



Published in final edited form as:

Biometrics. 2009 December ; 65(4): 1270–1278. doi:10.1111/j.1541-0420.2008.01183.x.

Group Testing Regression Models with Fixed and Random Effects

Peng Chen¹, Joshua M. Tebbs^{1,*}, and Christopher R. Bilder²

¹Department of Statistics, University of South Carolina, Columbia, SC 29208, U.S.A.

²Department of Statistics, University of Nebraska, Lincoln, NE 68583, U.S.A.

Summary

Group testing, where subjects are tested in pools rather than individually, has a long history of successful application in infectious disease screening. In this paper, we develop group testing regression models to include covariate effects which are best regarded as random. We present approaches to fit mixed effects models using maximum likelihood, investigate likelihood ratio and score tests for variance components, and evaluate small sample performance using simulation. We illustrate our methods using chlamydia and gonorrhea data collected by the state of Nebraska as part of the Infertility Prevention Project.

Keywords

Generalized linear mixed model; Latent binary response; Likelihood ratio test; Monte Carlo EM algorithm; Pooled testing; Score test

1. Introduction

The Infertility Prevention Project (IPP) is a national program, funded by the Centers for Disease Control and Prevention, aimed at providing routine screening, prevention strategies, and treatment for individuals with chlamydia and/or gonorrhea infection. Chlamydia and gonorrhea are the two most common sexually transmitted diseases (STDs) in the United States, and each year the US spends approximately \$4 billion on assessment, testing, and treatment for these infections (Screening and Treatment Guidelines, IPP, Region VII, 2003). Unlike most viral STDs, individuals with chlamydia and gonorrhea, both bacterial infections, are usually asymptomatic, and left untreated, both infections can result in a variety of long-term sequelae, such as pelvic inflammatory disease, ectopic pregnancy, and infertility (Kacena et al., 1998a, 1998b).

This research arises from our interaction with medical collaborators in Nebraska (Region VII of the IPP), where more than 30,000 individuals are tested for chlamydia and gonorrhea each year. With the current testing cost of about \$15 per individual expected to rise as public health costs increase nationally, our colleagues have expressed an interest in implementing group testing (also known as pooled testing) as a means of surveillance in the state. Group testing, where subjects are tested in pools rather than individually, has a long history of successful application in infectious disease screening, dating back to the work of Dorfman (1943). Today, pooling individual samples (e.g., blood, urine, etc.) is a common strategy to reduce testing costs and has been implemented in a variety of infectious disease applications involving HIV, hepatitis B/C (Cardoso, Koerner, and Kubanek, 1998; Pilcher et al., 2005), and elsewhere for

* email: tebbs@stat.sc.edu

chlamydia and gonorrhea screening (Kacena et al., 1998a, 1998b; Lindan et al., 2005; Rours et al., 2005).

Traditionally, statistical research in group testing has focused on homogeneous populations; that is, the positive/negative statuses of individuals tested are assumed to be independent and identically distributed (iid) random variables. However, in most infectious disease studies, there is covariate information available for each individual, and a major thrust of the analysis is to determine which covariates are associated with individual positivity. To address this issue, Vansteelandt, Goetghebeur, and Verstraeten (2000) and Xie (2001) have each proposed fixed effects regression approaches to account for individual covariate information with pooled binary responses. Vansteelandt et al. (2000) use a method by which maximum likelihood estimates are computed directly using the pooled responses, whereas Xie (2001) treats the individual responses as unobserved and uses the EM algorithm.

Unifying the approaches of Vansteelandt et al. (2000) and Xie (2001) is the assumption that individual (latent) statuses are independent. However, one aspect of the Nebraska study conspicuously calls this assumption into question. Because testing takes place at different clinic site locations throughout the state, individuals are inherently clustered by design and may share common characteristics. For example, the level of gonorrhea and chlamydia infection in Douglas County (which includes Omaha) has recently been described as at “epidemic” levels (Zagurski, 2006), whereas infection levels in other parts of the state are relatively lower. If the prevalence of chlamydia and gonorrhea does vary geographically, it would be more natural to regard the individual latent infection statuses as correlated and to conceptualize clinic locations as random effects. Taking this perspective, the analysis of group testing samples from the Nebraska IPP requires the development of new regression methods that allow for within-cluster correlation among the individual latent binary responses.

In this paper, we generalize the group testing regression modeling approach of Vansteelandt et al. (2000) to include random effects. In Section 2, we obtain maximum likelihood estimates and discuss large sample inference for fixed effects parameters. In Section 3, we present likelihood ratio and score tests for variance components which allow one to assess heterogeneity among clusters. In Section 4, we provide simulation evidence to show that our model fitting approaches behave well and that tests for variance components have very good size and power properties in finite samples. In Section 5, we illustrate the use of our new analytical methods with the Nebraska IPP data. In Section 6, we summarize and discuss extensions of this work.

2. Estimation and Inference

2.1. Notation and Assumptions

We are interested in developing mixed effects regression techniques to model the probability of a single infection (e.g., chlamydia) using pooled testing results. In doing so, we view individuals from a given clinic site as a cluster. Within the i th site, suppose that each individual is randomly assigned to one of n_i pools and let $Y_{ijk} = 1$ if the k th individual in the j th pool at site i is positive, $Y_{ijk} = 0$ otherwise, for $i = 1, 2, \dots, l$, $j = 1, 2, \dots, n_i$, and $k = 1, 2, \dots, c_{ij}$. We henceforth refer to c_{ij} as the pool size. We let $\mathbf{u} = (\mathbf{u}'_1, \mathbf{u}'_2, \dots, \mathbf{u}'_l)'$, where \mathbf{u}_i denotes a $q \times 1$ random effect vector for site i , and relate the latent status Y_{ijk} to the covariate information through the generalized linear mixed model

$$\Pr(Y_{ijk} = 1 | \mathbf{u}_i) = g^{-1}(\beta' \mathbf{x}_{ijk} + \mathbf{u}'_i \mathbf{z}_{ijk}), \quad (1)$$

where β is a $p \times 1$ vector of fixed effects parameters, \mathbf{x}_{ijk} is a $p \times 1$ covariate vector associated with the fixed effects, \mathbf{z}_{ijk} is a $q \times 1$ covariate vector associated with the site-specific random effects, and $g(\cdot)$ is a known, monotonic, differentiable link function. It is worth emphasizing that Y_{ijk} is not observed, unless $c_{ij} = 1$, which corresponds to testing subjects individually.

