Open Access Medical Statistics

Dovepress
open access to scientific and medical research

Open Access Full Text Article

METHODOLOGY

Comparison of various modeling approaches in the analysis of longitudinal data with a binary outcome: The Ontario Mother and Infant Study (TOMIS) III

This article was published in the following Dove Press journal:
Open Access Medical Statistics
18 July 2012
Number of times this article has been viewed

Yu Qing Bai[1]
Gary Foster[2]
Wendy Sword[3,4]
Paul Krueger[5]
Christine Kurtz Landy[6]
Susan Watt[7]
Lehana Thabane[2,3]

[1]Department of Mathematics and Statistics, McMaster University, Hamilton, ON, Canada; [2]Biostatistics Unit, St Joseph's Healthcare, Hamilton, ON, Canada; [3]Department of Clinical Epidemiology and Biostatistics, [4]School of Nursing, McMaster University, Hamilton, ON, Canada; [5]Department of Family and Community Medicine, University of Toronto, Toronto, ON, Canada; [6]School of Nursing, York University, Toronto, ON, Canada; [7]School of Social Work, McMaster University, Hamilton, ON, Canada

Background: Longitudinal studies are often used to investigate the developmental trends of outcomes over time. Several modeling strategies can be applied for the analyses of longitudinal data. In this study, various statistical approaches were discussed and compared using data from The Ontario Mother and Infant Study (TOMIS) III. TOMIS III was a longitudinal cohort study that assessed the associations between the method of delivery and health outcomes and service utilizations. The primary outcome of postpartum depression was used as an example.

Methods: Generalized estimating equations (GEE) assuming a serial correlation structure were used as the primary method of analysis to assess the association between the method of delivery and postpartum depression over 12 months. We performed sensitivity analyses using three other methods – namely, the (1) generalized linear mixed-effects model (GLMM), (2) hierarchical generalized linear model (HGLM), and (3) Bayesian hierarchical model (BHM), to compare the robustness of the results.

Results: The results from all four models indicated that the method of delivery had no significant effect on postpartum depression. However, GEE, GLMM, and BHM identified the following seven predictors of depression: annual household income; urinary incontinence (bladder problems); English or French (Canada's official languages) spoken at home; a lower SF-12 mental component score; unmet learning needs in the hospital; lower social support; and a lower SF-12 physical component score. HGLM showed similar results to the above three models with the exception of language spoken at home, which was not significant. GEE provided the good fit statistics for the data.

Conclusion: Method of delivery had no significant effect on postpartum depression, based on GEE analysis. This result remained robust under different methods of analyses. GEE demonstrated a good fit for the TOMIS III data.

Keywords: longitudinal data, generalized estimating equations, hierarchical model, TOMIS

Introduction

Longitudinal studies are widely used in health sciences research. In such studies, measurements of the same individual are performed repeatedly through time. Clusters are another distinct feature of some longitudinal studies.[1] The measurements conducted within clusters (eg, hospitals) typically exhibit some within-site correlations. In addition, patients that are nested in a hospital may also be nested in a region, giving the data a multilevel structure. All of these features should be considered when conducting analyses. A longitudinal study is the only way of capturing the within-individual change over time as it uses repeated measures on each individual. The primary goal

Correspondence: Lehana Thabane
Department of Clinical Epidemiology and Biostatistics, McMaster University, 1200 Main Street West, Hamilton, ON, L8N 3Z5 Canada
Tel +1 905 522 1155 ext 33720
Fax +1 905 308 7386
Email thabanl@mcmaster.ca

of a longitudinal study is to characterize the changes in response over time and to determine the factors that influence those changes.[1]

Generalized estimating equations (GEE) are a special type of generalized linear model for analyzing longitudinal data and have been discussed by many researchers.[1–5] In GEE, the correlations between individuals within a cluster can be accounted for using a correlation matrix. The generalized linear mixed-effects model (GLMM) is another technique for handling such correlations in longitudinal data. The GLMM can be considered a straightforward extension of the generalized linear model, adding random effects to the linear link function and expressing the response mean conditional on the random effects. A multilevel data structure is present when clusters exist in the data. This is especially true for health sciences data since individuals can be grouped in many different ways.[1] Therefore, the use of hierarchical models can also be an analytical option. Both the hierarchical generalized linear model (HGLM)[6] and the Bayesian hierarchical model (BHM)[7,8] can be introduced to the analysis of longitudinal data with cluster features. Although some studies have discussed different analytical methods for handling longitudinal data, there are very few published papers comparing these statistical methods.

