Statistics and Its Interface Volume 5 (2012) 237–251

Power of the Cochran-Armitage trend test when exposure scores are based on empirical quantiles of exposure

Huilin Li[*,†,‡] and Mitchell H. Gail[§]

Epidemiologists often categorize exposures based on quantiles of exposure and use the Cochran-Armitage trend test based on such categories to detect associations between disease and exposure. Power calculations typically assume that the population quantiles are known, but in practice quantiles are often estimated from the sample data. We evaluated the power of the Cochran-Armitage trend test for cohort designs and for case-control designs in which sample quantiles of exposure in the cohort or in controls from a case-control study, respectively, are used to define the cutpoints that separate exposure score categories. We give the asymptotic formulas for size and power for the Cochran-Armitage test based on empirical quantiles separately for cohort and case-control designs, together with efficient simulation methods to estimate size and power. Numerical results indicate that estimation of sample quantiles has only a slight effect on power for cohort studies with at least four categories or with more than 280 subjects. However, estimating quantiles can reduce power appreciably in smaller studies with fewer than four exposure categories. For case-control studies of rare diseases, the power loss is limited with more than 120 cases plus controls if the odds ratio comparing the highest exposure category to the lowest category is greater than 0.5. However, if that odd ratio is smaller than 0.5, only samples with more than 360 cases plus controls can guarantee a small loss of power, and increasing the number of exposure categories does not eliminate the loss of power.

AMS 2000 subject classifications: Primary 62F03, 60K35; secondary 92D30.

Keywords and phrases: Case-control design, Cochran-Armitage trend test, Cohort design, Exposure assessment, Power.

1. INTRODUCTION

The Cochran-Armitage test [2, 4] is widely used for testing for a trend in proportions. If there are k categories of exposure and the logit of the outcome event is linear in an exposure score, d_j for $j = 0, 1, \ldots, k-1$, then the Cochran-Armitage score test is locally efficient. Epidemiologists often use scores $d_j = j$, as we shall do in this paper, and they often take a continuous exposure, such as micronutrient intake, and categorize it into k groups based on population quantiles ξ_t. For example, for $k = 4$, exposure would be categorized into four groups defined by the quantiles, namely $[0, \xi_{0.25}], (\xi_{0.25}, \xi_{0.5}], (\xi_{0.5}, \xi_{0.75}]$ and $(\xi_{0.75}, 0]$, and the corresponding scores could be $d_j = 0, 1, 2$ and 3. Distribution theory based on conditioning on both margins of a $2 \times k$ table [7] implies that the test is applicable to case-control data as well as to cohort data. This method of categorization has two advantages, compared to using the continuous exposure in the logistic model: 1) The continuous measurement may be subject to error, including outliers, and this recoding reduces the impact of error; and 2) the recoded data may fit the logistic model as a linear trend better than the original continuous data, just as a log transformed exposure may fit better than a raw exposure.

In practice, the distribution of the exposure in the general population is seldom known. Instead, for a cohort study, the population quantiles are estimated from the sample quantiles in the cohort. For example, [9] used this approach to relate glycemic index, estimated from food frequency questionnaires and grouped by quintiles, to the risk of uterine leiomyoma in a prospective cohort study ($N = 21,861$). For a case-control study, the sample quantiles from controls are often used. For example, [1] conducted a case-control study of 91 cases of esophageal squamous cell carcinoma and 103 controls and categorized staining intensity, a continuous measure, based on quintiles in the controls. Yet sample size calculations typically assume that the population quantiles are known. In this paper, we investigate the effects of using estimated quantiles on the size and power of tests for trend in cohort and case-control studies.

We give notations and asymptotic formulas for size and power for the Cochran-Armitage test based on empirical quantiles separately for cohort and case-control designs in Section 2. We present numerical results on power, including

[*] Corresponding author.

[†] Part of the work was done during Dr. Li's stay at the National Cancer Institute.

[‡] Part of Dr. Li's research was supported by the Intramural Research Program of the Division of Cancer Epidemiology and Genetics of the National Cancer Institute.

[§] Dr. Gail's research was supported by the Intramural Research Program of the Division of Cancer Epidemiology and Genetics of the National Cancer Institute.

Table 1. $2 \times k$ cross-classification of cohort outcomes and exposure scores, X, based on known quintiles of a continuous exposure, Z

	Exposure scores X				Total
	d_0	d_1	$\dots$	d_{k-1}	
$Y=1$	r_0	r_1	$\dots$	r_{k-1}	R
$Y=0$	s_0	s_1	$\dots$	s_{k-1}	S
Total	n_0	n_1	$\dots$	n_{k-1}	N

Table 2. $2 \times k$ cross-classification of cohort outcomes and exposure scores, W, based on known quintiles of a continuous exposure, Z

	Estimated exposure scores W				Total
	d_0	d_1	$\dots$	d_{k-1}	
$Y=1$	r'_0	r'_1	$\dots$	r'_{k-1}	R
$Y=0$	s'_0	s'_1	$\dots$	s'_{k-1}	S
Total	N/k	N/k	$\dots$	N/k	N

both theoretical calculations and simulations, in Section 3. We discuss these findings in Section 4.

2. NOTATIONS AND ASYMPTOTIC THEORY

2.1 Cohort study

2.1.1 Cochran-Armitage trend test for cohort data with known quantiles

Table 1 describes the outcomes of a cohort study of N subjects, R of whom develop a disease ($Y=1$) and S of whom remain free of the disease ($Y=0$). Exposures are categorized into k levels based on known population quantiles of a continuous exposure, Z, with distribution F. The quantiles satisfy $F(\xi_t)=t$. Without loss of generality, we can take Z as uniformly distributed, because if F is known, we can transform the distribution of Z to uniformity by setting $Z^* = F(Z)$; then Z^* is uniformly distributed. Hence, hereafter, we take Z as uniformly distributed and note $\xi_t = t$. Then, an observation Z_i is assigned a score X_i for $i=1,2,\dots,N$ according to $X_i = d_0$ for $0 \le Z_i \le (1/k)$; and $X_i = d_j$ for $(j/k) < Z_i \le ((j+1)/k)$ with $j=1,\dots,k-1$. The disease outcome Y is related to X through the logistic regression:

$$(1) \qquad p_j = \mathrm{P}(Y=1|X=d_j) = \{1+\exp(-\mu-\beta d_j)\}^{-1}$$

We want to assess the association between X and Y by testing $H_0 : \beta = 0$ using the Cochran-Armitage score test.

The test can be based on the statistic $U = \sum_{i=1}^{N} Y_i(X_i - \bar{X}) = \sum_{j=0}^{k-1} r_j(d_j - \bar{d})$, where $\bar{d} = \sum_{j=0}^{k-1} n_j d_j/N$. Assuming that the variability in $\bar{d}$ is negligible (Nam, 1987), we calculate $V = \mathrm{Var}(U) = \sum_{j=0}^{k-1} n_j p_j(1-p_j)(d_j-\bar{d})^2$. Under the null hypothesis, $p_j = p \equiv \exp(\mu)/(1+\exp(\mu))$ for $j=0,1,\dots,k-1$. From the maximum likelihood estimate $\hat{p} = R/N$, a valid estimate of V is given by $\hat{V}_0 = RS\sum_{j=0}^{k-1} n_j(d_j-\bar{d})^2/N^2$. Hence, the commonly used Cochran-Armitage score test can be written as

$$(2) \quad Q = U(\hat{V}_0)^{-1/2} = \left\{\sum_{j=0}^{k-1} r_j(d_j-\bar{d})\right\}\left\{RS\sum_{j=0}^{k-1} n_j(d_j-\bar{d})^2/N^2\right\}^{-1/2}$$

Under the alternative hypothesis, $H_1 : \beta \neq 0$, the asymptotic power of the two-sided trend test is

$$(3) \quad \begin{aligned} &\mathrm{P}(|U(\hat{V}_0)^{-1/2}| > z_{1-\alpha/2}) \\ &\quad = 1 - \Phi\{(z_{1-\alpha/2}\sigma_* - \mathrm{E}_{H_1}U)/\sigma_1\} \\ &\qquad + \Phi\{(-z_{1-\alpha/2}\sigma_* - \mathrm{E}_{H_1}U)/\sigma_1\} \end{aligned}$$

where $\sigma_1 = \sqrt{V}$, $\mathrm{E}_{H_1}U = N/k\sum_{j=0}^{k-1} p_j(d_j - \bar{d}^*)$, $\bar{d}^* = \sum_{j=0}^{k-1} d_j/k$, $\sigma_* = \sqrt{\mathrm{E}_{H_1}(\hat{V}_0)} = \sqrt{\bar{p}(1-\bar{p})N/k\sum_{j=0}^{k-1}(d_j-\bar{d}^*)^2}$, and $\bar{p} = \sum_{j=0}^{k-1} p_j/k$.

2.1.2 Cochran-Armitage trend test for cohort data with estimated quantiles

Because in practice, the theoretical quantiles of Z are not known, we categorize exposure based on the empirical quantiles (order statistics) of the sample $Z_1, Z_2, \dots, Z_N$. The j^{th} order statistic $Z_{(j)}$ is the j^{th} smallest value of the sample. We assign an estimated score W_i for observed exposure Z_i for $i=1,2,\dots,N$, according to $W_i = d_0$ for $Z_i \le Z_{(N/k)}$; $W_i = d_j$ for $Z_{(jN/k)} < Z_i \le Z_{((j+1)N/k)}$ with $j=1,\dots,k-2$; and $W_i = d_{k-1}$ for $Z_{((k-1)N/k)} < Z_i$. This new data set can be represented as in Table 2. Note that in this table, the column marginal totals are fixed at N/k.

The naive Cochran-Armitage score test for this table is obtained by replacing the true exposure score X by the estimated exposure score W in equation (2) to obtain.

$$(4) \quad Q' \equiv U'(\hat{V}'_0)^{-1/2} = \left\{\sum_{j=0}^{k-1} r'_j(d_j-\bar{d}^*)\right\}\left\{RS\sum_{j=0}^{k-1}(d_j-\bar{d}^*)^2/(kN)\right\}^{-1/2}$$

We now examine whether Q' has valid size. The variance of $U' = \sum_{j=0}^{k-1} r'_j(d_j - \bar{d}^*)$ is $V' = (N/k)\sum_{j=0}^{k-1} P_j^W(1-P_j^W)(d_j-\bar{d}^*)^2$, where $P_j^W \equiv \mathrm{P}(Y=1|W=d_j) = \mathrm{E}_{X|W=d_j}[\mathrm{P}(Y=1|X)]$, as shown in Section 2.1.3 under the assumption $\mathrm{P}(Y=1|W,X) = \mathrm{P}(Y=1|X)$. Under the null hypothesis H_0, which implies that all the P_j^Ws are equal to p for $j=0,1,\dots,k-1$, we have $\mathrm{E}_0(U')=0$ and $V'_0 = (N/k)\sum_{j=0}^{k-1} p(1-p)(d_j-\bar{d}^*)^2$. Plugging in the maximum likelihood estimate $\hat{p} = R/N$, we obtain a consistent estimate of V_0', given by $\hat{V}_0' = RS\sum_{j=0}^{k-1}(d_j - \bar{d}^*)^2/(kN)$. This estimate is the same as the variance term in equation (4). Thus, under the null hypothesis, Q' in equation (4) is centered and properly standardized and therefore yields nominal size, at least according to asymptotic theory.

