Determinants of Contraceptive Practice in Oman by Using Multilevel Modeling

Moza S. Al-Balushi *, Mohammed S. Ahmed and M. Mazharul Islam

Department of Mathematics and Statistics, College of Science, Sultan Qaboos University, P.O. Box: 36, PC 123, Al-Khod, Muscat, Sultanate of Oman. *Email: p890778@student.squ.edu.om.

ABSTRACT: In this paper, multilevel logistic regression models are developed for examining the hierarchical effects of contraceptive use and its selected determinants in Oman using the 2008 Oman National Reproductive Health Survey (ONRHS). Comparison between single level and multilevel logistic regression models has been made to examine the plausibility of multilevel effects of contraceptive use. From the multilevel logistic regression model analysis, it was found that there is real multilevel variation among contraceptive users in Oman. The results indicate that a multilevel logistic regression model is the best fit over ordinary multiple logistic regression models. Generally, this study revealed that women's age, education, number of living children and region of residence are important factors that affect contraceptive use in Oman. The effect of regional variation for age of women, education of women and number of living children further implies that there exists considerable differences in modern contraceptive use among regions, and a model with a random coefficient or slope is more appropriate to explain the regional variation than a model with fixed coefficients or without random effects. The study suggests that researchers should use multilevel models rather than traditional regression methods when their data structure is hierarchal.

Keywords: Multilevel modelling; Hierarchical; Logistic regression; Contraceptive.

<div dir="rtl">

محددات تطبيق وسائل منع الحمل في عمان باستخدام نموذج متعدد المستويات

موزه البلوشية، محمد أحمد، م. مزهرول اسلام

ملخص: في هذه الورقة تم تطوير نموذج متعددة المستويات للانحدار اللوجستي لدراسة آثار الهرمية للاستخدام وسائل منع الحمل في عمان وتحديد العوامل التي تؤثر على هذا المتغير وباستخدام مسح 2008 العمانية الوطنية للصحة الإنجابية (ONRHS). وقد تم مقارنة بين مستوى واحد ومتعددة المستويات لنماذج الانحدار اللوجستي لفحص آثار التعدد في المستوى لاستخدام وسائل منع الحمل. من تحليل نموذج متعددة المستويات للانحدار اللوجستي تبين أن هناك تباين لتعدد المستويات بين المستخدمين وسائل منع الحمل في سلطنة عمان. وتشير النتائج إلى ان نموذج متعددة المستويات للانحدار اللوجستي هو الأنسب على نموذج الانحدار اللوجستي التقليدي. عموماً، فقد كشفت هذه الدراسة أن عمر المرأة، والتحصيل العلمي للمرأة، وعدد الأطفال الباقيين على قيد الحياة ومنطقة الإقامة تؤثر بشكل ملحوظ على النساء اللاتي يستخدمن وسائل منع الحمل في سلطنة عمان. تأثير التباين بين المناطق لعمر المرأة، والتحصيل العلمي للنساء وعدد الأطفال الباقيين على قيد الحياة يعني أن هناك وجود اختلاف ملحوظ في استخدام وسائل منع الحمل الحديثة بين المناطق حيث ان النموذج الذي يحتوي على المعامل العشوائي أو المنحدر هو أكثر ملاءمة لتفسير التباين بين المناطق من النموذج الذي يحتوي على المعاملات الثابتة أو من دون آثار عشوائية. وتقترح الدراسة للباحثين إلى أن استخدام نماذج متعددة المستويات هو الأنسب من الطرق التقليدية الانحدار عندما تكون بنية البيانات هرمية.

كلمات مفتاحية: نموذج تعدد المستويات ، الهرمية ، الانحدار اللوجستي و موانع الحمل.

</div>

1. Introduction

Stratified multistage cluster sampling designs are the norms for most national level socioeconomic, demographic and health surveys because of cost, time and efficiency considerations. As these survey designs are hierarchical in nature, data collected in these surveys often introduce multilevel dependency or correlation among the observations that can have implications for model-based statistical inference, although these samples provide efficient estimates of the descriptive population quantities [1-4]. The individual observations in hierarchical data structures are not completely independent, and the results of the analysis can be affected by this clustered structure of the underlying data. For example, households from the same community or geographical area are more homogenous than households in different communities.

Standard statistical techniques such as regression models have problems dealing with the hierarchical data structure because they assume independent and normally distributed errors with a constant variance (homoscedasticity)

[1, 5-8]. Analyzing variables from different levels without taking into account the hierarchical data structure leads to misleading estimation results and interpretation because one faces the problem of heteroscedasticity. To overcome these problems, a methodology of multilevel modeling has been developed recently, taking the hierarchical data structure into account and satisfying the assumption of independence and heteroscedasticity, which obviously leads to correct analysis and proper interpretation of the data. Multilevel models are useful in determining the direct effects of the individual and the group explanatory variables. There is an increasing use of multilevel models for the analysis of clustered data. These models are also referred to as mixed effects models, random effects models, or hierarchical models in the literature [1-4,9,10]. Many popular statistical packages, such as HLM, MLwiN, R, SAS, Stata, and SPSS have the capacity to fit multilevel models. Methodological work on analyzing multilevel models has been done by many [e.g., 5-7,11], who gives an illustrative introduction to multilevel models.

