}_j, V_j)$, $j = 1, \ldots, s$, given in Section 2 imply that, for specified ρ, $\boldsymbol{\delta}^{\mathrm{T}} = (\delta_0, \delta_1)$ can quite generally be estimated as the solution $\hat{\boldsymbol{\delta}}$ to

$$\sum_{j=1}^s D_j^{\mathrm{T}} V_j^{-1} (\mathbf{x}_j - \boldsymbol{\mu}_j) = 0, \tag{A1}$$

as $n_1 \to \infty$ and $s \to \infty$, where $\delta_0 = \log\{\alpha(1-\alpha)^{-1}\}$, $\delta_1 = \beta$, and $D_j^T = \partial \mu_j^T / \partial \delta$. These equations simplify to

$$\sum_{i=0}^{k} (x_{1i} - n_{1i}p_i) + \sum_{j=2}^{s} (x_j - n_j p_0)\{1 + (n_j - 1)\rho\}^{-1} = 0,$$

$$\sum_{i=0}^{k} d_i(x_{1i} - n_{1i}p_i) = 0. \tag{A2}$$

The results of Liang and Zeger show that $s^{1/2}(\hat{\delta} - \delta)$ will generally have an asymptotic normal distribution with mean zero, as $s \to \infty$ and $n_1 \to \infty$, even if ρ is replaced by an $s^{1/2}$-consistent estimator $\hat{\rho}(\hat{\delta})$. The requirement that $n_1 \to \infty$ arises since β appears in the mean and variance expressions only for the current experiment. The left side of the second equation in (A2), evaluated at $\beta = 0$, gives the trend statistic T, while solution of the first equation gives expression (4) for $\hat{\alpha}_0$. Under the assumed mean and variance expressions,

$$s\left(\sum_{1}^{s} D_j^T V_j^{-1} D_j\right)^{-1} \tag{A3}$$

provides a consistent estimator of the variance of $s^{1/2}(\hat{\delta} - \delta)$, where all quantities are evaluated at $\hat{\delta}$ and $\hat{\rho}(\hat{\delta})$, from which variance estimator (5) can be derived. The $s^{1/2}$-consistent estimator $\hat{\rho}_0 = \hat{\rho}_0(\hat{\alpha}_0)$, given by (6), arises by noting that each S_j $(j = 2, \ldots, s)$ has expectation ρ, so that $\sum w_j S_j / \sum w_j$ also has expectation ρ, for suitably defined weights w_j $(j = 2, \ldots, s)$. The weights w_j given in Section 2 are the reciprocals of the variances of S_j $(j = 2, \ldots, s)$ at $\rho = 0$, and hence should be close to optimal if ρ is close to zero.

A suitable iterative procedure for the calculation of $\hat{\delta}$ solving (A1) replaces a trial value δ_0 by an updated value δ_1 via

$$\delta_1 = \delta_0 + \left(\sum_{j=1}^{s} D_j^T V_j^{-1} D_j\right)^{-1} \left\{\sum_{j=1}^{s} D_j^T V_j^{-1}(\mathbf{x}_j - \mu_j)\right\},$$

with all quantities on the right side evaluated at δ_0, $\hat{\rho}_0(\delta_0)$. Iteration between this expression and (6) then gives $(\hat{\delta}, \hat{\rho})$.

Appendix 2

Likelihood and Estimating Equation Inference with Overdispersed Counts in the Current Experiment

Consider the log-likelihood function $l = \log L$, where $L = L(\alpha, \gamma, \beta)$ is given in Section 4. Following interchange of integration and differentiation in L_1 and evaluation at $\beta = 0$, $\partial l/\partial\gamma$, $-\partial^2 l/\partial\alpha\partial\gamma$, and $-\partial^2 l/\partial\gamma^2$ are exactly as given in Appendix 1 with j ranging from 1 to s. Similarly, $\partial l/\partial\alpha$ and $-\partial^2 l/\partial\alpha^2$ are obtained by dropping the contribution from L_1 to the expressions given in Appendix 1 and extending the range of j from 1 to s. Following some algebraic simplification, the other elements of the observed information matrix at $\hat{\alpha}_0$, $\hat{\gamma}_0$, and $\beta_0 = 0$ can be written