The Ontario Mother and Infant Study (TOMIS) III, funded by the Canadian Institutes of Health Research, used a prospective cohort design to study the associations between the method of delivery and health, health services utilization, and costs of care in the first year following postpartum discharge from the hospital. We hypothesized that the mode of delivery would be associated with the risk of postpartum depression. TOMIS III data have a typical longitudinal structure with multilevel features. The measurements were conducted at baseline and at three time points after discharge – at 6 weeks, 6 months, and 12 months. GEE was used as a primary modeling approach for this study while GLMM, HGLM, and BHM were used as sensitivity analyses to compare and verify the GEE results.

Methods

Overview of the TOMIS III study

A total of 2560 women were recruited from 11 hospitals across Ontario, Canada. The following comprise the eligibility criteria: ≥16 years of age; delivery of a live singleton infant; ≥37 weeks of gestation; mother assuming care of infant when discharged; mother competent to give consent; and mother could be contacted by telephone. Women were ineligible to participate if their infant required admission to a neonatal intensive care or special care nursery for more than 24 hours or were unable to communicate in one of the four study languages (English, French, Chinese, and Spanish).[9] Data were collected using a self-report questionnaire (baseline measurements) in the hospital and during scheduled telephone interviews at 6 weeks, 6 months, and 12 months after discharge. There was an attrition rate of 30%[9] over the course of this study, resulting in a substantial amount of missing data in the final dataset. The outcomes in this study included postpartum depression, maternal health, and infant health. The primary outcome of postpartum depression is used as an example in the analyses presented in this paper. Independent variables used in this study were adapted from a previously published study in which the procedure for variable selection is described.[10]

Outcome of postpartum depression

Postpartum depression (PPD) was measured using the 10-item Edinburgh Postnatal Depression Scale (EPDS). Individuals scoring greater than 12 on the EPDS were considered to have depression. PPD in this study was treated as a binary outcome, with 0 denoting "no depression" and 1 denoting "depression".

Intraclass correlation coefficient

The intraclass correlation coefficient (ICC) can be used to measure the similarity of individuals within the same hospital. The design effect then can be obtained from the ICC using the formula[11] Design effect = $1 + (m-1)\rho$, where m is the average size per cluster and ρ is the ICC. The design effect was used to measure the magnitude of the effect of clustering.[12]

Statistical methods

GEE was used as the primary analytical approach for the analysis of the outcome of PPD. Three other models – GLMM, HGLM, and BHM – were applied to the same dataset and the robustness of the results was evaluated by comparing the clinical similarity and differences, and comparing the fit statistics of these three models with GEE. All classical models were conducted using SAS 9.1 (SAS Institute Inc, Cary, NC) and the Bayesian hierarchical models were fitted using WinBUGS 1.4 (Medical Research Council, London/Swindon, UK). For all classical models, the results were reported as odds ratios (OR), corresponding two-sided 95% confidence intervals (CI), and associated *P*-values. The results from Bayesian analysis were reported as posterior estimates along with 95% credible intervals (CrI).

Generalized estimating equations

Let Y_{ij} denote the response of PPD for the *ith* patient in the *jth* hospital in this case. Y_{ij} is binary and follows a Bernoulli distribution and the mean, μ_{ij}, is related to the covariates by the logit link function:

$$logit\left(\mu_{ij}\right) = X_{ij}^{'}\beta,$$

where, X_{ij} is a vector of covariates. The regression coefficient β can be estimated by solving the GEE.

The GEE can account for the correlations between patients within each hospital using a working correlation matrix. An exchangeable working correlation matrix was used because we assumed that the correlation between any two patients within one hospital was equal. This was performed using the SAS procedure of GENMOD by specifying the REPEATED statement.

Generalized linear mixed-effects model

Compared to GEE, which uses the working correlation matrix to capture the within-cluster correlation, GLMM accounts for this correlation by adding a random effect, b_j, which is assumed to follow a multinormal distribution having a mean of zero and variance Ψ, ie, $b_j \sim N(0,\sigma^2)$. The linear predictor for a binary outcome can be expressed as:

$$logit\left(\mu_{ij}\right) = X_{ij}^{'}\beta + b_j$$

As with GEE, X_{ij} is a vector of covariates. The estimation for β using the maximum likelihood method is not straightforward due to the analytically intractable integrals in the likelihood function. The *glimmix* procedure with the RANDOM statement in SAS can be introduced to estimate β by applying the technique of restricted pseudo-likelihood based on the residual likelihood.[13,14]

Hierarchical generalized linear model

A three-level structure was considered for the TOMIS III data (Figure 1). The response, Y_{ijk}, denotes the depression measured for the *ith* patient at *jth* time in the *kth* hospital. A linear regression model was applied for each level and the random effect of each model was assumed to follow a normal distribution with a zero mean. Combining all models and ignoring the random effects on slopes, the final model was obtained as follows:

$$Y_{ijk} = \left\{\gamma_{000} + \sum_{p=1}^{p} x_{pijk}\left(\gamma_{p00} + \gamma_{p10}Time_j\right) + \gamma_{01k}Time_j\right\} + \left\{\left(\mu_{010}Time_j + \mu_{00k} + r_{0jk} + \varepsilon_{ijk}\right)\right\}.$$

The final model is expressed as the sum of two parts: a fixed part, which contains three fixed effects (for the intercept, for effects of patient-level factors, and for the effect of time) and a random part, which contains four random effects (for the intercept, ie, hospital-level residual μ_{00k}, time-level residual r_{0jk}, within-patient residual ε_{ijk}, and for time slope μ_{010}). In this equation, γ_{000}, γ_{p00}, and γ_{p10} are intercepts at the hospital level, and $Time_j$ is the time of measurement.

