ORIGINAL RESEARCH

# Identifying determinants and estimating the risk of inadequate and excess gestational weight gain using a multinomial logistic regression model

Binod Neupane[1]
Sarah D McDonald[1,2]
Joseph Beyene[1]

[1]Department of Clinical Epidemiology and Biostatistics, [2]Department of Obstetrics and Gynecology and Radiology, McMaster University, Hamilton, ON, Canada

**Abstract:** When there are three or more nominal categories of a response variable, the binomial logistic regression approach is widely used to model the relationships of exposure variables with different binomial responses one at a time. However, some of the separate binomial comparisons would be redundant. This approach is also suboptimal because of the loss of information that will result when only a subset of the data is analyzed at a time and the multiple testing problems arising from analysis of several pairs of categories. These drawbacks of fitting separate binomial regression models to a multicategory nominal outcome variable can be overcome using a single multinomial regression modeling framework. In this study, we compared the results using a multinomial regression with the separate two binomial regressions to determine factors associated with excess and inadequate weight gain during pregnancy in a data set from a gestational weight gain study involving a cross-sectional survey of 312 women with singleton pregnancies. We found that both approaches identified the same set of predictors, ie, higher neuroticism, planning to gain more weight than the recommended level, and bedtime television watching, with $P$-values $\leq 0.05$ of the excessive (versus appropriate) weight gain, for which the subgroup size was moderate. The final list of significant predictors of inadequate (versus appropriate) weight gain identified by multinomial regression were planned weight gain below the recommended range, overweight or obese women, and bedtime television watching, while those by a separate binomial approach were self-efficacy towards achieving healthy weight, lack of weight satisfaction, and bedtime television watching, which differed between the two approaches where the final set of predictors were identified by a variable selection process and the comparisons were made in a small subgroup. A multinomial approach is a useful analytical framework that researchers may consider when they have multinomial response categories because this approach allows nonredundant comparisons to be made, avoiding the need to analyze a subset of the data one at a time and also allows for risk prediction of multinomial categories from a well validated multinomial model, and will not lead to multiple testing problems.

**Keywords:** gestational weight gain, pregnancy, multinomial logistic regression, binomial logistic regression, risk factors

## Introduction

In epidemiological research, response variables can have multiple nominal (unordered) nonoverlapping categories or classes (called "multinomial" or "polytomous" responses), and investigators are interested in the relationships between one or more predictor (independent) variables and these response categories. The most common approach for assessing these relationships is to fit separate binomial logistic regression models for the log-odds of each of the response categories versus any other chosen reference category. However, this approach is suboptimal for many reasons.

Correspondence: Joseph Beyene
Department of Clinical Epidemiology and Biostatistics, McMaster University, 1280 Main Street West, Hamilton, ON L8S 4K1, Canada
Tel +1 905 525 9140 extn 21333
Fax +1 905 528 2814
Email beyene@mcmaster.ca

First, if there are $K$ nominal response categories ($k$=1, 2, …, $K$; $K{\geq}3$), there are $\binom{K}{2}=(K-1)K/2$ separate binomial logistic regression models one can possibly fit. For example, for $K$=4 response categories $A$, $B$, $C$, $D$, one can fit $\binom{4}{2}=6$ separate binary logistic regression models for the comparisons of $A$ versus $B$, $A$ versus $C$, $A$ versus $D$, $B$ versus $C$, $B$ versus $D$, and $C$ versus $D$. Multinomial responses can also be compared with any one reference category fitting ($K-1$) equivalent binomial logistic regression models under a single analysis framework of "multinomial (or polytomous) logistic regression model".[1,2] It is a generalization of binomial regression when there are more than two nominal outcome categories. In the above example, fitting three models for $A$ versus $D$, $B$ versus $D$, and $C$ versus $D$ (assuming $D$ is the "baseline" or "reference" category) provides sufficient information about all possible pairwise relationships. Here, we do not obtain any additional information from certain comparisons given the ($K-1$) by fitting separate binomial regression models, and hence only ($K-1$) logits models are nonredundant.[1] Second, only a subset of the data is being used for a given pairwise binomial comparison, while all data from all subjects belonging to all possible response categories are utilized in multinomial analysis. Thirdly, a multinomial approach allows us to estimate the net effects of a set of predictor variables.[3] Lastly, and most importantly, a multinomial model can be the only effective and reliable way of estimating or predicting the probability or risk of a person having each of the $K$ possible nominal response categories given his or her set of risk factors,[4,5] where separate binomial comparisons do not enable us to predict such a probability or risk.

The binomial model is well known among clinicians and epidemiologists because the model is simple to understand and the results are easy to interpret. Although not as widely used, the multinomial approach is not a new topic either. Nevertheless, there have been limited data[6,7] generated from a multinomial approach in the gestational weight gain (GWG) literature, because typically GWG is examined as excess ("above") or inadequate ("below") weight gain compared with adequate ("within" the recommended range of the clinical guidelines) in a binomial fashion.[8–12]

In a study designed to assess the associations between clinical, psychological, lifestyle knowledge, and demographic factors and excess or inadequate GWG compared with appropriate gain,[13] we used the binomial logistic regression approach for each of the binomial outcomes of excess versus appropriate, and inadequate versus appropriate weight gain, separately. In this study, our objective was to identify the determinants of the trinomial response categories of inadequate, appropriate, and excessive weight gain simultaneously by fitting a single multinomial model to the data. We used a multinomial regression model to reanalyze the same data with trinomial response categories of inadequate, appropriate, and excessive weight gain, and compared the results with two separate binomial regressions previously reported to determine factors associated with excess and inadequate weight gain during pregnancy. We were interested to investigate if there is any practical added value of performing multinomial modeling in one analysis framework over separate binomial analyses, besides making nonredundant comparisons. In particular, we sought to determine if the estimation or inference (estimated effect sizes or association $P$-values or conclusions) for an exposure variable at a prespecified significance level is markedly different from the results of a previous analysis in small to moderate sizes of subsets of the data. We were also interested in estimating the risk of gaining weight excessively, inadequately, and appropriately for a pregnant woman given her set of characteristics.