Under the alternative hypothesis $H_1 : \beta \neq 0$, the asymptotic power of the two-sided trend test is

$$\begin{aligned}(5)\qquad & \mathrm{P}(|U'(\hat{V'}_0)^{-1/2}| > z_{1-\alpha/2}) \\ &= 1 - \Phi\{(z_{1-\alpha/2}\sigma_*' - \mathrm{E}_{H_1}U')/\sigma_1'\} \\ &\quad + \Phi\{(-z_{1-\alpha/2}\sigma_*' - \mathrm{E}_{H_1}U')/\sigma_1'\}\end{aligned}$$

where $\sigma_1' = \sqrt{V'}$, $\mathrm{E}_{H_1}U' = N/k\sum_{j=0}^{k-1} P_j^W(d_j - \bar{d}^*)$, $\bar{d}^* = \sum_{j=0}^{k-1} d_j/k$, $\sigma_*' = \mathrm{E}_{H_1}(\hat{V}_0') = \{\bar{p}'(1-\bar{p}')N/k\sum_{j=0}^{k-1}(d_j - \bar{d}^*)^2\}^{1/2}$, and $\bar{p}' = \sum_{j=0}^{k-1} P_j^W/k$. Equation (5) can be used to show numerically that the power based on Q' is less than that based on Q (Section 3), a result that is also confirmed by simulations.

2.1.3 Calculation of P_j^W

We calculate P_j^W which is needed for equation (5), as follows. When the quantiles of Z are estimated from the cohort data, the estimated exposure scores satisfy $\mathrm{P}(W_i = d_j) = 1/k$. Moreover, under the assumption that Y is conditionally independent of W given X, we obtain

$$\begin{aligned}P_j^W &\equiv \mathrm{P}(Y = 1|W_i = d_j) \\ &= \sum_{h=0}^{k-1} \mathrm{P}(Y = 1, X = d_h|W = d_j) \\ &= \sum_{h=0}^{k-1} \mathrm{P}(Y = 1|X = d_h, W = d_j)\mathrm{P}(X = d_h|W = d_j) \\ &= \sum_{h=0}^{k-1} \mathrm{P}(Y = 1|X = d_h)\mathrm{P}(X = d_h|W = d_j) \\ &= \mathrm{E}_{X|W=d_j}[\mathrm{P}(Y = 1|X)]\end{aligned}$$

To compute P_j^W, we need $\mathrm{P}(Y = 1|X = d_h)$ from equation (1) and $\mathrm{P}(X = d_h|W = d_j) = \mathrm{P}(X = d_h, W = d_j)/\mathrm{P}(W = d_j) = k\mathrm{P}(X = d_h, W = d_j)$, where

$$\begin{aligned}&\mathrm{P}(X = d_h, W = d_j) \\ &\quad = \mathrm{P}(Z_{(jN/k)} < Z \leq Z_{((j+1)N/k)}, \\ &\qquad (h/k) < Z \leq \{(h+1)/k\}) \\ &\quad = \sum_{i=(jN/k)+1}^{(j+1)N/k} \mathrm{P}(Z = Z_{(i)}, (h/k) < Z \leq (h+1)/k) \\ &\quad = \sum_{i=(jN/k)+1}^{(j+1)N/k} \mathrm{P}((h/k) < Z_{(i)} \leq (h+1)/k) \\ &\quad = \sum_{i=(jN/k)+1}^{(j+1)N/k} \int_{(h/k)}^{(h+1)/k} \binom{N-1}{i} u^{i-1}(1-u)^{N-i}du\end{aligned}$$

2.2 Case-control study

2.2.1 Cochran-Armitage trend test for case-control data with known quantiles

The Cochran-Armitage statistic for case-control data (Table 1) is based on $U = \sum_{i=1}^N X_i(Y_i - \bar{Y}) = \sum_{j=0}^{k-1}(Sr_j - Rs_j)/N = (Sd'r - Rd's)/N$, where $d' = (d_0, d_1, \ldots, d_{k-1})$, $r' = (r_0, r_1, \ldots, r_{k-1})$ and $s' = (s_0, s_1, \ldots, s_{k-1})$. From the retrospective sampling, the cell counts r' and s' follow independent multinomial distributions with total trial numbers R and S and respective probabilities $p' = (p_0, p_1, \ldots, p_{k-1})$ and $q' = (q_0, q_1, \ldots, q_{k-1})$. For equally spaced uniform quintiles, we have $\mathrm{P}(X = d_j) = 1/k$. Hence, from Bayes theorem, $p_j = \mathrm{P}(X = d_j|Y = 1) = (1 + \exp(-\mu - \beta d_j))^{-1}(\sum_{l=0}^{k-1}(1 + \exp(-\mu - \beta d_l))^{-1})^{-1}$ and $q_j = \mathrm{P}(X = d_j|Y = 0) = (1 + \exp(\mu + \beta d_j))^{-1}(\sum_{l=0}^{k-1}(1 + \exp(\mu + \beta d_l))^{-1})^{-1}$. The mean of U is $\mathrm{E}(U) = SRd'(p - q)/N$ and the variance of U is $V = \mathrm{Var}(U) = SR(Sd'\Sigma_p d + Rd'\Sigma_q d)/N^2$, where $(\Sigma_p)_{jj} = p_j(1 - p_j)$, $(\Sigma_p)_{jh} = -p_j p_h$, $(\Sigma_q)_{jj} = q_j(1 - q_j)$, and $(\Sigma_q)_{jh} = -q_j q_h$ for $j \neq h$.

The null hypothesis $H_0 : \beta = 0$ implies $p_j = q_j$ for $j = 0, 1, \ldots, k-1$. Thus $\mathrm{E}_{H_0}(U) = 0$ and $V_0 = \mathrm{Var}_{H_0}(U) = SRd'\Sigma_p d/N$. A consistent estimate of V_0 can be obtained by plugging in estimates $\hat{p}_j = n_j/N$ to yield $\hat{V}_0 = SRd'\hat{\Sigma}d/N$, where $\hat{\Sigma}_{jj} = n_j(N - n_j)/N^2$ and $\hat{\Sigma}_{jj} = -n_i n_j/N^2$. Thus the Cochran-Armitage trend test for case-control data is

$$\begin{aligned}(6)\qquad Q &\equiv U(\hat{V}_0')^{-1/2} \\ &= ((Sd'r - Rd's)/N)\left(\frac{RS}{N}\left(\sum_{j=0}^{k-1} d_j^2 \frac{n_j}{N}\left(1 - \frac{n_j}{N}\right)\right.\right. \\ &\qquad \left.\left. - \sum_{j \neq k}\sum d_j d_k \frac{n_j n_k}{N^2}\right)\right)\end{aligned}$$

For an alternative hypothesis $H_1 : \beta \neq 0$, which implies $p_j \neq q_j$, for some j, the asymptotic power of the two-sided trend test Q is given by

$$\begin{aligned}(7)\qquad & \mathrm{P}(|U(\hat{V}_0)^{-1/2}| > z_{1-\alpha/2}) \\ &= 1 - \Phi\{(z_{1-\alpha/2}\sigma_* - \mathrm{E}_{H_1}U)/\sigma_1\} \\ &\quad + \Phi\{(-z_{1-\alpha/2}\sigma_* - \mathrm{E}_{H_1}U)/\sigma_1\}\end{aligned}$$

where $\sigma_1 = \sqrt{V}$, $\mathrm{E}_{H_1}U = SRd'(p - q)/N$, $\sigma_* = \{\mathrm{E}_{H_1}(\hat{V}_0)\}^{1/2} = (SRd'\mathrm{E}_{H_1}(\hat{\Sigma})d/N)^{1/2}$, $\mathrm{E}_{H_1}(\hat{\Sigma}_{ii}) = (Rp_i + Sq_i)/N - [Rp_i(1 - p_i) + Sq_i(1 - q_i) + (Rp_i + Sq_i)^2]/N^2$ and $\mathrm{E}_{H_1}(\hat{\Sigma}_{ij}) = -[(R^2 - R)p_i p_j + (S^2 - S)q_i q_j + RS(p_i q_j + p_j q_i)]/N^2$.

Table 3. $2 \times k$ cross-classification of case-control outcomes and exposure scores, W, based on known quintiles of a continuous exposure, Z

	Estimated exposure scores W				Total
	d_0	d_1	$\dots$	d_{k-1}	
$Y = 1$	$\tilde{r}_0$	$\tilde{r}_1$	$\dots$	$\tilde{r}_{k-1}$	R
$Y = 0$	S/k	S/k	$\dots$	S/k	S
Total	$\tilde{n}_0$	$\tilde{n}_1$	$\dots$	$\tilde{n}_{k-1}$	N