Because identifying positive individuals is not our goal, similarly to Vansteelandt et al. (2000), we consider the case where only the initial pooled responses are observed; that is, subsets of positive pools are not retested further. Define $T_{ij} = 1$ if the j th pool at site i tests positive, $T_{ij} = 0$ otherwise, for $i = 1, 2, \dots, l$ and $j = 1, 2, \dots, n_i$. We assume that the statuses of individuals and pools from different sites are independent and that, conditional on \mathbf{u}_i , the statuses of individuals and pools within site i are independent so that

$$\Pr(T_{ij}=1|\mathbf{u}_i) = \gamma_1 + (1 - \gamma_1 - \gamma_2) \prod_{k=1}^{c_{ij}} \{1 - g^{-1}(\beta' \mathbf{x}_{ijk} + \mathbf{u}'_i \mathbf{z}_{ijk})\},$$

where γ_1 and γ_2 denote assay sensitivity and specificity, respectively. It is assumed that γ_1 and γ_2 are constants close to 1 and do not depend on c_{ij} , a reasonable assumption with modern diagnostic assays based on nucleic acid technology (NAT). For a number of chlamydia studies, NATs have been shown to have high sensitivity and specificity for pools of up to size 10 when pooling urine or cervical swabs (see, e.g., Kacena et al., 1998b). Similar results have been observed in gonorrhea studies (Kacena et al., 1998a).

We treat the site-specific $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_l$ as iid multivariate normal random vectors with mean 0 and covariance matrix $\mathbf{D} \equiv \mathbf{D}(\boldsymbol{\varphi})$, where $\boldsymbol{\varphi}$ is an $m \times 1$ vector of variance components. Denote by $\boldsymbol{\theta} = (\beta', \boldsymbol{\varphi}')$ the $(p + m) \times 1$ vector of parameters. Under these assumptions, the log-likelihood of $\boldsymbol{\theta}$ based on the observed pool responses $\mathbf{T} = (T_{ij})$ is

$$l(\boldsymbol{\theta}|\mathbf{T}) = l(\beta, \boldsymbol{\varphi}|\mathbf{T}) = \sum_{i=1}^l \log \left\{ \int_{\mathcal{R}^m} \prod_{j=1}^{n_i} f_{ij}(c_{ij}, \mathbf{x}_{ij}, \mathbf{u}_i, \beta) f(\mathbf{u}_i|\boldsymbol{\varphi}) d\mathbf{u}_i \right\}, \tag{2}$$

where $\mathbf{x}_{ij} = (\mathbf{x}'_{ij1}, \mathbf{x}'_{ij2}, \dots, \mathbf{x}'_{ijc_{ij}})'$,

$$f_{ij}(c_{ij}, \mathbf{x}_{ij}, \mathbf{u}_i, \beta) = \left[\gamma_1 + (1 - \gamma_1 - \gamma_2) \prod_{k=1}^{c_{ij}} \{1 - g^{-1}(\beta' \mathbf{x}_{ijk} + \mathbf{u}'_i \mathbf{z}_{ijk})\} \right]^{T_{ij}} \times \left[1 - \gamma_1 - (1 - \gamma_1 - \gamma_2) \prod_{k=1}^{c_{ij}} \{1 - g^{-1}(\beta' \mathbf{x}_{ijk} + \mathbf{u}'_i \mathbf{z}_{ijk})\} \right]^{1-T_{ij}}, \tag{3}$$

and $f(\mathbf{u}_i|\boldsymbol{\varphi}) = (2\pi)^{-q/2} |\mathbf{D}|^{-1/2} \exp(-\mathbf{u}'_i \mathbf{D}^{-1} \mathbf{u}_i/2)$ is the $\mathcal{N}_q(0, \mathbf{D})$ probability density function.

2.2. Maximum Likelihood Estimation

We consider two approaches to find the MLE $\hat{\boldsymbol{\theta}} = (\hat{\beta}', \hat{\boldsymbol{\varphi}})'$ and evaluate the relative merits of each. First, we use adaptive Gauss-Hermite quadrature (Pinheiro and Bates, 1995) to approximate the integral in (2) and then maximize the approximated log-likelihood with respect to $\boldsymbol{\theta}$ using a Newton-Raphson procedure. This approach is straightforward to implement when

q is small; however, its use may be ill-advised for large q as the number of quadrature points needed to confer reasonable precision increases exponentially (Booth, Hobert, and Jank, 2001). As a second approach, we formulate a general Monte Carlo expected maximization (MCEM) algorithm (McCulloch, 1997). While this approach is computationally more expensive, it could be modified to allow for other random effects distributions (Chen, Zhang, and Davidian, 2002), and it has the potential to be applicable with other pooling strategies not considered here; see Section 6.

The MCEM approach is now described in our hierarchical setting. Treating the random effect \mathbf{u} as missing, the complete data log-likelihood can be written as

$$\begin{aligned} l_c(\theta|\mathbf{T}, \mathbf{u}) &= \log\{P(\mathbf{T}, \mathbf{u}|\beta, \varphi)\} = \log\{P(\mathbf{T}|\mathbf{u};\beta) f(\mathbf{u}|\varphi)\} \\ &= \sum_{i=1}^l \sum_{j=1}^{n_i} \log f_{ij}(c_{ij}, \mathbf{x}_{ij}, \mathbf{u}_i, \beta) - \frac{1}{2} \log|\mathbf{D}| - \frac{1}{2} \sum_{i=1}^l \mathbf{u}_i' \mathbf{D}^{-1} \mathbf{u}_i + c_0, \end{aligned} \quad (4)$$

where $P(\mathbf{T}, \mathbf{u}|\beta, \varphi)$ is the density function of (\mathbf{T}, \mathbf{u}) , $P(\mathbf{T}|\mathbf{u}; \beta)$ is the conditional density of \mathbf{T} given \mathbf{u} , $f_{ij}(c_{ij}, \mathbf{x}_{ij}, \mathbf{u}_i, \beta)$, is defined as in (3), and $c_0 = -q \log(2\pi)/2$ is free of θ . We define

$$\begin{aligned} I_1 &= \sum_{i=1}^l \sum_{j=1}^{n_i} \log f_{ij}(c_{ij}, \mathbf{x}_{ij}, \mathbf{u}_i, \beta) \\ I_2 &= -\frac{1}{2} \log|\mathbf{D}| - \frac{1}{2} \sum_{i=1}^l \mathbf{u}_i' \mathbf{D}^{-1} \mathbf{u}_i, \end{aligned} \quad (5)$$

noting that I_1 depends only on β and that I_2 depends only on φ through $\mathbf{D} = \mathbf{D}(\varphi)$. With the observed data \mathbf{T} , the E-step involves calculating $E(I_1|\mathbf{T})$ and $E(I_2|\mathbf{T})$ for a given estimate $\theta^{(b)} \equiv (\beta^{(b)}, \varphi^{(b)})$. These expectations can not be calculated in closed form, so we approximate them using Monte Carlo simulation; that is, we generate a large number (M) of random draws from the conditional distribution $f(\mathbf{u}|\mathbf{T}; \theta^{(b)})$ and use the sample means of I_1 and I_2 to estimate $E(I_1|\mathbf{T})$ and $E(I_2|\mathbf{T})$, respectively. Because $f(\mathbf{u}|\mathbf{T}; \theta^{(b)})$ is not available in closed form, we use the Metropolis-Hastings (MH) algorithm to sample from it. Since pools from different sites are independent, it suffices to generate samples from $f(\mathbf{u}_i|\mathbf{T}; \theta^{(b)})$, for each $i = 1, 2, \dots, l$. We choose the $\mathcal{N}_q(0, \mathbf{D}(\varphi^{(b)}))$ distribution as the proposal function for the MH algorithm. Summarizing, the MCEM algorithm is implemented as follows:

1. Choose a starting value for $\theta^{(0)} = (\beta^{(0)}, \varphi^{(0)})$.
2. (E-step). For a given $b = 0, 1, 2, \dots$, approximate $E(I_1|\mathbf{T})$ and $E(I_2|\mathbf{T})$ by

$$\begin{aligned} \widehat{I}_1^{(b)} &= \frac{1}{M} \sum_{h=1}^M \sum_{i=1}^l \sum_{j=1}^{n_i} \log f_{ij}(c_{ij}, \mathbf{x}_{ij}, \mathbf{u}_i^{(h)}, \beta) \\ \widehat{I}_2^{(b)} &= -\frac{1}{2} \log|\mathbf{D}| - \frac{1}{2M} \sum_{h=1}^M \sum_{i=1}^l \mathbf{u}_i^{(h)'} \mathbf{D}^{-1} \mathbf{u}_i^{(h)}, \end{aligned}$$

respectively, where $\mathbf{u}_i^{(h)}$, $h = 1, 2, \dots, M$, are M draws from the conditional distribution $f(\mathbf{u}_i|\mathbf{T}; \theta^{(b)})$, $i = 1, 2, \dots, l$, using the MH algorithm (see Web Appendix A).

3. (M-step). Maximize $\widehat{I}_1^{(b)}$ with respect to, β obtaining a new estimate $\beta^{(b+1)}$. Maximize $\widehat{I}_2^{(b)}$ with respect to φ , obtaining a new estimate $\varphi^{(b+1)}$. Set $b = b + 1$.
4. Repeat Steps 2 and 3 until $\|\beta^{(b+1)} - \beta^{(b)}\|$ and $\|\varphi^{(b+1)} - \varphi^{(b)}\|$ are sufficiently small.

2.3. Covariance Matrix Estimation

When one uses quadratures to find $\hat{\theta}$, the negative inverse Hessian at the last iteration of the Newton-Raphson maximization procedure provides an estimate of the covariance matrix of $\hat{\theta}$. With MCEM, an appeal to the missing information principle and the method of Louis (1982) gives the observed information matrix

$$I(\theta) = \frac{\partial^2 l(\theta|\mathbf{T})}{\partial\theta\partial\theta'} = -E \left\{ \frac{\partial^2 l_c(\theta|\mathbf{T}, \mathbf{u})}{\partial\theta\partial\theta'} \middle| \mathbf{T} \right\} - \text{cov} \left\{ \frac{\partial \log l_c(\theta|\mathbf{T}, \mathbf{u})}{\partial\theta} \middle| \mathbf{T} \right\} \\ = -E_{\mathbf{u}|\mathbf{T}} \left[\begin{array}{cc} \frac{\partial^2 I_1}{\partial\beta\partial\beta'} + \frac{\partial I_1}{\partial\beta} \frac{\partial I_1}{\partial\beta'} & \frac{\partial I_1}{\partial\beta} \frac{\partial I_2}{\partial\varphi'} \\ \frac{\partial I_2}{\partial\varphi} \frac{\partial I_1}{\partial\beta'} & \frac{\partial^2 I_2}{\partial\varphi\partial\varphi'} + \frac{\partial I_2}{\partial\varphi} \frac{\partial I_2}{\partial\varphi'} \end{array} \right],$$

where $l(\theta|\mathbf{T})$ is the observed data log-likelihood defined in (2), $l_c(\theta|\mathbf{T}, \mathbf{u})$ is the complete data log-likelihood in (4), and the functions I_1 and I_2 are as given in (5). We estimate $I(\theta)$ using

$$I(\hat{\theta}) = -\frac{1}{M} \sum_{h=1}^M \left[\begin{array}{cc} \frac{\partial^2 I_1}{\partial\beta\partial\beta'} + \frac{\partial I_1}{\partial\beta} \frac{\partial I_1}{\partial\beta'} & \frac{\partial I_1}{\partial\beta} \frac{\partial I_2}{\partial\varphi'} \\ \frac{\partial I_2}{\partial\varphi} \frac{\partial I_1}{\partial\beta'} & \frac{\partial^2 I_2}{\partial\varphi\partial\varphi'} + \frac{\partial I_2}{\partial\varphi} \frac{\partial I_2}{\partial\varphi'} \end{array} \right],$$

where the matrix expression is evaluated at $\theta = \hat{\theta}$ using random samples $\mathbf{u}_i^{(h)}$, $i = 1, 2, \dots, l$, $h = 1, 2, \dots, M$, generated from the conditional density $P(\mathbf{u}|\mathbf{T}; \hat{\theta})$ via the MH algorithm described in Web Appendix A. Regardless of which model fitting technique is used, standard errors are obtained from $I(\hat{\theta})^{-1}$, making the construction of large sample Wald confidence intervals possible. Based on the asymptotic results from Nie (2007), one would expect such confidence intervals to be adequate for l large.

3. Tests for Variance Components

We now consider the problem of testing whether individual random effects are present using the pooled responses. This is of practical interest because if there is no additional variation among sites, the analyst may wish to collapse over the sites and use the simpler regression model of Vansteelandt et al. (2000), which regards all effects as fixed. Herein, we concentrate on the $q = 1$ case and take $\mathbf{z}_{ijk} = 1$, for all i, j , and k , so that the test of interest is

$$H_0: \sigma^2 = 0 \quad \text{versus} \quad H_1: \sigma^2 > 0,$$

where $\sigma^2 = \text{var}(u_i)$. This is sufficient for our purposes as the Nebraska IPP data to be analyzed in Section 5 involves only potential site effects and no site-specific covariates. Generalizations for $q > 1$ are discussed in Section 6. In the literature, the test of H_0 versus H_1 has been described as “nonstandard” (Self and Liang, 1987), because the value of σ^2 under H_0 corresponds to a boundary point of the parameter space. The implementation of one-sided tests in a constrained parameter space has recently garnered attention from Molenberghs and Verbeke (2007). Based on their recommendations, we investigate likelihood ratio and score tests.