$$-\frac{\partial^2 l}{\partial\hat{\alpha}_0\partial\beta_0} = n_1\bar{d}(n_1\hat{\gamma}_0 + 1)^{-1}, \quad -\frac{\partial^2 l}{\partial\hat{\gamma}_0\partial\beta_0} = n_1\bar{d}(x_1 - n_1\hat{\alpha}_0)(n_1\hat{\gamma}_0 + 1)^{-2},$$

and

$$-\frac{\partial^2 l}{\partial\beta_0^2} = (n_1\hat{\gamma}_0 + 1)\{(n_1 + 1)\hat{\gamma}_0 + 1\}^{-1}\tilde{p}(1 - \tilde{p})\{\sum n_{1i}d_i^2 - (n_1\bar{d})^2\hat{\gamma}_0(n_1\hat{\gamma}_0 + 1)^{-1}\}.$$

Upon noting that the expectation of $-\partial^2 l/\partial\hat{\gamma}_0\partial\beta_0$ is zero, one can write a consistent estimator of the variance of T as

$$V = -\partial^2 l/\partial\beta_0^2 - (-\partial^2 l/\partial\hat{\alpha}_0\partial\beta_0 \quad 0)\begin{pmatrix} -\partial^2 l/\partial\hat{\alpha}_0^2 & -\partial^2 l/\partial\hat{\alpha}_0\partial\hat{\gamma}_0 \\ -\partial^2 l/\partial\hat{\alpha}_0\partial\hat{\gamma}_0^0 & -\partial^2 l/\partial\hat{\gamma}_0^2 \end{pmatrix}^{-1}\begin{pmatrix} -\partial^2 l/\partial\hat{\alpha}_0\partial\beta_0 \\ 0 \end{pmatrix}$$

$$= -\partial^2 l/\partial\beta_0^2 - (n_1\bar{d})^2(n_1\hat{\gamma}_0 + 1)^{-2}\text{var}\ \hat{\alpha}_0,$$

as was used in Section 4. An alternate variance estimator could be defined by replacing $-\partial^2 l/\partial\beta_0^2$ in V by its expectation, evaluated at $\hat{\alpha}_0$, $\hat{\gamma}_0$—namely,

$$\hat{V} = \hat{\alpha}_0(1 - \hat{\alpha}_0)(1 + \hat{\gamma}_0)^{-1}\{\sum n_{1i}d_i^2 - (n_1\bar{d})^2\hat{\gamma}_0(n_1\hat{\gamma}_0 + 1)^{-1}\},$$

which also reduces to V_0 at $\hat{\gamma}_0 = 0$. Note that likelihood-based inference away from $\beta = 0$ is difficult under this model, since the calculation of the likelihood function and its derivatives requires numerical integration.

General inference on α, β, and ρ is possible under the estimating equation approach to this problem, though β then has a marginal dose response parameter interpretation, rather than a conditional interpretation given the control group rate. The mean and variance specifications $(\boldsymbol{\mu}_j, V_j)$, $j = 1, \ldots, s$, given in Section 4 suggest that $\boldsymbol{\delta} = (\delta_0, \delta_1)$, where $\delta_0 = \log\{\alpha(1 - \alpha)^{-1}\}$, $\delta_1 = \beta$, be estimated as solutions to equations (A1) of Appendix 1. Upon noting that

$$V_1^{-1} = (1 - \rho)^{-1}\text{diag}\{(n_{10}p_0q_0)^{-1}, \ldots, (n_{1k}p_0q_0)^{-1}\} - \rho(1 - \rho)^{-1}\{1 + (n_1 - 1)\rho\}^{-1}\mathbf{f}\mathbf{f}^{\mathrm{T}},$$

where $\mathbf{f}^{\mathrm{T}} = \{(p_0q_0)^{-1/2}, \ldots, (p_kq_k)^{-1/2}\}$ and evaluating at $\beta = 0$, the left side of the second of these equations becomes

$$T = \sum x_{1i}d_i - \tilde{p}\sum n_{1i}d_i$$

after rescaling by the factor $(1 - \rho)$, where $\tilde{p}$ is as given in Section 4. It is natural to again consider (A3) as variance estimator for $s^{1/2}(\hat{\boldsymbol{\delta}} - \boldsymbol{\delta})$, from which the variance estimator for T given in Section 4 derives. As estimator of $\hat{\rho}(\boldsymbol{\delta})$, one can again use (6) with summations from 1 to s and with S_1 and w_1 defined by inserting x_{10} and n_{10} in place of x_1 and n_1 in the definitions of S_1 and w_1. Note, however, that the asymptotic results of Liang and Zeger (1986) do not apply to these estimation equation procedures since the parameter β is found only in the mean and variance expressions for a single independent response variable—namely, the response vector from the current experiment. Hence, a specialized development of the asymptotic distribution theory for the trend statistic $X^2_{\mathrm{OE}} = T^2V^{-1}$ and for related inference procedures is required.