## Materials and methods
### Multinomial logistic regression model
Suppose that the observations for a response variable $Y$ and $P$ predictor variables $\mathbf{X}$=($X_1$, $X_2$, …, $X_p$)′; ($j$=1, 2, …, $p$) are available from $N$ independent subjects, where the $i$th ($i$=1, 2, …, $N$) subject's response $Y_i$ can be one of the $K(\geq3)$ nonoverlapping nominal categories ($k$=1, 2, …, $K$). In a multinomial logistic regression framework, we can fit a "baseline-category logit model" for ($K-1$) logits simultaneously.[1] Here, if the $K$th category is the reference category, then ($K-1$) binomial logistic regression models are fitted simultaneously where the log of odds of having the response $k$($k$=1, 2, …, $K-1$) to the baseline response $K$ is modeled as the linear function of independent variables as (generalized logit model)[1]

$$\log\left(\frac{\Pr(Y=k\mid\mathbf{X})}{\Pr(Y=K\mid\mathbf{X})}\right)=\beta_{0k}+\beta_{1k}X_1+\beta_{2k}X_2+\cdots+\beta_{pk}X_p$$

(1)

under the condition that $\Pr(Y=1|\mathbf{X})$ + $\Pr(Y=2|\mathbf{X})$ + ⋯ + $\Pr(Y=K|\mathbf{X})$=1. A predictor variable can be either continuous or categorical. For ($K-1$) binomial comparisons with the reference category $K$, we have a vector of ($K-1$) regression (slopes) parameters $\boldsymbol{\beta}_j$=($\beta_{j1}$, $\beta_{j2}$, …, $\beta_{j,K-1}$)′ associated with the variable $X_j$ ($j$=1, 2, …, $p$). The $\beta_{jk}$ is interpreted as the increase in log-odds ratio (OR) of the response being in category $k$ compared with the reference category $K$ for a one unit

increase in the $j$th predictor variable $X_j$, for constant values of all other variables $X_{j'}$'s ($j' \neq j \neq 1, 2, \ldots, p$) and $\exp(\beta_{jk})$ is the corresponding OR for $X_j$. For $(K-1)$ logit models, there is a total of $(K-1)p$ regression slopes for $X_1, X_2, \ldots, X_p$ plus $(K-1)$ intercepts to be estimated. The parameters are usually estimated by the method of maximum likelihood. This method maximizes the likelihood of reproducing the data from the estimated parameters.[5]

Once the model is fitted, the conditional probability of the $i$th ($i=1, 2, \ldots, N$) subject's response in the category $k$ ($k=1, 2, \ldots, K-1$) and baseline category $K$ given his or her set of covariate values $\mathbf{x}_i = (x_{1i}, x_{2i}, \ldots, x_{pi})'$ can be estimated as[1]

$$\hat{p}_{ik}(\mathbf{x}_i) = \hat{\Pr}(Y_i = k \mid \mathbf{x}_i)$$

$$= \frac{\exp(\hat{\beta}_{0k} + \hat{\beta}_{1k}x_{1i} + \cdots + \hat{\beta}_{pk}x_{pi})}{1 + \sum_{k=1}^{K-1} \exp(\hat{\beta}_{0k} + \hat{\beta}_{1k}x_{1i} + \cdots + \hat{\beta}_{pk}x_{pi})}$$

$$\hat{p}_{iK}(\mathbf{x}_i) = \hat{\Pr}(Y_i = K \mid \mathbf{x}_i)$$

$$= \frac{1}{1 + \sum_{k=1}^{K-1} \exp(\hat{\beta}_{0k} + \hat{\beta}_{1k}x_{1i} + \cdots + \hat{\beta}_{pk}x_{pi})} \quad (2)$$

such that $\hat{p}_{ik}(\mathbf{x}_i) \geq 0$ for all $k=1, 2, \ldots, K$ and $\sum_{k=1}^{K} \hat{p}_{ik}(\mathbf{x}_i) = 1$.

When the response $Y$ has more than two categories, a separate binomial logistic model can also be fitted for response category "$k$" versus the reference category "$K$", where the model is the same as Equation 1 for the same set of predictors $X_1, X_2, \ldots, X_p$. However, such a binomial model assumes that "$k$" and "$K$" are the only two possible response categories. That is, it models under the condition $\Pr(Y=k|\mathbf{X}) + \Pr(Y=K|\mathbf{X})=1$, assuming the probability of any other category of the multinomial response to be zero.

## Reanalysis of a GWG study data

Data were collected from a self-administered cross-sectional survey of pregnant women from obstetrics and midwifery clinics in Hamilton and Brantford, ON, Canada, who were attending for prenatal care during May to June 2012. Women were eligible for participation if they already had at least one prenatal visit, could read English sufficiently well to complete the survey, and had a live, singleton pregnancy. Three hundred and twelve eligible women completed the survey questionnaires at the median gestational age of 32 weeks with an interquartile range of 11.3 weeks (first quartile 25.3 weeks, third quartile 36.6 weeks) providing information on the outcome and exposures of interest. The primary outcome variable was the actual weight gained at the time

of survey completion, calculated as the difference between self-reported recent weight at the time of survey completion and the self-reported prepregnancy weight. Information on several sociodemographic, pregnancy care, psychological, and lifestyle factors that potentially affect or are associated with weight gain were collected. These included maternal age, height, prepregnancy weight, first-time mother, nurse or obstetrician as the primary care health care professional, amount of weight planned to gain oneself or recommended by a health care professional, neuroticism score measuring level of emotional instability, self-esteem, self-efficacy towards achieving healthy weight, self-efficacy towards controlling food intake, self-efficacy towards regular exercise, watching television at bed before sleeping at least some nights in a week, taking soda daily, and others.

