International Journal of Health Sciences and Research
www.ijhsr.org ISSN: 2249-9571

Original Research Article

An Application of Multinomial Logistic Regression to Assess the Factors Affecting the Women to Be Underweight and Overweight: A Practical Approach

Shivam Dixit[1], Bhushan Kumar[2], Abhay Singh[3], Ramkumar Ashoka[4]

[1]Statistician cum Assistant professor, [2]Professor,
Dept. of Community Medicine, Gold Field Institute of Medical Science & Research, Chhainsa, Ballabgarh, Faridabad, Haryana, India
[3]Assistant Professor, Dept. of Community Medicine Vardhman Mahavir Medical College, New Delhi.
[4]Professor & HOD, Dept. of Anatomy, Gold Field Institute of Medical Sciences & Research Chhainsa, Ballabgarh, Faridabad, Haryana, India

Corresponding Author: Bhushan Kumar

Received: 08/09/2015 *Revised: 25/09/2015* *Accepted: 28/09/2015*

ABSTRACT

Objective: To evaluate the application of Multinomial Logistic Regression to assess the factors affecting the women to be underweight and overweight.
Methods: This was a cross-sectional study. The socio-demographic detail with height and weight of the women was recorded. The simple random sampling was adopted in the selection of women. The pregnant women were excluded from the study. Women's weight status, indicated by their BMI category, was used as the outcome variable in the analyses. A total of 435 women were interviewed. To assess the net effect of exposure variables on the outcome measures, multinomial logistic regression analysis was contemplated to be suitable as the outcome measure is polychotomous by nature.
Results: Most of the women were between 20-30 years of age (44.4%). More than one third of the women had family income between Rs. 5001-10, 000 (57.7%) per month and were illiterates (68.5%). The results of multinomial logistic regression showed that overweight increased with age and education with higher prevalence among urban women.
Conclusion: The growing demand which now appears before the Government or urban health planners is to address this rising urban epidemic with equal importance as given to other issues in the past.

Key words: Multinomial regression, Body mass index, underweight, Overweight,

INTRODUCTION

In recent years, logistic regression has been applied extensively in numerous disciplines including medical and social sciences. In statistics, when the variables of interest have only two possible responses, these are represented as binary outcome. In a study of obesity, selected subjects have either high (>30 kg/m^2) body mass index (BMI) or normal BMI (<30 kg/m^2). With independent variables such as age and gender, the response variable Y is defined to have two possible outcomes namely subjects having a high BMI or not having a high

BMI. Subsequently, these are coded as 1 and 0 respectively. The extension of the binary logistic regression model is multinomial logistic regression model, in which the response variable has more than two levels. In the study of obesity, BMI id divided into three different levels (underweight, normal and overweight), the multinomial logistic regression model is build with age and gender as covariates. These are labeled as 1, 2 and 3 respectively and defined 1 as the reference category.

In multinomial logistic model, the associations of genetic or environmental factors with all disease subtypes are estimated simultaneously and inferences comparing associations across subtypes can be performed (Rosenblatt et al, 2004; Agalliu et al, 2009; Morris et al, 2010). Multinomial logistic models have also been employed in prospective cohort studies to compare the associations between exposure variables and occurrence of multiple disease subtypes (Potter et al, 1995; Cui et al, 2008; Kenfield et al, 2008).

A landmark review of studies on socioeconomic status and obesity supports the view that obesity in the developing world would be essentially a non-communicable disease of the socioeconomic elite (Mendez et al, 2005). The problem relating to body mass index (BMI) for both men and women should receive equal attention. However, concerns relating to women in developing countries deserve extra attention because of cultural and economic backdrops, which hinder the blanching between male and female counterparts. An augmented number of literatures assert that an increased BMI of women is independently associated with increasing risk of adverse obstetric and neonatal outcome (Heslehurst et al, 2008; Oteng-Ntim et al, 2013). The risks of overweight also include diabetes mellitus, increased risk of cardiovascular disease, cancer, hypertension, and other medical problems (Torloni et al, 2009; O'Brien et al, 2003; Nagai et al, 2012).

This study was aimed to evaluate the application of Mutinomial Logistic Regression to assess the factors affecting the women to be underweight and overweight.