3.1. Likelihood ratio test

The likelihood ratio test statistic for H_0 versus H_1 is given by

$$T_{LR} = 2 \ln \left\{ \frac{\max_{H_0 \cup H_1} l(\boldsymbol{\beta}, \sigma^2)}{\max_{H_0} l(\boldsymbol{\beta}, \sigma^2)} \right\},$$

where $l(\boldsymbol{\beta}, \sigma^2)$ is the log-likelihood defined in (2) with $\boldsymbol{\varphi} = \sigma^2$ and $\boldsymbol{\theta} = (\boldsymbol{\beta}', \sigma^2)'$. Computing T_{LR} requires one to fit both the mixed model in Section 2 and the fixed effects model of Vansteelandt et al. (2000). For the scalar variance case, the results of Self and Liang (1987)

apply so that the asymptotic distribution of T_{LR} is the two-component mixture $\frac{1}{2}\chi_0^2 + \frac{1}{2}\chi_1^2$, where χ_0^2 is a point mass distribution at 0 and χ_1^2 denotes the χ^2 distribution with 1 degree of freedom. Large values of T_{LR} are evidence against H_0 .

3.2. Score test

We also propose a score test analogous to Liang (1987) who considers individual response data but adjust the test to acknowledge the one sided nature of H_1 . Specifically, we reparameterize the random intercept for the i th site in (1) as $u_i \equiv \alpha_i = \tau^{1/2}v_i$ and rewrite (1) as

$$\Pr(Y_{ijk} = 1 | v_i) = g^{-1}(\boldsymbol{\beta}' \mathbf{x}_{ijk} + \tau^{1/2}v_i),$$

where v_1, v_2, \dots, v_l are assumed to be iid random variables with unspecified distribution $F(\cdot)$, $E(v_i) = 0$, and $\text{var}(v_i) = 1$. The parameter τ is interpreted as the variance of u_i in (1) so that under this parametrization, the test for homogeneity among sites becomes $H_0 : \tau = \sigma^2 = 0$. The log-likelihood contribution from the i th site is

$$l_i \equiv l_i(\boldsymbol{\beta}, \tau) = \log \left\{ \int_{\mathcal{R}} f_i(\mathbf{x}_i, \tau^{1/2}v_i, \boldsymbol{\beta}) dF(v) \right\},$$

where $\mathbf{x}_i = (\mathbf{x}'_{i1}, \mathbf{x}'_{i2}, \dots, \mathbf{x}'_{ini})'$,

$$f_i(\mathbf{x}_i, \tau^{1/2}v_i, \boldsymbol{\beta}) = \prod_{j=1}^{n_i} f_{ij}(c_{ij}, \mathbf{x}_{ij}, \tau^{1/2}v_i, \boldsymbol{\beta}),$$

and $f_{ij}(c_{ij}, \mathbf{x}_{ij}, \tau^{1/2}v_i, \boldsymbol{\beta})$ is defined as in (3) with $\mathbf{u}_i = u_i = \tau^{1/2}v_i$ and $\mathbf{z}_{ijk} = 1$. Define

$$S(\boldsymbol{\beta}) = \frac{1}{2} \sum_{i=1}^l \left[\left\{ \frac{\partial}{\partial \alpha_i} \log f_i(\mathbf{x}_i, 0, \boldsymbol{\beta}) \right\}^2 + \frac{\partial^2}{\partial \alpha_i^2} \log f_i(\mathbf{x}_i, 0, \boldsymbol{\beta}) \right].$$

Expressions for $(\partial/\partial \alpha_i) \log f_i(\mathbf{x}_i, 0, \boldsymbol{\beta})$ and $(\partial^2/\partial \alpha_i^2) \log f_i(\mathbf{x}_i, 0, \boldsymbol{\beta})$ are provided in Web Appendix

B. The asymptotic variance of $S(\boldsymbol{\beta})$ is equal to $\text{var}\{S(\boldsymbol{\beta})\} = I_{\tau\tau} - \mathbf{I}'_{\tau\beta} \mathbf{I}_{\beta\beta}^{-1} \mathbf{I}_{\tau\beta}$, where

$$I_{\tau\tau} = \sum_{i=1}^l E \left\{ \left(\frac{\partial l_i}{\partial \tau} \right)^2 \right\} \quad \mathbf{I}_{\tau\beta} = \sum_{i=1}^l E \left(\frac{\partial l_i}{\partial \tau} \frac{\partial l_i}{\partial \boldsymbol{\beta}} \right) \quad \mathbf{I}_{\beta\beta} = \sum_{i=1}^l E \left(\frac{\partial l_i}{\partial \boldsymbol{\beta}} \frac{\partial l_i}{\partial \boldsymbol{\beta}'} \right). \tag{6}$$

A score statistic for $H_0 : \sigma^2 = 0$ versus $H_1 : \sigma^2 > 0$ is thus given by

$$T_s = \begin{cases} \{S(\hat{\beta})\}^2 / \widehat{\text{var}}\{S(\hat{\beta})\}, & S(\hat{\beta}) > 0 \\ 0, & S(\hat{\beta}) \leq 0, \end{cases}$$

where $\hat{\beta}$ is the MLE computed under H_0 and $\widehat{\text{var}}\{S(\hat{\beta})\}$ is computed by evaluating $\partial l_i / \partial \tau$, $\partial l_i / \partial \beta$, and the expectations in (6) at $\tau = 0$ and $\beta = \hat{\beta}$. Closed form expressions for $\partial l_i / \partial \tau$ and $\partial l_i / \partial \beta$ are provided in Web Appendix B. Closed form expressions for $I_{\tau\tau}$ and $I_{\tau\beta}$ do not exist, so we approximate them using Monte Carlo simulation to sample pooled responses under H_0 while preserving the original covariate values among all pools. Applying the results in Silvapulle and Silvapulle (1995), the asymptotic distribution of T_S is the same two-component mixture $\frac{1}{2}\chi_0^2 + \frac{1}{2}\chi_1^2$, with large values of T_S being evidence against H_0 .

One small technical detail warrants a remark. In our formulation of the score statistic, we calculate the estimated variance $\widehat{\text{var}}\{S(\hat{\beta})\}$ using expected information, whereas Molenberghs and Verbeke (2007) use observed information. We use expected information because calculating the second derivative of the log-likelihood $l(\theta)$ numerically can produce unstable results; in addition, the observed information matrix can be non-positive definite.