The Canadian clinical guidelines,[14] which are based on the recommendation from the US Institute of Medicine (IOM),[15] recommend that underweight women (body mass index [BMI] $< 18.5$ kg/m²) gain 0.44–0.58 kg/week, normal weight women (BMI 18.5–24.9 kg/m²) gain 0.35–0.50 kg/week, overweight women (BMI 25.0–29.9 kg/m²) gain 0.23–0.33 kg/week, and obese women (BMI $> 30$ kg/m²) gain 0.17–0.27 kg/week on average in the second or third trimester for healthy pregnancy. The average GWG per week for a woman during her second or third trimester is the standardized GWG because it accounts for the differences in gestational ages between mothers. Therefore, we calculated this amount for each woman by dividing her weight gain in the second or third trimester (total weight gain since conceiving the baby in 1 kg –2 kg) by her number of gestational weeks in the second or third trimesters (gestational age in weeks –13), where 2 kg was subtracted from total weight gain because it is the amount expected GWG to be gained by a pregnant mother during her first trimester (first 13 weeks). For each woman, the response variable was created as three categories of inadequate, appropriate, or excessive GWG if her average GWG per week was below, within, or above, respectively, the range of GWG recommended for her BMI category by the guidelines. The details of the study design and list of outcomes and potential exposure variables and their categorization schemes can be found in McDonald et al.[13]

## Statistical analysis

We had three possible GWG response categories ($K=3$) of inadequate ($I$), appropriate ($A$), and excess ($E$) weight gain. We used the multinomial logistic regression model to simultaneously fit two binomial logistic regression models for the logit of excess versus appropriate ($E$ versus $A$), and

logit of inadequate versus appropriate (*I* versus *A*) against the covariates, simultaneously, as:

$$\log\left(\frac{\Pr(GWG = E)}{\Pr(GWG = A)}\right) = \beta_{0E} + \beta_{1E}X_1 + \beta_{2E}X_2 + \cdots + \beta_{pE}X_p$$

$$\log\left(\frac{\Pr(GWG = I)}{\Pr(GWG = A)}\right) = \beta_{0I} + \beta_{1I}X_1 + \beta_{2I}X_2 + \cdots + \beta_{pI}X_p$$

(3)

The parameter estimates were obtained by the method of maximum likelihood.

In the univariate analysis (*p*=1), we assessed the association of each of the *J* potential predictor variables with the trinomial outcome fitting multinomial logistic regression model for one variable at a time. The joint association of the *j*th variable $X_j$($j$=1, 2, …, *J*) with the trinomial outcome was assessed using a joint *P*-value corresponding to the test of null hypothesis that $X_j$ was not associated with any of binomial outcomes *E* versus *A* or *I* versus *A*, versus the alternative hypothesis that $X_j$ was associated with at least one of these binomial outcomes (ie, $H_0$:$\boldsymbol{\beta}_j$=($\beta_{jE}$, $\beta_{jI}$)′=(0,0)′ versus $H_1$: at least one of $\beta_{jE}$ or $\beta_{jI}$ was different from zero). We used Wald's chi-square test for the joint association testing with *K*-category response (*K*=3 in our problem), where the corresponding test statistic has the $\chi^2_{(K-1)}$ distribution for testing the association of a continuous exposure variable, and $\chi^2_{(K-1)(q-1)}$ distribution for testing the association of a *q*-category exposure variable, under $H_0$. We considered a variable for the multivariable multinomial regression if its joint *P*-value was ≤0.10 in the above test. However, if the estimate of correlation coefficient (Pearson correlation for continuous variables and polychoric correlation for binary or ordinal variables) between any two variables satisfying the criteria of univariate analysis was ≥0.70 we considered only one with better biological plausibility. Also, we did not consider any such significant variable to be a candidate variable for multivariable regression if the variable had more than 10% missing data. This would prevent us from losing a lot of observations in multivariable analysis since we performed the analysis with a complete case scenario (ie, we did not attempt to impute for missing data).

We first included these variables in the multivariable multinomial regression model and manually eliminated one variable at a time that had the largest joint *P*-value in the partial test of association obtained using the same test statistic used in the univariate case above until all variables in the model had joint *P*-values ≤0.10 with the multinomial outcome, and an individual association *P*-value ≤0.05 with

at least one binomial outcome *E* versus *A* or *I* versus *A*. Here, an individual association *P*-value for $X_j$ with a single binomial outcome, say, *E* versus *A*, was obtained for the test of the hypothesis $H_0$:$\beta_{jE}$=0 versus $H_1$:$\beta_{jE}$≠0 using Wald test, where the test statistic $\boldsymbol{T} = (\hat{\beta}_{jE}/s_{jE})^2 \sim \chi^2_1$ (asymptotically), and $\hat{\beta}_{jE}$ is the estimate of $\beta_{jE}$ and $s_{jE}$ is the corresponding standard error.

Once the final multinomial model of *p* included variables $X_1$, $X_2$, …, $X_p$ was identified, we assessed goodness of the overall fit of the model using the deviance statistic, $D \sim \chi^2_{N-p}$ (asymptotically). For each of $X_j$ included in the final model for each of the binomial outcomes *E* versus *A* and *I* versus *A*, we obtained and reported the estimate of OR $[\exp(\hat{\beta}_{jE})]$, and corresponding 95% confidence interval (CI, that is based on the Wald statistic) and the corresponding *P*-value. For any woman with a set of predictor values $\mathbf{x}$=($x_1$, $x_2$, …, $x_p$)′, the predicted probabilities of gaining excessive, inadequate, or appropriate weight were predicted from the fitted multinomial model as:

$$\Pr\{GWG = \boldsymbol{E} \mid \mathbf{x}\}$$
$$= \frac{\exp(\hat{\beta}_{0E} + \hat{\beta}_{1E}x_1 + \cdots + \hat{\beta}_{pE}x_p)}{1 + \exp(\hat{\beta}_{0E} + \hat{\beta}_{1E}x_1 + \cdots + \hat{\beta}_{pE}x_p) + \exp(\hat{\beta}_{0I} + \hat{\beta}_{1I}x_1 + \cdots + \hat{\beta}_{pI}x_p)},$$
$$\Pr\{GWG = \boldsymbol{I} \mid \mathbf{x}\}$$
$$= \frac{\exp(\hat{\beta}_{0I} + \hat{\beta}_{1I}x_1 + \cdots + \hat{\beta}_{pI}x_p)}{1 + \exp(\hat{\beta}_{0E} + \hat{\beta}_{1E}x_1 + \cdots + \hat{\beta}_{pE}x_p) + \exp(\hat{\beta}_{0I} + \hat{\beta}_{1I}x_1 + \cdots + \hat{\beta}_{pI}x_p)},$$
$$\Pr\{GWG = \boldsymbol{A} \mid \mathbf{x}\}$$
$$= \frac{1}{1 + \exp(\hat{\beta}_{0E} + \hat{\beta}_{1E}x_1 + \cdots + \hat{\beta}_{pE}x_p) + \exp(\hat{\beta}_{0I} + \hat{\beta}_{1I}x_1 + \cdots + \hat{\beta}_{pI}x_p)}$$

(4)

We analyzed the data using SAS version 9.2 (SAS Institute, Cary, NC, USA).