MATERIALS AND METHODS

Description of data

This was a cross-sectional study conducted in the rural and urban health catchment area of a tertiary care hospital in Haryana state in India among the married women of reproductive age group. The study was approved by the ethical committee of the institute. The consent was taken from each woman before enrolling in the study. The socio-demographic details with height and weight of the women were recorded. Simple random sampling was adopted in the selection of women. Women having at least one living child were included in the study, however pregnant women were excluded.

Women's weight status, indicated by their BMI category, was used as the outcome variable in the analyses. BMI was calculated as weight in kg divided by height in meter squared. This measurement of BMI is generally considered an appropriate method for epidemiological studies where objective measurement is less feasible. Weight was measured using an electronic scale with a precision of 0.1 kg, and height was measured with an adjustable measuring-board designed for use in survey settings, which can provide accurate measurements to the nearest 0.1 cm. A total of 435 women were interviewed.

Statistical Analysis

To assess the net effect of exposure variables on the outcome measures, multinomial logistic regression analysis was contemplated to be suitable as the outcome measure is polychotomous by nature. The multinomial logistic model is the extension of the binary logistic regression model to outcome measure with j=1, 2, 3, …,k nominal outcomes. In its general form, the probability of vector i belonging to category j is given by the following formula (Long, 1997):

$$\Pr(y_i = j|x_i) = \frac{\exp(\chi_i \beta_k)}{\sum_{k=1}^{j} (\chi_i \beta_k)}$$

where χ_i is a vector containing the values of m covariates for person i, and β_k is a vector of m +1 parameters ($\beta_{0k}, \beta_{1k}, \beta_{2k}$ …, β_{mk}) for each k=1, 2, 3, …, j. To identify the parameters, it is common to choose one reference category and set the corresponding vector of parameters equal to a vector of zeroes. The BMI <18.50 was labeled as "underweight", ≥25 as "overweight" and 18.50-24.99 as "normal". The results of the multinomial logistic regression analysis have been shown by odds ratios (ORs) with 95% confidence interval (CI) for easy understanding. The level of significance was set at 5%. Data were analyzed using SPSS (version 16.0) (SPSS Inc., Chicago, IL, USA).

RESULTS

Table-1: Distribution of women across categories of BMI according to different predictors

Demographic variables	No. of women		Underweight		Normal		Overweight		p-value[1]
	No.	%	No.	%	No.	%	No.	%	
Age in years									
<19	78	17.9	7	9.0	57	73.1	14	17.9	0.001*
20-30	193	44.4	81	42.0	90	46.6	22	11.4	
>31	164	37.7	27	16.5	100	61.0	37	22.6	
Family income in Rs. Per month									
<5000	135	31.0	45	33.3	66	48.9	24	17.8	0.16
5001-10,000	251	57.7	59	23.5	153	61.0	39	15.5	
>10, 000	49	11.3	11	22.4	28	57.1	10	20.4	
Education									
Illiterate	298	68.5	83	27.9	170	57.0	45	15.1	0.31
Literate	137	31.5	32	23.4	77	56.2	28	20.4	
Occupation									
Service	127	29.2	37	29.1	44	34.6	46	36.2	0.0001*
Labor	164	37.7	58	35.4	92	56.1	14	8.5	
Self employed	144	33.1	20	13.9	111	77.1	13	9.0	
Family size									
<10	360	82.8	97	26.9	205	56.9	58	16.1	0.68
>=10	75	17.2	18	24.0	42	56.0	15	20.0	
No. of living children									
≤2	49	11.3	15	30.6	27	55.1	7	14.3	0.74
>2	386	88.7	100	25.9	220	57.0	66	17.1	
Place of residence									
Rural	213	49.0	60	28.2	123	57.7	30	14.1	0.30
Urban	222	51.0	55	24.8	124	55.9	43	19.4	
Total	435	100.0	115	26.4	247	56.8	73	16.8	

[1]Chi-square test, *Significant

Table-1 presents the distribution of women across categories of BMI according to different predictors. Most of the women were between 20-30 years of age (44.4%). More than one third of the women had family income between Rs. 5001-10, 000 (57.7%) per month and were illiterates (68.5%). More than one third of the women belonged to labor class. Majority of the women were having family size <10. Numbers of living children were above 2 among majority of the women. About half (51%) of the women belonged to urban area. The percentage of overweight was higher among the older women and the association was significant (p=0.001). There was no significant (p>0.05) association of BMI category with family income, education, family size, number of living children and place of residence.