4. Simulation Evidence

4.1. Maximum Likelihood Estimation

We have performed simulations to assess the characteristics of the maximum likelihood estimates in a variety of situations encountered in large scale infectious disease studies. We first generated individual binary responses Y_{ijk} according to the simple logit model

$$\text{logit}\{\Pr(Y_{ijk}=1|u_i)\} = \beta_0 + \beta_1 x_{ijk} + u_i, \quad (7)$$

where $u_i \sim \mathcal{N}(0, \sigma^2)$, $\theta = (\beta_0, \beta_1, \sigma) = (-5, 1, 1)$, and $x_{ijk} \sim \mathcal{N}(0, 0.64)$. These configurations provide a mean prevalence of about 1.4 percent (range: 0.1 to 5.0 percent). We generated $N = 2000$ (10000) individuals with $l = 10$ (40) sites and randomly assigned individual responses to pools within site to create simulated group responses T_{ij} . For simplicity, we take $c_{ij} = c$ and $n_i = n$, for all i and j , so that the overall number of individuals is $N = lnc$. We assume throughout this section that $\gamma_1 = \gamma_2 = 1$ only because the results obtained from imperfect testing ($\gamma_1 < 1$, $\gamma_2 < 1$) were similar. Imperfect tests are considered in Section 5. To fit the model in (7), both adaptive quadratures and MCEM were used. Differences between the fits were negligible so we present the results from quadratures. For comparative purposes, we also fit the model in (7) to the individual simulated data (i.e., $c = 1$) using the glmm package in R. For each (N, l, n, c) combination, 500 data sets were simulated.

The results from Table 1 demonstrate that, in terms of bias, the fixed effects estimates $\hat{\beta}_0$ and $\hat{\beta}_1$ from group testing do about as well as those from individual testing. The variance component estimate $\hat{\sigma}$ from group testing tends to slightly underestimate the true σ , but this occurs for individual testing as well. We compare group testing and individual testing on the basis of N fixed, so it is not surprising that the relative mean squared error (RMSE); that is, the ratio of the MSE from individual testing to the MSE from group testing, is often less than unity. However, it is interesting to note that the RMSE for $\hat{\sigma}$ is not largely affected by the pool size c . This is likely explained by the fact that the variance component σ is a site-level

parameter, and the loss of information (if any) is small when pooling individuals within site. The estimated 95 percent Wald coverage probabilities are also given based on the 500 Monte Carlo data sets. These are mostly within the margin of error, suggesting that the variance estimates from Section 2.3 are adequate.

We now present the results from a second simulation illustrating the performance of the MCEM algorithm with the model

$$\text{logit} \left\{ \Pr \left(Y_{ijk} = 1 \mid \mathbf{u}_i \right) \right\} = \beta_0 + \beta_1 x_{ijk} + u_{i1} + u_{i2} z_{ijk}, \tag{8}$$

where $\mathbf{u}_i = (u_{i1}, u_{i2})'$ follows a bivariate normal distribution with mean $0 = (0, 0)'$, $\text{var}(u_{i1}) = \sigma_1^2$, $\text{var}(u_{i2}) = \sigma_2^2$, and $\text{corr}(u_{i1}, u_{i2}) = \rho$. Additionally, we take $x_{ijk} \sim \mathcal{N}(0, 1)$, $z_{ijk} \sim \mathcal{U}(-3, 3)$, $\boldsymbol{\beta} = (\beta_0, \beta_1)' = (-4, 1)'$, $\boldsymbol{\varphi} = (\sigma_1, \sigma_2, \rho)' = (0.7, 0.5, 0)'$, and $\boldsymbol{\theta} = (\boldsymbol{\beta}', \boldsymbol{\varphi}')$. These configurations provide a mean prevalence of about 4.3 percent (range: 0.1 to 17.4 percent). We generated individual responses according to the model in (8) and created pools by random assignment as before; 200 simulated data sets were used for each (N, l, n, c) combination. For each data set, we initially took the number of Monte Carlo draws to be $M = 1500$ and increased M incrementally based on the recommendations in Booth and Holbert (1999). To prevent premature stopping, we used the convergence criteria

$$\| \boldsymbol{\beta}^{(b+1)} - \boldsymbol{\beta}^{(b)} \| \leq \epsilon \times \| \boldsymbol{\beta}^{(b+1)} \| \quad \text{and} \quad \| \boldsymbol{\varphi}^{(b+1)} - \boldsymbol{\varphi}^{(b)} \| \leq \epsilon \times \| \boldsymbol{\varphi}^{(b+1)} \|$$

for 5 consecutive iterations with $\epsilon = 0.0001$. For each iteration, 10 percent of the simulated random effects were discarded for “burn in” purposes. When convergence was reached, 100,000 random draws from $P(\mathbf{u} \mid \mathbf{T}; \boldsymbol{\theta}^{\wedge})$ were used to estimate $I(\boldsymbol{\theta})$; see Web Appendix A. For a given simulated data set, we terminated the algorithm if our convergence criteria was not met after 80 iterations. Non-convergence mainly occurs when the pool size c is large and the number of sites l is small. For the worst case, presented in Table 2 with $(l, n, c) = (20, 10, 10)$, we observed non-convergence and/or a non-positive definite information matrix for about 8 percent of the data sets and removed these from consideration. This behavior is unfortunate, but not unexpected, given the stochastic nature of the MCEM approach. Furthermore, additional investigation reveals that non-convergence is fueled by the relatively low prevalence settings used when simulating individual data. Of course, group testing is most applicable in low prevalence settings.

The results from Table 2 show that the fixed effects estimates $\hat{\beta}_0$ and $\hat{\beta}_1$ from MCEM generally have small bias. The random effects standard deviations $\hat{\sigma}_1$ and $\hat{\sigma}_2$ do tend to underestimate the true standard deviations, but not by much. The correlation estimate $\hat{\rho}$ is generally on target but is somewhat more variable than the fixed effect estimates. We also report in Table 2 numerical evidence that the information matrix estimate $I(\hat{\boldsymbol{\theta}})$ is adequate. This can be seen by the relative closeness of “SD” and “SE,” the sample standard deviation of the 200 maximum likelihood estimates $\hat{\boldsymbol{\theta}}$ and the average estimated standard error from $I(\hat{\boldsymbol{\theta}})^{-1}$, respectively. With few exceptions, the estimated 95 percent Wald interval confidence coefficients for all parameters are within the margin of Monte Carlo error.

4.2. Tests for Homogeneity

We next characterize the performance of the likelihood ratio and score tests for homogeneity discussed in Section 3. Using the model in (7) with $\boldsymbol{\beta} = (\beta_0, \beta_1)' = (-4, 1)'$, $u_i \sim \mathcal{N}(0, \sigma^2)$, and $x_{ijk} \sim \mathcal{N}(0, 0.64)$, Table 3 shows the estimated size and power of the $\alpha = 0.05$ tests for $\sigma = 0, 0.2, 0.3, \dots, 0.6$. For each (N, l, n, c, σ) combination, we generated 500 data sets and created pools

using random assignment as before; the estimated power in Table 3 is the proportion of times $H_0 : \sigma^2 = 0$ is rejected out of 500. For the score test, the expectations $I_{\tau\tau}$ and $\mathbf{I}_{\tau\beta}$ in (6) were approximated using 2000 replicates as described in Section 3.2. The results suggest that both tests confer the correct size; estimated nominal $\alpha = 0.05$ sizes are all within the margin of Monte Carlo error. Powers for both tests increase in nearly the same manner, although the score test has a slight advantage for σ close to 0. This finding is congruous with the locally most powerful result described in Lin (1997) for individual testing when $q = 1$.