## Results

As previously reported,[13] response data were available from 312 women, of whom 68(21.8%), 54(17.3%), and 190 (60.9%) inadequately, appropriately, or excessively gained weight, respectively. Descriptive statistics, count (%) for categorical and mean (standard deviation) for continuous exposure variables of interest in each of the three GWG response categories and their joint association *P*-values with trinomial GWG responses obtained using univariate multinomial approach are displayed in Table 1. Among the significant variables in univariate analysis, "self-efficacy towards achieving healthy weight" was highly correlated with "self-efficacy towards controlling food intake" (*r*=0.78), and with "self-efficacy towards regular

**Table 1** Descriptive statistics and associations of variables with actual weight gain during pregnancy in univariate analysis using multinomial logistic regression

| Variables | Gestational weight gain categories | | | Joint P-value[e] |
|---|---|---|---|---|
| | Inadequate (n=68) | Appropriate (n=54) | Excessive (n=190) | |
| **Categorical variables, count (%)** | | | | |
| Race: white | 56 (21.1) | 43 (16.2) | 166 (62.6) | 0.303 |
| Current smoker | 4 (11.1) | 4 (11.1) | 28 (77.8) | 0.102 |
| First time giving birth | 32 (26.4) | 14 (11.6) | 75 (62) | 0.060 |
| Education: college or higher | 54 (20.7) | 46 (17.6) | 161 (61.7) | 0.565 |
| Household income ≥$20,000 | 53 (21.5) | 43 (17.5) | 150 (61.0) | 0.973 |
| Pre-BMI group[a] | | | | 0.054 |
|    Underweight | 3 (23.1) | 2 (15.4) | 8 (61.5) | |
|    Normal weight | 30 (17.8) | 38 (22.5) | 101 (59.8) | |
|    Overweight or obese | 35 (27.1) | 13 (10.1) | 81 (62.8) | |
| Chronic health problem[b] | 15 (20.5) | 9 (12.3) | 49 (67.1) | 0.366 |
| Referral to obstetrician | 20 (23.3) | 15 (17.4) | 51 (59.3) | 0.920 |
| Referral to dietician[d] | 3 (23.1) | 2 (15.4) | 8 (61.5) | 0.980 |
| Referral to nurse practitioner[d] | 0 (0) | 0 (0) | 4 (100) | 0.999 |
| Referral to midwife[d] | 1 (25) | 0 (0) | 3 (75) | 0.998 |
| HCP recommended gestational weight gain | 19 (21.3) | 15 (16.9) | 55 (61.8) | 0.983 |
| HCP recommended calories intake | 8 (17.8) | 7 (15.6) | 30 (66.7) | 0.733 |
| HCP recommended appropriate calories intake[d] | 0 (0) | 2 (25) | 6 (75) | 0.940 |
| HCP discussed nutrition and healthy eating | 41 (22.8) | 32 (17.8) | 107 (59.4) | 0.873 |
| HCP discussed appropriate weight gain | 31 (21.1) | 30 (20.4) | 86 (58.5) | 0.460 |
| HCP discussed risks of too much weight gain | 17 (17.2) | 18 (18.2) | 64 (64.6) | 0.421 |
| HCP discussed risks of too little weight gain | 12 (20) | 14 (23.3) | 34 (56.7) | 0.314 |
| HCP discussed exercise during pregnancy | 42 (22.5) | 36 (19.3) | 109 (58.3) | 0.391 |
| HCP discussed breastfeeding | 37 (23.3) | 29 (18.2) | 93 (58.5) | 0.699 |
| HCP discussed vitamins intake during pregnancy | 67 (22.6) | 50 (16.8) | 180 (60.6) | 0.447 |
| Planned weight gain | | | | <0.001 |
|    Below | 30 (51.7) | 12 (20.7) | 16 (27.6) | |
|    Within | 26 (23.0) | 32 (28.3) | 55 (48.7) | |
|    Excess | 12 (9.2) | 7 (5.3) | 112 (85.5) | |
| Prepregnancy weight considered just right or excessive | 65 (21.7) | 53 (17.7) | 181 (60.5) | 0.998 |
| Satisfied with prepregnancy weight | 38 (18.1) | 47 (22.4) | 125 (59.5) | 0.002 |
| Thinking exercise is important during pregnancy | 61 (21.3) | 52 (18.2) | 173 (60.5) | 0.999 |
| High view on self-esteem somewhat true or very true | 53 (20.3) | 48 (18.4) | 160 (61.3) | 0.035 |
| Spending at least 90 minutes at a screen daily | 56 (24.9) | 33 (14.7) | 136 (60.4) | 0.036 |
| No meals usually eaten in front of a screen[c] | 22 (20.2) | 22 (20.2) | 65 (59.6) | 0.595 |
| Exercise at least 30 minutes a day | 33 (19.3) | 33 (19.3) | 105 (61.4) | 0.315 |
| Sleeping at least 6 hours a day | 60 (22.2) | 50 (18.5) | 160 (59.3) | 0.346 |
| Having television in bedroom | 35 (23.8) | 22 (15) | 90 (61.2) | 0.448 |
| Bedtime TV watching | 34 (26.8) | 14 (11) | 79 (62.2) | 0.008 |
| Soda drinking (at least one glass a week) | 40 (21.4) | 26 (13.9) | 121 (64.7) | 0.099 |
| **Continuous variables, mean (SD)** | | | | |
| Maternal age, years | 28.1 (5.5) | 30.8 (4.6) | 30 (5.5) | 0.017 |
| Gestational age, weeks | 26.5 (8.9) | 31.8 (7.2) | 31.2 (6.6) | <0.001 |
| Prepregnancy BMI[a] | 27.7 (8.2) | 23.6 (4.9) | 24.8 (4.9) | 0.001 |
| Locus of control | 9.8 (2.7) | 9.5 (2.2) | 10.3 (2.9) | 0.108 |
| Attitude score | 17.8 (3.4) | 17.9 (2.8) | 17.3 (3.4) | 0.503 |
| Neurotic score | 4.9 (3.2) | 3.6 (2.6) | 5.4 (3.1) | 0.001 |
| Lie score | 6 (2.2) | 6.4 (2.6) | 6.2 (2.3) | 0.576 |
| Self-efficacy towards healthy weight | 25.5 (4.5) | 28 (6.1) | 27 (5.6) | 0.024 |
| Self-efficacy towards controlling food intake | 10.7 (2.4) | 11.9 (2.2) | 11.5 (2.2) | 0.006 |
| Self-efficacy towards regular exercise | 6.8 (1.8) | 7.4 (1.7) | 7.4 (1.7) | 0.060 |
| Self-efficacy towards weight management | 8 (2.3) | 8.8 (3.5) | 8.1 (3.1) | 0.287 |