The primary objective of this paper is to develop an analytical framework of multilevel analysis for studying contraceptive use in Oman utilizing the 2008 Oman National Reproductive Health Survey (ONRHS) data, which is hierarchical in nature as it followed a multistage cluster sampling design. Although the hierarchical effect of contraceptive use in Oman is still unknown, information regarding this would be very important for policy purposes.

2. Data and Methods

2.1 Data

The data used in this study come from the 2008 Oman National Reproductive Health Survey (ONRHS). The 2008 ONRHS was tagged on to the World Health Survey in Oman (OWHS) in 2008. The 2008 OWHS was designed by the World Health Organization (WHO) as part of comprehensive standardized data collection on population health in different countries. Both the ONRHS and OWHS were implemented by the Ministry of Health of Oman. The details of the OWHS may be seen in [12].

The ONRHS used a sub-sample of 3,703 Omani households out of 5,464 households selected for the 2008 WHSO. The ONRHS considered ever-married Omani women of reproductive age of 15 to 49 years. Ultimately 4,560 women were successfully interviewed, who constitute our study sample. The sample was calculated to provide national estimates, as well as separate estimates for both rural and urban areas.

The sample of the 2008 ONRHS was selected following a multistage stratified cluster sampling design. Stratification was made on two factors: level of urbanization (urban/rural) and geographical distribution. Administratively, Oman was divided into 10 regions at the time of survey (currently 11 regions). The survey covered all the 10 regions of Oman, dividing into 10 urban and 10 rural strata. Equal samples were selected from all strata, with increased samples from urban Muscat. From each stratum, 10 clusters were selected randomly. The 2003 population census 'Enumeration Areas' (EAs) were considered as clusters. The 2003 census has defined the EA as the area assigned for one enumerator. There are a total of 3167 EAs in the country. Each EA contains on average 110 households (both Omani and non-Omani) arranged in two or more enumeration blocks (EBs). Twenty four households were selected systematically from the selected EAs. EAs were used as the primary sampling units (PSUs).

As the 2008 ONRHS data is of hierarchical structure, the hierarchy for our study follows individuals as level-1 and regions as level-2. The survey considered non-pregnant currently married women as the respondents for current contraceptive use. After excluding currently pregnant and widowed or divorced women, we were left with 3056 women as our study population.

2.2 Study Variables

The study considered current use of contraceptive methods by currently married non-pregnant women of reproductive age as the dependent variable. The current use covers both modern and traditional methods of contraceptive. The variable is categorical which we dichotomous as follows: those women currently using any of the methods as 1 and those not currently using any method coded as 0.

The explanatory variables considered for this analysis include age of the women, region of residence, number of living children, level of education, place of residence, type of family and economic status or wealth quintiles. A shortcoming of the dataset used is that it contains no information on household income or wages, household expenditure patterns and community characteristics. However, the dataset contains information about house fixed and durable assets such as TV, refrigerator, car, house etc. These variables were used to construct a proxy index of household economic status or wealth index using the principal components analysis. It is a composite measure of the cumulative living standard of a household, which places individual households on a continuous scale of relative wealth. The wealth index is divided into four categories, with the lowest one representing the poorest 25 percent and the highest one representing the wealthiest 25 percent of households see [13]. This wealth index has the advantage of providing a reasonably reliable measure of the household's economic status, and it is not affected by the endogeneity and transitory nature of labour income.

2.3 Analytical Methods

2.3.1 Multilevel logistic regression model

As mentioned earlier, as our basic data structure is of two-levels, with individual women as level-1 and region of residence as level-2, we concentrate on a two-level logistic regression model. Our outcome variable y_{ij} is a binary response for individual i $(i=1, 2,..., n_j)$ and group j $(j = 1,2,..., k)$ and we have one explanatory variable at individual level x_{ij}. Then $p_{ij} = \Pr(y_{ij} = 1)$ can be defined as the probability of the response equal to one and p_{ij} can be modeled using a logit link function where y_{ij} has a Bernoulli distribution. Then we can write the two level model as

$$\log\left[\frac{p_{ij}}{1 - p_{ij}}\right] = \beta_0 + \beta_1 x_{ij} + u_j. \qquad \text{(combined model)} \qquad (1)$$

This is a formal model, where j refers to the level 2 unit (region) and i to the level 1 unit (women), such that u_j is the random effect of level 2 and a standard logistic regression model will be without u_j. Conditional on u_j and y_{ij} s being assumed to be independent, Model (1) can be written as follows, splitting up into two models: one for level 1 and the other for level 2.

$$\log\left[\frac{p_{ij}}{1 - p_{ij}}\right] = \beta_{0j} + \beta_{1j} x_{ij}, \qquad \text{(Level-1 model)} \qquad (2)$$

and

$$\beta_{0j} = \beta_0 + u_j \qquad \text{(Level-2 model)} \qquad (3)$$

and, u_j is normally distributed with $E(u_j) = 0$ and $\text{var}(u_j) = \sigma_u^2$. Model (1) is the simplest possible multilevel model for binary data.

