



THE AGA KHAN UNIVERSITY

eCommons@AKU

Community Health Sciences

Department of Community Health Sciences

February 2017

Multilevel modeling of binary outcomes with three-level complex health survey data.

Shafquat Rozi

Aga Khan University, shafquat.rozi@aku.edu

Sadia Mahmud

Aga Khan University

Gillian Lancaster

Lancaster University, Lancashire, UK.

Wilbur Hadden

University of Maryland, College Park, USA.

Gregory Pappas

Health Care Utilization, Complex Health Survey with Sampling Weights, Simulations for Complex Survey, Pseudo Likelihood, Three-Level Data

Follow this and additional works at: http://ecommons.aku.edu/pakistan_fhs_mc_chs_chs



Part of the [Family Medicine Commons](#), and the [Public Health Commons](#)

Recommended Citation

Rozi, S., Mahmud, S., Lancaster, G., Hadden, W., Pappas, G. (2017). Multilevel modeling of binary outcomes with three-level complex health survey data.. *Open Journal of Epidemiology*, 7(1), 27-43.

Available at: http://ecommons.aku.edu/pakistan_fhs_mc_chs_chs/280

Multilevel Modeling of Binary Outcomes with Three-Level Complex Health Survey Data

Shafquat Rozi^{1,2}, Sadia Mahmud³, Gillian Lancaster¹, Wilbur Hadden⁴, Gregory Pappas⁵

¹Department of Mathematics and Statistics, Lancaster University, Lancashire, UK

²Department of Community Health Sciences, Aga Khan University, Karachi, Pakistan

³Department of Medicine, Visiting Faculty, Aga Khan University, Karachi, Pakistan

⁴Department of Sociology, University of Maryland, College Park, USA

⁵George Washington University, District of Columbia Washington DC, USA

Email: shafquat.rozi@aku.edu

How to cite this paper: Rozi, S., Mahmud, S., Lancaster, G., Hadden, W. and Pappas, G. (2017) Multilevel Modeling of Binary Outcomes with Three-Level Complex Health Survey Data. *Open Journal of Epidemiology*, 7, 27-43.

<https://doi.org/10.4236/ojepi.2017.71004>

Received: November 15, 2016

Accepted: January 22, 2017

Published: January 25, 2017

Copyright © 2017 by authors and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

Complex survey designs often involve unequal selection probabilities of clusters or units within clusters. When estimating models for complex survey data, scaled weights are incorporated into the likelihood, producing a pseudo likelihood. In a 3-level weighted analysis for a binary outcome, we implemented two methods for scaling the sampling weights in the National Health Survey of Pakistan (NHSP). For NHSP with health care utilization as a binary outcome we found age, gender, household (HH) goods, urban/rural status, community development index, province and marital status as significant predictors of health care utilization (p -value < 0.05). The variance of the random intercepts using scaling method 1 is estimated as 0.0961 (standard error 0.0339) for PSU level, and 0.2726 (standard error 0.0995) for household level respectively. Both estimates are significantly different from zero (p -value < 0.05) and indicate considerable heterogeneity in health care utilization with respect to households and PSUs. The results of the NHSP data analysis showed that all three analyses, weighted (two scaling methods) and un-weighted, converged to almost identical results with few exceptions. This may have occurred because of the large number of 3rd and 2nd level clusters and relatively small ICC. We performed a simulation study to assess the effect of varying prevalence and intra-class correlation coefficients (ICCs) on bias of fixed effect parameters and variance components of a multilevel pseudo maximum likelihood (weighted) analysis. The simulation results showed that the performance of the scaled weighted estimators is satisfactory for both scaling methods. Incorporating simulation into the analysis of complex multilevel surveys allows the integrity of the results to be tested and is recommended as good practice.

Keywords

Health Care Utilization, Complex Health Survey with Sampling Weights,

1. Introduction

Multilevel modeling (MLM) of complex survey data is an approach increasingly being used in public health research. The influence of contextual factors on public health has been considered as important in recent public health research [1]. The standard multilevel modeling approach can properly estimate parameters and standard errors for clustered data that have resulted from equal probability sampling [2]. However, with unequal probabilities of selection the standard MLM does not estimate the model parameters and their standard errors appropriately [3] [4].

