Published in final edited form as:

Stat Modelling. 2010 December ; 10(4): 421–439. doi:10.1177/1471082X0901000404.

A Bayesian model for repeated measures zero-inflated count data with application to outpatient psychiatric service use

Brian H. Neelon^{1,*}, A. James O'Malley¹, and Sharon-Lise T. Normand^{1,2}

¹ Department of Health Care Policy, Harvard Medical School, Boston, USA

² Department of Biostatistics, Harvard School of Public Health, Boston, MA

Abstract

In applications involving count data, it is common to encounter an excess number of zeros. In the study of outpatient service utilization, for example, the number of utilization days will take on integer values, with many subjects having no utilization (zero values). Mixed-distribution models, such as the zero-inflated Poisson (ZIP) and zero-inflated negative binomial (ZINB), are often used to fit such data. A more general class of mixture models, called hurdle models, can be used to model zero-deflation as well as zero-inflation. Several authors have proposed frequentist approaches to fitting zero-inflated models for repeated measures. We describe a practical Bayesian approach which incorporates prior information, has optimal small-sample properties, and allows for tractable inference. The approach can be easily implemented using standard Bayesian software. A study of psychiatric outpatient service use illustrates the methods.

Keywords

Bayesian inference; hurdle model; repeated measures; zero-altered model; zero-inflated model

1 Introduction

In modeling count data, it is common to encounter a large number of zeros. For example, in studies examining the utilization of psychiatric outpatient services, many patients will report no utilization (zero counts). When the number of zeros is greater (less) than expected under a standard count distribution with a fixed mean μ the data are said to be zero-inflated (deflated).

Several classes of discrete mixture models have been proposed for zero-modified data. These include hurdle models (Mullahy, 1986; Heilbron, 1989) and zero-inflated count models (Lambert, 1992; Greene, 1994). Hurdle models comprise a mixture of a point mass at zero and a truncated count model for the nonzero values. Heilbron (1994) proposed a special type of hurdle model, called the zero-altered model, which is convenient for testing for zero inflation or deflation.

Zero-inflated count models contain a mixture of a point mass at zero and an untruncated count distribution. These include the zero-inflated Poisson (ZIP) and zero-inflated negative binomial (ZINB) models. Because each part of the mixture accommodates zeros, zero-inflated models explicitly partition the zeros into two types: “structural” zeros—those that

* Address for Correspondence: Brian Neelon, Department of Health Care Policy, Harvard Medical School, 180 Longwood Avenue, Boston, MA 02115-5899, USA. Tel: +1 617-432-3138. E-mail: neelon@hcp.med.harvard.edu.

occur, for example, because a subject is ineligible for outpatient services; and “chance” zeros—those that occur by chance among eligible patients. Because the probability of structural zeros is non-negative, these models do not allow for zero deflation. For a comprehensive review of zero-inflated models, see Ridout et al. (1998).

There is a well-developed literature on frequentist approaches to fitting repeated measures, zero-inflated models. Hall (2000), Yau and Lee (2001), and Min and Agresti (2005) have proposed various random intercept zero-inflated models. Liu and Power (2007) proposed a zero-inflated growth model, which allowed for correlated random intercepts and slopes for both components. Roeder et al. (1999), Dalrymple et al. (2003), and Min and Agresti (2005) have developed finite-mixture zero-inflated models that allow for clustering of subjects based on their latent response trajectories.

Several authors have recently proposed Bayesian alternatives to fitting zero-inflated models. The advantages of Bayesian inference are well-known and include elicitation of prior beliefs, avoidance of asymptotic approximations, and practical estimation of functions of parameters (Gelman et al., 2004; Congdon, 2005). Rodriques (2003) and Ghosh et al. (2006), for example, have developed Bayesian ZIP models for cross-sectional data, using Markov chain Monte-Carlo (MCMC) with data augmentation to obtain posterior samples. Ghosh et al. (2006) found that, for small datasets with many zeros, the Bayesian estimation outperformed maximum likelihood in terms of both bias and precision. More recently, Fahrmeir and Osuna Echavarría (2006) and Gschlößl and Gzado (2008) have introduced zero-inflated models for spatially correlated data by specifying region-specific random effects to account for geographical dependencies. In addition, Fahrmeir and Osuna-Echavarría (2006) develop semiparametric ZIP models using a generalized additive structure for the Poisson component.

We present a flexible Bayesian modeling strategy for longitudinal zero-inflated count data. This approach allows for correlated random effects in both components of the model. For ease of exposition, we focus on Poisson models, but the methods can easily be extended to other count distributions, such as the negative binomial. Following Min and Agresti (2005), we describe three zero-inflated models: the Poisson hurdle model, the ZIP model, and the zero-altered Poisson (ZAP) model. The models can be fit in standard Bayesian software packages such as WinBUGS (Spiegelhalter et al., 2003).

The remainder of the paper is organized as follows: Section 2 describes the Bayesian zero-inflated count models and discusses commonly used priors; Section 3 outlines posterior computation and reviews strategies for model assessment and comparison; Section 4 applies the models to an analysis of outpatient psychiatric service use; and Section 5 provides a discussion.

2 Bayesian regression models for zero-inflated count data

2.1 Hurdle model

The hurdle Poisson model is a two-part mixture model consisting of a point mass at zero followed by a truncated Poisson for the nonzero observations. For independent and identically distributed (i.i.d.) responses, the hurdle model can be written as

$$P(Y_i=0)=1-p, \quad 0 \leq p \leq 1 \quad (2.1)$$

$$P(Y_i=k)=p \frac{\mu^k e^{-\mu}}{k!(1-e^{-\mu})}, \quad k=1, \dots, \infty, 0 < \mu < \infty, \quad (2.2)$$

where Y_i denotes the response for subject $i = 1, \dots, n$, and μ is the mean for an untruncated Poisson distribution. When $(1-p) > e^{-\mu}$ the data are zero inflated; when $(1-p) < e^{-\mu}$ there is zero deflation; and when $p = 1$ there are no zero counts and the model reduces to a truncated Poisson.