**Notes:** [a]Tested as both categorical and continuous predictor; [b]any of chronic depression, anxiety, eating disorder, high blood pressure, diabetes, asthma; [c]none of the most commonly eaten snacks at breakfast, lunch, dinner; [d]the association testing may be inappropriate for these exposures variables as some observed (and expected) cell frequencies in individual outcome categories are 0s or too small with total exposure rate small, these variables were not considered as candidate variables for multivariable analysis; [e]joint P-value: P-value for association (using Wald's chi-square test) with at least one comparison of excess versus appropriate or inadequate versus appropriate weight gain in multinomial analysis.
**Abbreviations:** BMI, body mass index; HCP, health care professional; SD, standard deviation; TV, television.

exercise" (*r*=0.75). We choose "self-efficacy towards achieving healthy weight" as this variable is directly relevant to achieving a healthy weight (which could be achieved through controlling food intake and regular exercise in turn, and hence the latter two variables are more like surrogate variables). Finally, we tried the following list of variables as the candidate for multivariable multinomial regression model: psychological variables, namely "neuroticism" score, "self-efficacy towards achieving healthy weight", high view on self-esteem (categorized as somewhat true or higher versus not very true), and other sociodemographic and lifestyle factors, namely maternal age, first time giving birth, prepregnancy BMI category, planning to gain weight excessively or inadequately, satisfaction with prepregnancy weight, spending at least 90 minutes time on screen (eg, computer, television) in a day, watching television in bed before sleeping at least some nights in a week, and taking soda daily. After manual (backward) elimination of the least significant variables one at a time, the variables of neurotic scale, categorical planned weight gain, categorical prepregnancy BMI, and binary bedtime watching television were retained in the final multivariable multinomial model. The deviance chi-square *P*-value was 0.984 for the goodness-of-fit of the final multinomial model, indicating that our model fitted very well to the data.

The estimates of OR and their 95% CIs and *P*-values of associations for the final set of predictors with excessive and inadequate weight gain from the multinomial analysis and separate binomial analyses (copied from a previously reported paper[13] with permission from the publisher) are presented in Table 2. In multinomial analysis, the predictors of excessive weight gain were: the neurotic score per unit increase (OR 1.24, 95% CI 1.10, 1.41; *P*-value =0.001), excessively planning weight gain compared with appropriately planning (OR 12.47, 95% CI 4.57, 34.03; *P*-value <0.001), and watching television compared with not watching television (OR 2.38, 95% CI 1.11, 5.11; *P*-value =0.027). The previous binomial analysis of excess versus appropriate weight gain also identified the same set of significant variables: neurotic score (OR 1.26, 95% CI 1.10, 1.44; *P*-value =0.001), planning excessive weight gain (OR 11.18, 95% CI 4.45, 28.06; *P*-value <0.001) and watching television in bed (OR 2.38, 95% CI 1.08, 5.23; *P*-value =0.031) in the adjusted analysis.

In multinomial analysis, inadequately planning weight gain compared with appropriately planning (OR 2.87, 95% CI 1.13, 7.26; *P*-value =0.026), overweight or obese woman compared with normal weight woman (OR 4.31, 95% CI 1.69, 11.00; *P*-value =0.002) and watching television before going to bed (OR 2.50, 95% CI 1.08, 5.79; *P*-value =0.033) were associated with inadequate weight gain, while satisfaction with their prepregnancy weight (OR 4.84, 95% CI 1.56, 15.02; *P*-value =0.006), watching television at bed (OR 3.92, 95% CI 1.50, 10.30; *P*-value =0.006), and having self-efficacy towards achieving a healthy weight (OR 0.91, 95% CI 0.83,

**Table 2** Odds ratios and associated *P*-values from the final multinomial and separate binomial logistic regression models

| Predictors | Comparison | Analysis approach[a] | | | |
|---|---|---|---|---|---|
| | | **Multinomial model[b]** | | **Separate binomial models[c]** | |
| | | **OR (95% CI)** | ***P*-value** | **OR (95% CI)** | ***P*-value** |
| **For excessive weight gain** | | | | | |
| Neuroticism | Unit increase in score | 1.24 (1.10, 1.41) | 0.001 | 1.26 (1.10, 1.44) | 0.001 |
| Planned weight gain[d] | Excess versus appropriate | 12.47 (4.57, 34.03) | <0.001 | 11.18 (4.45, 28.06) | <0.001 |
| | Inadequate versus appropriate | 0.76 (0.30, 1.93) | 0.569 | 0.69 (0.26 to 1.80) | 0.449 |
| Bedtime TV watching | Yes versus no | 2.38 (1.11, 5.11) | 0.027 | 2.38 (1.08, 5.23) | 0.031 |
| **For inadequate weight gain** | | | | | |
| Planned weight gain[d] | Excess versus appropriate | 1.16 (0.36, 3.71) | 0.809 | – | – |
| | Inadequate versus appropriate | 2.87 (1.13, 7.26) | 0.026 | – | – |
| Prepregnancy BMI[d] | Overweight or obese versus normal weight | 4.31 (1.69, 11.00) | 0.002 | – | – |
| | Underweight or obese versus normal weight | 1.78 (0.26, 12.27) | 0.559 | – | – |
| Bedtime TV watching | Yes versus no | 2.50 (1.08, 5.79) | 0.033 | 3.92 (1.50, 10.30) | 0.006 |
| Self-efficacy towards achieving healthy weights | Unit increase in score | – | – | 0.91 (0.83, 0.99) | 0.033 |
| Satisfaction with prepregnancy weight | Yes versus no | – | – | 4.84 (1.56, 15.02) | 0.006 |

**Notes:** [a]Final models were identified by variable selection approach; [b]multinomial logistic regression fits two binomial models of excessive versus appropriate weight gain and inadequate versus appropriate weight gain simultaneously; [c]Copyright © 2013. Society of Obstetricians and Gynaecologists of Canada. Data reproduced from McDonald SD, Park CK, Timm V, Schmidt L, Neupane B, Beyene J. What psychological, physical, lifestyle, and knowledge factors are assowith excess or inadequate weight gain during pregnancy? A cross-sectional survey. *J Obstet Gynaecol Can.* 2013;35(12):1071–1082.[13] [d]dummy variables.
**Abbreviations:** OR, odds ratio; CI, confidence interval; BMI, body mass index; TV, television.