Table-2: Adjusted odds ratios (95% CI) for the study variables and covariates from multinomial logistic model predicting underweight and overweight for the ever married women

Study variables	BMI category (Adjusted OR (95%CI), p-value)	
	Underweight vs Normal weight	Overweight vs Normal weight
Age in years		
<19	0.45 (0.17-1.14), 0.09	0.53 (0.24-1.72), 0.11
20-30	3.75 (2.15-6.54), 0.0001*	0.75 (0.39-1.44), 0.39
>31	Reference	Reference
Family income in Rs. Per month		
<5000	1.19 (0.50-2.83). 0.68	0.99 (0.38-2.56), 0.98
5001-10,000	1.06 (0.46-2.46), 0.88	0.78 (0.31-1.93), 0.59
>10, 000	Reference	Reference
Education		
Illiterate	1.10 (0.64-1.88), 0.73	1.01 (0.55-1.86), 0.96
Literate	Reference	Reference
Occupation		
Service	5.65 (2.78-11.46), 0.0001*	8.54 (4.06-17.94), 0.0001*
Labor	3.49 (1.81-6.75), 0.0001*	1.13 (0.48-2.67), 0.60
Self employed	Reference	Reference
Family size		
<10	1.30 (0.66-2.58), 0.44	0.82 (0.39-1.73), 0.61
>=10	Reference	Reference
No. of living children		
≤2	1.22 (0.57-2.59), 0.60	0.83 (0.31-2.19), 0.71
>2	Reference	Reference
Place of residence		
Rural	0.94 (0.57-1.56), 0.60	0.83 (0.31-2.19), 0.71
Urban	Reference	Reference

OR-Odds ratio, CI-Confidence Interval, *Significant

Table-2 shows the adjusted odds ratios (95% CI) for the study variables and covariates from multinomial logistic model predicting underweight and overweight for the ever married women. Prevalence of underweight was 3.75 times significantly higher among the women of age 20-30 years compared with women above 31 years (Adjusted OR=3.75, 95%CI=2.15-6.54, p=0.0001). Prevalence of underweight was lower among the women of age <19 years than >31 years, however it was statistically not significant (p>0.05). Prevalence of overweight was insignificantly (p>0.05) lower among the women <19 years and 20-30 years as compared to >31 years. There was no significant (p>0.05) association of BMI category with family income, education, family size, number of living children and place of residence. Prevalence of underweight was higher among the women who were in service (Adjusted OR=5.65, 95%CI=2.78-11.46, p=0.0001) and labor (Adjusted OR=3.49, 95%CI=1.81-6.75, p=0.0001) than self employed. Prevalence of overweight was also higher among the women of service class (Adjusted OR=8.54, 95%CI=4.06-17.94, p=0.0001) as compared to self employed.

Table-3: Model Fitting Information

Model	Model Fitting Criteria	Likelihood Ratio Tests		
	-2 Log Likelihood	Chi-Square	d.f.	p-value
Intercept Only	472.275			
Final	345.693	126.581	18	0.0001*

R^2=0.25, d.f: Degree of freedom, *Significant

The developed model was observed to be significant (p=0.0001) which explained 25% variability in the model (Table-3).

DISCUSSION

Multinomial logistic regression is used to predict categorical placement in or the probability of category membership on a dependent variable based on multiple independent variables. Independent variables can be either dichotomous (i.e., binary) or continuous (i.e., interval or ratio in scale). Multinomial logistic regression is a simple extension of binary logistic regression that allows for more than two categories of the dependent or outcome variable. Like binary logistic regression, multinomial logistic regression uses maximum likelihood estimation to evaluate the probability of categorical membership.

Multinomial logistic regression does necessitate careful consideration of the sample size and examination for outlying cases. Like other data analysis procedures, initial data analysis should be thorough and include careful univariate, bivariate, and multivariate assessment. Specifically, multicollinearity should be evaluated with simple correlations among the independent variables. Also, multivariate diagnostics (i.e. standard multiple regression) can be used to assess for multivariate outliers and for the exclusion of outliers or influential cases.

Sample size guidelines for multinomial logistic regression indicate a minimum of 10 cases per independent variable (Schwab, 2002).