5. Nebraska IPP Data

The state of Nebraska takes part in the nationwide IPP through its Sexually Transmitted Diseases and Infertility Control Program. At $l = 78$ clinic sites throughout the state, urine or swab (cervical or male urethra) specimens are collected on each individual and are transported to the Nebraska Public Health Laboratory (NPHL) in Omaha for chlamydia and gonorrhea testing. More than 30,000 individual tests are performed annually by the NPHL; we use the individual testing results from the first quarter of 2006 to test our group testing methods. The data set consists of chlamydia and gonorrhea infection statuses (infected/not) for 6,138 subjects, as well as several risk covariates. The number of subjects within site varies from 1 to 540. The sample prevalence for chlamydia and gonorrhea is 7.8 percent and 1.7 percent, respectively, making group testing potentially attractive as a means of surveillance.

We consider chlamydia and gonorrhea infections separately. While group testing is not used currently in Nebraska, it is being used elsewhere for chlamydia and gonorrhea screening, as noted in Section 1. In fact, Lindan et al. (2005) estimates that approximately 12 percent of laboratories in the US use group testing for detecting chlamydia. When compared to individual testing, Kacena et al. (1998b) and Kacena et al. (1998a) report that pooling can reduce testing costs by 39 and 60 percent for chlamydia and gonorrhea, respectively.

We use pool sizes $c = 2, 5, 8$ and construct pools in the following way. If site i has n_i subjects, we first create $\lfloor n_i/c \rfloor$ pools by assigning subjects to pools at random. Any remaining subjects are assigned to a smaller-sized pool. For example, if there are 23 subjects available for one site and $c = 5$, 4 pools of size 5 and 1 pool of size 3 are created for that site. To incorporate the effects of imperfect testing, pooled responses are recorded assuming that assay sensitivity and specificity are $\gamma_1 = 0.95$ and $\gamma_2 = 0.98$, respectively. These values are reasonable for both infections based on the empirical findings of Kacena et al. (1998a) and Kacena et al. (1998b). To cover a large number of the possible arrangements of individuals within site, we simulate 100 sets of pools. Among the available covariates from the Nebraska data set, we chose age, gender, urethritis status, and infection symptoms status, based on a variable selection process with the individual data. For each data set and for each infection, we fit the model

$$\text{logit} \left\{ \Pr \left(Y_{ijk} = 1 | u_i \right) \right\} = \beta_0 + \beta_1 \text{Age}_{ijk} + \beta_2 \text{Gender}_{ijk} + \beta_3 \text{Urethritis}_{ijk} + \beta_4 \text{Symptoms}_{ijk} + u_i, \quad (9)$$

where $u_i \sim \mathcal{N}(0, \sigma^2)$ is the random clinic site effect. We do not consider variable selection issues in this paper, although this is a good topic for future research with group testing models.

Table 4 displays the results from fitting the model in (9). For each infection, we provide the parameter estimates based on the raw individual data. Probability values for the score and likelihood ratio tests of $H_0 : \sigma^2 = 0$ were highly significant for both infections ($p < 0.01$) based on the individual data. Estimates and standard errors from pooled testing in Table 4 are obtained by averaging over the 100 simulated sets of pools. There is a general agreement between the estimates from individual testing and pooled testing; those from pooled testing are more variable, but this is expected. For each set of pools, we performed the $\alpha = 0.05$ level likelihood

ratio and score tests as described in Section 3. For each $c > 1$, the proportion of times H_0 is rejected is also given in Table 4 for the score test (likelihood ratio test results were similar). Above all, geographic heterogeneity is apparent for both infections using the pooled responses.

6. Discussion

We have generalized the work of Vansteelandt et al. (2000) to incorporate random effects into group testing regression models and have illustrated the usefulness of this extension using chlamydia and gonorrhea data collected by the state of Nebraska. We have presented methods for maximum likelihood estimation, for large sample inference, and for assessing homogeneity in the latent binary responses that arise with group testing. For future research, it may be of interest to investigate different pool composition strategies, as was done by Vansteelandt et al. (2000) and Bilder and Tebbs (2008) in fixed effects models. Vansteelandt et al. (2000) showed that forming \mathbf{x} -homogeneous pools (i.e., pools whose covariate values are as similar as possible) minimizes the amount of lost information due to pooling and provides the best fixed effects parameter estimates. In models with random effects, it is not clear how \mathbf{x} -homogeneous and/or \mathbf{z} -homogeneous compositions would affect the parameter estimates.

Tests for variance components in Section 3 have been examined in the scalar variance case, but more general tests of variance components could be formulated using the results from Self and Liang (1987) and Silvapulle and Silvapulle (1995). For example, in (8), one might wish to test $H_0 : \sigma_1^2 = \sigma_2^2 = 0$ or possibly $H_0 : \sigma_2^2 = 0$, while leaving σ_1^2 unspecified. Such examples fall into a larger class of tests involving the unique components of \mathbf{D} . In higher dimensional settings, likelihood ratio and score statistics continue to follow mixture χ^2 distributions, but with possibly a larger number of components.

Our hierarchical formulation requires that individuals are pooled within site, but this assumption may not be practical in some applications. A useful extension of this work would be to allow for subjects from different sites to be pooled together. In this setting, fitting the model using quadratures would most likely be computationally infeasible, but the MCEM approach may prove fruitful. Even so, when compared to pooling subjects within site, we would expect both the loss in precision of the random effects estimates and the loss in power of the homogeneity tests to be significant. Although we have introduced random effects models for group testing using only the initial pool results, our work can be extended to handle the multistage pooling designs presented in Brookmeyer (1999). When $\gamma_1 = \gamma_2 = 1$ (perfect testing), this extension is almost automatic, but is perhaps not practical in infectious disease contexts. When imperfect tests are used, conceptually our methodology is applicable, although the likelihood function is far more complicated. We would expect this to be also true for other pooling strategies such as array-based pooling (see, e.g., Kim et al., 2007).