3. Results

3.1 Differentials of contraceptive use: bivariate analysis

Table 1 presents the percentage distribution of currently married and non-pregnant women using any of the family planning methods by selected socio-economic and demographic factors. The survey data indicate that 29% of the eligible women were using any family planning methods. Women's age shows significant ($p < 0.01$) association with contraceptive use. The association is curvilinear, as the rate of contraceptive increases with the age of women until 35 years and then decreases. Contraceptive use rate is higher among women living in urban areas than their rural counterparts (30% vs. 26%) but the difference is not statistically significant. However, the contraceptive use rate varies significantly among the 10 geographical regions considered in this study. Musandam (45%) followed by Muscat (38%) show higher prevalences of contraceptive use and Al-Dhakhlia (18%) the lowest. Women's education and number of living children show significant association with contraceptive use in Oman.

3.2 Single level logistic regression analysis

The single level standard logistic regression analysis results are reported in Table 2. The results indicate women's age, educational level, number of living children, and region of residence as significant predictors of contraceptive use in Oman. After controlling for other factors, the age of women shows significant negative association with contraceptive use. Young women of age 15-29 are 2 times more likely to use some family planning method than their older counterparts.

Married women who had secondary education were 2.8 times more likely to use contraception than uneducated woman. Similarly, woman who had college, university and higher education were 3.1 times more likely to use contraception than uneducated woman. The likelihood of contraceptive use by married women who had (5-7) children was 3.7 times greater than by married women who had (0-2) children. Similarly, married women who had more than seven children were 3.4 times more likely to use contraception than married women who had (0-2) children.

3.3 Multilevel logistic regression analysis of contraceptive use

In this study we consider multilevel models allowing between-region variance of modern contraceptive use. The data have a two-level hierarchical structure with women at level 1, nested within 10 regions at level 2. The multilevel logistic regression analysis was used for comparisons and analysis of the effects of socio-economic and demographic variables on modern contraceptive use among women, and to estimate any residual correlations across regions in predicting modern contraceptive use in Oman. The multilevel modeling process was performed stepwise. The first step

examined the null model (or empty model) of overall probability of modern contraceptive use without adjustment for predictors. The second step included both the analysis of single and multilevel models for random intercept and fixed slope multilevel analysis. The third step considered a model for two level random intercept and random slope (random coefficient) multilevel logistic regression analysis.

Table 1. Percentage distribution of women currently using contraceptive by background characteristics 15-49 (n =3056).

Background Characteristics	Categories	Using contraceptive	
		Yes	No
Total		28.9	71.1
Age of women*	15-24	22.5	77.5
	25-29	28.1	71.9
	30-34	35.1	64.9
	35-39	31.6	68.4
	40-44	32.2	67.8
	45-49	18.9	81.1
Place of residence	Urban	30.1	69.9
	Rural	25.6	74.4
Region*	Muscat	37.5	62.5
	Dhofar	26.4	73.6
	Al-Dakhlyah	17.8	82.2
	North Sharqiah	19.2	80.8
	South Sharqiah	23.1	76.9
	North Batinah	32.7	67.3
	South Batinah	28.8	71.2
	Al-Dhahirah	31.5	68.5
	Musandam	44.7	55.3
	Al-Wasta	31.3	68.8
Women Education*	Illiterate	19.2	80.8
	Primary+ Literate	28.3	71.7
	Preparatory completed	35.1	64.9
	Secondary completed	30.1	69.9
	College/university	35.1	64.9
Number of living children*	(0-2) children	21.8	78.2
	(3-4) children	33.0	67.0
	(5-7) children	38.2	61.8
	more than seven children(ref)	26.3	73.8
Economic status	Poor	31.2	68.8
	Lower middle	26.7	73.3
	Upper middle	27.1	72.9
	Rich	30.5	69.5
Living arrangement	Nuclear family	31.8	68.2
	Extended family/Composite	26.3	73.7

3.3.1 Multilevel logistic regression model comparison

We compare the three possible multilevel logistic regression models (empty model, random intercept model and random slope model) using a deviance-based chi-square value and the Akaike information criterion (AIC). The deviances of the empty model with random intercept (deviance = 3634) and the random intercept and fixed slope model (deviance = 3403) indicate that the random intercept and fixed slope model is better than the empty model with regard to random intercept. In addition to this, the AIC value of the empty model with random intercept (AIC = 3638) is larger than that for the random intercept and fixed coefficient model (AIC = 3419), which implies that random intercept and fixed slope is better than the empty model with random intercept in predicting contraceptive use across regions.