Unconditional logistic regression (Breslow and Day, 1980) refers to the modeling of strata with the use of dummy variables (to express the strata) in a traditional logistic model. Here, one model is applied to all the cases and the strata are included in the model in the form of separate dummy variables, each reflecting the membership of cases to particular strata. Conditional logistic regression (Breslow and Day, 1980; Vittinghoff et al, 2005) refers to applying the logistic model to each of the strata individually. The coefficients of the predictors (of the logistic model) are conditionally modeled based on the membership of cases to particular strata. Marginal logistic modeling (Vittinghoff et al, 2005) refers to an aggregation of the strata so that the coefficients reflect the population values averaged across the strata.

Primary objective of this study was to develop a multinomial logistic model to predict the prevalence of underweight and overweight among the women of reproductive age groups, with special reference to their economic status. Study found that higher proportions of women in urban India are either overweight or obese than their counterparts from rural area. This generally contrasts with the findings of other studies conducted on the similar issues in Western and African countries where the percentage of urban and rural subjects are similar (Ziraba et al, 2009; Shafique et al, 2007). Nevertheless, a number of studies conducted in developing countries, especially in Asia, support the findings of this study (Ramachandran et al, 2004; Mendez et al, 2005).

In the mega cities, the situation is alarming. Many women are either overweight in the selected cities studied in India. This condition could well be compared with many other developed nations where the prevalence of overweight is accumulating steadily (World Bank, 2011; Gruebner et al, 2011).

Along with a number of studies, this study equally opined that fraction of overweight increases with age and education (Ziraba et al, 2009; Agrawal and Mishra, 2004). The multinomial analysis found that women aged <19 years are at low risk of overweight. Many studies have attempted to determine the causes behind this association between overweight and demographic covariates. Among all, physical activity declines along with metabolic rate, in the middle years of women. On the other hand, the energy requirement decreases; therefore, even regular or routine eating may lead to weight gain. In addition, the established cultural or social values with respect to care and diet given during and after pregnancy help women to gain more weight than ever. Furthermore, newly-married women at young age are more health-conscious and involved in more physical activity than women at older ages with children. This might be another important reason for weight gain after childbirth among women (Sinha and Kapoor, 2010).

CONCLUSION

The study found that the problem of overweight is more of an urban concern. The flagship programme National Health Rural Mission (NRHM), funded by central government has a number of building blocks or measures to address anaemia and undernutrition prevalent among women and children in rural India. Yet, the programme does not recognize the growing epidemic of overweight among women in urban India. With this backdrop, the growing demand which appears before the Government or the urban health planners is to address this rising epidemic with equal importance. A timely prevention will reduce the burden of many chronic co-morbidities, like diabetes, cardiovascular diseases, hypertension, and infertility on the health system in India. This can be achieved either through undertaking separate urban health programme or incorporating special clause in the proposed National Urban Health Program, citing the importance of healthy diet and physical exercise.

REFERENCES

- Rosenblatt K, Daling JR, Chen C, Sherman KJ, Schwartz S. Marijuana use and risk of oral squamous cell carcinoma. Cancer Research. 2004;64: 4049–4054.
- Agalliu I, Gern R, Leanza S, Burk RD. Associations of high-grade prostate cancer with BRCA1 and BRCA2 founder mutations. Clin Cancer Res. 2009; 15:1112–1120.
- Morris AP, Lindgren CM, Zeggini E, Timpson NJ, Frayling TM, Hattersley AT, et al. A powerful approach to sub-phenotype analysis in population-based genetic association studies. Genet Epidemiol. 2010; 34(4):335–343.
- Potter J, Cerhan J, Sellers TA, McGovern PG, Drinkard C, Kushi LR, et al. Progesterone and estrogen receptors and mammary neoplasia in the Iowa Women's Health Study: How many kinds of breast cancer are there? Cancer Epidemiol Biomarkers & Prev. 1995;4:319–326.
- Cui Y, Shikany JM, Liu S, Shagufta Y, Rohan T. Selected antioxidants and risk of hormone receptor-defined invasive breast cancers among postmenopausal women in the Women's Health Initiative Observational Study. Am J Clin Nutr. 2008;87:1009–18.
- Kenfield SA, Wei EK, Stampfer MJ, Rosner RA, Colditz GA. Comparison of aspects of smoking among the four histological types of lung cancer. Tobacco Control. 2008; 17:198–204.
- Mendez MA, Monteiro CA, Popkin BM. Overweight exceeds underweight among women in most developing countries. Am J Clin Nutr. 2005; 81: 714–21.
- Heslehurst N, Simpson H, Ells LJ, Rankin J, Wilkinson J, Lang R, et al. The impact of maternal BMI status on pregnancy outcomes with immediate short-term obstetric resource implications: a meta-analysis. Obes Rev. 2008; :635–83.
- Oteng-Ntim E, Kopeika J, Seed P, Wandiembe S, Doyle P. Impact of obesity on pregnancy outcome in different ethnic groups: calculating population attributable fractions. PLoS One. 2013.
- Torloni MR, Betrán AP, Horta BL, Nakamura MU, Atallah AN, Moron AF, et al. Prepregnancy BMI and the risk of gestational diabetes: a systematic review of