2.2.2 Cochran-Armitage trend test for case-control data with estimated quantiles

In the case-control study, one can obtain an estimated score W by classifying Z with the empirical quantiles in controls. This data set can be represented as in Table 3. Let $Z_1^{\mathrm{CO}}, Z_2^{\mathrm{CO}}, \dots, Z_S^{\mathrm{CO}}$ and $Z_1^{\mathrm{CA}}, Z_2^{\mathrm{CA}}, \dots, Z_R^{\mathrm{CA}}$ be the observed continuous exposures in controls and cases respectively. Based on the order statistics and corresponding quantiles of $Z_1^{\mathrm{CO}}, Z_2^{\mathrm{CO}}, \dots, Z_S^{\mathrm{CO}}$, the estimated scores for controls are defined as $W_i = d_0$ for $Z_i^{\mathrm{CO}} \le Z_{(S/k)}^{\mathrm{CO}}$; $W_i = d_j$ for $Z_{(jS/k)}^{\mathrm{CO}} < Z_i^{\mathrm{CO}} \le Z_{((j+1)S/k)}^{\mathrm{CO}}$ for $j = 1, \dots, k-2$; and $W_i = d_{k-1}$ for $Z_{((k-1)S/k)}^{\mathrm{CO}} < Z_i^{\mathrm{CO}}$. The estimated scores for cases are $W_i = d_0$ for $Z_i^{\mathrm{CA}} \le Z_{(S/k)}^{\mathrm{CO}}$; $W_i = d_j$ for $Z_{(jS/k)}^{\mathrm{CO}} < Z_i^{\mathrm{CA}} \le Z_{((j+1)S/k)}^{\mathrm{CO}}$ for $j = 1, \dots, k-2$; and $W_i = d_{k-1}$ for $Z_{((k-1)S/k)}^{\mathrm{CO}} < Z_i^{\mathrm{CA}}$. Note that the elements of vector s in Table 3 are fixed at S/k, and $\tilde{r}_j = \sum_{i=1}^{N} Y_i I[W_i = d_j]$. The Cochran-Armitage statistic is $\tilde{U} = \sum_{j=0}^{k-1} d_j(S\tilde{r}_j - RS/k)/N = (Sd'\tilde{r} - RS\bar{d}^*)/N$, where $\tilde{r} = (\tilde{r}_0, \tilde{r}_1, \dots, \tilde{r}_{k-1})$. Conditionally on the order statistics $Z_{(1)}^{\mathrm{CO}}, Z_{(2)}^{\mathrm{CO}}, \dots, Z_{(S)}^{\mathrm{CO}}$ the cell counts $\tilde{r}$ follow a multinomial distribution with index R and cell probabilities $\tilde{p} = (\tilde{p}_0, \tilde{p}_1, \dots, \tilde{p}_{k-1})$ with $\tilde{p}_j = \mathrm{P}(W = d_j|Y = 1) = F_1(Z_{((j+1)S/k)}^{\mathrm{CO}}) - F_1(Z_{(jS/k)}^{\mathrm{CO}})$. Here F_1, the cumulative distribution function of Z in cases, can be computed as in Section 2.2.3. Because the samples of cases and controls are independent, the unconditional mean of $\tilde{U}$ can be computed from results in Section 2.2.3 as $\tilde{\mu} = \mathrm{E}(\tilde{U}) = RS\sum_{j=0}^{k-1} d_j[\mathrm{E}(\tilde{p}_j) - 1/k]/N$. Likewise, the unconditional variance is $\tilde{V} \equiv \mathrm{Var}(\tilde{U}) = \mathrm{Var}(S\sum_{j=0}^{k-1} d_j\tilde{r}_j/N) = (S/N)^2[\sum_{j=0}^{k-1} d_j^2 \mathrm{Var}(\tilde{r}_j) + \sum_{j \ne k} d_j d_k \mathrm{Cov}(\tilde{r}_j, \tilde{r}_k)]$, where the unconditional variances and covariances of $\tilde{r}_j$ are $\mathrm{Var}(\tilde{r}_j) = R^2\mathrm{Var}(\tilde{p}_j) + R\mathrm{E}[\tilde{p}_j(1 - \tilde{p}_j)]$ and $\mathrm{Cov}(\tilde{r}_j, \tilde{r}_k) = (R^2 - R)\mathrm{E}(\tilde{p}_j\tilde{p}_k) - R^2\mathrm{E}(\tilde{p}_j)\mathrm{E}(\tilde{p}_k)$. All the moments calculations for $\tilde{p}_j$ are presented in Section 2.2.3.

The naive Cochran-Armitage trend test is obtained by replacing the entries $\tilde{r}_j$ and $s_j = S/k$ from Table 3 in the formula (6), which is appropriate for known scores (Table 2). The resulting naive Cochran-Armitage trend test is

$$(8) \quad \tilde{Q} \equiv \tilde{U}(\hat{\tilde{V}}_0^*)^{-1/2} = ((Sd'\tilde{r} - RS\bar{d}^*)/N)(RSd'\hat{\tilde{\Sigma}}d/N)^{-1/2}$$

where $\hat{\tilde{\Sigma}}_{jj} = \tilde{n}_j(N - \tilde{n}_j)/N^2$ and $\hat{\tilde{\Sigma}}_{jk} = -\tilde{n}_j\tilde{n}_k/N^2$. We are interested in whether $\tilde{Q}$ is properly standardized and therefore has nominal size, at least in large samples.

Define $\tilde{q}_i = \mathrm{P}(W = d_i|Y = 0)$. Under $H_0 : \beta = 0$, which implies $\tilde{p}_j = \tilde{q}_j$ for $j = 0, 1, \dots, k-1$, we find $\mathrm{E}_0(\tilde{U}) = RS(d_{k-1} - \sum_{j=0}^{k-1} d_j/k)/(N(S+1))$ which does not equal to zero. In particular when $d_j = j$, we have $\mathrm{E}_0(\tilde{U}) = (R(k-1)/2N)(S/(S+1)$. Actually $\mathrm{E}_0(\tilde{U})$ increases as the number of categories increases. However, $\mathrm{E}_0(\tilde{U})$ is of order $O(1)$, which is negligible compared to the variance for large samples. We have $\tilde{V}_0 \equiv \mathrm{Var}_{H_0}(\tilde{U}) = (S/N)^2[\sum_{j=0}^{k-1} d_j^2\mathrm{Var}_0(\tilde{r}_j) + \sum_{j \ne k} d_j d_k \mathrm{Cov}_0(\tilde{r}_j, \tilde{r}_k)]$, where $\mathrm{Var}_0(\tilde{r}_j) = R\tilde{p}_j[1 - (S/(S+1))\tilde{p}_j](S/(S+1))((R+S+1)/(S+2))$, for $j = 0, 1, \dots, k-2$; $\mathrm{Var}_0(\tilde{r}_{k-1}) = R(\tilde{p}_{k-1}S/(S+1) + 1/(S+1))((1-\tilde{p}_{k-1})(S/(S+1))((R+S+1)/(S+2)))$; $\mathrm{Cov}_0(\tilde{r}_j, \tilde{r}_h) = -R\tilde{p}_j\tilde{p}_h(S/(S+1))^2((R+S+1)/(S+2))$, for $j < h \le k-2$; and $\mathrm{Cov}_0(\tilde{r}_j, \tilde{r}_{k-1}) = -R\tilde{p}_j(\tilde{p}_{k-1} + 1/S)(S/(S+1))^2((R+S+1)/(S+2))$ for $j < k-2$. Thus $\tilde{V}_0$ is of order $O(N)$, and asymptotically we have $\tilde{V}_0 \to V_0$. The estimate, $\hat{\tilde{V}}_0$, of $\tilde{V}_0$ obtained by plugging in the estimates $\tilde{p}_i = \tilde{n}_i/N$, is consistent; thus $\hat{\tilde{V}}_0 \to \tilde{V}_0^*$. It follows that $\tilde{Q}$ is asymptotically standard normal and has proper size in large samples.

Under the alternative hypothesis, $H_1 : \beta \ne 0$, which implies $\tilde{p}_j \ne \tilde{q}_j$ for some j, the asymptotic power of the two-sided trend test $\tilde{Q}$ is given by

$$(9) \quad \begin{aligned} &\mathrm{P}(|\tilde{U}(\hat{V}^*)^{-1/2}| > z_{1-\alpha/2}) \\ &\quad = 1 - \Phi\{(z_{1-\alpha/2}\tilde{\sigma}_* - \mathrm{E}_{H_1}\tilde{U})/\tilde{\sigma}_1\} \\ &\qquad + \Phi\{(-z_{1-\alpha/2}\tilde{\sigma}_* - \mathrm{E}_{H_1}\tilde{U})/\tilde{\sigma}_1\} \end{aligned}$$

where $\tilde{\sigma}_1 = \sqrt{\tilde{V}}$, $\mathrm{E}_{H_1}\tilde{U} = RS\sum_{j=0}^{k-1} d_j[\mathrm{E}_{H_1}(\tilde{p}_j) - 1/k]/N$, $\tilde{\sigma}_* = \{\mathrm{E}_{H_1}(\hat{\tilde{V}}_0^*)\}^{1/2} = (SRd'\mathrm{E}_{H_1}(\hat{\tilde{\Sigma}})d/N)^{1/2}$, $\mathrm{E}_{H_1}(\hat{\tilde{\Sigma}}_{jj}) = (R\mathrm{E}_{H_1}\tilde{p}_j + S/k)/N - (R^2\mathrm{E}_{H_1}\tilde{p}_j^2 + 2SR\mathrm{E}_{H_1}\tilde{p}_j/k + S^2/k^2)/N^2$, and $\mathrm{E}_{H_1}(\hat{\tilde{\Sigma}}_{jh}) = -((R^2 - R)\mathrm{E}_{H_1}\tilde{p}_j\tilde{p}_h + SR\mathrm{E}_{H_1}(\tilde{p}_j + \tilde{p}_h)/k + S^2/k^2)/N^2$. The needed moments are given next.

2.2.3 Calculations of moments of $\tilde{p}_j$ and the distributions F_0 and F_1 of Z in cases and controls respectively

In the case-control study, the estimated exposure score W_i is defined in Section 2.2.2 for cases and controls separately based on the empirical quantiles of the control sample $Z_1^{\mathrm{CO}}, Z_2^{\mathrm{CO}}, \dots, Z_S^{\mathrm{CO}}$. In the power calculation (9), we need to know the following moments of $\tilde{p}_j$: $\mathrm{E}(\tilde{p}_j)$, $\mathrm{E}(\tilde{p}_j^2)$ and $\mathrm{E}(\tilde{p}_j\tilde{p}_h)$. To compute these moments, we need the distributions F_0 and F_1 of Z in cases and controls respectively, as given at the end of this Section, and the joint distribution of pairs and quadruplets of order statistics that correspond to quantiles of the control distribution. Note that F_0 is not uniform even if Z is uniformly distributed in the general population, though it is approximately uniform if the disease is rare. The needed moments are given by

$$\begin{aligned}\mathrm{E}(\tilde{p}_j) &= \mathrm{E}\{F_1(Z^{\mathrm{CO}}_{((j+1)S/k)}) - F_1(Z^{\mathrm{CO}}_{(jS/k)})\}\\ &= \int_0^1 \int_0^{\xi_1} \{F_1(\xi_1) - F_1(\xi_2)\}\\ &\quad \times f_{((j+1)n/k),(jn/k)}(\xi_1,\xi_2)d\xi_2 d\xi_1,\\ \mathrm{E}(\tilde{p}_j^2) &= \mathrm{E}\{F_1(Z^{\mathrm{CO}}_{((j+1)S/k)}) - F_1(Z^{\mathrm{CO}}_{(jS/k)})\}^2\\ &= \int_0^1 \int_0^{\xi_1} \{F_1(\xi_1) - F_1(\xi_2)\}^2\\ &\quad \times f_{((j+1)n/k),(jn/k)}(\xi_1,\xi_2)d\xi_2 d\xi_1,\end{aligned}$$

and assuming $h - j > 1$,

$$\begin{aligned}&\mathrm{E}(\tilde{p}_j\tilde{p}_h)\\ &\quad = \mathrm{E}[\{F_1\big(Z^{\mathrm{CO}}_{((j+1)S/k)}\big) - F_1(Z^{\mathrm{CO}}_{(jS/k)})\}\\ &\qquad \times \{F_1(Z^{\mathrm{CO}}_{((h+1)S/k)}) - F_1(Z^{\mathrm{CO}}_{(hS/k)})\}]\\ &\quad = \int_0^1 \int_0^{\xi_4} \int_0^{\xi_3} \int_0^{\xi_2} \{F_1(\xi_2) - F_1(\xi_1)\}\{F_1(\xi_4) - F_1(\xi_3)\}\\ &\qquad \times f_{(\frac{jn}{k}),(\frac{(j+1)n}{k}),(\frac{hn}{k}),(\frac{(h+1)n}{k})}(\xi_1,\xi_2,\xi_3,\xi_4)d\xi_1 d\xi_2 d\xi_3 d\xi_4.\end{aligned}$$