Finally, we are currently pursuing the development of techniques to simultaneously model the prevalence of multiple infections within a group testing regression framework, thereby extending the work of Hughes-Oliver and Rosenberger (2000) to incorporate covariate information. Such regression models have the potential to find widespread use, because many epidemiological investigations using group testing do involve multiple infections and disease statuses are often correlated.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors are grateful to the Editor, the Associate Editor, and the two anonymous referees for their helpful comments. We especially thank Dr. Peter Iwen, Dr. Steven Hinrichs, and Philip Medina for their consultation on the IPP. We also thank Drs. Xianzheng Huang, Kerrie Nelson, and Geert Verbeke for their comments on an earlier draft of this manuscript. This research was supported by Grant R01 AI067373 from the National Institutes of Health.

References

- Bilder C, Tebbs J. Bias, efficiency, and agreement for group-testing regression models. *Journal of Statistical Computation and Simulation*. 2008 in press.
- Booth J, Hobert J. Maximizing generalized linear mixed model likelihoods with an automated Monte Carlo EM algorithm. *Journal of the Royal Statistical Society* 1999;61:265–285. Series B
- Booth J, Hobert J, Jank W. A survey of Monte Carlo algorithms for maximizing the likelihood of a two-stage hierarchical model. *Statistical Modelling* 2001;1:333–349.
- Brookmeyer R. Analysis of multistage pooling studies of biological specimens for estimating disease incidence and prevalence. *Biometrics* 1999;55:608–612. [PubMed: 11318222]
- Cardoso M, Koerner K, Kubanek B. Mini-pool screening by nucleic acid testing for hepatitis B virus, hepatitis C virus, and HIV: Preliminary results. *Transfusion* 1998;38:905–907. [PubMed: 9767739]
- Chen J, Zhang D, Davidian M. A Monte Carlo EM algorithm for generalized linear mixed models with flexible random effects distribution. *Biostatistics* 2002;3:347–360. [PubMed: 12933602]
- Dorfman R. The detection of defective members of large populations. *Annals of Mathematical Statistics* 1943;14:436–440.
- Hughes-Oliver J, Rosenberger W. Efficient estimation of the prevalence of multiple rare traits. *Biometrika* 2000;87:315–327.
- Kacena K, Quinn S, Hartman S, Quinn T, Gaydos C. Pooling of urine samples for screening for *Neisseria gonorrhoeae* by ligase chain reaction: Accuracy and application. *Journal of Clinical Microbiology* 1998a;36:3624–3628. [PubMed: 9817885]
- Kacena K, Quinn S, Howell M, Madico G, Quinn T, Gaydos C. Pooling urine samples for ligase chain reaction screening for genital *Chlamydia trachomatis* infection in asymptomatic women. *Journal of Clinical Microbiology* 1998b;36:481–485. [PubMed: 9466763]
- Kim H, Hudgens M, Dreyfuss J, Westreich D, Pilcher C. Comparison of group testing algorithms for case identification in the presence of testing error. *Biometrics* 2007;63:1152–1163. [PubMed: 17501946]
- Liang K. A locally most powerful test for homogeneity with many strata. *Biometrika* 1987;74:259–264.
- Lin X. Variance component testing in generalised linear models with random effects. *Biometrika* 1997;84:309–326.
- Lindan C, Mathur M, Kumta S, Jerajani H, Gogate A, Schachter J, Moncada J. Utility of pooled urine specimens for detection of *Chlamydia trachomatis* and *Neisseria gonorrhoeae* in men attending public sexually transmitted infection clinics in Mumbai, India, by PCR. *Journal of Clinical Microbiology* 2005;43:1674–1677. [PubMed: 15814983]
- Louis T. Finding observed information using the EM algorithm. *Journal of the Royal Statistical Society* 1982;44:98–130. Series B
- McCulloch E. Maximum likelihood algorithms for generalized linear mixed models. *Journal of the American Statistical Association* 1997;92:162–170.
- Molenberghs G, Verbeke G. Likelihood ratio, score, and Wald tests in a constrained parameter space. *The American Statistician* 2007;61:22–27.
- Nie L. Convergence rate of MLE in generalized linear and nonlinear mixed-effects models: Theory and applications. *Journal of Statistical Planning and Inference* 2007;137:1787–1804.
- Pilcher C, Fiscus S, Nguyen T, Foust E, Wolf L, Williams D, Ashby R, O'Dowd J, McPherson J, Stalzer B, Hightow L, Miller W, Eron J, Cohen M, Leone P. Detection of acute infections during HIV testing in North Carolina. *New England Journal of Medicine* 2005;352:1873–1883. [PubMed: 15872202]
- Pinheiro J, Bates D. Approximations to the log-likelihood function in the nonlinear mixed-effects model. *Journal of Computational and Graphical Statistics* 1995;4:12–35.

- Rours G, Verkooyen R, Willemse H, van der Zwaan E, van Belkum A, de Groot R, Verbrugh H, Ossewaarde J. Use of pooled urine samples and automated DNA isolation to achieve improved sensitivity and cost-effectiveness of large-scale testing for *Chlamydia trachomatis* in pregnant women. *Journal of Clinical Microbiology* 2005;43:4684–4690. [PubMed: 16145127]
- Screening and Treatment Guidelines, Infertility Prevention Project, Region VII. 2003. Available at <http://www.devsys.org/html/ipp/index.html>
- Self S, Liang K. Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions. *Journal of the American Statistical Association* 1987;82:605–610.
- Silvapulle M, Silvapulle P. A score test against one-sided alternatives. *Journal of the American Statistical Association* 1995;90:342–349.
- Vansteelandt S, Goetghebeur E, Verstraeten T. Regression models for disease prevalence with diagnostic tests on pools of serum samples. *Biometrics* 2000;56:1126–1133. [PubMed: 11129470]
- Xie M. Regression analysis of group testing samples. *Statistics in Medicine* 2001;20:1957–1969. [PubMed: 11427952]
- Zagurski K. Omaha World Herald Feb 2;2006 :08B.