Complex survey designs often involve unequal selection probabilities of clusters and/or people within clusters. The survey includes design (sampling) weights to account for unequal selection probabilities. When estimating multilevel models that are based on complex survey data, sampling weights are incorporated into the likelihood, producing a pseudo likelihood [1] [4] [5] [6].

In this paper we explore the utility of multilevel modeling of three-level survey data with a binary outcome by incorporating weights in the estimation procedure to investigate the determinants of health seeking behavior of Pakistani population. It is important to understand a comprehensive picture of patterns of disease, conditions, and risk factors which affect the health of Pakistanis and their uptake of services.

We apply this methodology on the National Health Survey of Pakistan (NHSP) as our example of real data in survey research. The NHSP was a cross-sectional survey and provides a comprehensive picture of patterns of disease, conditions, and risk factors which affect the health of the people of Pakistan and their uptake of services. In the NHSP data primary sampling units (PSUs) are at level 3, households at level 2 and individuals at level 1. The outcome of interest in this paper is health care utilization coded as a binary outcome.

The most important recommendation for precise estimation of model parameters in weighted analysis is to have larger cluster sizes. Simulation studies have revealed that the regression coefficients are biased for small cluster sizes [7] [8]. In the literature a cluster size greater than 20 is considered as reasonably large cluster size for 2 level data for weighted multilevel analyses [1]. It is also important to normalize the weights at each level of the data. Including weights without scaling those produces biased parameter estimates especially with small cluster sizes. The weights scaling method is considered to be a bias reduction technique (Asparouhov, 2008).

Two scaling methods are commonly used, but there is no general method available which can deal with all types of design and data issues [7] [9]. The scaling of the top level weights will not influence the estimates of slopes and standard error [3] [7] [8]. Simulation studies have reported that the size of clus-

ter, intra-class correlation, informativeness of weights (e.g. where the design weights correlate with the outcome), and the type of outcome variable all affect the results of a scaling method [1].

Our literature search observed that three-level weighted analyses have been reported rarely. We found two important methodological papers [1] [3] that have discussed and implemented multilevel analysis for complex survey designs using real data. The applications and simulations that they report are limited to two-level data only. Jenkins [10] discusses weighted and un-weighted analysis for three-level data based on a simulation study. However, their simulation study for a binary outcome was rather modest. We are demonstrating methodology for analyzing real three-level binary complex survey data. In addition we conducted and report a thorough simulation study.

The objective of the study is to conduct MPML approach to estimate model parameters for determinants of health care utilization taking sampling weight into account for a three-level complex survey design with healthcare utilization as a binary outcome. These parameter estimates will leads to determination of contextual and individual level predictors of health care utilization. In addition we assessed the performance of MPML for three-level binary complex survey data by conducting a simulation study that has focused on varying the prevalence of the binary outcome and the intraclass correlation (ICC) at the third (PSU) and the second (household) level simultaneously.

The structure of this paper is as follows. In methods section we describe the methodology for a multilevel analysis of three-level binary response data. We study two approaches for conducting the multilevel analysis using data from the NHSP. The first approach uses multilevel maximum marginal likelihood (MML) estimation (un-weighted analysis). The second approach uses MPML estimation with sampling weights (weighted analysis). In the weighted analysis section we apply two weight scaling methods, suggested in literature by [1] and [3], for three-level data. The scaling methods are used to normalize weights at each level. In application section we apply the proposed methodology to the NHSP data. In simulation section we report a simulation study that was conducted to assess the impact of different prevalence of the binary outcome, and ICCs at different clustering levels on the bias of estimators of multilevel pseudo maximum likelihood (weighted analysis) using the two scaling methods. Finally we conclude this paper with a detailed discussion.

2. Methods

2.1. Sampling Weight Adjustment

The design probability for a sampling unit is a feature of the survey design and is assumed to be known before data analysis. If the design probabilities are informative they are correlated with the response. The weights are defined as the inverse of the probability of selection for the sampling unit.