The SAS GLIMMIX procedure was used to fit HGLM by specifying Time and Intercept in the RANDOM statement.

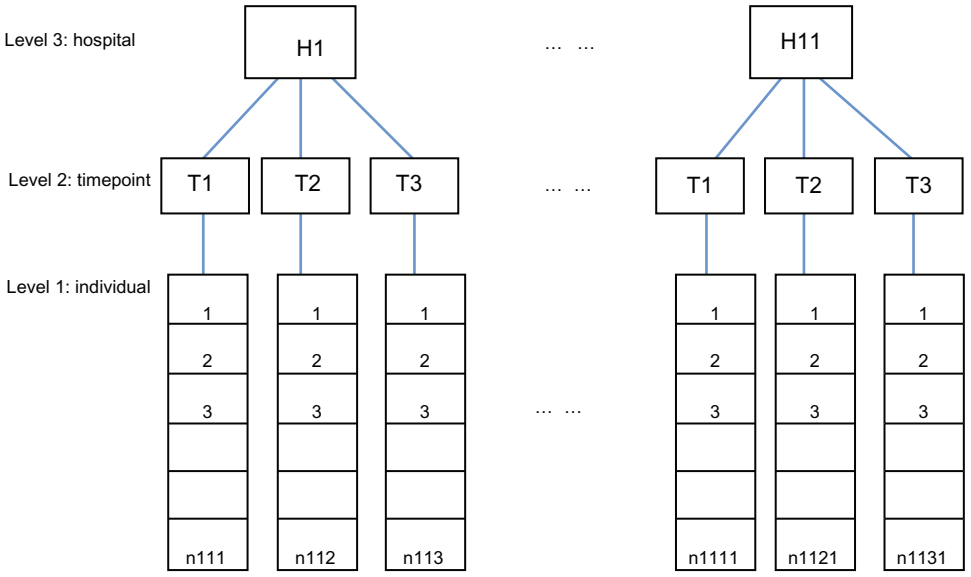

Figure 1 Three-level data structures for TOMIS III data.
Abbreviation: TOMIS III, The Ontario Mother and Infant Study.

Bayesian hierarchical model

Let Y_{ij} denote the binary outcome of depression on the *i*th patient at *j*th time. The Bayesian hierarchical model is

$$Y_{ij} \mid \pi_{ij} \sim Bernoulli(\pi_{ij})$$

with link function

$$logit(\pi_{ij}) = X_{ij}^{r}\beta + b_i,$$

where π_{ij} is the mean of Y_{ij}. The random effect b_i was assumed to follow the normal distribution $b_i \sim N(0,\sigma^2)$. The uncertainty of the estimated between-cluster variance σ^2 was taken into account by assuming a prior in the Bayesian approach.[15,16] The observed data Y_{ij} were treated as fixed and known quantities and our interests were the distributions of the parameter β. To minimize or eliminate the researchers' pre-beliefs or ex ante information, a non-informative prior, Uniform (0, 10), was used for σ. One of the Markov Chain Monte Carlo (MCMC) methods, namely, the Gibbs sampling algorithm, was introduced to summarize posterior distributions. The number of iterations was set at 20,000 with a burn-in number of 5,000. The seed was 0500485. The convergence of MCMC was assessed by comparing Monte Carlo errors and standard deviations. In addition, both dynamic trace plots and quantile plots were checked to detect the convergence.[17,18]

Impacts of priors for Bayesian analysis

In this study, we applied a uniform distribution as a noninformative prior for the Bayesian approach. However, an improper prior may influence the posterior. To investigate the impact of different priors, we performed sensitivity analyses using a variety of uniform priors such as U(0, 5), U(0,15), U(0, 20), U(0, 25), and U(0, 50) and conjugate priors such as inverse gamma (0.001, 0.001), (0.01, 0.01), and (0.1, 0.1).