Model (2.2) can be easily extended to allow for predictors and repeated responses via mixed effects regression:

$$\text{logit}[\Pr(Y_{ij} > 0 | \mathbf{b}_{1i})] = \text{logit}(p_{ij} | \mathbf{b}_{1i}) = \mathbf{x}'_{1ij} \boldsymbol{\beta}_1 + \mathbf{z}'_{1ij} \mathbf{b}_{1i} \quad (2.3)$$

$$\log(\mu_{ij} | \mathbf{b}_{2i}) = \mathbf{x}'_{2ij} \boldsymbol{\beta}_2 + \mathbf{z}'_{2ij} \mathbf{b}_{2i}, \quad j=1, \dots, n; i=1, \dots, n, \quad (2.4)$$

where $\mathbf{x}_{kij}$ and $\mathbf{z}_{kij}$ ($k = 1, 2$) denote covariate vectors, $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$ are vectors of fixed effect coefficients, and $\mathbf{b}_{1i}$ and $\mathbf{b}_{2i}$ denote random effect parameters. One might allow different sets of predictors in the two components if the purpose is to obtain the most parsimonious model, or if there is *a priori* scientific reason to believe that the two components are affected by unique predictors. For illustrative purposes, we will assume throughout that the fixed effect covariates are identical for the two components (i.e., $\mathbf{x}_{1ij} = \mathbf{x}_{2ij} = \mathbf{x}_{ij}$). We will also assume that random intercepts are sufficient to capture the heterogeneity between subjects; that is, $\mathbf{b}_{1i} = b_{1i}$, $\mathbf{b}_{2i} = b_{2i}$ and $\mathbf{z}_{1ij} = \mathbf{z}_{2ij} = 1$. Generalizations to multivariate random effects for the two components are straightforward.

Under a fully Bayesian approach, prior distributions are assumed for all model parameters. A common prior distribution for the fixed effects is an independent, multivariate normal distribution: $\boldsymbol{\beta}_1$, and $\boldsymbol{\beta}_2 \sim N_p(\boldsymbol{\beta}_0, \sigma_{\boldsymbol{\beta}}^2 \mathbf{I}_p)$, where $\boldsymbol{\beta}_0$ and $\sigma_{\boldsymbol{\beta}}^2$ are known hyperparameters and $\mathbf{I}_p$ denotes the $p \times p$ identity matrix. A large prior variance, such as $\sigma_{\boldsymbol{\beta}}^2 = 10^3$, can be used to express lack of prior information about $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$. For the random effects, a typical prior is the bivariate normal distribution with unstructured covariance:

$$\mathbf{b}_i = \begin{pmatrix} b_{1i} \\ b_{2i} \end{pmatrix} \sim N_2(\mathbf{0}, \boldsymbol{\Sigma}), \quad \boldsymbol{\Sigma} = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix},$$

where ρ denotes $\text{corr}(b_{1i}, b_{2i})$. Prior specification is completed by assigning a prior to $\boldsymbol{\Sigma}$, such as the diffuse inverse-Wishart (I-W) distribution, $\boldsymbol{\Sigma} \sim \text{I-W}(2, \mathbf{I}_2)$. Some authors have noted that inverse-Wishart priors can yield biased and imprecise model estimates, particularly if the number of clustering units is small (Browne and Draper, 2000; Lambert et al., 2005). As an alternative, one can specify a product normal parameterization for the random effects (Spiegelhalter, 1998; Cooper et al., 2007)

$$b_{1i} \sim N(0, \eta_1^2) \text{ and } b_{2i} | b_{1i} \sim N(\psi b_{1i}, \eta_2^2),$$

with diffuse hyper-priors such as $\psi \sim N(0, 10^3)$ and $\eta_1^2, \eta_2^2 \sim \text{Inv} - \text{Gamma}(0.001, 0.001)$. The components of Σ can then be expressed as

$$\sigma_1^2 = \eta_1^2, \quad \sigma_2^2 = \psi^2 \eta_1^2 + \eta_2^2, \quad \rho = \frac{\psi \eta_1^2}{\sqrt{\eta_1^2 (\psi^2 \eta_1^2 + \eta_2^2)}}.$$

Because $|\psi| \in [0, \infty)$, the product normal prior makes no assumptions about the relative magnitudes of σ_1^2 and σ_2^2 .

Under prior independence, the joint posterior distribution of the model parameters is:

$$\begin{aligned} \pi(\beta_1, \beta_2, \mathbf{b}, \Sigma | \mathbf{y}) &\propto f(\mathbf{y} | \beta_1, \beta_2, \mathbf{b}, \Sigma) \pi(\beta_1) \pi(\beta_2) \pi(\mathbf{b} | \Sigma) \pi(\Sigma) \\ &= \prod_{i=1}^n \prod_{j=1}^{n_i} (1 - p_{ij})^{1-d_{ij}} \left(\frac{p_{ij} \mu_{ij}^{y_{ij}} e^{-\mu_{ij}}}{y_{ij}! (1 - e^{-\mu_{ij}})} \right)^{d_{ij}} \times \\ &\quad \pi(\beta_1) \pi(\beta_2) \pi(\mathbf{b} | \Sigma) \pi(\Sigma) \end{aligned} \quad (2.5)$$

where $f(\mathbf{y} | \cdot)$ denotes the likelihood for $\mathbf{y}$ given the model parameters; d_{ij} is an indicator that $y_{ij} > 0$; $\mathbf{b}$ is a $2n$ stacked-vector of random effects with prior $\pi(\mathbf{b} | \Sigma) \stackrel{d}{=} N_{2n}(\mathbf{0}, \mathbf{I}_n \otimes \Sigma)$; $p_{ij} = p_{ij}(\mathbf{x}_{ij} | \beta_1, b_{1i})$ and $\mu_{ij} = \mu_{ij}(\mathbf{x}_{ij} | \beta_2, b_{2i})$ are functions of the covariates and the fixed and random effects; and $\pi(\beta_1)$, $\pi(\beta_2)$ and $\pi(\Sigma)$ denote prior densities for β_1 , β_2 and Σ , respectively.

2.2 ZIP model

The ZIP model (Lambert, 1992) consists of a degenerate distribution at zero mixed with an untruncated Poisson distribution. For i.i.d. data, the ZIP model can be written as

$$P(Y_i=0) = (1 - p) + p e^{-\mu}, \quad 0 \leq p \leq 1 \quad (2.6)$$

$$P(Y_i=k) = p \frac{\mu^k e^{-\mu}}{k!}, \quad k=1, \dots, \infty, 0 < \mu < \infty, \quad (2.7)$$

where the quantity $1 - p$ denotes the probability of a structural zero. When $p = 1$, the probability of zero-inflation is zero, and the model reduces to the ordinary Poisson with $\Pr(Y_i = 0) = e^{-\mu}$. Conversely, when $p < 1$, $e^{-\mu} < (1 - p) + p e^{-\mu}$ and the zeros are inflated. Thus, unlike the hurdle model, the ZIP model allows only for zero inflation. In fact, because the ZIP model can be rewritten as a hurdle model (2.2) with mixing probability $\theta = p(1 - e^{-\mu})$, it is a special case of the Poisson hurdle model in which only zero inflation is allowed.