Table 1

Simulation study in (7) with $\theta = (\beta_0, \beta_1, \sigma)' = (-5, 1, 1)'$ and $\gamma_1 = \gamma_2 = 1$. Performance of maximum likelihood estimates computed using 500 simulated data sets for each case. Sample means (Mean) and sample standard deviations (Std dev) of the 500 estimates are provided. RMSE is the ratio of the mean squared errors from individual testing (IND) to group testing (POOL). Estimated 95 percent Wald coverage probabilities (Cov) are included only for group testing. The margin of error for the Wald coverage, assuming a 99 percent confidence level, is 0.03. $N = lrc$ is the total number of individuals, l is the number of sites, n is the number of pools per site, and c is the common pool size

		β_0						β_1						σ					
N	l	n	c	IND	POOL	IND	POOL	IND	POOL	IND	POOL	IND	POOL	IND	POOL	IND	POOL		
2000	10	40	5	Mean	-5.04	-5.06	1.02	1.02	1.02	1.02	1.02	0.91	0.92						
				Std dev	0.48	0.49	0.26	0.26	0.26	0.37	0.36								
				RMSE/Cov		0.95/0.94		0.64/0.96		1.07/0.97									
	20	10	10	Mean	-5.04	-5.11	1.02	1.02	1.02	1.02	0.91	0.94							
				Std dev	0.48	0.59	0.26	0.26	0.26	0.37	0.37								
				RMSE/Cov		0.76/0.94		0.26/0.95		1.07/0.97									
				0.65/0.96		0.12/0.91													
10000	40	50	5	Mean	-4.99	5.00	1.00	1.00	1.00	1.00	0.96	0.96							
				Std dev	0.22	0.23	0.11	0.11	0.11	0.17	0.18								
				RMSE/Cov		0.94/0.94		0.62/0.93		0.99/0.92									
	25	10	10	Mean	-4.99	-4.99	1.00	1.00	1.00	1.00	0.96	0.96							
				Std dev	0.22	0.25	0.11	0.11	0.11	0.17	0.18								
				RMSE/Cov		0.82/0.94		0.30/0.96		0.98/0.93									
				0.71/0.94		0.14/0.95													

Simulation study in (8) with $\theta = (\beta_0, \beta_1, \sigma_1, \sigma_2, \rho)' = (-4, 1, 0.7, 0.5, 0)'$ and $\gamma_1 = \gamma_2 = 1$. Performance of group testing maximum likelihood estimates computed using 200 simulated data sets for each case. MCEM was used to fit each data set. Sample means (Mean) and sample standard deviations (SD) of the 200 estimates are provided. The averaged approximate standard errors (SE) obtained from the information matrix in Section 2.3 are also provided. The margin of error for the 95 percent Wald coverage probabilities (Cov), assuming a 99 percent confidence level, is 0.04. $N = lrc$ is the total number of individuals, l is the number of sites, n is the number of pools per site, and c is the common pool size

Table 2

N	l	n	c	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\sigma}_1$	$\hat{\sigma}_2$	$\hat{\rho}$
				Mean	1.01	0.65	0.48	0.03
				SD/SE	0.34/0.27	0.17/0.16	0.24/0.23	0.13/0.13
				Cov	0.92	0.93	0.98	0.97
				Mean	-4.04	1.02	0.64	0.48
				SD/SE	0.31/0.33	0.24/0.23	0.22/0.25	0.17/0.18
				Cov	0.95	0.95	0.98	0.98
				Mean	-4.14	1.04	0.63	0.50
				SD/SE	0.43/0.45	0.40/0.34	0.23/0.28	0.21/0.27
				Cov	0.97	0.92	0.96	0.96
				Mean	-4.01	1.01	0.69	0.49
				SD/SE	0.19/0.19	0.12/0.11	0.15/0.16	0.09/0.09
				Cov	0.95	0.93	0.96	0.96
				Mean	-3.99	0.99	0.67	0.48
				SD/SE	0.23/0.23	0.17/0.16	0.18/0.17	0.12/0.13
				Cov	0.92	0.93	0.95	0.94
				Mean	-3.99	0.99	0.66	0.47
				SD/SE	0.30/0.30	0.25/0.23	0.19/0.19	0.16/0.20
				Cov	0.93	0.94	0.97	0.94

Simulation study in (7) with $\theta = (\beta_0, \beta_1, \sigma)' = (-4, 1, \sigma)'$ and $\gamma_1 = \gamma_2 = 1$. Estimated size and power of the = 0.05 likelihood ratio test T_{LR} and score test T_S from Section 3, computed using 500 simulated data sets. The margin of error for the estimated size when $\sigma = 0$, assuming a 99 percent confidence level, is 0.03. $N = lnc$ is the total number of individuals, l is the number of sites, n is the number of pools per site, and c is the common pool size

Table 3

N	l	n	c	$\sigma = 0$	0.2	0.3	0.4	0.5	0.6
2000	2	50	T_{LR}	0.03	0.10	0.30	0.55	0.78	0.89
			T_S	0.04	0.13	0.36	0.63	0.83	0.92
	5	20	T_{LR}	0.04	0.08	0.24	0.45	0.75	0.88
			T_S	0.05	0.12	0.28	0.51	0.80	0.88
4000	10	10	T_{LR}	0.03	0.08	0.19	0.38	0.61	0.74
			T_S	0.05	0.11	0.26	0.45	0.69	0.80
	2	50	T_{LR}	0.04	0.22	0.53	0.83	0.94	0.99
			T_S	0.06	0.27	0.59	0.88	0.97	1.00
4000	5	20	T_{LR}	0.05	0.14	0.38	0.75	0.93	0.98
			T_S	0.05	0.18	0.44	0.80	0.95	0.99
	10	10	T_{LR}	0.03	0.11	0.32	0.57	0.80	0.92
			T_S	0.05	0.16	0.38	0.67	0.88	0.97

Table 4

Nebraska IPP data. Parameter estimates and standard errors from the model in (9) for different pool sizes c with $\gamma_1 = 0.95$ and $\gamma_2 = 0.98$. When $c = 1$, estimates and standard errors are based on the individual data. For $c > 1$, estimates and standard errors are averaged over 100 sets of pools. The "Reject" column gives the proportion of level $\alpha = 0.05$ score tests rejecting H_0 , based on the 100 sets of pools

Chlamydia							
c	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	$\hat{\sigma}$	Reject
1	-2.45(0.26)	-0.05(0.01)	0.68(0.14)	0.87(0.25)	0.46(0.12)	0.46(0.09)	—
2	-2.46(0.35)	-0.05(0.01)	0.65(0.18)	0.86(0.31)	0.50(0.16)	0.46(0.09)	1.00
5	-2.65(0.44)	-0.03(0.02)	0.47(0.24)	0.93(0.46)	0.57(0.24)	0.39(0.10)	0.88
8	-2.89(0.45)	-0.01(0.02)	0.31(0.26)	0.96(0.60)	0.58(0.29)	0.32(0.11)	0.77
Gonorrhoea							
c	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	$\hat{\sigma}$	Reject
1	-4.88(0.50)	-0.02(0.01)	0.25(0.26)	1.57(0.34)	0.95(0.23)	0.99(0.24)	—
2	-5.42(0.62)	-0.01(0.02)	0.48(0.34)	1.39(0.39)	1.04(0.30)	0.90(0.27)	1.00
5	-6.21(0.73)	0.00(0.02)	0.79(0.45)	1.20(0.51)	1.19(0.42)	0.83(0.27)	1.00
8	-6.62(0.81)	-0.01(-0.03)	0.93(0.49)	1.01(0.65)	1.42(0.50)	0.83(0.27)	0.96