The deviances of a fixed slope and random intercept (deviance = 3403) and random slope model (deviance = 3386) show that the random slope model is better than the random intercept and fixed slope model. The AIC value of the random coefficient model (AIC = 3406) is smaller than that of the random intercept and fixed coefficient model (AIC = 3419), implying that the random coefficient model is better compared to the random intercept and fixed slope model in describing contraceptive use status (see Table 3).

3.3.2 Empty random intercept logistic regression analysis

The simplest non-trivial specification of the hierarchical linear model is one in which only the intercept varies between level two units and no predictors (explanatory) variables are entered in the model. A random intercept or variance components model allows the overall probability of contraceptive use to vary across regions. An intercept-only multilevel model has been estimated to predict the probability of women who use contraception; this is described in Equation (8), where j refers to the level 2 unit (region) and i to the level 1 unit (women) and which is still a single level model.

Table 2. Maximum likelihood estimate for binary logistic regression of predicting contraceptive use.

Covariate	$\widehat{\beta}$	S.E.	P-value	OR	95% C.I. OR Lower	Upper
Age of women			0.003			
15-24	0.80	0.251	0.001	2.226	1.360	3.641
25-29	0.73	0.215	0.001	2.076	1.361	3.166
30-34	0.63	0.202	0.002	1.876	1.263	2.785
35-39	0.33	0.184	0.077	1.385	0.966	1.986
40-44	0.57	0.181	0.001	1.777	1.247	2.533
45-49 (ref)				1.000		
Place of residence						
Urban	-.068	0.111	0.538	0.934	0.752	1.160
				1.000		
Rural						
Region			0.000			
Muscat	0.83	0.217	0.000	2.285	1.494	3.496
Dhofar	0.45	0.225	0.044	1.573	1.012	2.444
Al-Dakhila	-0.23	0.237	0.339	0.797	0.501	1.268
South Sharqiah	0.19	0.239	0.421	1.212	0.192	0.239
North Batina	0.74	0.207	0.000	2.106	0.745	0.207
South Batina	0.45	0.215	0.035	1.572	0.452	0.215
Al-Dhahirah/	0.57	0.240	0.018	1.765	0.568	0.240
Musandam	1.33	0.400	0.001	3.776	1.329	0.400
Al-Wasta/	0.87	0.427	0.042	2.387	0.870	0.427
North Sharqiah (ref)				1.000		
Women Education			0.000			
Illiterate (ref)				1.00		
Primary/ Literate	0.51	0.148	0.001	1.664	1.246	2.222
Preparatory completed	1.03	0.191	0.000	2.814	1.936	4.089
Secondary completed	1.02	0.183	0.000	2.787	1.945	3.992
College/university	1.14	0.203	0.000	3.127	2.102	4.651
Number of children			0.000			
(0-2) children (ref)				1.000		
(3-4) children	0.73	0.136	0.000	2.065	1.582	2.695
(5-7) children	1.31	0.161	0.000	3.694	2.696	5.061
>7 children	1.24	0.196	0.000	3.446	2.345	5.064
Economic status			0.151			
Poor	0.28	0.128	0.027	1.327	1.032	1.705
Lower middle	0.11	0.132	0.402	1.117	0.863	1.446
Upper middle	0.08	0.124	0.542	1.079	0.845	1.376
Rich(Ref)						
Family type						
Nuclear family	0.27	0.089	0.002	1.221	1.017	1.465
Extended/Composite(ref)						
Constant	-3.59	.319	126.641	1	0.000	

ref = reference category

Table 3. Summary of three possible multilevel logistic regression model selection criteria 3.

Model selection Criteria	Multilevel Empty Model	Multilevel Random Intercept Model	Multilevel Random slope Model
Deviance	3634	3403	3386
AIC	3638	3419	3406

The results of the multilevel empty model are presented in Table 4. The ratio of women who use contraception to women who do not by using ML estimate from the standard logit model is exp (-0.900) = 0.407 that is the same sample ratio of 884 women who used contraception to 2173 women who did not . This is the odds-ratio when no predictors have been considered in the model. This ratio is estimated to be exp(-0.917) = 0.400, exp(-0.938) = 0.391, exp(-0.928) = 0.395, and exp(-0.931) = 0.394 from multilevel model by MQL-1, MQL-2, PQL-1 and PQL-2 multilevel estimation methods respectively. By comparing the odds-ratio, it can be said that under all multilevel methods, the standard logistic model odds-ratio has overestimated.