The ZIP model (2.7) can be extended to accommodate repeated measurements and covariates by modeling p_{ij} and μ_{ij} (now indexed by subject i and observation j) using fixed and random effects, as in equations (2.3) and (2.4). Note that in the ZIP model, $1 - p_{ij}$ represents the probability of a structural zero for the ij^{th} observation, not the overall probability of zero, which is represented by $(1 - p_{ij}) + p_{ij} e^{-\mu_{ij}} = 1 - p_{ij} (1 - e^{-\mu_{ij}}) = 1 - \theta_{ij}$.

For Bayesian analysis, prior distributions are similar to those for the hurdle model—e.g., multivariate normal priors for the fixed and random effects, and an inverse-Wishart or product normal prior for the random effect covariance.

2.3 Zero-altered model

Heilbron (1994) developed a “zero-altered” hurdle model—also called a “compatible” model—to test for zero-inflation. The model assumes a single distribution for the i^{th} response, Y_i , but uses separate parameters for the two parts of the model: $\Pr(Y_i = 0; \mu_{1i})$ and $\Pr(Y_i = y_i | y_i > 0; \mu_{2i})$. For example, the zero-altered Poisson (ZAP) model takes the form

$$\Pr(Y_i=0; \mu_{1i})=e^{-\mu_{1i}}, \quad \log(\mu_{1i})=\mathbf{x}_i'\beta_1 \quad (2.8)$$

$$\Pr(Y_i=y_i | y_i > 0; \mu_{2i})=\frac{y_i^{\mu_{2i}} e^{-\mu_{2i}}}{y_i!(1 - e^{-\mu_{2i}})}, \quad \log(\mu_{2i})=\mathbf{x}_i'\beta_2. \quad (2.9)$$

If we let $\mathbf{x}_i'\beta_1=\gamma+\mathbf{x}_i'\beta_2$, then testing for zero-inflation reduces to testing $\gamma = 0$. When $\gamma < 0$, the data are zero-inflated; when $\gamma > 0$, the zeros are deflated; and when $\gamma = 0$, the model reduces to a standard Poisson model.

Min and Agresti (2005) extended the ZAP model to repeated-measures data. As they point out, a complementary log-log link can be used to model $\Pr(Y_{ij} > 0) = p_{ij}$; hence, the random-intercept ZAP model can be written as

$$\text{cloglog}(p_{ij}|b_i)=\log[-\log(1 - p_{ij}|b_i)]=\gamma+\mathbf{x}_{ij}'\beta+b_i \quad (2.10)$$

$$\log(\mu_{ij}|b_i)=\mathbf{x}_{ij}'\beta+b_i, \quad (2.11)$$

where $b_i \sim N(0, \sigma_b^2)$ denotes a shared random intercept between the two components. A common prior for σ_b^2 is an inverse-gamma(ε, ε) prior, with ε set at a small value, such as $\varepsilon = 0.001$. However, this choice can be deceptively informative because it concentrates most of the prior mass near zero, which, as Gelman (2006) notes, can lead to unstable posterior estimates since the posterior becomes improper as $\varepsilon \rightarrow 0$. As an alternative, Gelman (2006) recommends a diffuse but proper uniform prior on the random effect standard deviation, σ_b , restricted to the positive domain (e.g., $U(0,100)$).

3 Posterior computation, model selection and model checking

3.1 Posterior computation

Posterior computation for the Bayesian zero-inflated model proceeds via an MCMC algorithm such as Gibbs sampling, which draws iteratively from the full conditional distributions of the model parameters. The full conditionals for the above models do not have closed forms; however, since the likelihoods are log-concave, adaptive rejection sampling (ARS) or adaptive rejection Metropolis sampling (ARMS) can be used (Gilks and Wild, 1992; Gilks et al., 1995). For models with random slopes, a Metropolis-Hastings with

iteratively weighted least square (IWLS) proposal densities has been shown to efficiently sample vectors of random effects (Gamerman, 1997). The product normal prior can also be extended to accommodate random slopes (c.f., Cooper et al., 2007).

The Gibbs algorithm can be implemented in WinBUGS software (Spiegelhalter et al., 2003). Unfortunately, WinBUGS does not have a pre-designated truncated Poisson distribution, which is required to specify the likelihood for the hurdle model. One way to circumvent this problem in WinBUGS is to use the “zeros trick” to explicitly define the hurdle likelihood¹. This approach can also be used to fit the ZIP model; however, the ZIP model must first be re-parameterized as a hurdle model with mixing probabilities θ_{ij} as defined in Section 2.2. Alternatively, one can fit the ZIP model using the data-augmentation procedure described by Ghosh et al. (2006); however, this procedure does not generalize to hurdle models. Zero-inflated negative binomial models can be similarly fit within WinBUGS.

Convergence of the MCMC chains can be monitored using standard Bayesian diagnostic procedures, such as trace plots; autocorrelation statistics; the Geweke Z-diagnostic (Geweke, 1992), which compares the mean and variance of a parameter at various points in the chain; and the Brooks-Gelman-Rubin scale-reduction statistic, $\hat{R}$, which compares the total within- and between-chain variation to the within-chain variation (Brooks and Gelman, 1998; Gelman et al., 2004). At convergence, $\hat{R} = 1$, indicating that the initially dispersed chains have converged to a stationary distribution. As a practical rule of thumb, a 0.975 quantile for $\hat{R} \leq 1.2$ is sufficient to claim convergence (Smith, 2007). Convergence diagnostics may be performed using the R package *boa* (Smith, 2007).

3.2 Model selection

A variety of Bayesian model selection procedures can be used to determine the optimal model. One method is to estimate all models of interest simultaneously, averaging over the posterior probabilities that each model is the true model to obtain posterior parameter estimates (Hoeting et al., 1999). An alternative approach is to estimate each model separately and compare the models using Bayes factors, which are ratios of the posterior probabilities of competing models (see, e.g., Kass and Raftery, 1995). However, model averaging methods are computationally intensive, requiring procedures such as reversible jump MCMC (Green, 1995) to ensure coverage of the model space, while Bayes factors require proper priors.

As an alternative, we consider two computationally tractable Bayesian procedures to compare the fit of competing models: the deviance information criterion (DIC; Spiegelhalter, 2002) and a cross-validatory comparison based on an approximation to the marginal likelihood $f(y)$ (Gelfand, 1996). As with Bayes factors, these methods are evaluated one model at a time.

3.2.1 The deviance information criterion (DIC)—Because more complex models generally provide better fit, standard likelihood-based model selection procedures, such as the Akaike information criterion (AIC) and the Bayesian information criterion (BIC), include “penalty” terms to offset gains in model fit due solely to added complexity. For fixed effect models, the complexity—as measured by the number of model parameters—is easily determined. For random effect models, the number of parameters is less clear. If the random effect variance is large, each random effect will contribute approximately one parameter to the model; conversely, if the variance is small (suggesting a large amount of

¹For details on the use of the zeros trick, see “Tricks: Advanced Use of the Bugs Language” in the WinBUGS 1.4 User Manual (Spiegelhalter et al., 2003).

shrinkage), the number of random effect parameters will be close to zero (Elliott et al., 2005). The DIC was proposed to estimate the number of “effective” parameters in a Bayesian hierarchical model (Spiegelhalter et al., 2002).