Weighting Scheme of Three-Level Data

Let us first consider the weighting strategy for three-level data. The 3rd (PSU)

level weight can be computed as the reciprocal of the probability that PSU k was selected from the sampling frame. It is defined as $w_k = \frac{1}{p_k}$. Here p_k is the probability of selection of the k^{th} 3rd level cluster. The 2nd level weight is computed as the reciprocal of the probability that household j was selected given that PSU k was selected. Here we have $w_{jk} = \frac{1}{p_{jk}}$. Here, p_{jk} is the probability of selection of the j^{th} 2nd level cluster given that the k^{th} 3rd level cluster was selected. The first level weight is computed as the reciprocal of the probability that the i^{th} individual is selected, given that the j^{th} 2nd level cluster was selected in the k^{th} 3rd level cluster. $w_{ijk} = \frac{1}{p_{ijk}}$. Here p_{ijk} is the probability of selection of the i^{th} elementary unit in the j^{th} level cluster of k^{th} level cluster.

2.2. Multilevel Model for Three-Level Complex Survey Data

Consider a dichotomous outcome variable $y_{ijk} \sim \text{Bernoulli}(\pi_{ijk})$, then the logit link function is:

$$\eta_{ijk} = \beta_0 + \beta_1'x_{ijk} + \beta_2'x_{HH(jk)} + \beta_3'x_{PSU(k)} + \varepsilon_{j(k)}^{(2)} + \varepsilon_k^{(3)} \tag{1}$$

where $\eta_{ijk} = \ln\left(\frac{\pi_{ijk}}{1 - \pi_{ijk}}\right)$, π_{ijk} denotes the probability that the i^{th} subject in the j^{th} 2nd level cluster and the k^{th} 3rd level cluster uses a health care facility $\pi_{ijk} = P(Y_{ijk} = 1)$, x_{ijk} denotes the vector of the subject level variables, $x_{HH(jk)}$ denotes the vector of the second level variables, and $x_{PSU(k)}$ denotes the vector of 3rd level predictor variables. In addition β_1 denotes the vector of regression parameters for the subject level variables, β_2 denotes the vector of regression parameters for second level variables, and β_3 denotes the vector of regression parameters for the 3rd level variables. $\varepsilon_{j(k)}^{(2)}$ denotes the random effect for the j^{th} level cluster in the k^{th} level cluster and $\varepsilon_k^{(3)}$ denotes the random effect for k^{th} 3rd level cluster. We assume that $\varepsilon_{j(k)}^{(2)} \sim N(0, \sigma_{HH}^2)$ and $\varepsilon_k^{(3)} \sim N(0, \sigma_{PSU}^2)$ are independent.

2.3. Maximum Marginal Likelihood or Conventional Likelihood

The probability of selection of the k^{th} 3rd-level cluster at the initial stage is $p_k (k = 1, \dots, n^{(3)})$ and at the second stage j^{th} 2nd level cluster is sampled with a conditional probability of $p_{jk} (j = 1, \dots, n_k^{(2)})$. At first level an individual i is sampled with a conditional probability of $p_{ijk} (i = 1, \dots, n_{jk}^{(1)})$. n_{jk} is the number of individuals in the j^{th} 2nd level cluster in the k^{th} 3rd level cluster, n_k is the number of 2nd-level clusters in the k^{th} 3rd level cluster, and n is the total number of 3rd level clusters.

Let $y_{jk(2)}$ be the response vector for the j^{th} 2nd level cluster in the k^{th} 3rd level cluster.

The log-likelihood contribution of a level 2 cluster, conditional on the random effect at level 3 is:

$$L_{jk}^{(2)}\left(y_{jk(2)} \mid \varepsilon_k^{(3)}\right) = \log \left[\int_{-\infty}^{\infty} \exp \left\{ \sum_{i=1}^{n_{jk}^{(1)}} L_{ijk}^{(1)}\left(y_{ijk} \mid \varepsilon_{jk}^{(2+)}\right) \right\} f^{(2)}\left(\varepsilon_{jk}^{(2)}\right) d\varepsilon_{jk}^{(2)} \right] \quad (2)$$

where $L_{ijk}^{(1)}\left(y_{ijk} \mid \varepsilon_{jk}^{(2+)}\right)$ is the log-likelihood of a level 1 unit given all random effects and $f^{(2)}\left(\varepsilon_{jk}^{(2)}\right)$ is the normal density of the random effects at level 2.