From Table 4, it can be noticed that that there is a significant difference between the results of the standard logistic model and the multilevel logistic model. Failing to take into account the clustering within clusters (level 2), the standard logit model has overestimated the odds-ratio by about 1.85% (-0.900 - (-0.917)*100/0.917), 4.05%, 3.02% and 3.33% and compared to the multilevel model by using the corresponding methods MQL-1, MQL-2, PQL-1 and PQL-2. The fixed intercepts for all the multilevel estimated methods are similar methods and the random quantity including its standard error at cluster level for PQL-1 and PQL-2 are high compared to all other methods.

Table 4. Estimation of parameters and standard errors of an intercept-only, fixed effect, and random effect logit model for predicting the probability of contraceptive use.

Model	Logit	MQL-1	MQL-2	PQL-1	PQL-2
Fixed effect	-0.900	-0.917	-0.938	-0.928	-0.931
Intercept	(0.062)	(0.096)	(0.101)	(0.099)	(0.102)
Random Effect		0.094	0.095	0.106	0.107
Intercept		(0.036)	(0.037)	(0.042)	(0.042)
N	3056	3056	3056	3056	3056
significant ($p < 0.05$)					

The likelihood ratio statistic for testing the null hypothesis that $\sigma_{u0}^2 = 0$ can be calculated by comparing the two-level model with the single-level model without the level 2 random effects. The difference in -2*log likelihood between an empty model with random effect from Table 4 and an empty model without random effect from (single level logistic regression analysis) is (3979 -3634)=345, which constitutes strong evidence that the between-regions variance is non-zero. This implies that the empty model with random effect is better than the empty model without random effects. The decrease in deviance is 345, showing that the multilevel model fits the data dramatically better than the standard logistic regression. Conversely, the variance of the random effect of the random intercepts (σ_{u0}^2 =0.107, p-value =0.0116) reveals that there is significant difference in contraceptive use across the regions.

The empty model with random effect also helps to calculate the between-region variations, by the help of the intra-class correlation coefficient (ICC). ICC is estimated by $\rho = \sigma_u^2 / (\sigma_u^2 + \sigma_e^2)$, where $\sigma_e^2 = \pi^2 / 3$ is the variance of the standard logistic distribution. Then ICC = 0.0315, which implies that 3.15% of the variation in the contraceptive use is attributable to unobserved regional characteristics.

3.3.3 Multilevel random intercept and fixed coefficient model analysis

The random intercept and fixed coefficient logistic regression is the same as the empty random intercept logistic regression model, but with additional explanatory variables to allow for the effects of women's age, educational level, region, number of living children, living arrangements and economic status.

In the random intercept and fixed coefficient multilevel logistic regression model, we allowed the probability of contraceptive use to vary across regions, but we assumed that the effects of the explanatory variables are the same for each region. That is, the random intercept varies across region, but women's level explanatory variables are fixed across regions. The results are presented in Table 5.

For the single level model, place of residence and women's education were found to be significant determinants of variation in contraceptive use in all regions with respect to reference categories, and while it is not with the multilevel model (see Table 5). The estimated coefficients and odds-ratio have a similar interpretation, as in an/the ordinary logistic regression discussed above. The p-value differs for the same explanatory variables, such as place of residence and economic status (Poor), between single level and multilevel ;these were not significant for the single model but are significant under the multilevel model. This is because analyzing variables from different levels without taking into account the hierarchical data structure leads to misleading estimation results, as one faces the problem of heteroscedasticity. The individual observations in a hierarchical data structure are not completely independent and the results if analysis can be effected by this cluster structure of the underlying data. Women in the same region are more homogenous than women in different regions.

The results in Table 5 indicate that women's age, educational level, number of children, and living arrangements were found to be significant determinants of variation in contraceptive use in all regions, with the respecting reference categories and place of residence was not significant under both modeling. On the other hand, the 'poor' category for economic status was found to be a significant determinant of variation in contraceptive use in all regions with the respecting reference categories under the multi-level, while it was not significant under the single-level model.

The random part of the random intercept and fixed slope model showed that the intercept variance of the random effect was 0.123, whereas the variance of the intercept for the empty model was 0.107. The variance of the intercept for the empty multilevel model decreased compared to that under random effect of the intercept and fixed slope model. The reduction of the random effects of the intercept variance was due to the inclusion of fixed explanatory variables. That is, taking into account the fixed independent variables can provide extra predictive value on contraceptive use in each region. The significance of random effect intercept variance indicates that there is significant regional random effect variation on contraceptive use among women (see Table 5). This implies that there is still unexplained variation in contraceptive use across regions.

It is observed that there exist significant differences between the $\bar{\beta}$ coefficients of these two models for each of the explanatory variables. Also the $\bar{\beta}$ coefficients of the standard model have been either under or overestimated for $\bar{\beta}$ coefficients in comparison with those of the multilevel model. The difference in $\bar{\beta}$ coefficients estimated from a multilevel and from a standard level model arises because of the addition of the random effects, and implies that a single level model for this outcome is not appropriate.