When $h = j + 1$, a similar triple integral results. In these expressions $f_{(m),(n)}(\xi_1, \xi_2)$, the joint density of two order statistics from the control population is given by $\frac{N!}{(m-1)!(n-m-1)!(N-n)!}F_0^{n-1}(\xi_1)f_0(\xi_1)\{F_0^{n-1}(\xi_2) - F_0^{n-1}(\xi_1)\}^{n-m-1}f_0(\xi_2)\{1 - F_0^{n-1}(\xi_2)\}^{N-n}$, where $\xi_1 \le \xi_2$. Likewise, the joint density $f_{(m),(n),(s),(t)}(\xi_1, \xi_2, \xi_3, \xi_4)$ of four order statistics in the control population is $\frac{N!}{(m-1)!(n-m-1)!(s-n-1)!(t-s-1)!(N-t)!}F_0^{m-1}(\xi_1)f_0(\xi_1) \times \{F_0(\xi_2) - F_0(\xi_1)\}^{n-m-1}f_0(\xi_2)\{F_0(\xi_3) - F_0(\xi_2)\}^{s-n-1} \times f_0(\xi_3))\{F_0(\xi_4) - F_0(\xi_3\}^{t-s-1}f_0(\xi_4)\{1 - F_0^{n-1}(\xi_4)\}^{N-t}$, where $\xi_1 \le \xi_2 \le \xi_3 \le \xi_4$. The related order statistics formula can be found in [5]. In these formulas, the required distributions and densities in cases and controls are given by:

$$\begin{aligned}F_1(Z) &\equiv \mathrm{P}(Z \le z|Y=1)\\ &= \mathrm{P}(Z \le z, Y=1)/\mathrm{P}(Y=1)\\ &= \begin{cases} z(1+\exp(-\mu))^{-1}/\mathrm{P}(Y=1) \quad \text{for } z \le 1/k;\\ \Big(\sum_{x=0}^{j-2}(1+\exp(-\mu-\beta x))^{-1}/k + (1+\exp(-\mu\\ \quad -(j-1)\beta))^{-1}\{z-(j-1)/k\}\Big)/\mathrm{P}(Y=1)\\ \quad \text{for } (j-1)/k < z \le j/k, 2 < j \le (k-1);\\ \Big(\sum_{x=0}^{k-2}(1+\exp(-\mu-\beta x))^{-1}/k + (1+\exp(-\mu\\ \quad -(k-1)\beta))^{-1}\{z-(k-1)/k\}\Big)/\mathrm{P}(Y=1)\\ \quad \text{for } (k-1)/k < z \end{cases}\end{aligned}$$

and by

$$\begin{aligned}F_0(Z) &\equiv \mathrm{P}(Z \le z|Y=0)\\ &= \begin{cases} z(1+\exp(\mu))^{-1}/\mathrm{P}(Y=0) \quad \text{for } z \le 1/k;\\ \Big(\sum_{x=0}^{j-2}(1+\exp(\mu+\beta x))^{-1}/k + (1+\exp(\mu\\ \quad +(j-1)\beta))^{-1}\{z-(j-1)/k\}\Big)/\mathrm{P}(Y=0)\\ \quad \text{for } (j-1)/k < z \le j/k, 2 < j \le (k-1);\\ \Big(\sum_{x=0}^{k-2}(1+\exp(\mu+\beta x))^{-1}/k + (1+\exp(\mu\\ \quad +(k-1)\beta))^{-1}\{z-(k-1)/k\}\Big)/\mathrm{P}(Y=0)\\ \quad \text{for } (k-1)/k < z \end{cases}\end{aligned}$$

The corresponding densities are

$$\begin{aligned}&f_1(Z|Y=1)\\ &\quad = \begin{cases} (1+\exp(-\mu))^{-1}/\mathrm{P}(Y=1) \quad \text{for } z \le 1/k;\\ (1+\exp(-\mu-(j-1)\beta))^{-1})/\mathrm{P}(Y=1)\\ \quad \text{for } (j-1)/k < z \le j/k, 2 < j \le (k-1);\\ (1+\exp(-\mu-(k-1)\beta))^{-1})/\mathrm{P}(Y=1)\\ \quad \text{for } (k-1)/k < z \end{cases}\end{aligned}$$

and

$$\begin{aligned}&f_0(Z|Y=0)\\ &\quad = \begin{cases} (1+\exp(\mu))^{-1}/\mathrm{P}(Y=0) \quad \text{for } z \le 1/k;\\ (1+\exp(\mu+(j-1)\beta))^{-1})/\mathrm{P}(Y=0)\\ \quad \text{for } (j-1)/k < z \le j/k, 2 < j \le (k-1);\\ (1+\exp(\mu+(k-1)\beta))^{-1})/\mathrm{P}(Y=0)\\ \quad \text{for } (k-1)/k < z \end{cases}\end{aligned}$$

Here $\mathrm{P}(Y = 1) = 1 - \mathrm{P}(Y = 0) = (1/k)(\sum_{j=0}^{k-1}(1+\exp(-\mu-\beta d_j))^{-1}$is the probability of disease in the source population.

3. NUMERICAL RESULTS

We first present data on estimated size and power obtained from simulations and later compare these results with calculations from the asymptotic formulas in Section 2. Because simulated results usually agreed very well with the asymptotic theory (see Figures 5–8), we only present Figures for the simulated results. For cohort designs, we simulated by drawing Z from a uniform $U(0, 1)$ distribution, determining X from known quantiles for Z, and drawing Y from the Bernoulli distribution in equation (1). After N such triplets were generated, values W for each triplet were computed from the order statistics of Z as described in Section 2.1.2. For case-control designs, we used the distributions $F_1(z) \equiv \mathrm{P}(Z \le z|Y = 1)$ and $F_0(z) \equiv \mathrm{P}(Z \le z|Y = 0)$ given in Section 2.2.3 to generate Z for cases and controls.

For cases, applying the inverse probability transformation to a uniform $U(0,1)$ random variable yielded $Z = F_1^{-1}(U)$. A random sample of Z values for controls were likewise generated from $Z = F_0^{-1}(U)$. Values of X were determined by comparing Z to quantiles of the uniform $U(0,1)$ distribution, both for cases and controls. Values of W were determined for controls by the order statistics of in controls, and values of W for cases were obtained by comparing the Z values for cases with the order statistics of Z in controls (Section 2.2.2).

The simulation results were based on $10,000$ repetitions independently for each of the parameter settings studied. We conducted simulations for $N = \{80, 120, 160, 200, 240, 280, 320, 360\}$, $k = \{2, 3, 4, 5\}$ and for a range of odds ratios comparing highest to lowest categories. Using these data we sought to determine combinations of N and k that assured that power loss was no more than 5% from estimating quantiles.

Figure 1 depicts the simulated power for cohort studies with $k = 4$ categories (upper panels) or $k = 2$ categories (lower panels) as a function of the odds ratio between the highest and the lowest categories, denoted as $\exp(\beta^*)$, and shown on a log scale. Values of $\mu \in \{-2, 0, 1\}$ are shown for sample sizes $N = 120$ (circles) and $N = 280$ (triangles). Power for the test Q based on known exposure scores, X, is shown in open symbols, and for the test Q' based on estimated exposure scores, W, in solid symbols. Estimated size was close to the nominal 0.05 level for an odds ratio of 1.0 in each case. The power of the test Q' is less than that of the test Q. For small N, this difference is appreciable for large values of $|\beta^*|$, namely both for large and small odds ratios. For example, for $\exp(\beta^*) = 4$ the powers of the test Q are respectively 0.63, 0.94, 0.87, and 1.00 for $\{\mu = -2, N = 120, k = 4\}$, $\{\mu = -2, N = 280, k = 4\}$, $\{\mu = -2, N = 120, k = 2\}$ and $\{\mu = -2, N = 280, k = 2\}$. The corresponding powers of Q' are 0.60, 0.93, 0.81, 0.99. For $k = 4$, over a range of values of $\exp(\beta^*)$ in $[0.1, 8]$ and μ in $\{-2, 0, 1\}$, the differences in power (power of Q minus power of Q') ranged from -0.004 to 0.058 for $N = 120$ and from -0.003 to 0.026 for $N = 280$. For $k = 2$, over a range of values of $\exp(\beta)$ in $[0.1, 4]$ and of μ in $\{-2, 0, 2\}$, the differences in power ranged from -0.005 to 0.08 for $N = 120$ and from 0 to 0.05 for $N = 280$.

To get a clearer view of which combinations of k and N lead to a power loss greater than 5% when quantiles are estimated in cohort data, we plotted the power loss in percent against the odds ratio $\exp(\beta^*)$ (Figure 2). For $k > 3$, the power loss from using empirical quantiles was less than 5% even for $N = 80$ (with the exception of the single point $\exp(\beta^*) = 0.1$, $\mu = -2$, and $k = 4$). When $k = 2$ or 3, using estimated quantiles can lead to a decrease in power exceeding 5% with a small sample size such as $N = 80$ or 120, both for large and small odds ratios.

For case-control data, we are most interested in the rare disease scenario, i.e. $\mu = -6$, but for completeness, we include $\mu \in \{-6, 0, 1\}$ (Figure 3). The number of cases R equals the number of controls S in these numerical studies. Estimated size was close to the nominal 0.05 level for an odds ratio of 1.0 in each case. For a rare disease ($\mu = -6$) with $k = 4$, the power of $\tilde{Q}$ nearly coincides with that of Q for odds ratios above 1.0. Indeed, the simulated power of $\tilde{Q}$ can exceed that of Q very slightly, especially if odds ratio is slightly above 1.0. For $N = 360$, there is very little discrepancy for an odds ratio above 1.0. For an odds ratio below 1.0, the power of Q can exceed that of $\tilde{Q}$ appreciably, especially for $N = 120$. For example, with $N = 120$, $k = 4$, and odds ratio 0.3 comparing the highest to the lowest category, the power of Q is 0.6689, the power of $\tilde{Q}$ is 0.580, and the reduction in power is 0.088. Differences ranged from -0.0195 to 0.088 for $N = 120$ and from -0.017 to 0.049 for $N = 360$. For $k = 2$ with $\mu = -6$, the estimated sizes of the tests Q and $\tilde{Q}$ were respectively 0.054 and 0.051 for $N = 120$, and 0.051 and 0.052 for $N = 360$. The power of Q tended to exceed that of $\tilde{Q}$ both for odds ratios above 1.0 and for odds ratios below 1.0 (upper left panel, Figure 3). There is a small region of positive odds ratios near 1.0 for which $\tilde{Q}$ had slightly greater power than Q, but the differences were so small as to be imperceptible in Figure 3. Larger losses in power from the use of $\tilde{Q}$, compared to Q, are evident for $\mu = 0$ or 2 (lower panels, Figure 3). Although case-control designs are usually used for rare diseases ($\mu = -6$), they could be employed for common diseases to avoid the need for prospective follow-up or to reduce costs if exposure assessment is expensive.