The log-likelihood contribution of a 3rd level cluster is:

$$L_k^{(3)}\left(y_{k(3)}\right) = \log \left[\int_{-\infty}^{\infty} \exp \left\{ \sum_{j=1}^{n_k^{(2)}} L_{jk}^{(2)}\left(y_{jk(2)} \mid \varepsilon_k^{(3)}\right) \right\} f^{(3)}\left(\varepsilon_k^{(3)}\right) d\varepsilon_k^{(3)} \right] \quad (3)$$

where $L_{jk}^{(2)}\left(y_{jk(2)} \mid \varepsilon_k^{(3)}\right)$ is the log-likelihood of a level 2 unit given the random effect of level 3, $f^{(3)}\left(\varepsilon_k^{(3)}\right)$ is the normal density of the random effects at level 3 and $y_{k(3)}$ is the response vector for the k^{th} 3rd level cluster. Since 3rd level clusters (PSUs) are independent the full log-likelihood is:

$$L(y) = \sum_{k=1}^{n^{(3)}} L_k^{(3)}\left(y_{k(3)}\right) \quad (4)$$

where y is the vector of all responses.

2.4. Pseudo Maximum Likelihood Estimation Method

The PML estimates are obtained by maximizing the weighted log-likelihood. The sampling weights are incorporated into the likelihood for estimation of parameters and their standard errors. We maximize the weighted pseudo likelihood with respect to the regression parameters.

The log-likelihood contribution of a level 2 unit, conditional on the random effect at level 3 is:

$$L_{wjk}^{(2)}\left(y_{jk(2)} \mid \varepsilon_k^{(3)}\right) = \log \left[\int_{-\infty}^{\infty} \exp \left\{ \sum_{i=1}^{n_{jk}^{(1)}} w_{ijk} L_{ijk}^{(1)}\left(y_{ijk} \mid \varepsilon_{jk}^{(2+)}\right) \right\} f^{(2)}\left(\varepsilon_{jk}^{(2)}\right) d\varepsilon_{jk}^{(2)} \right]. \quad (6)$$

The log-likelihood contribution of a 3rd level cluster is:

$$L_{wk}^{(3)}\left(y_{k(3)}\right) = \log \left[\int_{-\infty}^{\infty} \exp \left\{ \sum_{j=1}^{n_k^{(2)}} w_{jk} L_{wjk}^{(2)}\left(y_{jk(2)} \mid \varepsilon_k^{(3)}\right) \right\} f^{(3)}\left(\varepsilon_k^{(3)}\right) d\varepsilon_k^{(3)} \right]. \quad (7)$$

Let $y_{k(3)}$ be the response vector in the k^{th} 3rd level cluster. The full log-pseudo likelihood is:

$$L_w(y) = \sum_{k=1}^{n^{(3)}} w_k L_{wk}^{(3)}\left(y_{k(3)}\right). \quad (8)$$

The Equation (7) and (8) differ from (3) and (4), respectively, only by addition of the weights as factors in the summation.

2.5. Scaling Method 1

In scaling method 1 weights are standardized by summing the scaled weights over the sample size of the corresponding cluster such that the sum of the scaled weights becomes equal to the effective cluster size. It is suggested by Carle [1] that scaling method 1 provides the least biased estimate for between cluster variance.

$$a_1^{(1)} = \frac{\sum_{i=1}^{n_{jk}} w_{ijk}}{\sum_{i=1}^{n_{jk}} (w_{ijk})^2} \text{-----Scale of level 1 weights}$$

$$a_1^{(2)} = \frac{\sum_{j=1}^{n_k} w_{jk}}{\sum_{j=1}^{n_k} (w_{jk})^2} \text{-----Scale of level 2 weights}$$

where $a_1^{(1)}$ and $a_1^{(2)}$ are the two scaling factors for levels 1 and level 2 respectively.