Table 5. Results of random intercept and fixed coefficient logistic regression analysis.

	Single level				Multilevel			
Covariates	$(\hat{\beta})$	S.E.	p-value	OR	$(\hat{\beta})$	S.E.	p-value	OR
Age								
15-24	0.672	0.246	0.006	1.958	0.725	0.259	0.005	2.064
25-29	0.598	0.210	0.004	1.819	0.690	0.221	0.002	1.995
30-34	0.563	0.197	0.004	1.756	0.604	0.207	0.004	1.829
35-39	0.297	0.181	0.101	1.346	0.275	0.188	0.145	1.316
40-44	0.567	0.179	0.002	1.763	0.571	0.184	0.002	1.771
45-49 (ref)								
Place of residence								
Urban	-0.024	0.102	0.816	1.024	-0.025	0.116	0.829	0.975
Rural								
Education								
Illiterate (ref)								
Primary/ Literate	0.524	0.144	0.000	1.690	0.603	0.153	0.000	1.828
Preparatory C	1.056	0.186	0.000	2.875	1.145	0.197	0.000	3.141
Secondary C	1.060	0.176	0.000	2.885	1.145	0.192	0.000	3.142
College/univ	1.226	0.195	0.000	3.409	1.279	0.279	0.000	3.595
No. of living children								
(0-2) ch (ref)								
(3-4) ch	0.715	0.134	0.000	2.044	0.741	0.140	0.000	2.099
(5-7) ch	1.245	0.159	0.000	3.475	1.354	0.165	0.000	3.874
8+	1.120	0.193	0.000	3.066	1.318	0.201	0.000	3.734
Wealth quintile								
Poor	0.151	0.118	0.199	1.163	0.369	0.132	0.005	1.446
Lower middle	-0.036	0.122	0.768	.965	0.165	0.136	0.223	1.180
Upper middle	0.003	0.119	0.978	1.003	0.111	0.127	0.383	1.118
Rich (ref)								
Family type								
Nuclear family	0.169	0.091	0.064	1.185	0.167	0.095	0.081	1.181
Extended family (ref)								
Constant	-2.976	0.261	0.000	0.051	-3.399	0.304	0.000	0.35
Random intercept					Variance C.	S.E.		
$\sigma_{u0}^2 = \mathrm{var}(u_j)$					0.123	0.054	0.022	
Ref = reference category								

3.3.4 Multilevel random slope model analysis

So far we have allowed the probability of contraceptive use to vary across regions, assuming that the effects of the explanatory variables are the same for each region. We will now modify this assumption by allowing the difference between urban and rural families within region to vary across regions. To allow for this effect, we will need to introduce a random coefficient for place of residence. The model will be fitting has the same form as the random slopes model considered in the previous section, but since the variable 'place of residence' has only two categories, we use the more general term 'coefficient' rather than 'slope' to describe its effect. The results are presented in Table 6.

The results indicate that the women's age, educational leveland number of children are significant determinates of variation in contraceptive use in all regions with respect to the corresponding reference categories under multilevel modeling (see Table 6).

Table 6 also shows that the values of $\sigma_{u0}^2 = \mathrm{var}(u_{0j})$ and $\sigma_{u1}^2 = \mathrm{var}(u_{1j})$ are the estimated variance of intercept and slope of place of residence. The Wald test has been used to test the significant of the parameters σ_{u0}^2, σ_{u1}^2 and σ_{u10}. The test statistics of the first parameter is 4.883, which we compare to a chi-squared distribution on 1 d.f. We conclude that there is a significant difference between the regions. The test statistics of the second and third parameters is 3.574, which is approximately chi-squared distributed on 2 d.f. (p =0.167). At 10% we conclude that both parameters are non –zero, which implies that the effect of urban and rural middle does indeed vary across the region.

On average (after adjusting for the effects of age and the other explanatory variables), the log-odds of using contraception are 0.035 higher for urban areas than for rural areas. Depending on the value of σ_{u1}^2, the difference in a given region will be larger or smaller than 0.035.

The last column of Table 6 shows that the $\hat{\beta}$ coefficient under single level analysis corresponding to place of residence covariates has been under estimated about 42% compared with the multilevel estimate. Similarly, the $\hat{\beta}$ coefficients for the women's age group (25-29) of the single level model have been under estimated by almost 31%. The $\hat{\beta}$ coefficient under single level analysis corresponding to the women's age covariate for its all categories has been under

estimated with different percentage compared with multilevel estimate. On the other hand the $\widehat{\beta}$ coefficients for the women's education group (college/university) of the single level model have been over estimated by almost 13%.

Thus it is evident that if the multilevel effect is not taken into account in multivariate modeling, the estimates will be either underestimated or overestimated considerably. These results imply that a single-level multivariate model for this outcome variable is not appropriate.