To identify values of k and N for which the power of $\tilde{Q}$ is appreciably less than that of Q, we plotted the loss of power against $\exp(\beta^*)$ for case-control data (Figure 4). For the rare disease setting $\mu = -6$, and with $k \geq 3$, power loss from estimating quantiles is within the range $[-5\%, 5\%]$ even for $N = 80$ for odds ratio $\exp(\beta^*) > 0.5$. Power loss can exceed 5% for odds ratios below 0.5, however, if $N = 80$ or 120. For $k = 2$, the power loss exceeds 5% for a range of odds ratios $\exp(\beta^*)$ greater than 3 for $N = 80$ and 120 and also for odds ratios < 0.7 (Figure 3). For $\mu = 0$ and $\mu = 2$ (Figure 4), some very large power losses are identified with $N = 80$ or 120, even for $k = 5$, when the odds ratio is less than 0.7. For odds ratios above 1.6, the loss of power is appreciable for $k = 2$ and 3, but less for $k = 4$ or 5.

We also compared these simulated results with the results based on asymptotic theory for cohort data (equations (3) and (5)) and for case-control data (equations (7) and (9)). The agreement between theory and simulations was excellent for $N = 120$ and $N = 360$, but we noted small differences for $N = 80$. Figures 5–8 for $N = 80$, 120 and 360 provide details respectively for cohort designs for X, cohort designs for W, case-control designs for X and case-control designs for W. For cohort data theoretical power agrees well with simulations for X, but for W, with $N = 80$

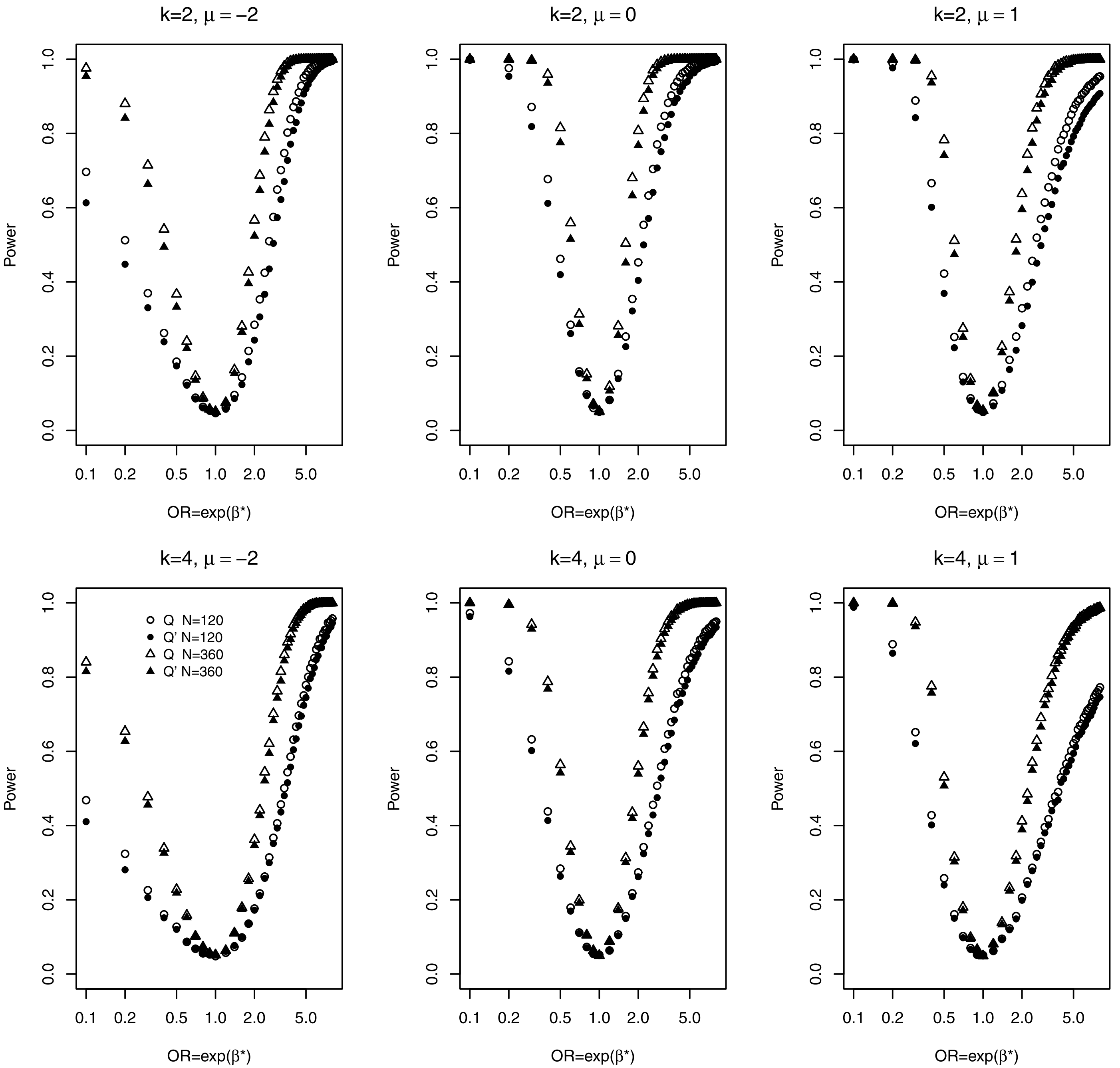

Figure 1. Estimated power from simulations in cohort designs for the Cochran-Armitage trend test Q with known quantiles (open symbols) and for the trend test Q' with estimated quantiles (solid symbols). In each picture, circles and triangles correspond to total sample size $N = 120$ and 280 respectively. The upper panel is for $k = 2$ categories and the bottom panel is for $k = 4$ categories.

and $k = 2$, theoretical power is lower by 0.007 to 0.051 (absolute power difference) for small odds ratios $\exp(\beta^*) < 1$ when $\mu = -2$ and by 0.046 for large odds ratios when $\mu = 1$. For cohort data with $N = 80$ and $k = 4$, theoretical power for Q' exceed estimates from simulations by 0 to 0.072 for small odds ratios when $\mu = -2$ and agrees well with simulations for large odds ratios. For case-control data, theoretical power agrees well with simulations for X and in most scenarios for W except for small N with $\mu = -6$ and $k = 2$. The discrepancy between theory and simulations for $N = 80$ ranges from -0.056 to 0.039 and for $N = 120$ ranges from -0.041 to 0.035.

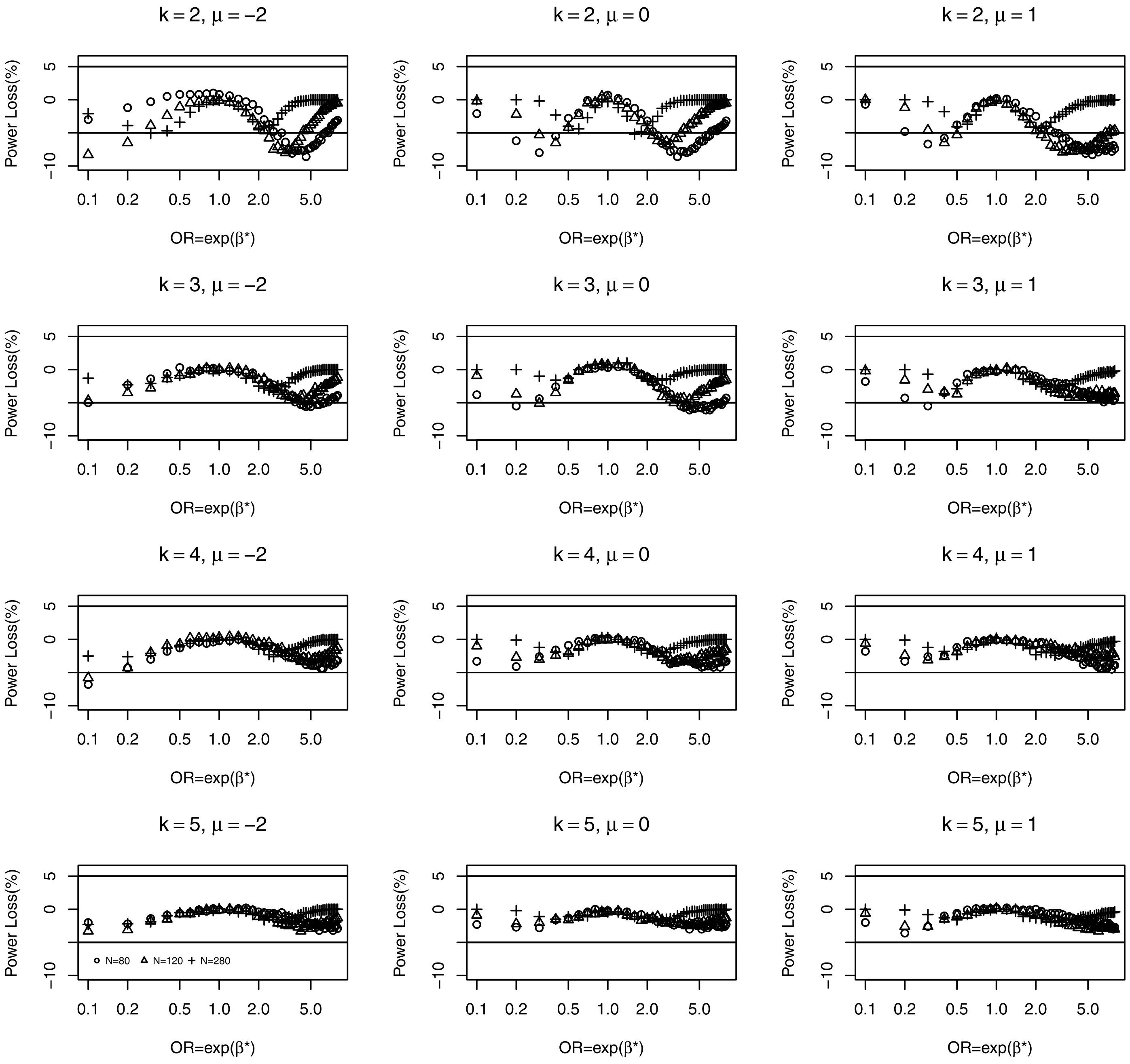

Figure 2. Estimated power loss in percent for cohort data from using estimated quantiles of risk instead of known population quantiles as a function of odds ratio (comparing the highest to lowest category; abscissa; on log scale) and sample size $N = 80, 120, 280$. *Results are shown for* $k = 2, 3, 4$ *or* 5 *exposure categories and for logistic intercepts* $\mu = -2, 0,$ *or* 1. *Power losses were estimated from simulations.*

4. DISCUSSION

We evaluated the power of Cochran-Armitage trend test for cohort designs and for case-control designs in which sample quantiles of exposure in the cohort or in controls from a case-control study, respectively, are used to define the cutpoints that separate exposure score categories. In fact, many studies proceed in this manner, even though power calculations for these studies often assume that the population quantiles for the exposure categories are known, and that disease risk depends on the known exposure scores through equation (1). It was therefore of interest to examine the procedures based on sample quantiles and see if the results based on known population quantiles were misleading.