2.6. Scaling Method 2

In the scaling method 2 the scaled weights add up to sample sizes; the corresponding totals in method 1 are smaller [4].

$$a_2^{(1)} = \frac{n_{jk}}{\sum_{i=1}^{n_{jk}} (w_{ijk})} \text{-----Scale for level 1 weights}$$

$$a_2^{(2)} = \frac{n_k}{\sum_{j=1}^{n_k} (w_{jk})} \text{-----Scale for level 2 weights}$$

Method 2 provides least biased estimates for estimating slopes, intercepts and odds ratios [8] [11].

2.7. Intraclass Correlation

The ICC for three level binary data can be defined for each level separately:

$$ICC_{(PSU)} = \frac{\sigma_{PSU}^2}{\sigma_{PSU}^2 + \sigma_{HH}^2 + \frac{\pi^2}{3}} \text{-----ICC attributable to level 3-----(1)}$$

$$ICC_{(HH)} = \frac{\sigma_{cluster}^2}{\sigma_{PSU}^2 + \sigma_{HH}^2 + \frac{\pi^2}{3}} \text{-----ICC attributable to level 2-----(2)}$$

3. Application

The NHSP was a cross-sectional survey, conducted in 1990-1994. In the present application the outcome we are considering is health care utilization; if an individual had sought any medical care in the last 14 days the person was considered as utilizing health care. In the urban sampling frame a city/town was divided into enumeration blocks of 200 - 250 households. The sample of the NHSP for urban areas was drawn from the list of these enumeration blocks. A sample of 2400 households was considered sufficient to obtain estimates of important characteristics at the national level with an acceptable sampling error [12].

A multistage cluster sampling design with stratification was employed to collect data in the NHSP. There were 8 strata corresponding to the four provinces in Pakistan divided into urban and rural areas. In the NHSP a PSU is a block of households (HHs) in urban areas, and a village of households in rural areas. At the first stage PSUs were selected from each stratum with probability proportionate to size with respect to the number of households in the urban strata and to the population of the village in the rural strata. In total 80 PSUs were sampled, and out of these 32 were drawn from urban areas and 48 from rural areas. In the 2nd stage, on average 30 households were selected through systematic random sampling from each PSU. All subjects residing in the 2400 households were se-

lected for the study. The household non-response rate was 3.1% in the NHSP and the overall individual non-response rate was 7.6%. A total of 18,315 subjects were interviewed. For our analysis we considered health care utilization by subjects aged 14 years and older. As the objective is to determine the association of individual, household and PSU level characteristics with health care utilization, subjects younger than 14 years were excluded as their health seeking behavior would be determined by their guardians.

Out of the total of 18,315 subjects interviewed in NHSP, 9856 were aged 14 years and older. However, the number of individuals in our analysis is smaller because of missing data. We excluded subjects with missing values and performed a complete case analysis. The socio-demographic characteristics differ by less than 3% between the sample $n = 9856$ and the subset ($n = 8454$) used for this study suggesting that our findings will be representative of the Pakistani population aged 14 years and above. Two separate weight adjustments were applied for nonresponse at the household level and for individual level, respectively.

Nine explanatory variables were considered across the three levels as potential predictors of health care utilization. The PSU level explanatory variables are urban/rural status, province, and community development index for community development index we used the results provided by Hadden *et al.* [13] to assign an index value to each PSU. The household level variables are ownership of household goods and household size and the individual level explanatory variables are age, gender, marital status and literacy.

3.1. Statistical Analysis

Three-level un-weighted multivariable logistic regression model was regressed on covariates selected in the univariate analysis. The specified significance level for a variable to remain in the multivariable model was 0.05. We conducted scale examination for continuous predictor variables in the multivariable model. All possible interactions were assessed for significance (p -values < 0.05). We developed a final model conducting the un-weighted analysis using these rules for model selection. We then fitted the same model using the weighted analyses to compare the un-weighted and weighted analyses. The pseudo-maximum likelihood estimation for generalized linear mixed models with two or more than two levels using adaptive quadrature is implemented in GLLAMM a user written program included in STATA [14].