Results

Participant characteristics

A total of 2560 women were recruited from 11 hospitals across the province of Ontario. Among these participants, 85.6% were 25 years or older and 67.7% had had a vaginal delivery. Approximately 70.8% were born in Canada and 81.9% spoke English or French at home. Most participants, 85.1%, had a college or university education and 89.7% had a total annual income more than Canadian $20,000. The percentage of women having their first pregnancy was 41.9% (n = 1071). At 6 weeks postdischarge, the percentage of subjects experiencing PPD, as indicated by an EPDS score ≥12, was 7.6%.

Results from different modeling approaches

Prior to modeling, multicollinearity diagnostics were conducted using logistic regression detections. The results indicated that total social score, social support, and instrumental support, as well as history of depression and any previous postpartum depression exhibited collinearity. We retained total social score and history of depression in the analysis and removed the others as we were most interested in the influence of these two variables on the primary outcome. Normally, when the ICC is greater than 0.1 or the design effect is larger than 2,[19,20] the correlations within clusters should be accounted for in the modeling approaches. The results from the ICC calculation showed that the ICC and 95% CI within hospitals were 0.01 (0.00, 0.04). The design effect along with the 95% CI was calculated as 3.59 (1.76, 9.51). Even though the ICC was small, the design effect still indicated that the outcome variable had slight correlations within each cluster.

Comparisons of interest

The results from the primary analytical modeling of GEE showed that the OR (postpartum depression versus mode of delivery) along with the 95% CI and *P*-value was 0.99 (0.73, 1.34) ($P = 0.9375$), which demonstrated that the association between the mode of delivery and postpartum depression was not significant. The estimates from the other models were 1.00 (0.71, 1.40) ($P = 0.9814$) for GLMM, 1.03 (0.75, 1.41) ($P = 0.8522$) for HGLM, and 0.98 (0.71, 1.34) for BHM, which gave similar results as GEE at $\alpha = 0.05$. However, for continuous covariates (eg, the SF-12 mental component score), similarities were observed between the estimates from GLMM [0.24 (0.20, 0.28), $P < 0.0001$] and BHM [0.22 (0.19, 0.26)] and between the estimates from GEE [0.84 (0.82, 0.85), $P < 0.0001$] and HGLM [0.84 (0.82, 0.85), $P < 0.0001$] (Figures 2 and 3).

Fit statistics

The GEE model is estimated based on a quasi-likelihood function; therefore, a modified Akaike information criterion (AIC) based on a quasi-likelihood can be computed to measure the goodness of fit of GEE.[21] The results from fit statistics demonstrated that GEE provided the smallest AIC (1403.73) and largest likelihood value (–690.20) among the three classical models. The AICs were 27151.05 and 28317.57 and the logarithm likelihood values were –13579.53 and –14156.88 for GLMM and HGLM, respectively. AIC cannot be applied for the Bayesian approach.