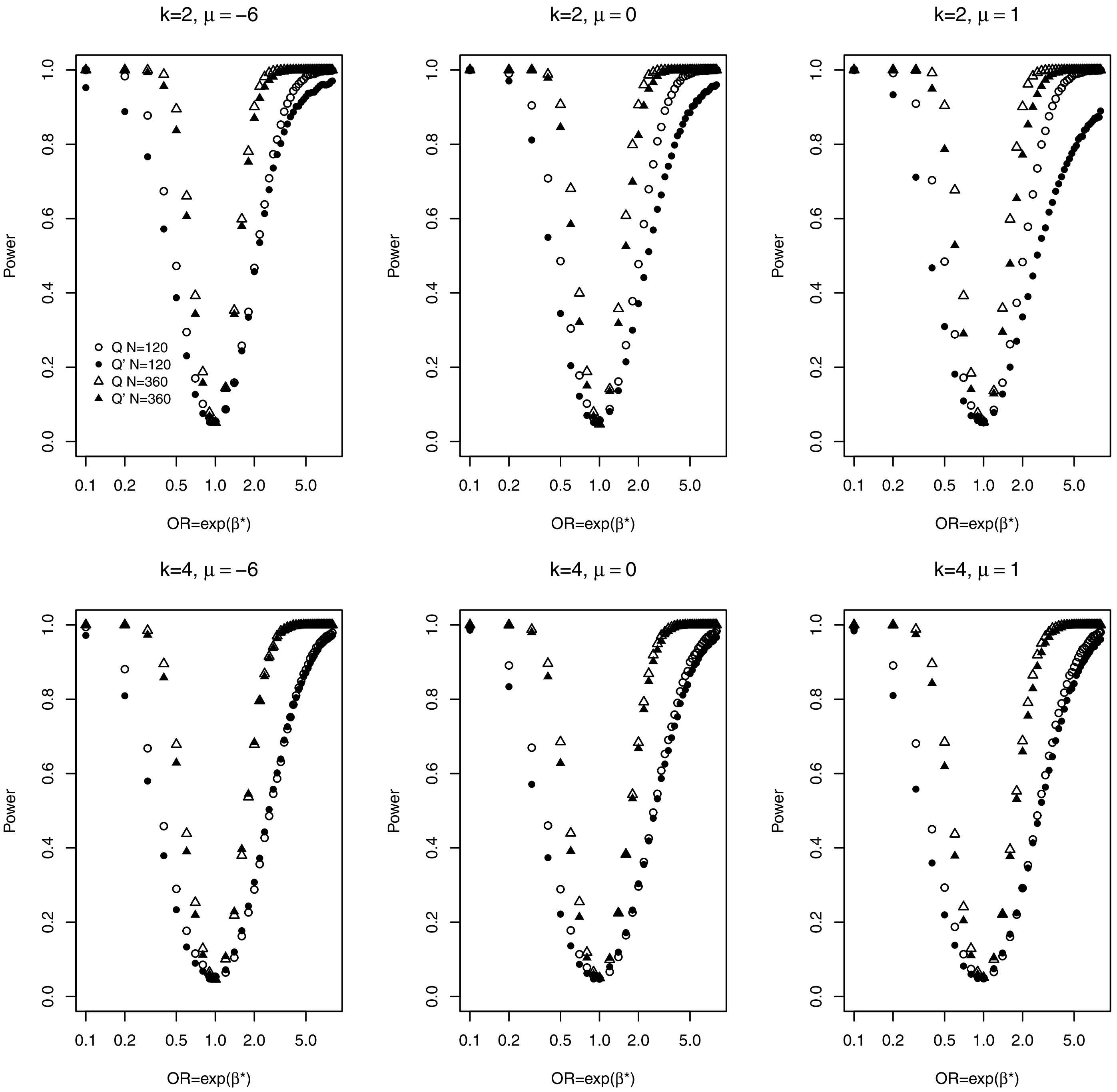

Figure 3. Estimated power from simulations in case-control designs for the Cochran-Armitage trend test Q with known quantiles (open symbols) and for the trend test $\tilde{Q}$ with estimated quantiles (solid symbols). In each picture, circles and triangles correspond to total sample size $N = 120$ and 360 respectively. The upper panel is for $k = 2$ categories and the bottom panel is for $k = 4$ categories.

For cohort studies, the trend tests based on estimated quantiles had near nominal size. For $k \geq 4$ power loss was less than 0.05 except for an uncommon outcome $\mu = -2$, $N = 80$, and a very small odds ratio, 0.1, comparing highest to lowest category (Figure 2). With $k = 2$ or 3, power losses can exceed 0.05 with $N = 80$ or 120 for large and small odds ratios. Therefore to avoid power loss above 5% from using sample quantiles in cohort studies with $N = 120$ or fewer subjects, we recommend using at least $k = 4$ categories. For larger cohorts such as $N = 280$ or greater, loss of power is minimal, even with $k = 2$.

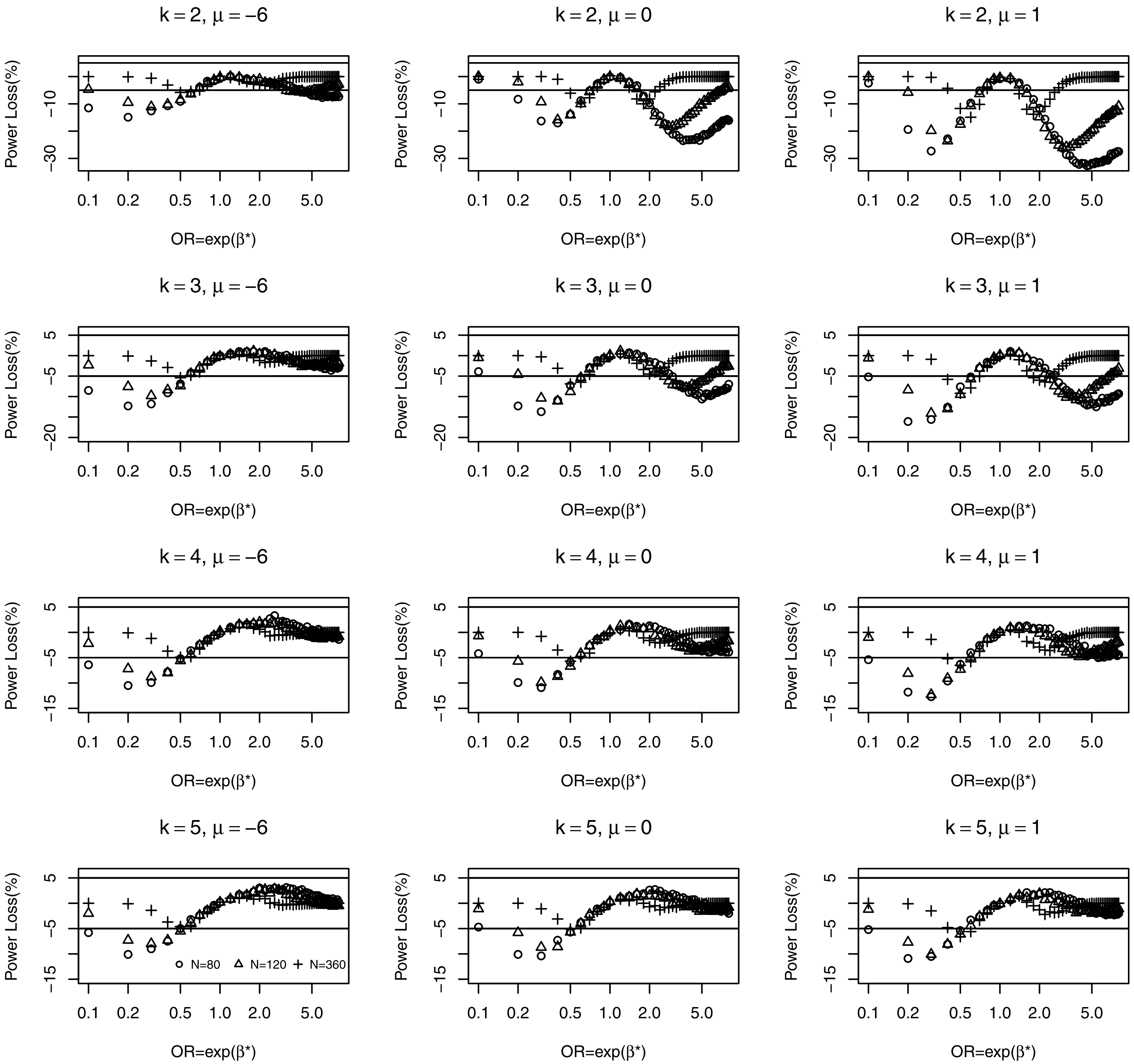

Figure 4. Estimated power loss in percent for case-control data from using estimated quantiles of risk instead of known population quantiles as a function of odds ratio (comparing the highest to lowest category; abscissa; on log scale) and sample size $N = 80, 120, 360$. Results are shown for $k = 2, 3, 4$ or 5 exposure categories and for logistic intercepts $\mu = -6, 0,$ or 1. Power losses were estimated from simulations.