3.2. Results

About 22.1% subjects aged 14 years and older sought medical care. The mean age of the respondents was 35.4 (± 17.3) years. Most of the respondents were illiterate; 64.6% individuals could not read and write, and only 12.2% were educated to 10 or more years of education. About 28.6% of individuals reported that they had never been married while 63.9% were currently married. About 21.4% belonged to the Sindh province, 51.1% to Punjab, 10.0% to Baluchistan and 17.9% to North West Frontier Province (NWFP). Thirty seven percent of people

belonged to urban and 63% to rural communities. The average number of durable goods owned per household was 3.1.

The results of analysis of the NHSP data showed that the findings of the scaled weighted analysis generally agree with the un-weighted analysis (Table 1). We found age, gender, marital status, household ownership of durable goods, urban/rural status, community development index, and province as significant predictors of health care utilization (p-value < 0.05). We also found two significant

Table 1. MLM estimates and MPML estimates un-weighted and weighted analyses using scaling methods 1 and 2.

| Variables* | Un-weighted maximum likelihood | | | Weighted maximum likelihood (scaling method 1) | | | Weighted maximum likelihood (scaling method 2) | | |
|--|--------------------------------|--------|---------------------|--|---------|---------------------|--|--------|---------------------|
| | Estimate | SE | p-value | Estimate | SE | p-value | Estimate | SE | |
| Individual Level Variable | | | | | | | | | |
| <u>Gender</u> | | | | | | | | | |
| Female | -0.0113 | 0.1151 | 0.922 | 0.05048 | 0.1451 | 0.728 | 0.0528 | 0.1564 | 0.730 |
| <u>Age</u> | | | | | | | | | |
| | | | [‡] <0.001 | | | [‡] 0.04 | | | [‡] 0.03 |
| 19 - 35 | 0.3107 | 0.1044 | 0.003 | 0.3263 | 0.1394 | 0.019 | 0.3201 | 0.1451 | 0.028 |
| 36 or more | 0.4789 | 0.1189 | <0.001 | 0.4531 | 0.1912 | 0.018 | 0.4629 | 0.2005 | 0.022 |
| <u>Marital Status</u> | | | | | | | | | |
| | | | [‡] 0.811 | | | [‡] 0.320 | | | [‡] 0.382 |
| Married | -0.0123 | 0.1116 | 0.912 | 0.0906 | 0.1658 | 0.585 | 0.0745 | 0.1634 | 0.651 |
| Widow/divorced/separated | 0.0802 | 0.2169 | 0.711 | 0.1585 | 0.2779 | 0.569 | 0.1429 | 0.3003 | 0.622 |
| Household Level Variables | | | | | | | | | |
| <u>Ownership of household goods</u> | | | | | | | | | |
| | | | [‡] 0.001 | | | [‡] <0.001 | | | [‡] <0.001 |
| 1 - 4 | 0.3974 | 0.1066 | <0.001 | 0.4410 | 0.1260 | <0.001 | 0.4792 | 0.1329 | <0.001 |
| 5 or more | 0.4843 | 0.1245 | <0.001 | 0.6423 | 0.1585 | <0.001 | 0.7101 | 0.1688 | <0.001 |
| Community Level Variables | | | | | | | | | |
| <u>Urban/rural status</u> | | | | | | | | | |
| Urban | 1.4047 | 0.5323 | 0.008 | 1.3216 | 0.3645 | 0.001 | 1.3429 | 0.3783 | <0.001 |
| <u>CDI</u> | | | | | | | | | |
| | | | [‡] 0.01 | | | [‡] <0.001 | | | [‡] <0.001 |
| Middle | 0.2165 | 0.1438 | 0.132 | 0.2615 | 0.1822 | 0.151 | 0.2702 | 0.1963 | 0.169 |
| High | 0.7772 | 0.3330 | 0.020 | 0.8644 | 0.1262 | <0.001 | 0.9213 | 0.1333 | <0.001 |
| <u>Province</u> | | | | | | | | | |
| | | | [‡] <0.001 | | | [‡] <0.001 | | | [‡] <0.001 |
| NWFP | 0.9149 | 0.2238 | <0.001 | 0.6543 | 0.4025 | 0.104 | 0.6692 | 0.4216 | 0.107 |
| Sindh | 0.7797 | 0.2149 | <0.001 | 0.6493 | 0.3894 | 0.095 | 0.6732 | 0.4056 | 0.091 |
| Punjab | 1.1056 | 0.1983 | <0.001 | 0.9754 | 0.3828 | 0.011 | 1.0230 | 0.3979 | 0.010 |
| <u>Interactions</u> | | | | | | | | | |
| <u>CDI *urban/rural</u> | | | | | | | | | |
| | | | [‡] 0.045 | | | [‡] 0.002 | | | |
| Middle community* urban | -1.3616 | 0.5552 | 0.014 | -1.3103 | 0.4082 | 0.001 | -1.3453 | 0.4267 | [‡] 0.001 |
| High community* urban | -1.9071 | 0.6271 | 0.002 | -1.8898 | 0.3689 | <0.001 | -1.9680 | 0.3839 | 0.002 |
| <u>Marital status* gender</u> | | | | | | | | | |
| | | | [‡] 0.005 | | | [‡] 0.036 | | | [‡] 0.032 |
| Married* female | 0.3201 | 0.1345 | 0.017 | 0.2278 | 0.1568 | 0.146 | 0.2519 | 0.1611 | 0.118 |
| Widow/divorced/separated* female | 0.4998 | 0.2466 | 0.043 | 0.5523 | 0.3845 | 0.151 | 0.5719 | 0.4204 | 0.159 |
| log likelihood | -4267.95 | | | -2,098,251.4 | | | -2,326,415 | | |
| [†] σ ² _{PSU} | 0.1355 | 0.0348 | 0.001 | 0.0961 | 0.03393 | 0.001 | 0.11204 | 0.0381 | 0.001 |
| [†] σ ² _{HH(PSU)} | 0.2244 | 0.0625 | 0.001 | 0.2726 | 0.09952 | 0.001 | 0.28098 | 0.1052 | 0.001 |
| ICC(HH) | 0.06 | | | 0.080 | | | 0.076 | | |
| ICC(PSU) | 0.04 | | | 0.030 | | | 0.030 | | |