The deviance information criterion is defined as $DIC = \bar{D}(\theta) + p_D$, where $\bar{D}(\theta) = E[D(\theta)|y]$ is the posterior mean of the deviance, $D(\theta)$, and $p_D = \bar{D}(\theta) - \hat{D}(\theta) = E[D(\theta)|y] - D(E[\theta|y])$ is the difference in the posterior mean of the deviance and the deviance evaluated at the posterior mean of the parameters. The deviance, typically taken as negative twice the log-likelihood, is a measure of the model’s relative fit, whereas p_D denotes a penalty for the model’s complexity. As a rule of thumb, if two models differ in DIC by more than three, the one with the smaller DIC is considered the best fitting (Spiegelhalter et al., 2002).

With the exception of the ordinary Poisson model, the default DIC option in WinBUGS is not available for the mixture models presented in this paper. While $\bar{D}(\theta)$ can be easily computed within WinBUGS, $\hat{D}(\theta)$ must be calculated outside of WinBUGS by importing the posterior draws of the fixed and random effects into a software package such as R (R Development Core Team, 2007), taking the averages across the posterior samples, and then evaluating the deviance at the posterior means. Note that this is a “conditional” DIC akin to DIC #7 in Celeux et al. (2006). That is, we compute $\bar{D}(\theta)$ and $\hat{D}(\theta)$ conditional on the random effects rather than by first integrating them out. Although Celeux et al. (2006) recommend alternatives to this form of DIC, our approach is a direct extension of the DIC used by WinBUGS, and thus is the most familiar to users.

3.2.2 Model comparison based on pseudo-marginal likelihoods—To overcome the computational difficulties of Bayes factors, several authors have recommended approximating the Bayes factor using a cross-validation estimate of the marginal likelihood $f(y)$ (Berger and Pericchi, 1993; Gelfand and Dey, 1994; Gelfand, 1996). The essential idea is to treat $f(y)$ as a “prior predictive density” and to then approximate this density using a cross-validation predictive density estimated from a holdout subset of the data. If the holdout sample consists of a single observation, with $y_{[i]}$ denoting the response vector with the i^{th} observation deleted, then the predictive distribution

$$f(y_i|y_{[i]}) = \int_{\theta} f(y_i|\theta, y_{[i]})\pi(\theta|y_{[i]}) d\theta \quad (3.1)$$

is called the conditional predictive ordinate (CPO) for observation i (Geisser and Eddy, 1979). A Monte Carlo estimate of CPO_i is given by the harmonic mean of the i^{th} likelihood

$$\widehat{CPO}_i = \frac{L}{\sum_l [1/f(y_i|\theta^{(l)})]}, \quad (3.2)$$

where $\theta^{(l)}$ denotes the parameter vector sampled at iteration $l = 1, \dots, L$ (following a burn-in period), and $f(y_i|\theta^{(l)})$ denotes the i^{th} likelihood evaluated at iteration l . While extreme CPO values can occur if $f(y_i|\theta^{(l)}) \approx 0$, Congdon (2005) concludes that the harmonic mean estimate of CPO will be stable as long as individual log-likelihoods generally exceed -10 in value. An alternative estimate of CPO can be obtained via a weighted resampling approach (Dey et al., 1997), but the computation is more challenging.

The product of the CPOs, called the pseudo-marginal likelihood (PsML), has been proposed as a proxy for the marginal likelihood $f(y)$ (Gelfand and Dey, 1994; Gelfand, 1996). The

negative log of PsML is referred to as the negative cross-validated log likelihood (NLL). The difference in NLLs, $NLL_1 - NLL_2$, can be used to compare the fit between two models, $\mathcal{M}_1$ and $\mathcal{M}_2$ (Cooper et al., 2003). A difference greater than 0 favors $\mathcal{M}_2$ and a difference less than 0 favors $\mathcal{M}_1$. Exponentiating this difference yields a pseudo-Bayes factor (PsBF). Unlike traditional Bayes factors, PsBFs can be used with improper priors. The CPOs can also be used to assess model fit by first scaling the CPOs (dividing by the maximum value) and identifying outliers as observations with low values. If few CPOs are less than 0.01, the model is considered to fit adequately (Congdon, 2005).

While the DIC and NLL both compare the predictive fit of competing models, they do so in different ways. The DIC compares models conditional on all unknown parameters—including fixed and random effects—and adds a penalty for models with more complexity. The NLLs compare models marginally, having averaged out all unknown parameters. Thus, as Spiegelhalter et al. (2002) note, each measure has a unique “focus” within a Bayesian hierarchical model. The choice of measure depends on the aim of the analysis. If the aim is to compare models conditional on the model parameters, the DIC is more appropriate; if the focus is to compare the predictive ability of two models marginally, the NLL is preferable. To examine model fit from multiple viewpoints, we use both measures.

3.3 Model checking

Once an optimal model has been determined, its fit to the data should be evaluated. After all, the best-fitting model may still fit the data poorly. As noted above, an informal strategy for assessing model fit is to identify whether there are many scaled-CPO values less than 0.01. A second approach is to conduct posterior predictive checks, in which the observed data are compared to data replicated from the posterior predictive distribution (Gelman et al., 1996). If the model fits well, the replicated data, $\mathbf{y}^*$, should resemble the observed data, $\mathbf{y}$. To quantify the similarity, we can choose a discrepancy measure, $T = T(\mathbf{y}, \boldsymbol{\theta})$, that takes an extreme value if the model is in conflict with the observed data. Popular choices for T include sample moments and quantiles, and residual-based measures.

The Bayesian p-value (P_B) denotes the probability that the discrepancy measure based on the predictive sample, $T^* = T(\mathbf{y}^*, \boldsymbol{\theta})$, is more extreme than the observed measure T . A p-value close to 0.50 represents adequate model fit, while p-values near 0 or 1 indicate lack of fit. Generally, p-values between 0.20 and 0.80 represent adequate fit. A Monte Carlo estimate of P_B can be computed by evaluating the proportion of draws in which $T^* > T$.

For zero-inflated count models, we recommend three test statistics to assess the fit of both the Bernoulli and Poisson components of the model: the proportion of observations equal to zero, the over-dispersion index, defined as $\text{Var}(\mathbf{y}^*)/\text{E}(\mathbf{y}^*)$, and the sample skewness.