0.99; $P$-value =0.033) were associated with inadequate weight gain in separate binomial analysis. Thus, not only the set of predictors differed, but also the estimated effect sizes and $P$-values of the common predictors changed from the previous analysis for inadequate weight gain. Overall, the results were very similar for the excessive weight gain analysis but differed for the inadequate weight gain analysis.

The estimates of the parameters (beta coefficients) in Equation 4 for the two binomial models of excessive versus appropriate and inadequate versus appropriate weight gain fitted simultaneously using multinomial logistic regression are provided in Table 3. Estimated risks (predicted probabilities) of gaining excessive, inadequate, and appropriate weight for some combinations of predictor values, which are estimated using Equation 4, are presented in Table 4 with a median score of 5 for neuroticism for illustrative purposes. These risks are also plotted in Figure 1 for a different set of predictors. Figure 1 shows that if the final multinomial model we identified is the true model, then normal weight women who plan to gain more weight than that recommended by the IOM are at very high risk of gaining excessive weight, and their risk is even higher if they have a habit of watching television in bed before sleeping. Table 4 shows that the probabilities of actually gaining weight more than, within, and less than the recommended range are estimated to be, for instance, 0.49, 0.36, and 0.14, respectively, for a woman of normal weight who scores 5 on a scale measuring neuroticism, who does not watch television in bed, and who plans to gain weight that is within the range recommended by the IOM guidelines; however, these probabilities are estimated to be 0.92, 0.05, and 0.02, respectively, if a similar woman plans to gain weight that is above the recommended range.

## Discussion

We used a multinomial logistic regression approach to assess the factors associated with the trinomial outcome of excessive, appropriate, and inadequate GWG. In the analysis

of excessive versus appropriate weight gain, not only were the conclusions about the significance of the predictors the same, but the estimates of effect sizes and $P$-values were also very similar in the multinomial analysis we performed here and in separate binomial analyses reported earlier,[13] although the multinomial approach produced a smaller $P$-value (or narrower CI interval) for one predictor. However, the set of predictors differed, and the estimated effect sizes and $P$-values for the common predictors also changed from the previous analysis for inadequate weight gain, where the subgroup size for separate binomial analysis was small and we derived the final set of predictors by variable selection.

The smaller $P$-value for a predictor from multinomial analysis for the outcome of excessive weight gain may be due to using data from inadequate weight gain simultaneously for multinomial analysis, whereas only a subgroup of data corresponding to excessive and appropriate weight gain categories were used for binomial analysis. It could also be due to using some data points in one analysis but not in other analysis, as we used a variable selection process to identify the final model under both the multinomial and separate binomial approach.

For analysis of inadequate weight gain, there may be several reasons why there were differences in the final sets of predictors identified by the two approaches. One explanation might be that the event rates were quite small and the sample size was also small (a total of 122 women gained either inadequate or appropriate weight) for this analysis. Unstable results in logistic regression analyses can be obtained in small-sized studies.[16,17] Next, the maximum likelihood estimates of log-ORs in both analyses are biased in a small-to-moderate-sized study.[18] In particular, complete separation of binary or multinomial cases in a profile or cell of multiple predictors is likely, and that can result in a too strong or unrealistic maximum likelihood estimate of the regression parameter.[19] The maximum likelihood estimates are only asymptotically normal around a true mean,

**Table 3** Selected variables in the final multinomial logistic regression model and corresponding parameter estimates

| Variables | Variables notation | Possible values | Estimated parameters for | |
|---|---|---|---|---|
| | | | Excess gain | Inadequate gain |
| Intercept | – | – | $\hat{\beta}_{0E}$=−0.787 | $\hat{\beta}_{0I}$=−1.554 |
| Neuroticism | $X_1$ (continuous) | $X_1$=0 to 12 score | $\hat{\beta}_{1E}$=0.219 | $\hat{\beta}_{1I}$=0.125 |
| Planned GWG | $X_2$ (dummy) | $X_2$=1 if excess, 0=else | $\hat{\beta}_{2E}$=2.523 | $\hat{\beta}_{2I}$=0.144 |
| category | $X_3$ (dummy) | $X_3$=1 if inadequate, 0=else | $\hat{\beta}_{3E}$=−0.268 | $\hat{\beta}_{3I}$=1.053 |
| Prepregnancy | $X_4$ (dummy) | $X_4$=1 if overweight or obese, 0=else | $\hat{\beta}_{4E}$=−0.280 | $\hat{\beta}_{4I}$=1.461 |
| BMI category | $X_5$ (dummy) | $X_5$=1 if underweight, 0=else | $\hat{\beta}_{5E}$=0.346 | $\hat{\beta}_{5I}$=0.575 |
| Bedtime TV watching | $X_{6=p}$ (binary) | $X_6$=1 if yes, 0=no | $\hat{\beta}_{6E}$=0.866 | $\hat{\beta}_{6I}$=0.915 |