The statistics X^2_{OB} and X^2_{T} are based on

$$\begin{aligned} T &= \sum x_{1i}d_i - \tilde{p}\sum n_{1i}d_i \\ &= \{\sum x_{1i}d_i - x_1n_1^{-1}\sum n_{1i}d_i\} + (x_1n_1^{-1} - \tilde{p})n_1\bar{d} \\ &= T_* + R_1, \end{aligned}$$

where T_* is the numerator of the Cochran–Armitage statistic X_{CA}. Now assuming sn_1^{-1} approaches a nonzero constant as $n_1 \to \infty$, so that $n_1^{1/2}(\hat{\gamma}_0 - \gamma)$ has an asymptotic normal distribution, one obtains

$$\begin{aligned} |n_1^{-1/2}R| &= |n_1^{-1/2}(x_1n_1^{-1} - \hat{\alpha}_0)(n_1\hat{\gamma}_0 + 1)^{-1}n_1\bar{d}| \\ &\leq |\bar{d}(1)\{n_1^{1/2}\gamma + n_1^{1/2}(\hat{\gamma}_0 - \gamma) + n_1^{-1/2}\}^{-1}| \\ &\to_p 0. \end{aligned}$$

Similarly, X^2_{OE} is based on $T = T_* + R$, where

$$R = (x_1n_1^{-1} - \hat{p}_0)(1 - \hat{\rho}_0)\{1 + (n_1 - 1)\hat{\rho}_0\}^{-1}n_1\bar{d},$$

from which

$$|n_1^{-1/2}R| \leq |(1)(2)\{n_1^{1/2}\rho + n_1^{1/2}(\hat{\rho}_0 - \rho) + n_1^{-1/2}(1 - \hat{\rho}_0)\}^{-1}\bar{d}| \to_p 0,$$

provided $\hat{\rho}_0$ is an $s^{1/2}$-consistent estimator of ρ. Hence, these statistics are all asymptotically equivalent to T_*.

Now given the current study control response rate θ_1, $n_1^{-1/2}T_*$ has an asymptotic normal distribution with mean 0 and variance

$$n_1^{-1}\theta_1(1 - \theta_1)\{\sum n_{1i}d_i^2 - (n_1\bar{d})^2n_1^{-1}\}.$$

This variance is consistently estimated by the denominator terms in X^2_{CA}, X^2_{T}, and X^2_{OB} since $x_1n_1^{-1}$ and $\tilde{p}$ are consistent estimators of θ_1. Asymptotic χ^2_1 distributions then follow for these statistics as

indicated in expression (8). On the other hand, the unconditional variance of T_* under a beta–binomial mixing distribution is

$$V_* = \alpha(1 - \alpha)(1 + \gamma)^{-1}\{\sum n_{1i}d_i^2 - (n_1\bar{d})^2 n_1^{-1}\},$$

as may be obtained by first considering the mean and variance of T_* given θ_1. It follows that the asymptotic distribution of $X_*^2 = T_*^2 V_*^{-1}$ is given by the mixture (9). It is straightforward to show that the proposed variance "estimator" for the overdispersed estimating equation statistics is also a consistent estimator of V_*, so that the asymptotic distribution of X_{OE}^2 is also given by (9). The same asymptotic result would hold if the expected value $\tilde{V}$ were substituted for $-\partial^2 l/\partial\beta_0^2$ in X_{T}^2 or X_{OB}^2.

Note that the asymptotic equivalence of both X_{OB}^2 and X_{OE}^2 and the Cochran–Armitage statistic may appear to conflict with the results of Hoel and Yanagawa (1986). However, these authors consider a different situation in which the beta–binomial parameters are not fixed, but rather $a + b \to \infty$ or, equivalently, the pairwise correlation $\rho \to 0$, as $n_1 \to \infty$. These assumptions require the historical control response rates to approach the current control response rate as $n_1 \to \infty$, and hence force the historical control data to make a nonnegligible contribution to trend testing as $n_1 \to \infty$.