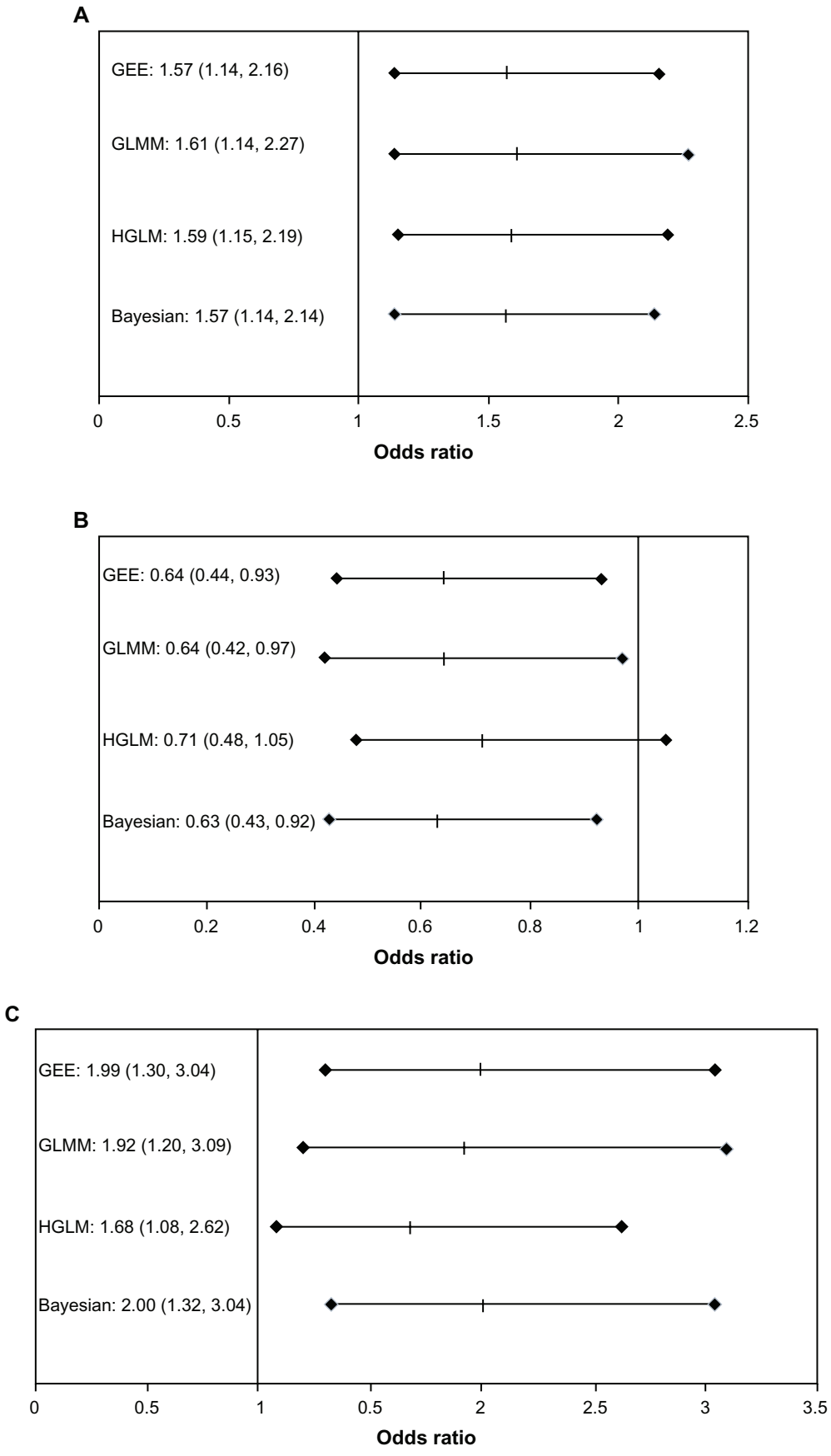

Figure 2 (*Continued*)