For case-control studies of rare diseases ($\mu = -6$), the trend tests based on estimated quantiles also had near nominal size. For case-control studies with $N = 360$ cases plus controls, power loss from estimating quantiles from controls is minimal. For smaller studies (e.g. $N = 80$ or 120) the power loss can exceed 0.05 for odds ratios of 0.7 or less, even with $k = 5$; for odds ratios of 3.5 or more, power loss can exceed 0.05 if k is less than 4. In the latter case, using $k = 4$ or 5 largely eliminates this power loss. Based on these analyses, we would recommend that to avoid power loss from estimating quantiles, case-control studies of rare diseases have a sample size of 360 or more cases plus controls. Smaller sample sizes won't lead to appreciable power loss from estimating quantiles if k is 4 or

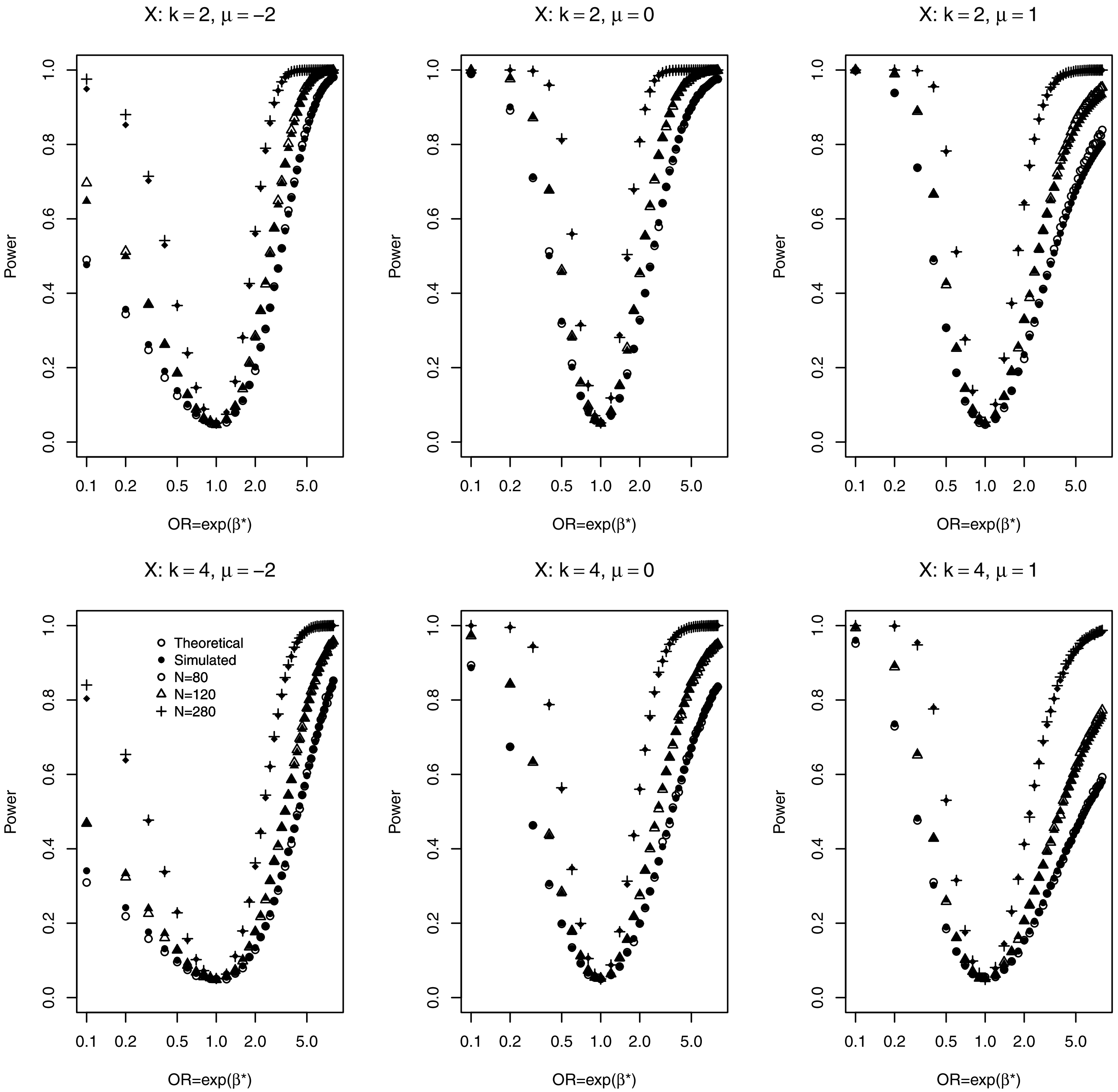

Figure 5. Theoretical (open symbols) and simulated (solid symbols) power estimates in cohort designs for the Cochran-Armitage trend test with known quantiles X with $k = 2$ (upper panel) and with $k = 4$ (lower panel) with $k = 2$. The left, middle and right columns correspond to $\mu = -2, 0, 1$ respectively.

more and the odds ratio comparing highest to lowest exposure categories exceeds one. For case-control studies of common events, such as progression of previously diagnosed macular degeneration, power losses can be more extreme but can be controlled by using the guidelines for rare disease.

In the Introduction, we discussed the cohort study of [9], which was designed to relate glycemic index, grouped by quintiles, to the risk of uterine leiomyoma. Because $N = 21,861$ in this study, there is no power loss from estimating quintiles. On the other hand, [6] reported on a cohort study of renal allograft rejection in 91 transplanted patients.

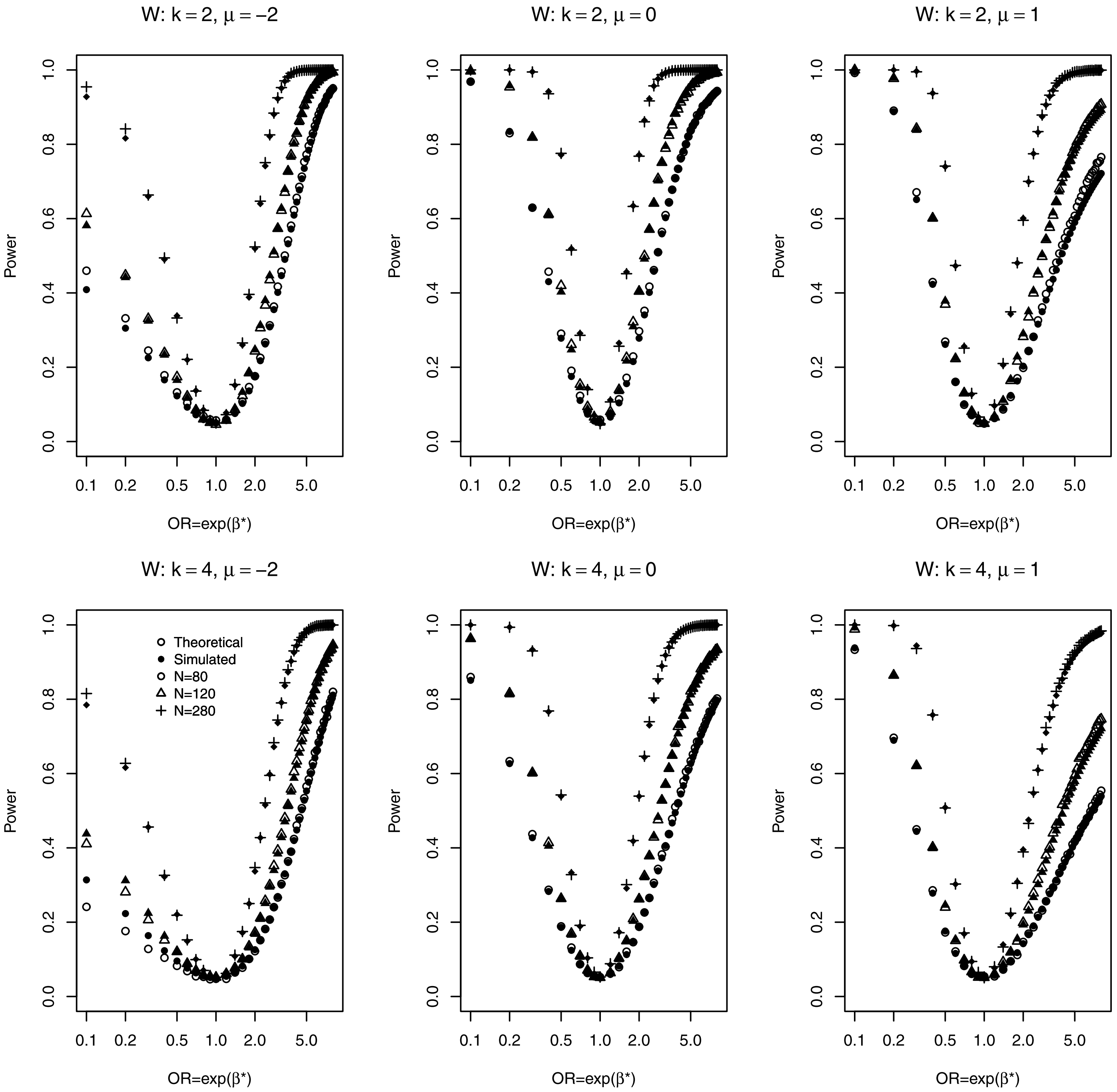

Figure 6. Theoretical (open symbols) and simulated (solid symbols) power estimates in cohort designs for the Cochran-Armitage trend test with known quantiles W with $k = 2$ (upper panel) and with $k = 4$ (lower panel) with $k = 2$. The left, middle and right columns correspond to $\mu = -2, 0, 1$ respectively.

They were interested in studying the effects of continuous immunologic parameters, such as the percentage of lymphocytes that expressed an immunologic marker, on the risk of rejection. Their primary analysis compared marker percentages in those who did or did not have allograft rejection. Had they instead fit a logistic model for risk of rejection in this small cohort based on quantiles of the distribution of marker percentage, it would be necessary to use at least $k = 4$ categories to avoid power loss from estimating quantiles. Likewise consider the small case-control study of staining intensity in 91 cases of esophageal cancer and 103 controls [1]. Because the investigators used $k = 5$ categories based