*Reference categories for individual level variables: male, age 14 - 18 years, never married respectively; *Reference categories for household level variable: none; *Reference categories for community level variables: rural, low, Baluchistan respectively. †Note: Estimates and standard errors for random intercept variance; ‡Note: over all p-value of covariates with more than two categories.

interactions; between gender and marital status, and between the community development index and urban/rural status respectively. The estimates of intraclass correlation from hierarchical modeling are useful and help to interpret the data results.

The estimated variance (unobserved heterogeneity) of the random intercepts using an un-weighted analysis are 0.1355 (standard error 0.0348) for PSU level, and 0.2244 (standard error 0.0625) for household level respectively. Both estimates are significantly different from zero and indicate considerable heterogeneity in health care utilization with respect to households and PSUs that is unaccounted for by the predictor variables and should be adjusted for an adequate analysis. For the analysis of the NHSP data, the household level variability had not been considered and adjusted for in previous research reported in the literature [15] [16]. The estimated intra-class correlations, from the un-weighted analysis, for level two (households) and level three (PSUs) showed that about 6% of the overall variation in the response is due to variation between households and 4% is due to variation between PSUs.

The estimated variance of the random effects using scaling method 1 are 0.096 (standard error 0.033) for PSU level and 0.272 (standard error 0.099) for household level respectively. The variance of the random effects using scaling method 2 are estimated as 0.112 (standard error 0.033) for PSU level and 0.280 (standard error 0.105) for household level. Estimated intra-class correlation for scaling methods 1 and scaling method 2 for household level and PSU level are: household; $ICC_{method1} = 0.080$; $ICC_{method2} = 0.076$ and PSU; $ICC_{method1} = 0.030$; $ICC_{method2} = 0.031$, respectively. The estimates of regression coefficients and standard errors from the weighted analysis for a PSU level variable, province diverged somewhat from the un-weighted analysis. The effect of the PSU level variable province, was highly significant (p-value < 0.001) for all three provinces (relative to Baluchistan) in the un-weighted analysis but after weight adjustment this effect became marginal, except for the province Punjab (p-value = 0.01). The confidence intervals were narrow for the un-weighted analysis but due to larger standard errors the confidence intervals become wider under the two scaled weighted analyses (methods 1 and 2).