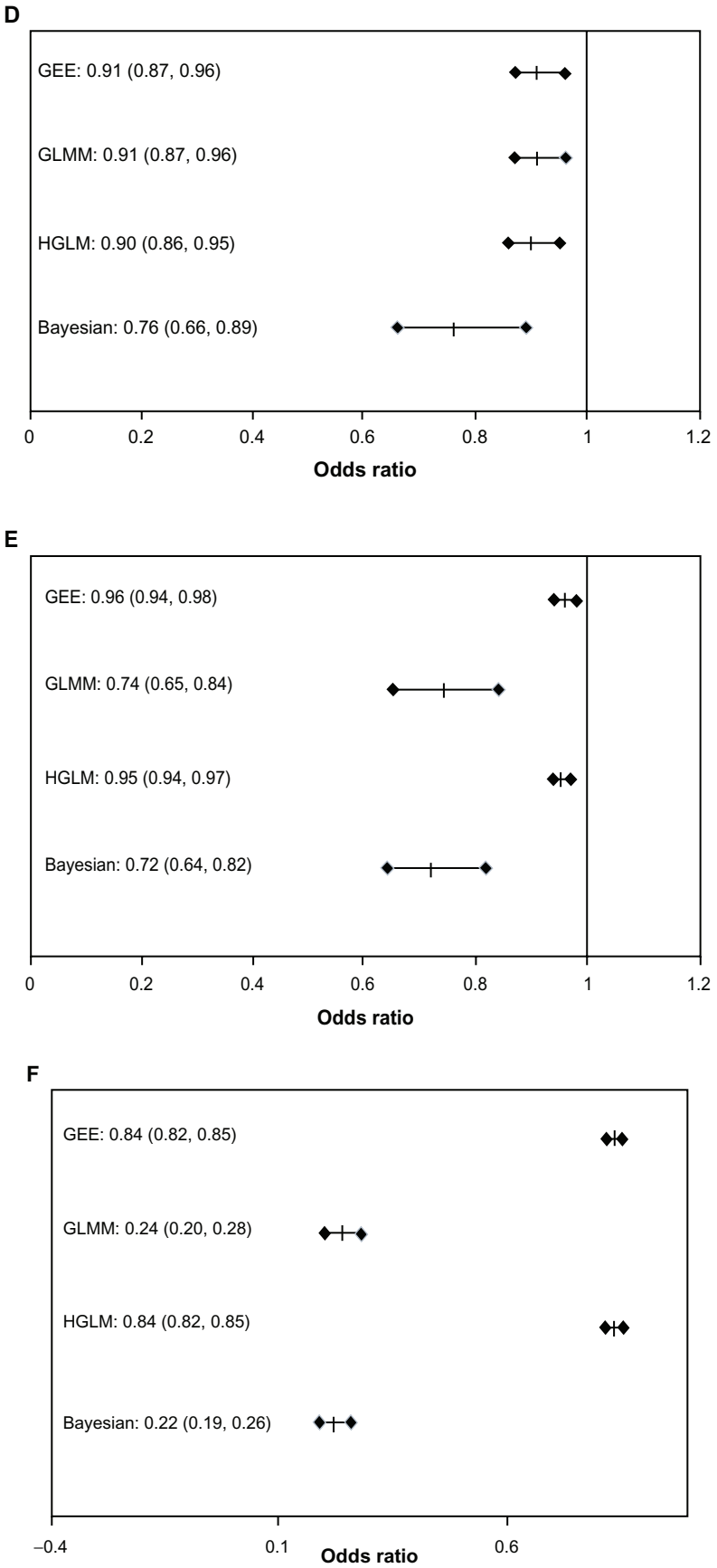

Figure 2 (*Continued*)

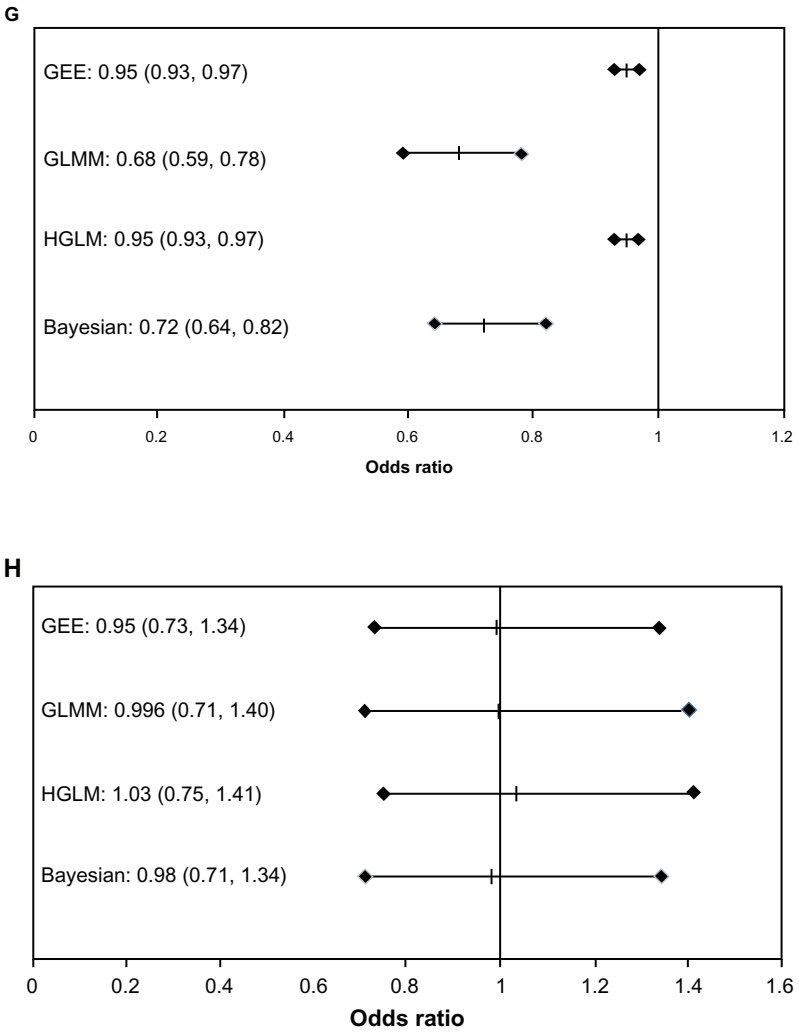

Figure 2 Forest plots of estimates of variables from different models. (**A**) Bladder problems; (**B**) language spoken at home; (**C**) total income; (**D**) number of met learning needs in the hospital; (**E**) SF-12 physical component score; (**F**) SF-12 mental component score; (**G**) total social support; and (**H**) delivery method.
Abbreviations: GEE, generalized estimating equations; GLMM, generalized linear mixed-effects model; HGLM, hierarchical generalized linear model.