4 Analysis of ACCESS data

We applied the zero-inflated models to data from the Access to Community Care and Effective Services and Support (ACCESS) study (Rosenheck et al., 2002). One aim of the study was to identify potential disparities in outpatient psychiatric service use among white, black and Latino homeless adults receiving case management. Due to small sample sizes, black and Latino were categorized into a single group representing minority races. Self-reports of service use were collected at baseline and at three- and 12-month follow-ups; at each occasion, respondents were asked to report the number of days in the previous two months that they used outpatient services.

Of special interest were adults with limited utilization, defined as 16 or fewer utilization days (these patients comprised over 90% of the sample). Our analysis was based on a 10%

random sample of subjects with limited utilization who attended all three follow-up visits. The sample included 680 subjects and 2,040 observations. While zero-inflated count distributions do not provide a perfect fit to these data, the histograms presented below (Figure 1) suggest that they offer a good approximation. Concern about the truncation at 16 is allayed by the fact that, for the ZIP model presented below, the average posterior probability of observing a value greater than 16 is 0.006 (range:[0.002, 0.016]), indicating that the fitted distribution has very little mass in this upper tail. Similar upper-tail probabilities were obtained for the other distributions fitted below.

Table 1 provides the percentage of zeros and mean non-zero counts by visit and race, and Figure 1 presents histograms of utilization days by race and visit. As the histograms illustrate, the sample was heavily right-skewed. At baseline, whites had a slightly higher percentage of zeros (55.52% vs. 51.44%). However, among subjects with some utilization, whites had a higher mean number of utilization days than minorities (4.29 vs. 3.59). Thus, while whites were less likely than minorities to use services at baseline, those who did use tended to do so more often than minorities. Table 1 also suggests non-linear trends over time. At the three-month follow-up, the percentage of whites with some utilization increased dramatically, but returned to near baseline by the one-year follow-up. Among whites with some utilization, the mean number of utilization days decreased over time. For minorities, the average number of utilization days increased at the second visit and then declined slightly at the final visit, although it remained above the baseline level.

We fit four Bayesian models for count data: the ordinary Poisson, the ZAP, the Poisson hurdle and the ZIP model. Each model included a dummy indicator for visits 3 and 12, an indicator for minority race (0=white, 1=black/Latino), minority race $\times$ visit interaction terms, gender, and an indicator of whether the patient participated in a substance abuse treatment program (SA Treatment). For all four models, we assigned a diffuse normal prior for the fixed-effect parameters. For the Poisson and ZAP models, we chose diffuse $U(0,100)$ priors for the random effect standard deviation; for the hurdle and ZIP models, we chose a product-normal prior for the random effect vector $\mathbf{b}_i = (b_{1i}, b_{2i})'$. As a sensitivity check, we re-ran the hurdle and ZIP models using an inverse-Wishart prior for the random effects covariance matrix and found similar results.

The models were fit in the software package WinBUGS (Spiegelhalter, 2003). We ran three, initially-dispersed MCMC chains for 110,000 iterations, discarding the first 10,000 as a burn-in to ensure convergence of the chains. We retained every 50th draw to reduce autocorrelation. MCMC diagnostics, such as trace plots, Geweke Z-statistics (Geweke, 1992), and Brooks-Gelman-Rubin scale reduction statistics (Brooks and Gelman, 1998), indicated rapid convergence and efficient mixing of the chains.

Table 2 provides the model comparison statistics for the four models. The estimate of γ from the ZAP model was -1.28 (95% CrI= $[-1.36, -1.21]$), suggesting strong evidence of zero-inflation. The ZIP model provided the best fit based on both the NLL and DIC statistics. In addition, for the ZIP model, 93% of the scaled CPOs were larger than 0.01, suggesting reasonable fit. Over 99% of the individual log-likelihood values exceeded -10 , indicating stability in the harmonic mean estimate (Congdon, 2005). As a computational check, we recalculated the NLLs using the weighted resampling approach proposed by Dey et al. (1997), and found qualitatively similar results: the ZIP model was most preferred, followed by the hurdle and ZAP models.

Figure 2 presents post-burn-in trace plots for several representative parameters from the ZIP model: β_{13} (minority race, binomial component), β_{25} (minority race $\times$ visit₁₂ interaction, Poisson component), σ_2 (standard deviation of b_{2i}) and ρ (the random effect correlation). For

clarity of presentation, we have graphed only two of the three MCMC chains. The overlapping trajectory lines suggest convergence and efficient mixing of the chains. The lag-5 autocorrelations for the four parameters ranged from -0.003 to 0.03 ; the Geweke Z-diagnostic p-values ranged from 0.40 to 0.68 , indicating no significant difference in posterior means across regions of the chains; and the 0.975 quantiles of the Brooks-Gelman-Rubin statistic were all less than 1.003 , again indicating convergence of each chain.

Table 3 presents the posterior estimates for the ZIP model. At baseline, minorities had a higher posterior probability of any utilization than whites (as measured on the log-odds scale by β_{13}), but a lower mean number of utilization days among users (as measured on the log scale by β_{23}). However, the credible intervals for both estimates contained the null value. While more precise estimates might be obtained by eliminating less predictive covariates, reducing the model to its most parsimonious form was not our primary objective.

Among whites, the log-odds of any utilization increased over time, as indicated by the positive estimates for β_{11} and β_{12} . These results are consistent the unadjusted results in Table 2, where the proportion of zeros decreased over time for whites. Conversely, for minorities, the log-odds of any utilization decreased across time (cf., β_{14} and β_{15}). This trend differs from the unadjusted results for minorities, reecting the impact of covariate adjustment as well as the correlation imposed on the two components of the ZIP model. To verify this, we fit the two ZIP components separately (i.e., with $\rho = 0$) and noticed some reduction in the negative trend estimates for minorities, suggesting that the Table 3 results—and specifically the reversal in trends compared to the unadjusted case—are due to both covariate adjustment and correlation between the outcomes.

Interestingly, the time trends were reversed for the mean number of utilization days: the average utilization counts for whites decreased over time, whereas the average counts for minorities increased. Thus, while minorities had a lower odds of any service use, those who did use tended to do so more often than their white counterparts. There was also a moderate negative correlation (ρ) between the random intercepts for the two parts of the model, indicating that—after adjusting for observed covariates—subjects who have a lower than average probability of any use also tend to report more utilization days. Taken together, these findings demarcate two distinct subpopulations: a group of regular but moderate users (predominately white), and a group of reluctant users—primarily minorities—who, when they do use services, do so with high frequency.