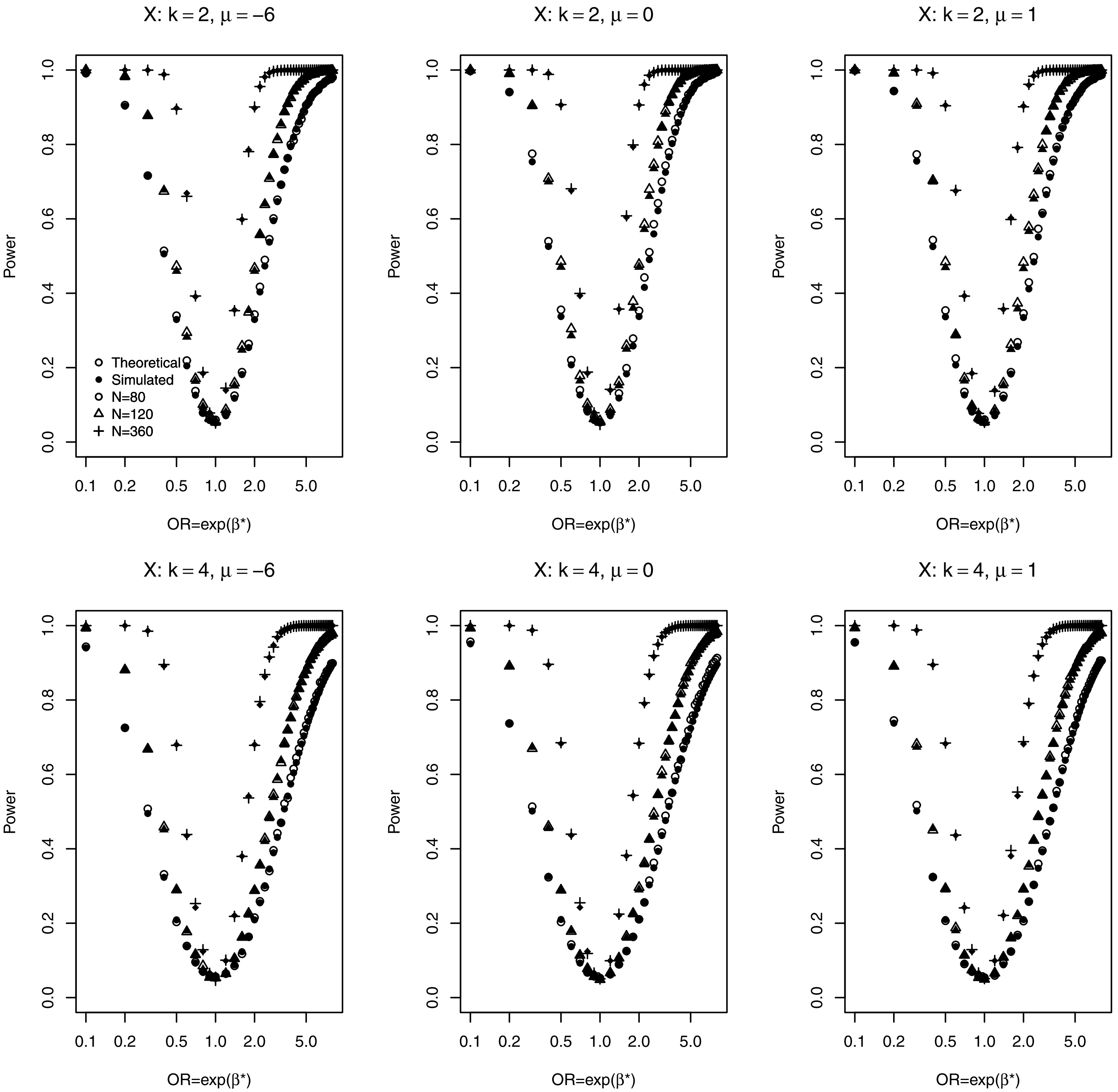

Figure 7. Theoretical (open symbols) and simulated (solid symbols) power estimates in case-control designs for the Cochran-Armitage trend test with known quantiles X with $k = 2$ (upper panel) and with $k = 4$ (lower panel) with $k = 2$. The left, middle and right columns correspond to $\mu = -6, 0, 1$ respectively.

on quintiles, power loss from estimating quintiles should be less than 0.05, provided that risk increased with screening intensity, as was the case in these data (Figure 4).

Our work addresses a different problem and uses different analytical techniques than the paper by [3]. They considered the joint distribution of two continuous exposures, calculated the quantiles of each marginal distribution, and developed the joint asymptotic distribution of counts in cells formed by jointly categorizing the two exposures based on their respective quantiles. The asymptotic theory has applications to problems such as calculating the distribution of the kappa statistic and other measures of agreement. [3] did not consider relating a categorized exposure to an outcome, as in the current paper, however.

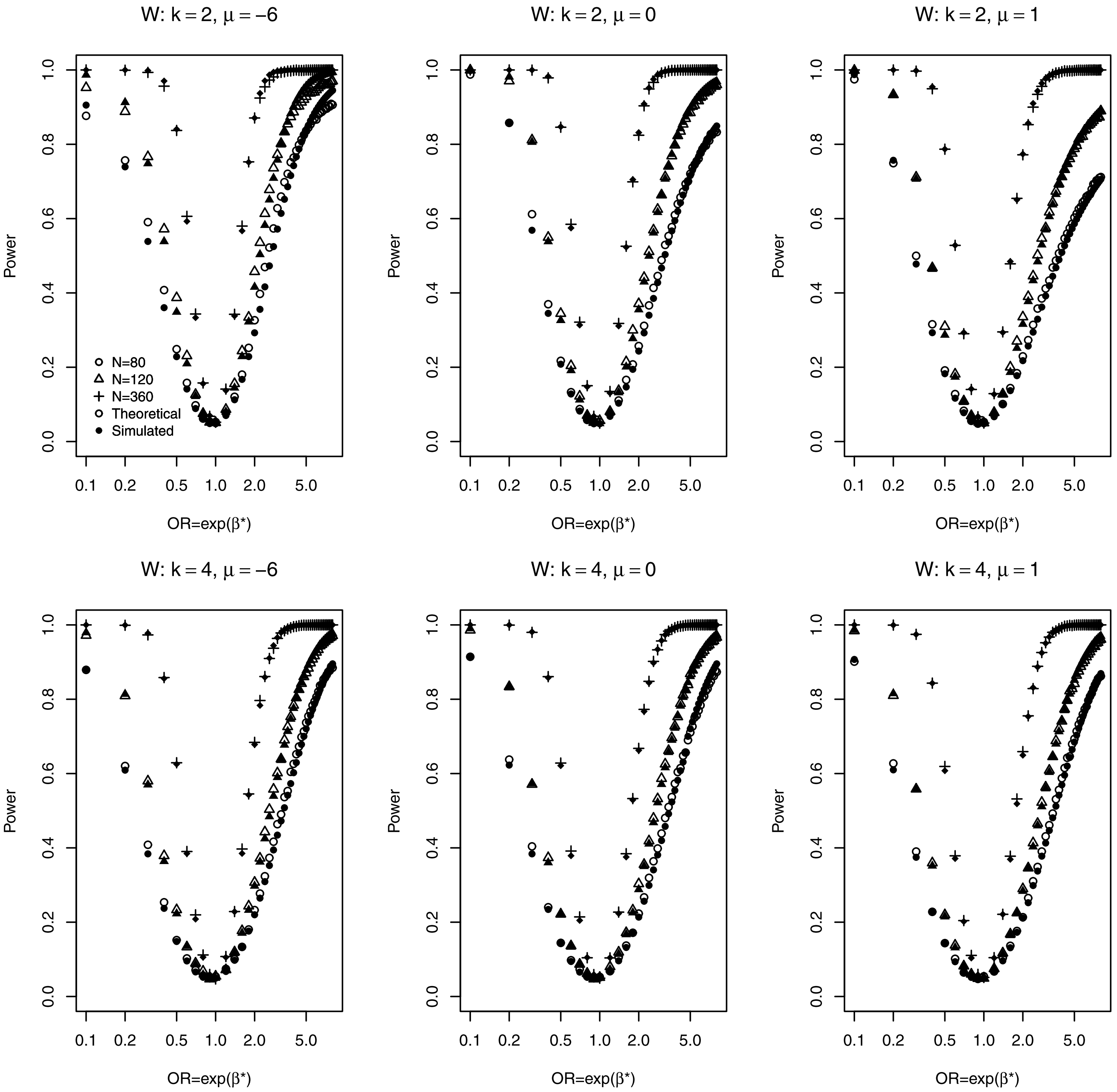

Figure 8. Theoretical (open symbols) and simulated (solid symbols) power estimates in case-control designs for the Cochran-Armitage trend test with known quantiles W with $k = 2$ (upper panel) and with $k = 4$ (lower panel) with $k = 2$. The left, middle and right columns correspond to $\mu = -6, 0, 1$ respectively.

Our methods for cohort data are simpler and quite different from those needed for case-control data. The simplicity of the approach for cohort data derives from the facts that quantiles are estimated from the entire sample and that the problem can be treated as a measurement error problem with errors in a baseline covariate. In contrast, for the case-control design, quantiles are estimated from controls only, and, with respect to the retrospective likelihood, categorization affects the outcomes, rather than the conditioning variable (case status).

The guidelines we present and Figures 1–4 give information on the effects of estimating quantiles on study power for a broad range of cohort and case-control designs. However, the practitioner who is concerned about the potential loss in power from using estimated quantiles, especially when the sample size is small (e.g. less than 120), can estimate power with simulations, as described in Section 3. This simulation method is applicable regardless of the actual distribution of exposure, because the category-based trend tests depend only on ranks. The asymptotic power computations in this paper are likewise general. They are based on theory for order statistics and require numerical integrations (see Section 2). Asymptotic theory and simulations agree well for $N = 120$ and $N = 360$, but simulations are preferred for sample sizes such as $N = 80$, for which the theoretical power deviates slightly from the simulated estimates.

ACKNOWLEDGEMENTS

The authors thank the Associate Editor and referee for many valuable and useful comments for improving the contents and clarity of this paper.

Received 3 March 2011

REFERENCES

[1] Abedi-Ardekani, B., Kamangar, F., Hewitt, S. M., Hainaut, P., Sotoudeh, M., Abnet, C. C., Taylor, P. R., Boffetta, P., Malekzadeh, R., and Dawsey, S. M. (2010). Polycyclic aromatic hydrocarbon exposure in oesophageal tissue and risk of oesophageal squamous cell carcinoma in north-eastern Iran. *Gut* **53** 1054–1069.

[2] Armitage, P. (1955). Test for linear trend in proportions and frequencies. *Biometrics* **11** 375–386.

[3] Borkowf, C. B., Gail, M. H., Carroll, R. J., and Gill, R. D. (1997). Analyzing bivariate continuous data grouped into categories defined by empirical quantiles of marginal distributions. *Biometrics* **53** 1054–1069.

[4] Cochran, W. G. (1954). Some methods for strengthening the common chi-squared tests. *Biometrics* **10** 417–451. MR0067428

[5] David, H. A. and Nagaraja, H. N. (2003). *Order Statistics*, Wiley, New York. MR1994955

[6] Hueso, M., Mestre, M., Benavente, Y., Bas, J., Griny, J. M., and Navarro, E. (2011). Pretransplant low CD3+CD25high cell counts or a low CD3+CD25high/CD3+ HL-DR+ ratio are associated with an increased risk to acute renal allograft rejection. *Transplantation* **92** 536–542.

[7] Mantel, N. (1963). Chi-square tests with one degree of freedom: Extensions of the Mantel-Haenszel procedure. *Journal of the American Statistical Association* **48** 690–700. MR0153079

[8] Nam, J. (1987). A simple approximation for calculating sample sizes for detecting linear trend in proportions. *Biometrics* **43** 701–705.

[9] Radin, R. G., Palmer, J. R., Rosenberg, L., Kumanyika, S.K., and Wise, L. A. (2010). Dietary glycemic index and load in relation to risk of uterine leiomyomata in the Black Women's Health Study. *American Journal of Clinical Nutrition* **91** 1281–1288.

Huilin Li
Division of Biostatistics
School of Medicine
New York University
650 First Ave, Room 547
New York, NY 10016
USA
E-mail address: huilin.li@nyumc.org

Mitchell H. Gail
Biostatistics Branch
Division of Cancer Epidemiology and Genetics
National Cancer Institute
Executive Plaza South, Room 8032
Bethesda, MD 20892-7244
USA
E-mail address: gailm@mail.nih.gov