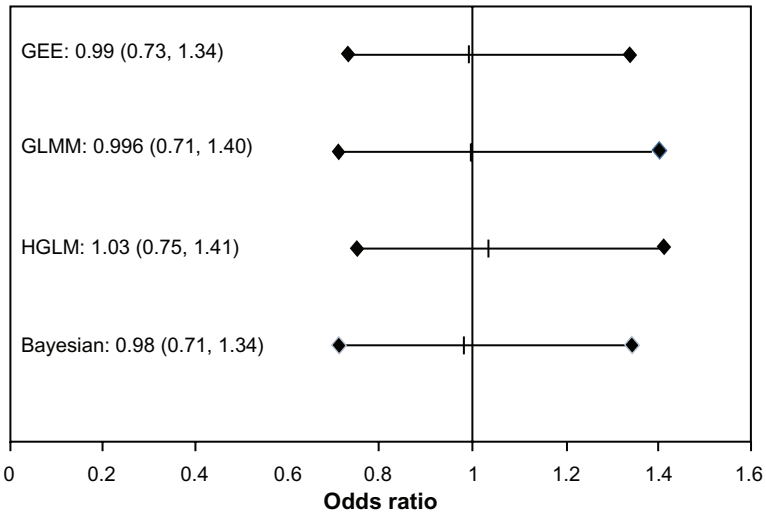

Figure 3 Forest plot of comparison of four modeling approaches.
Note: Odds ratio = postpartum depression versus mode of delivery.
Abbreviations: GEE, generalized estimating equations; GLMM, generalized linear mixed-effects model; HGLM, hierarchical generalized linear model.

Impacts of priors for Bayesian analysis

We applied five uniform priors, U(0, 5), U(0,15), U(0, 20), U(0, 25), and U(0, 50) and three conjugate priors, Inverse Gamma (0.001, 0.001), (0.01, 0.01), and (0.1, 0.1). The ORs (depression versus mode of delivery) and 95% CrI corresponding to the above priors were 0.98 (0.72, 1.33), 0.98 (0.72, 1.35), 0.98 (0.72, 1.34), 0.98 (0.71, 1.35), 0.98 (0.72, 1.34), 0.98 (0.72, 1.33), 0.98 (0.72, 1.34), and 0.98 (0.72, 1.35), respectively (Figure 4). The results of the sensitivity analyses for different priors were very similar; there were no differences among estimates for up to two decimal places.

Discussion

Summary of findings

GEE as a primary modeling approach was introduced to analyze the TOMIS III data. To evaluate the robustness of the analytical results, two classical models (GLMM and HGLM) and a Bayesian model were compared to the GEE model. The results of the association between the delivery method and postpartum depression were similar for all four models, though they had different ways to explain the between-cluster correlations. GEE captured the correlation by specifying a working correlation matrix, while HGLM used a combined form of linear models for all levels. Both GLMM and BHM can be considered random-effect logistic models for the binary outcome of depression and they explained the correlations using random effects. This can explain why the estimates from GLMM and BHM had similar results for some covariates but different results for others when compared to GEE and HGLM. The fit statistics of GEE provided the smallest AIC and largest likelihood values, indicating that the GEE model demonstrated an excellent fit for the analysis of the TOMIS III data.

Although we applied noninformative priors for Bayesian analysis to eliminate the effects of researchers' prebeliefs or external information on posterior distributions, sensitivity analysis is also necessary to ensure the impact of priors. The results from our sensitivity analyses showed that the ORs and 95% CI are quite similar for all uniform priors and inverse gamma priors, indicating that the sensitivity of the results to different prior assumptions exhibited weak information relative to the observed data.

Limitations and future work

The designed sample size for TOMIS III was 3774 women based on an attrition rate of 30% and an ICC of 0.018.[10] Thirty independent variables and one response variable were involved in the primary analysis for the outcome of postpartum depression. The missing data rate for the predictors of physical score and mental score were up to 50% at a follow-up time of 12 months. About half (50.7%) of the participants completed the 12-month interview with EPDS. These issues may reduce the power of the analysis.

For this study, only main effects models were considered for all analyses. In practice, some combined effects (ie, interactions) should be considered in regression models. For example, an interaction term of mode of delivery and country of birth was reported to have a significant influence on postpartum depression at 6 weeks.[9] Future work is recommended for analysis with interaction terms.

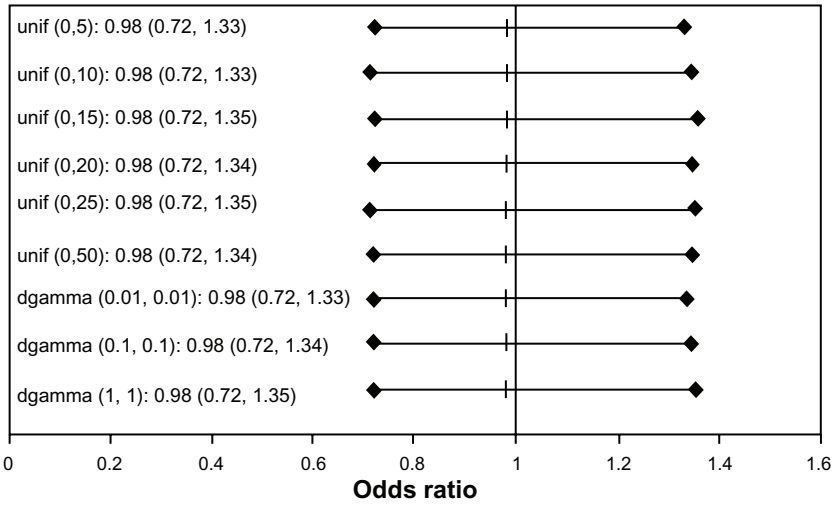

Figure 4 Forest plot of sensitivity analysis for various prior distributions.