Under the Bayesian approach, functions of parameters are estimated as empirical averages from the posterior samples, thus eliminating the need for analytical approximations or asymptotics. For example, suppose we wish to estimate the average number of utilization days at visit j ($j = 1, 2, 3$) for whites and minorities. Monte-Carlo estimates of these group means, denoted W_j and M_j for whites and minorities, respectively, are obtained as follows:

1. At iteration l ($l = 1, \dots, L$) following a sufficient burn-in period, draw posterior values $\beta_1^{(l)}, \beta_2^{(l)}$, and $\mathbf{b}_i^{(l)}$, $i = 1, \dots, n$;
2. Use the parameter draws in (1) to evaluate $p_{wij}^{(l)}, p_{Mij}^{(l)}, \mu_{wij}^{(l)}$ and $\mu_{Mij}^{(l)}$ ($j=1, \dots, 3; i=1, \dots, n$), where

$$\begin{aligned}
\text{logit}(p_{w_{ij}}^{(l)}) &= \mathbf{x}_{ij}^{*'} \boldsymbol{\beta}_1^{*(l)} \\
\log(\mu_{w_{ij}}^{(l)}) &= \mathbf{x}_{ij}^{*'} \boldsymbol{\beta}_2^{*(l)} \\
\text{logit}(p_{M_{ij}}^{(l)}) &= \mathbf{x}_{ij}^{*'} \boldsymbol{\beta}_1^{*(l)} + \beta_{13}^{(l)} + \beta_{14}^{(l)} \times \text{visit}_{3ij} + \beta_{15}^{(l)} \times \text{visit}_{12ij} \\
\log(\mu_{M_{ij}}^{(l)}) &= \mathbf{x}_{ij}^{*'} \boldsymbol{\beta}_2^{*(l)} + \beta_{23}^{(l)} + \beta_{24}^{(l)} \times \text{visit}_{3ij} + \beta_{25}^{(l)} \times \text{visit}_{12ij},
\end{aligned}$$

and where $\mathbf{x}_{ij}^*$ is a 5×1 subvector that includes all non-race covariates, and $\boldsymbol{\beta}_k^{*(l)}$ s the corresponding subvector of parameters for component k ($k = 1, 2$);

3. Repeat steps (1) and (2) for iterations $1, \dots, L$;
4. Estimate W_j and M_j ($j = 1, 2, 3$) as

$$\begin{aligned}
\widehat{W}_j &= \frac{1}{L} \sum_{l=1}^L \left[\frac{1}{n} \sum_{i=1}^n \frac{p_{w_{ij}}^{(l)} \mu_{w_{ij}}^{(l)}}{1 - e^{-\mu_{w_{ij}}^{(l)}}} \right] \\
\widehat{M}_j &= \frac{1}{L} \sum_{l=1}^L \left[\frac{1}{n} \sum_{i=1}^n \frac{p_{M_{ij}}^{(l)} \mu_{M_{ij}}^{(l)}}{1 - e^{-\mu_{M_{ij}}^{(l)}}} \right].
\end{aligned}$$

In addition, 95% credible intervals (CrIs) can be constructed by computing the 2.5 and 97.5 quantiles of the posterior samples for W_j and M_j .

Table 4 presents the average number of utilization days for whites and minorities at each time point. Whites and minorities averaged around two utilization days per visit. For both groups, the time trends were non-linear. Whites had higher mean counts than minorities during the first two visits, but a lower average count after one year. For both groups, the highest mean number of utilizations occurred over the first three months of follow up.

As a final assessment of model fit, we conducted posterior predictive checks using the proportion of zeros, the over-dispersion index and the sample skewness as test statistics. Figures 3(a) and 3(b) display the posterior predictive distributions for the first two statistics. The dashed vertical lines mark the observed sample values of 0.46 for the proportion of zeros and 5.12 for the over-dispersion index. The shaded regions represent the area corresponding to the Bayesian predictive p-values. The Bayesian predictive p-values were 0.21, 0.28 and 0.70 for the proportion of zeros, the over-dispersion index and the sample skewness, suggesting adequate fit of the ZIP model.

5 Discussion

We have described a Bayesian approach to fitting zero-inflated models for repeated measures count data. The advantages of the Bayesian approach include ease of interpretation, incorporation of prior information, practical estimation of any function of parameters or predictive values, and reduced small-sample bias compared to maximum-likelihood procedures (Ghosh et al., 2006).

We have considered four competing models: the ordinary Poisson, Poisson hurdle, ZIP and ZAP models. The choice of model should be informed by one's conception of the underlying data-generating mechanism as well as by model fit (Rose et al., 2006). If the belief is that the zero counts arise from both chance and structural components, then the ZIP model may be preferable. This appears to be the case in the ACCESS study, where it seems likely that some subjects are "never users," while others are potential users who, by chance, have no utilization days prior to a given visit. The ZIP model may also provide better fit

when there is strong evidence of zero-inflation and the Poisson mean is itself close enough to zero to allow for Poisson-generated zeros. However, as Min and Agresti (2005) note, the ZIP model can be unstable when zero deflation exists at a particular level of a categorical predictor, even if the data on the whole are zero inflated. Their simulations suggest that hurdle models generally provide better fit. For the ACCESS study, there appears to have been consistent evidence of zero-inflation across all covariate levels, so that the zero-inflation constraint inherent in the ZIP model conferred a model-fitting advantage.

The methods discussed here can be extended to “semi-continuous” two-part models, which assume a continuous (e.g., log-normal) rather than Poisson distribution for the non-zero values (Duan, 1983; Olsen and Schafer, 2001; Tooze et al., 2002). These models have been frequently used to analyze health expenditure data (Cooper et al., 2003; Cooper et al., 2007). The methods can also be extended to latent class models to identify underlying clusters of patients defined by their response trajectories. A basic latent class model might assume conditional independence between the Bernoulli and Poisson processes given class membership. Alternatively, one could include random effects as in our analyses, but allow for class-specific covariance structures, resulting in a zero-inflated growth mixture model (Muthen et al. 2002; Elliott et al., 2005). We are currently developing Bayesian methodology and model-fitting techniques for such models. Other extensions include allowing random slopes as well as random intercepts. Cooper et al. (2007) present a version of the product normal prior that allows for both a random intercept and slope in the second part of the model. Finally, more work is needed to determine the best practical approach to model comparison for random effect models. Celeux et al. (2006) have recommended a variety of DIC alternatives, but many of these are more computationally challenging than the default provided by WinBUGS.

In general, zero-inflated models have a wide range of application, and a Bayesian approach should appeal to investigators seeking more flexible alternatives to classical model-fitting procedures.

Acknowledgments

The authors would like to thank Robert Rosenheck, MD, PhD, of Yale University for generously providing the ACCESS data, Marcela Horvitz-Lennon, MD, of the University of Pittsburgh for clinical expertise with the ACCESS study and two anonymous referees. This work was supported by Grants R01-MH61434 and P50-MH73469 from the National Institute of Mental Health, Bethesda, MD.