the literature with meta-analysis. Obes Rev.2009;10:194–203.

- Nagai M, Kuriyama S, Kakizaki M, Ohmori-Matsuda K, Sone T, Hozawa A, et al. Impact of obesity, overweight and underweight on life expectancy and lifetime medical expenditures: the Ohsaki cohort study.BMJ Open. 2012;
- O'Brien TE, Ray JG, Chan WS. Maternal body mass index and the risk of preeclampsia: a systematic overview. Epidemiology. 2003;14:368–74.
- Tennant PWG, Rankin J, Bell R. Maternal body mass index and the risk of fetal and infant death: a cohort study from the North of England. Hum Reprod. 2011;26:1501–11.
- Sahu MT, Agarwal A, Das V, Pandey A. Impact of maternal body mass index on obstetric outcome. J Obstet Gynaecol Res. 2007;33:655–9.
- Long JS. Regression models for categorical and limited dependent variables. Thousand Oaks: Sage; 1997. p. 297. (Advanced quantitative techniques in the social sciences, V. 7).
- Ziraba AK, Fotso JC, Ochako R. Overweight and obesity in urban Africa: a problem of the rich or the poor? BMC Public Health 2009;9:465. doi: 10.1186/1471- 2458-9-465.
- Agrawal P, Mishra V. Covariates of overweight and obesity among women in North India. Honolulu: East-West Center, 2004. 25 p. (East-West Center working papers no. 116, population and health series).
- Sinha R, Kapoor AK. Cultural practices and nutritional status among premenopausal women of urban setup in India. Open Anthropol J 2010;3:168-71.
- Ziraba AK, Fotso JC, Ochako R. Overweight and obesity in urban Africa: a problem of the rich or the poor. BMC Public Health. 2009; 9: 465.
- Shafique S, Akhter N, Stallkamp G, de Pee S, Panagides D, Bloem MW. Trends of under- and overweight among rural and urban poor women indicate the double burden of malnutrition in Bangladesh. Int J Epidemiol. 2007; 36: 449–57.
- Ramachandran A, Snehalatha C, Baskar AD, Mary S, Kumar CK, Selvam S, et al. Temporal changes in prevalence of diabetes and impaired glucose tolerance associated with lifestyle transition occurring in the rural population in India. Diabetologia. 2004; 47: 860–5.
- Mendez MA, Monteiro CA, Popkin BM. Overweight exceeds underweight among women in most developing countries. Am J Clin Nutr. 2005; 81: 714–21.
- World Bank. The growing danger of non-communicable diseases: acting now to reverse course. Washington, DC: World Bank; 2011. p. 14 p.
- Gruebner O, Staffeld R, Khan MMH, Burkart K, Krämer A, Hostert P. Urban health in megacities extending the framework for developing countries. Hum Health Glob Environ Change. 2011:42–9.
- Schwab, J. A. (2002). Multinomial logistic regression: Basic relationships and complete problems. http://www.utexas.edu/courses/schwab/sw388r7/SolvingProblems/
- Breslow, N. E., and Day, N. E. (1980). Statistical Methods in Cancer Research. Lyon, UK: International Agency for Research on Cancer.
- Vittinghoff, E., Shiboski, S. C., Glidden, D. V., & McCulloch, C. E. (2005). Regression Methods in Biostatistics: Linear, Logistic, Survival, and Repeated Measures Models. New York: Springer Science+Business Media, Inc.

How to cite this article: Dixit S, Kumar B, Singh A et. al. An application of Multinomial Logistic Regression to assess the factors affecting the women to be underweight and overweight: A practical approach. Int J Health Sci Res. 2015; 5(10):11-17.

****