Conclusion

The GEE model was applied to detect the association between the method of delivery and postpartum depression. The sensitivity of analytical results was investigated by comparing the GEE estimates to three other modeling strategies. From all four analyses, we concluded that the method of delivery had no significant effect on postpartum depression. The results remained robust under all methods of analysis. However, GEE demonstrated the best fit for the data.

Authors' contributions

YQB and LT conceived the study. WS, PK, CKL, SW, and LT participated in the design and implementation of the TOMIS III study. Data cleaning was performed by GF. YQB conducted the data analysis and wrote the initial draft of the manuscript. Results of the data analysis were interpreted by YQB, GF, and LT. All authors reviewed and revised the draft version of the manuscript. All authors read and approved the final version of the manuscript.

Acknowledgments

The TOMIS III study was funded by a grant from the Canadian Institutes of Health Research. This study was funded in part by the funds from the CANNeCTIN Programme.

Disclosure

The authors report no conflicts of interest in this work.

References

1. Fitzmaurice GM, Laird NM, Ware JH. *Applied Longitudinal Analysis*. New York, NY: John Wiley & Son, Inc.; 2002.
2. Liang KY, Zeger L. Models for Longitudinal Data: A generalized estimating equation approach. *Biometrics*. 1988;44:1049–1060.
3. Ballinger GA. Using generalized estimating equations for longitudinal data analysis. *Organizational Research Methods*. 2004;7:127–150.
4. Liang KY, Zeger L. Longitudinal data analysis for discrete and continuous outcomes. *Biometrics*. 1986;44:121–130.
5. Twisk JW. Longitudinal data analysis. A comparison between generalized estimating equations and random coefficient analysis. *Eur J Epidemiol*. 2004;19(8):769–776.
6. Hardin JW, Hilbe JM. *Generalized Linear Models and Extensions*. 2nd ed. College Station, TX: Stata Press; 2001.
7. Gill J. Bayesian Methods: *A Social and Behavioral Sciences Approach*. Boca Raton: Chapman & Hall/CRC. 2002.
8. Gelman A, Carlin JB, Stem HS, Rubin DB. *Bayesian Data Analysis*. 2nd ed. New York, NY: Chapman & Hall/CRC; 2004.
9. Sword W, Watt S, Krueger P, et al. The Ontario Mother and Infant Study (TOMIS) III: a multi-site cohort study of the impact of delivery method on health, service use, and costs of care in the first postpartum year. *BMC Pregnancy Childbirth*. 2009;9:16.
10. Sword W, Landy CK, Thabane L, et al. Is mode of delivery associated with postpartum depression at 6 weeks: a prospective cohort study. *BJOG*. 2011;118(8):966–977.
11. Ukoumunne OC, Gulliford MC, Chinn S, Sterne JA, Burney PG. Methods for evaluating area-wide and organisation based interventions in health and health care: a systematic review. *Health Technol Assess*. 1999;3(5):iii–92.
12. Bland JM. Cluster randomised trials in the medical literature: two bibliometric surveys. *BMC Med Res Methodol*. 2004;4:21.
13. Wolfinger R, O'Connell M. Generalized linear mixed models: a pseudo-likelihood approach. *J Stat Compu Simul*. 1993;4:233–243.
14. *The GLIMMIX Procedure*. SAS Support Documentation. Cary, NC: SAS Institute, Inc; 2006.
15. Peters TJ, Richards SH, Bankhead CR, Ades AE, Sterne JA. Comparison of methods for analysing cluster randomized trials: an example involving a factorial design. *Int J Epidemiol*. 2003;32:840–846.
16. Ma JH, Thabane L, Kaczorowski J, et al. Comparison of Bayesian and classical methods in the analysis of cluster randomized controlled trials with a binary outcome: the Community Hypertension Assessment Trial (CHAT). *BMC Med Res Methodol*. 2009;9:37.
17. Cowles MK, Carlin BP. Markov chain Monte Carlo diagnostics: a comparative review. *J Am Stat Assoc*. 1995;91:883–904.
18. Brooks SP, Roberts GO. Assessing convergence of Markov chain Monte Carlo algorithms. *Stat Comput*. 1998;8:319–335.
19. Hox J. *Multilevel Analysis*: *Techniques and Applications*. Mahwah, NJ: Lawrence Erlbaum; 2002.
20. Sorra S, Dye N. Multilevel psychometric properties of the AHRQ hospital survey on patient safety culture. *BMC Health Serv Res*. 2010;10:199.
21. Pan W. Akaike's information criterion in generalized estimating equations. *Biometrics*. 2001;57:120–125.