References

- Berger J, Pericchi L. The intrinsic Bayes factor for model selection and prediction. *Journal of the American Statistical Association* 1996;91:109–22.
- Brooks SP, Gelman A. General methods for monitoring convergence of iterative simulations. *Journal of Computational and Graphical Statistics* 1998;7:434–55.
- Browne WJ, Draper D. Implementation and performance issues in the Bayesian and likelihood fitting of multilevel models. *Computational Statistics* 2000;15:391–420.
- Celeux G, Forbes F, Robert CP, Titterton DM. Deviance information criteria for missing data models. *Bayesian Analysis* 2006;1:651–74.
- Congdon, P. *Bayesian models for categorical data*. Chichester: John Wiley & Sons; 2005.
- Cooper NJ, Sutton AJ, Mugford M, Abrams KR. Use of Bayesian Markov chain Monte Carlo methods to model cost-of-illness data. *Medical Decision Making* 2003;23:38–53. [PubMed: 12583454]
- Cooper NJ, Lambert PC, Abrams KR, Sutton AJ. Predicting costs over time using Bayesian Markov chain Monte Carlo methods: an application to early inflammatory polyarthritis. *Health Economics* 2007;16:37–56. [PubMed: 16981192]

- Dalrymple ML, Hudson IL, Ford RPK. Finite mixture, zero-inflated Poisson and hurdle models with application to SIDS. *Computational Statistics and Data Analysis* 2003;41:491–04.
- Dey DK, Chen M-H, Chang H. Bayesian approach for nonlinear random effect models. *Biometrics* 1997;53:1239–52.
- Duan N. Smearing estimate: a nonparametric retransformation method. *Journal of the American Statistical Association* 1983;78:605–10.
- Elliott MR, Gallo JJ, Ten Have TR, Bogner HR, Katz IR. Using a Bayesian latent growth curve model to identify trajectories of positive affect and negative events following myocardial infarction. *Biostatistics* 2005;6:119–43. [PubMed: 15618532]
- Fahrmeir L, Osuna Echavarría L. Structured additive regression for overdispersed and zero-inflated count data. *Applied Stochastic Models in Business and Industry* 2006;22:351–69.
- Gamerman D. Efficient Sampling from the posterior distribution in generalized linear models. *Statistics and Computing* 1997;7:57–68.
- Geisser S, Eddy W. A predictive approach to model selection. *Journal of the American Statistical Association* 1979;74:153–60.
- Gelfand, AE. Model determination using sampling-based methods. In: Gilks, WR.; Richardson, S.; Spiegelhalter, DJ., editors. *Markov chain Monte Carlo in practice*. London: Chapman & Hall; 1996. p. 145–60.
- Gelfand AE, Dey D. Bayesian model choice: asymptotics and exact calculations. *Journal of the Royal Statistical Society: Series B* 1994;56:501–14.
- Gelman A. Prior distributions for variance parameters in hierarchical models. *Bayesian Analysis* 2006;1:515–33.
- Gelman, A.; Meng, XL. Model checking and model improvement. In: Gilks, WR.; Richardson, S.; Spiegelhalter, DJ., editors. *Markov chain Monte Carlo in practice*. London: Chapman & Hall; 1996. p. 191–201.
- Gelman A, Meng XL, Stern H. Posterior predictive assessment of model fitness via realized discrepancies. *Statistica Sinica* 1996;6:733–807.
- Gelman, A.; Carlin, JB.; Stern, HS.; Rubin, DB. *Bayesian data analysis*. 2. Boca Raton: Chapman & Hall; 2004.
- Geweke, J. Evaluating the accuracy of sampling-based approaches to calculating posterior moments. In: Bernardo, J.; Berger, JO.; Dawid, AP.; Smith, AFM., editors. *Bayesian Statistics*. Vol. 4. Oxford: Oxford University Press; 1992.
- Ghosh SK, Mukhopadhyay P, Lu JC. Bayesian analysis of zero-inflated regression models. *Journal of Statistical Planning and Inference* 2006;136:1360–75.
- Gilks WR, Wild P. Adaptive rejection sampling for Gibbs sampling. *Applied Statistics* 1992;41:337–48.
- Gilks WR, Wang CC, Yvonnet B, Coursaget P. Random-effects models for longitudinal data using Gibbs sampling. *Biometrics* 1993;49:441–53. [PubMed: 8369380]
- Gilks WR, Best NG, Tan KKC. Adaptive rejection Metropolis sampling within Gibbs sampling. *Applied Statistics* 1995;44:455–72.
- Green P. Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika* 1995;82:711–32.
- Greene, W. Working Paper EC- 94-10. Department of Economics, New York University; 1994. Accounting for excess zeros and sample selection in Poisson and negative binomial regression models.
- Gschlößl S, Gzado C. Modelling count data with overdispersion and spatial effects. *Statistical Papers* 2008;49:531–552.
- Hall D. Zero-inflated Poisson and binomial regression with random effects: a case study. *Biometrics* 2000;56:1030–39. [PubMed: 11129458]
- Heilbron, DC. SIMS Technical Report 9. Department of Epidemiology and Biostatistics, University of California; San Francisco: 1989. Generalized linear models for altered zero probabilities and overdispersion in count data.