Table 6. Results of random slope logistic regression model analysis of contraceptive use.

Covariate	Single level model				Multilevel model				Over/Under Estimation
	$\widehat{\beta}$	S.E.	p-value	OR	$\widehat{\beta}$	S.E.	p-value	OR	
Age									
15-24	0.672	0.246	0.006	1.958	0.856	0.345	0.019	2.354	-0.27
25-29	0.598	0.210	0.004	1.819	0.782	0.086	0.000	2.186	-0.31
30-34	0.563	0.197	0.004	1.756	0.664	0.214	0.002	1.943	-0.18
35-39	0.297	0.181	0.101	1.346	0.367	0.178	0.072	1.443	-0.24
40-44	0.567	0.179	0.002	1.763	0.612	0.187	0.001	1.844	-0.08
45-49 (ref)									
Place of residence									
Urban	0.024	0.109	0.816	1.024	0.034	0.218	0.559	1.035	-0.42
Rural (ref)									
Education									
Illiterate (ref)									
Primary/ Literate	0.524	0.144	0.000	1.690	0.472	0.120	0.000	1.603	0.10
Preparatory completed	1.056	0.186	0.000	2.875	0.978	0.249	0.000	2.955	0.07
Secondary completed	1.060	0.176	0.000	2.885	0.950	0.187	0.000	2.586	0.10
college/university	1.226	0.195	0.000	3.409	1.067	0.286	0.000	2.907	0.13
Children ever born									
(0-2) child (ref)									
(3-4) child	0.715	0.134	0.000	2.044	0.736	0.118	0.000	2.088	-0.03
(5-7) child	1.245	0.159	0.000	3.475	1.311	0.392	0.000	3.710	-0.05
>7	1.120	0.193	0.000	3.066	1.217	0.347	0.000	3.377	-0.09
HH economy									
Poor	0.151	0.118	0.199	1.163	0.302	0.241	0.129	1.353	-1.00
Lower middle	-0.04	0.122	0.768	0.965	0.132	0.170	0.279	1.141	4.67
Upper middle	0.003	0.119	0.978	1.003	0.099	0.220	0.568	1.104	-32.00
Rich (ref)									
Family type									
Nuclear family	0.169	0.091	0.064	1.185	0.217	0.137	0.166	1.242	-0.28
Extended family (ref)									
Constant	-2.98	0.261	0.000	0.051	-3.250	0.637	0.000	0.039	-0.09
Random intercept					Variance component	S.E.			
$\sigma_{u0}^2 = \text{var}(u_{0j})$					0.133	0.058			
$\sigma_{u1}^2 = \text{var}(u_{1j})$					0.243	0.126			
$\sigma_{u01} = \text{cov}(u_{0j}, u_{1j})$					-0.115	0.072			
Ref = reference category									

4. Discussion and Conclusion

This study analyzed determinants of contraceptive use in Oman using the multilevel modeling technique with application to the hierarchical data from the 2008 Oman National Reproductive Health Survey. Different models are fitted to the data to identify potential determinants of use of modern contraception among women in the reproductive age group. First, bivariate analysis with the chi-square test and the ordinary logistic regression model was fitted to the data and significant variables were considered for further investigation in multilevel models. Secondly, the multilevel models were fitted. The multilevel model was step wise; on the first step the intercept only or the empty model was fitted to check whether multilevel effects or heterogeneity exists among the hierarchies. In the next step the random intercept and fixed slope model, usually called the random intercept model, was fitted, and finally the random intercept and random slope (random slope model) was fitted.

Our study found that multilevel effects are all significant and have to be taken into consideration in logistic regression modelling. The results for ordinary logistic regression reveals that region of residence, women's age (in categories), educational level, number of living children, and type of family are the most important explanatory variables for use of contraception in Oman. Women use of modern contraception was significantly associated with regions of residence. This might be because of the fact that regions are highly correlated, and therefore the assumption of independence of higher level (regions) in classical regression analysis might seriously bias the standard errors of regression coefficients [3, 17]. The odds of using modern contraception in rural and urban areas were not significantly different. This might be due to the fact that the access to contraceptive commodities and women's knowledge, attitudes and practice are uniform in both rural and urban areas. It is also observed that the odds of contraceptive use do not vary significantly among the different economic groups in Oman, which contradicts the findings in other studies elsewhere.