**Abbreviations:** E, excess; I, inadequate; GWG, gestational weight gain; BMI, body mass index; TV, television.

**Table 4** Predicted probabilities of gaining excessive, appropriate, and inadequate gestational weight at a median value of neurotic score $(X_1)=5$

| Predictors | | | Predicted probabilities of weight gain[a] | | |
|---|---|---|---|---|---|
| Planned GWG category | Prepregnancy BMI category | Bedtime TV watching | Excess | Appropriate | Inadequate |
| Inadequate | Underweight | No | 0.33 | 0.22 | 0.45 |
| Inadequate | Underweight | Yes | 0.37 | 0.11 | 0.53 |
| Inadequate | Normal weight | No | 0.33 | 0.32 | 0.36 |
| Inadequate | Normal weight | Yes | 0.39 | 0.16 | 0.45 |
| Inadequate | Overweight or obese | No | 0.12 | 0.15 | 0.73 |
| Inadequate | Overweight or obese | Yes | 0.12 | 0.07 | 0.81 |
| Appropriate | Underweight | No | 0.53 | 0.28 | 0.19 |
| Appropriate | Underweight | Yes | 0.62 | 0.14 | 0.24 |
| Appropriate | Normal weight | No | 0.49 | 0.36 | 0.14 |
| Appropriate | Normal weight | Yes | 0.62 | 0.19 | 0.19 |
| Appropriate | Overweight or obese | No | 0.28 | 0.27 | 0.46 |
| Appropriate | Overweight or obese | Yes | 0.32 | 0.13 | 0.55 |
| Excess | Underweight | No | 0.93 | 0.04 | 0.03 |
| Excess | Underweight | Yes | 0.95 | 0.02 | 0.03 |
| Excess | Normal weight | No | 0.92 | 0.05 | 0.02 |
| Excess | Normal weight | Yes | 0.95 | 0.02 | 0.03 |
| Excess | Overweight or obese | No | 0.81 | 0.06 | 0.12 |
| Excess | Overweight or obese | Yes | 0.84 | 0.03 | 0.13 |

**Notes:** [a]Indicates predicted probabilities without accompanying confidence intervals. However, the predicted probabilities are just the estimates. Hence, there are certain levels of uncertainty in these, so they should not be viewed as fixed values.
**Abbreviations:** GWG, gestational weight gain; BMI, body mass index; TV, television.

and the $P$-values and CIs based on Wald's statistic can be poorly calculated when the estimates are far from zero,[20] and our sample sizes for inadequate versus appropriate weight gain analysis might not be large enough to hold asymptotic results. For analysis of excessive weight gain, for which the sample size was moderate (a total of 244 women gained either excessive or adequate gestational weight), both analyses produced consistent results. However, it is not well understood how multinomial and binomial approaches perform in terms of type I and type II error rates in small-to-moderate-sized studies involving complex survey data like ours, where the variable selection process took place prior to final model parameter estimations. In particular, when variable selection processes lead to different sets of predictors in the binomial and multinomial models, as in the case of inadequate versus appropriate analyses in our data, the resulting estimate, CI, or $P$-value even for the predictor selected in both analyses may not be comparable because the effects of different sets of predictors are adjusted in the two different models.

Given that our aim was to explore the contrast in the variables' significance and corresponding estimates from multinomial and separate binomial approaches, and we had already used the OR as an effect measure for the separate binomial approach in our previous publication, we used the OR as the effect measure in this study as well, while implementing a multinomial approach. However, for modeling the probability of common outcomes (>10% event rates), such as excessive or inadequate GWG like in our study, the log-binomial regression model fitting would have been more suitable than separate binomial logistic or multinomial logistic regression model fitting. However, this procedure would lead to estimation of relative risk as the effect measure as opposed to an odds ratio.[21] The log-binomial regression model produces less biased and more robust/stable estimates of relative risk for common outcomes than does the multinomial or binomial logistic regression model. It is also important to note that the final multinomial model we developed was fitted on a population/aggregate level and has not been well validated; therefore, it may not be clinically useful for predicting the risk of individuals gaining too much or too little weight during pregnancy.

In conclusion, the multinomial modeling framework enables us to simultaneously test for association of categorical and continuous independent variables with different nominal categories. It also enables us to estimate the risk associated with each of the multiple categories (ie, obtain the predicted probabilities of all response categories) given a woman's set of characteristics or behaviors from a carefully developed and well-validated model. However, comprehensive simulation studies are needed to fully understand the operating characteristics of these methods.
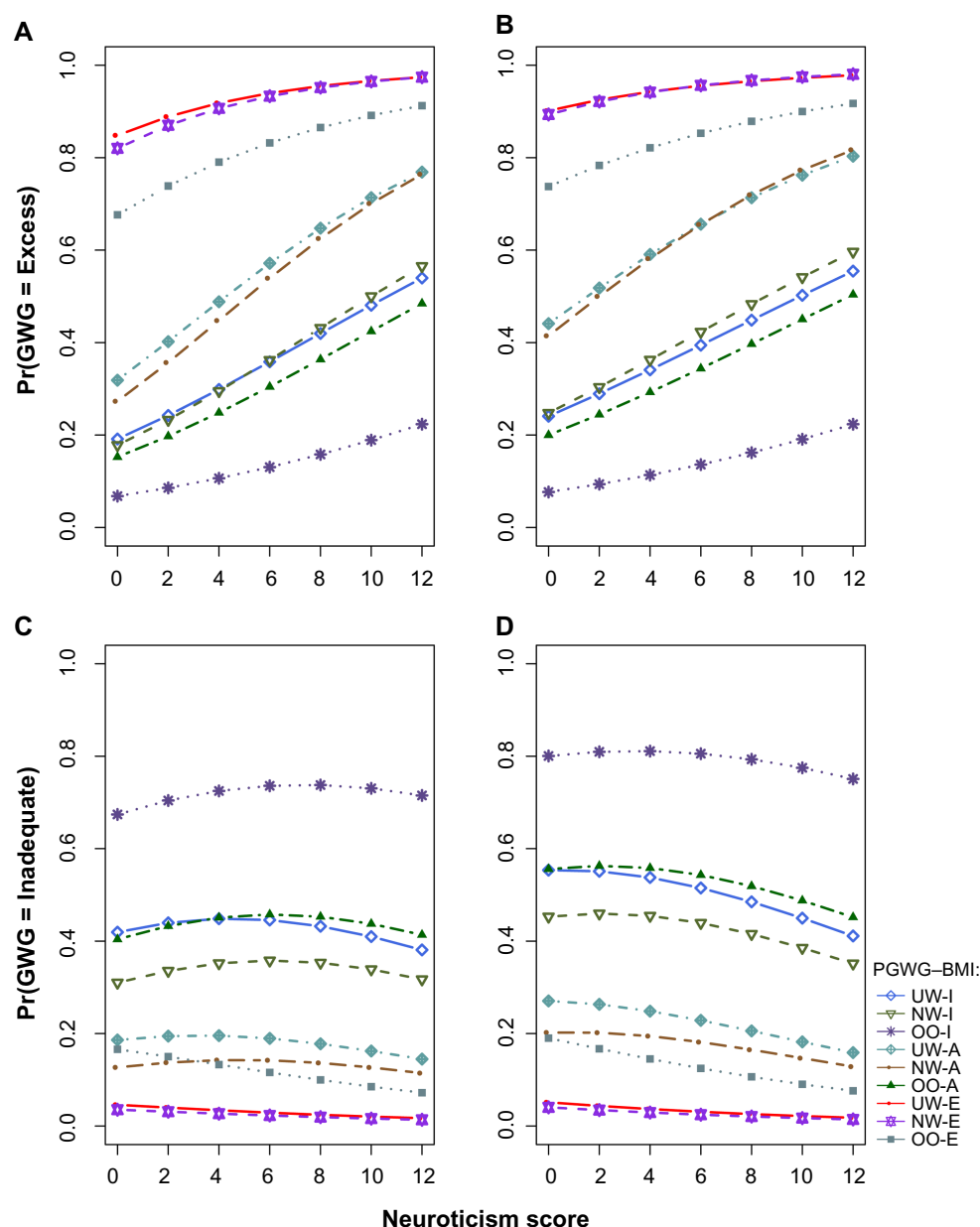
**Figure 1** Plots of predicted probabilities of excessive and inadequate weight gain against neurotic score for different combinations of predictors. (**A** and **B**) Predicted probabilities of excessive weight gain for women who do not watch television and who do, respectively, before bed. (**C** and **D**) Predicted probabilities of inadequate weight gain for women who do not watch television and who do, respectively, before bed. In each plot, there are nine lines plotted for nine combinations of body mass index (BMI) and planned gestational weight gain (PGWG) categories: UW-I, underweight women who planned inadequately; NW-I, normal weight women planned inadequately; OO-I, overweight or obese women who planned inadequately; UW-A, underweight women planned appropriately; NW-A, normal weight women who planned appropriately; OO-A, overweight or obese women who planned appropriately; UW-E, underweight women who planned excessively; NW-E, normal weight women who planned excessively; OO-E, overweight or obese women who planned excessively.
**Abbreviations:** GWG, gestational weight gain; Pr, probability.