- Heilbron DC. Zero-altered and other regression models for count data with added zeros. *Biometrical Journal* 1994;36:531–47.
- Hoeting JA, Madigan D, Raftery AE, Volinsky CT. Bayesian model averaging: a tutorial. *Statistical Science* 1999;14:382–417.
- Kass RE, Raftery AE. Bayes Factors. *Journal of the American Statistical Association* 1995;90:773–95.
- Lambert D. Zero-inflated Poisson regression, with an application to defects in manufacturing. *Technometrics* 1992;34:1–14.
- Lambert PC, Sutton AJ, Burton PR, Abrams KR, Jones DR. How vague is vague? A simulation study of the impact of the use of vague prior distributions in MCMC using WinBUGS. *Statistics in Medicine* 2005;24:2401–28. [PubMed: 16015676]
- Liu H, Power DA. Growth curve models for zero-inflated count data: an application to smoking behavior. *Structural Equation Modeling* 2007;14:247–79.
- Min Y, Agresti A. Random effect models for repeated measures of zero-inflated count data. *Statistical Modelling* 2005;5:1–19.
- Mullahy J. Specification and testing of some modified count data models. *Journal of Econometrics* 1986;33:341–65.
- Muthén B, Brown CH, Booil Jo KM, Khoo ST, Yang CC, Wang CP, Kellam SG, Carlin JB, Liao J. General growth mixture modeling for randomized preventive interventions. *Biostatistics* 2002;3:459–75. [PubMed: 12933592]
- Olsen MK, Schafer JL. A two-part random-effects model for semicontinuous longitudinal data. *Journal of the American Statistical Association* 2001;96:730–45.
- R Development Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing; Vienna, Austria: 2007. URL: <http://www.R-project.org>
- Ridout, M.; Demétrio, CGB.; Hinde, J. Models for count data with many zeros. *Proceedings from the International Biometric Conference*; Cape Town. December 1998; 1998.
- Rodrigues J. Bayesian analysis of zero-inflated distributions. *Communications in Statistics, Theory and Methods* 2003;32:281–89.
- Roeder K, Lynch KG, Nagin DS. Modeling uncertainty in latent class membership: a case study in criminology. *Journal of the American Statistical Association* 1999;94:766–76.
- Rose CE, Martin SW, Wannemuehler KA, Plikaytis BD. On the use of zero-inflated and hurdle models for modeling vaccine adverse event count data. *Journal of Biopharmaceutical Statistics* 2006;16:463–81. [PubMed: 16892908]
- Rosenheck RA, Lam JA, Morrissey JP, Calloway MO, Stolar M, Randolph F. Service systems integration and outcomes for mentally ill homeless persons in the ACCESS program. *Psychiatric Services* 2002;53:958–66. [PubMed: 12161669]
- Smith AFM, Gelfand AE. Bayesian statistics without tears: a sampling-resampling perspective. *American Statistician* 1992;46:84–88.
- Smith BJ. boa: An R package for MCMC output convergence assessment and posterior inference. *Journal of Statistical Software* 2007;21:1–37.
- Spiegelhalter DJ. Bayesian graphical modelling: a case study in monitoring health outcomes. *Applied Statistics* 1998;47:115–33.
- Spiegelhalter DJ, Best NG, Carlin BP, van der Linde A. Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B* 2002;64:583–39.
- Spiegelhalter, DJ.; Thomas, A.; Best, N.; Lunn, D. WinBUGs Version 1.4: User Manual. Cambridge: Medical Research Council Biostatistics Unit; 2003. URL: <http://www.mrc-bsu.cam.ac.uk/bugs>
- Tooze JA, Grunwald GK, Jones RH. Analysis of repeated measures data with clumping at zero. *Statistical Methods in Medical Research* 2002;11:341–55. [PubMed: 12197301]
- Yau KK, Lee AH. Zero-inflated Poisson regression with random effects to evaluate an occupational injury prevention programme. *Statistics in Medicine* 2001;20:2907–20. [PubMed: 11568948]

Figure 1.
Histograms of days utilization, by race and visit.

Figure 2.

Trace plots based on two MCMC chains for four representative parameters from the ZIP model: (a) β_{13} (minority race, binomial component); (b) β_{25} (minority race $\times$ visit₁₂ interaction, Poisson component); (c) σ_2 (standard deviation for b_{2i}); and (d) ρ (random effect correlation). Horizontal lines denote posterior means.

Figure 3.

Posterior predictive distributions for the ZIP model based on (a) the sample proportion of zeros and (b) the over-dispersion index. The dashed vertical lines denote the observed value of the test statistics; the shaded regions highlight the area corresponding to the Bayesian predictive p-values.

Table 1

Summary statistics for the ACCESS study.

Visit	Race	n	Percent Zeros	Mean Non-Zero Days of Utilization (SD)
Baseline	White	299	54.52	4.29 (3.82)
	Black/Latino	381	51.44	3.59 (3.59)
3 Months	White	299	37.12	4.16 (3.58)
	Black/Latino	381	40.66	4.08 (3.76)
12 Months	White	299	48.16	4.05 (3.81)
	Black/Latino	381	44.88	3.77 (3.22)

Table 2

Model comparison statistics for ACCESS study.

Model	NLL ($\Delta^{\dagger}$)	$\bar{D}$ (θ)	P_D	DIC ($\Delta^{\dagger}$)	% CPO > 0.01
Ordinary Poisson	4871.84 (—)	7924.27	531.44	8455.72 (—)	89.17
ZAP [‡]	4186.89 (684.95)	7353.00	439.14	7792.14 (663.58)	92.50
Poisson Hurdle	4065.15 (121.74)	6813.00	731.55	7544.55 (247.59)	92.79
ZIP	4039.93 (25.22)	6769.00	659.70	7428.70 (115.85)	93.14

[‡] Change from previous model.

[‡] Posterior mean for $\gamma = -1.28$ (95% CrI = $[-1.36, -1.21]$).

Table 3

Posterior estimates and 95% credible intervals (CrIs) from ZIP model for the ACCESS study.

Model Component	Parameter (Variable)	Posterior Mean (95% CrI)
Binomial	β_{10} (Intercept)	0.27 (−0.27, 0.88)
	β_{11} (Visit ₃)	1.25 (0.72, 1.80)
	β_{12} (Visit ₁₂)	0.64 (0.10, 1.17)
	β_{13} (Minority Race)	0.36 (−0.24, 0.97)
	β_{14} (Minority Race×Visit ₃)	−0.54 (−1.25, 0.18)
	β_{15} (Minority Race×Visit ₁₂)	−0.17 (−0.89, 0.53)
	β_{16} (Male Gender)	−0.15 (−0.61, 0.34)
	β_{17} (SA Treatment)	0.03 (−0.44, 0.51)
Poisson	β_{20} (Intercept)	1.23 (1.02, 1.44)
	β_{21} (Visit ₃)	−0.02 (−0.15, 0.11)
	β_{22} (Visit ₁₂)	−0.14 (−0.28, −0.003)
	β_{23} (Minority Race)	−0.20 (−0.40, 0.02)
	β_{24} (Minority Race×Visit ₃)	0.14 (−0.04, 0.32)
	β_{25} (Minority Race×Visit ₁₂)	0.20 (−0.0001, 0.38)
	β_{26} (Male Gender)	−0.05 (−0.23, 0.12)
	β_{27} (SA Treatment)	−0.04 (−0.21, 0.13)
Variance Components	σ_1	1.97 (1.56, 2.45)
	σ_2	0.79 (0.71, 0.90)
	ρ	−0.40 (−0.64, −0.14)

Table 4

Estimated mean utilization days, by race and visit.

Visit	Race	Posterior Mean (95% CrI)
Baseline	White	1.92 (1.63, 2.24)
	Black/Latino	1.78 (1.54, 2.04)
3 Months	White	2.69 (2.35, 3.06)
	Black/Latino	2.44 (2.16, 2.74)
12 Months	White	2.03 (1.75, 2.35)
	Black/Latino	2.14 (1.88, 2.42)