Multilevel models fitted to the data also indicate that contraceptive use was significantly associated with region of residence; women's educational attainment, age group, number of living children and type of family. The fixed parts of the effects of explanatory variables or factors included in the multilevel models have similar interpretations as those of the multiple logistic regressions analysis but the random parts of the intercept and the coefficients provide additional information. The random intercept in the random intercept and fixed Slope model is significantly different from zero (P <0.05) indicating that use of contraception differs from region to region. The deviance based chi-square test for random effects in the random intercept model is also high. This indicates that the random intercept model with the fixed slope is found to give a better fit than the empty model for predicting the use of modern contraceptives across regions of Oman. The variance component of random intercept is also large which further supports the fact that there is a high variability in modern contraceptive use in Oman across regions. Within region variation further implies that use of contraceptives within regions is less likely (or heterogeneous) than between regions. Thus multilevel analysis has demonstrated that different regions have significantly different mean effects. This is in line with the findings of other studies [e.g.,14-16].

The random coefficient estimates for intercepts and the slopes vary significantly at 5% significance level, which implies that there is a significant variation in the effects of religion women's age, educational level and number of living children; these variables differ significantly across the regions. The variance component for the random intercept is large compared to its standard error. Thus there remains some regional-level variance unexplained in the final model. Hence using ordinary logistic regressions is not advisable for predicting modern contraceptive use in such a data set without accounting for the effect of regions. On the other hand, it was found that the random intercept and explanatory variables provide additional information. First, the variances of the random components related to the random term were found to be statistically significant, implying the presence of differences in the use of modern contraceptives across the regions. Secondly, from those explanatory variables considered here, the effects of the religious beliefs of a woman, her place of residence and her exposure to family planning messages on the radio differs from region to region. Third, the interaction between random parts of religious beliefs and place of residence create significant differences in the use of modern contraception across regions. This result supports the findings of other studies [e.g.,15-17].

We observed considerable differences in β coefficients when the multilevel effects were not taken into consideration. Ordinary multiple logistic regressions underestimate or overestimate the coefficients compared to the multilevel regression. Hence, ignoring the multilevel effects will result in a serious bias of the estimates of the coefficients of the parameters. Generally, a slight reduction of standard errors in single level analysis might imply that ignoring the effect of regional variations will decrease the precision of the estimates [18].

5. Acknowledgement

References

1. Snijders, T.A.B. and Bosker, R.J. *Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling.* Sage Publications, London, 1999.
2. Raudenbush, S.W. and Bryk, A.S. *Hierarchical Linear Models: Applications and Data Analysis Methods.* Thousand Oaks, CA: Sage Publications, London, 2002.
3. Goldstein, H. *Multilevel Statistical Models.* Edward Arnold: London, Halsted Press, New York, 1995.
4. Austin, P.C., Goel, V., van Walraven C. An introduction to multilevel regression models. *Can J. Public Health*, 2001, **92,** 150-4.
5. Goldstein, H. *Multilevel Models in Educational and Social Research.* Griffin: London, Oxford University Press: London, 1987.
6. Goldstein, H. *Performance Indicators in Education. Statistics in Society*, D. Dorling and S. Simpson. Part VI, pp. 281-286, Arnold, London, 1999.

7. Hox, J.J. *Multilevel Analysis: Techniques and Applications*. Lawrence Erlbaum Associates, Inc., Mahwah, NJ, 2002.

8. Harttgen, K. and Misselhorn, M. *A Multilevel Approach to Explain Child Mortality and Undernutrition in South Asia and Sub-Saharan Africa*. Discussion Paper, Ibero-America Institute for Economic Research, No. 152, 2006.

9. Kreft, I. and De Leeuw, J. *Introducing Multilevel Modeling*. Sage Publications, London, 1998.

10. Goldstein, H. Nonlinear multilevel models, with an application to discrete response data. *Biometrika,* 1991, **78**(1), 45-51.

11. Bryk, A.S. and Raudenbush, S.W. *Hierarchical Linear Models*. Newbury Park, CA: Sage Publication, London, 1992.

12. Al-Riyami, A., Abdel El Aty, M.A., Jaju, S., Morsi M., Al-Kharusi, H. and Al-Shekaili, W. Oman National Reproductive Health Survey. Muscat: Ministry of Health, Sultanate of Oman, 2008.

13. Filmer, D., Pritchett and L.H. Estimating wealth effects without expenditure data or tears: An Application to Educational Enrolments in States of India. *Demography,* 2001, **38**(1), 115-132.

14. Khan, H.T.A. A hierarchical model of contraceptive use in urban and rural Bangladesh. *Contraception*, 1997, **55**: 91-96.

15. Khan, H.R. and Shaw, J.E.H. Multilevel Logistic regression analysis applied to binary contraceptive prevalence Data. *J. Data Science*, 2011, **9**, 93-110.

16. Kalam, E.M.N. and Khan, H.T.A. Modeling contraceptive prevalence in Bangladesh: A hierarchical approach. Asian Meta Centre, Research Paper Series, No. 21, 2002

17. Ferede, T. Multilevel modeling model of modern contraceptive use among rural and urban population of Ethiopia. *Am. J. Math. Stat.*, 2013, **3**(1), 1-16.

18. Goldstein, H. *Multilevel Statistical Models*. Edward Arnold, London, 2003.