## Acknowledgements

## Disclosure

The authors report no conflicts of interest in this work.

## References

1. Agresti A. *Categorical Data Analysis*. 2nd ed. New York, NY, USA: Wiley-Interscience; 2002.

2. Hosmer DW, Lemeshow S. *Applied Logistic Regression*. 2nd ed. New York, NY, USA: Wiley; 2000.

3. Morgan SP, Teachman JD. Logistic-regression – description, examples, and comparisons. *J Marriage Fam*. 1988;50(4):929–936.

4. Peng C-YJ, Nichols RN. Using multinomial logistic models to predict adolescent behavioral risk. *J Mod Appl Stat Methods*. 2003;2(1):16.

5. Peng C-YJ, Lee KL, Ingersoll GM. An introduction to logistic regression analysis and reporting. *J Educ Res*. 2002;96(1):3–14.

6. Chasan-Taber L, Schmidt MD, Pekow P, Sternfeld B, Solomon CG, Markenson G. Predictors of excessive and inadequate gestational weight gain in Hispanic women. *Obesity*. 2008;16(7):1657–1666.

7. Kowal C, Kuk J, Tamim H. Characteristics of weight gain in pregnancy among Canadian women. *Matern Child Health J*. 2012;16(3): 668–676.

8. Walker LO, Hoke MM, Brown A. Risk factors for excessive or inadequate gestational weight gain among Hispanic women in a US-Mexico border state. *J Obstet Gynecol Neonatal Nurs*. 2009;38(4):418–429.

9. Weisman CS, Hillemeier MM, Downs DS, Chuang CH, Dyer AM. Preconception predictors of weight gain during pregnancy: prospective findings from the Central Pennsylvania Women's Health Study. *Womens Health Issues*. 2010;20(2):126–132.

10. Herring SJ, Nelson DB, Davey A, et al. Determinants of excessive gestational weight gain in urban, low-income women. *Womens Health Issues*. 2012;22(5):e439–e446.

11. Brawarsky P, Stotland NE, Jackson RA, et al. Pre-pregnancy and pregnancy-related factors and the risk of excessive or inadequate gestational weight gain. *Int J Gynaecol Obstet*. 2005;91(2):125–131.

12. Mehta UJ, Siega-Riz AM, Herring AH. Effect of body image on pregnancy weight gain. *Matern Child Health J*. 2011;15(3):324–332.

13. McDonald SD, Park CK, Timm V, Schmidt L, Neupane B, Beyene J. What psychological, physical, lifestyle, and knowledge factors are associated with excess or inadequate weight gain during pregnancy? A cross-sectional survey. *J Obstet Gynaecol Can*. 2013;35(12): 1071–1082.

14. Health Canada. Prenatal nutrition guidelines for health professionals: gestational weight gain. Available from: http://www.hc-sc.gc.ca/fn-an/nutrition/prenatal/ewba-mbsa-eng.php. Accessed August 2, 2013.

15. Rasmussen KM, Yaktine AL. *Weight Gain during Pregnancy: Reexamining the Guidelines*. Washington, DC, USA: National Academies Press; 2009.

16. Peduzzi P, Concato J, Kemper E, Holford TR, Feinstein AR. A simulation study of the number of events per variable in logistic regression analysis. *J Clin Epidemiol*. 1996;49(12):1373–1379.

17. Nemes S, Jonasson JM, Genell A, Steineck G. Bias in odds ratios by logistic regression modelling and sample size. *BMC Med Res Methodol*. 2009;9:56.

18. Bull SB, Lewinger JP, Lee SSF. Confidence intervals for multinomial logistic regression in sparse data. *Stat Med*. 2007;26(4):903–918.

19. Albert A, Anderson JA. On the existence of maximum-likelihood estimates in logistic-regression models. *Biometrika*. 1984;71(1): 1–10.

20. Hauck WW Jr, Donner A. Wald's test as applied to hypotheses in logit analysis. *J Am Stat Assoc*. 1977;72(360a):851–853.

21. Camey SA, Torman VB, Hirakata VN, Cortes RX, Vigo A. [Bias of using odds ratio estimates in multinomial logistic regressions to estimate relative risk or prevalence ratio and alternatives]. *Cad Saude Publica*. 2014;30(1):21–29. Portuguese.