Another difference that was observed between the weighted and un-weighted analysis was the estimate of an interaction between two individual level variables (gender and marital status). The estimated regression coefficient and standard errors from the weighted analysis diverged slightly from the un-weighted analysis.

To explore the stability of the models we carried out a simulation study for three level complex survey data with a binary outcome to assess the impact of varying prevalence of the outcome, and ICC at each level on the accuracy of the estimates of MPML.

4. Simulation Study Involving Three Level Complex Survey Data for a Binary Outcome

We carried out a simulation study to assess the influence of different conditions on the parameter estimates using two weight scaling methods. It will determine

which method provides less biased estimates for three level complex survey data [8]. A simulation study is needed as both weight scaling methods could be biased. This simulation is based on a multilevel logistic regression model.

The multilevel logistic regression model is:

$$\eta_{ijk} = \beta_0 + \beta_1 x_{ijk} + \beta_{21} x_{1HH(jk)} + \beta_{22} x_{2HH(jk)} + \beta_3 x_{PSU(k)} \\ + \beta_{41} x_{1PSU(k)} + \beta_{42} x_{2PSU(k)} + \beta_{43} x_{3PSU(k)} + \varepsilon_{j(k)}^{(2)} + \varepsilon_k^{(3)}$$

where $\eta_{ijk} = \ln\left(\frac{\pi_{ijk}}{1 - \pi_{ijk}}\right)$ is the logit transformation of the probability of health

care utilization and $y_{ijk} \sim \text{Bernoulli}(\pi_{ijk})$. Here π_{ijk} is the probability that the i^{th} subject in the j^{th} household and the k^{th} PSU will utilize health care facilities, $\varepsilon_{j(k)}^{(2)}$ is the random intercept for the j^{th} household in the k^{th} PSU, $\varepsilon_{j(k)}^{(2)} \sim \text{NID}(0, \sigma_{HH}^2)$ and $\varepsilon_k^{(3)}$ is the random intercept for the k^{th} PSU, $\varepsilon_k^{(3)} \sim \text{NID}(0, \sigma_{PSU}^2)$. Also x_{ijk} is an explanatory variable gender (male/female) at the respondent level. Ownership of household goods (none, 1 - 4 goods, 5 or more goods) is an explanatory variable at the household level, $x_{1HH(jk)}$ and $x_{2HH(jk)}$ are the dummy variables for 1 - 4 goods and 5 or more goods respectively, and none is the reference category. There are two explanatory variables urban/rural status and province (Baluchistan, NWFP, Sindh, and Punjab) at PSU level. $x_{PSU(k)}$ is a dummy variable for urban area and rural is the reference category. $x_{1PSU(k)}$, $x_{2PSU(k)}$ and $x_{3PSU(k)}$ are dummy variables for NWFP, Sindh, and Punjab respectively and Balochistan is the reference category.

The data were generated using the same level of clustering, number of clusters and average number of observations in each cluster as we have in the NHSP. The number of PSUs was set at $n_k = 80$, the number of households in each PSU $n_{jk} = 30$, and the number of individuals in each household $n_{ijk} = 4$.

We set four scenarios for our simulation study to compare two weight scaling methods for each scenario: 1) mid-range prevalence of the outcome as 30% with low intra-class correlation at PSU level 0.05 and at household level 0.2, 2) low prevalence of the outcome as 10% with low ICC at PSU level 0.05 and at household level 0.2, 3) mid-range prevalence of the outcome as 30% with high intra-class correlation at PSU level 0.15 and at household level 0.3, and 4) low prevalence of the outcome as 10% with high ICC at PSU level 0.15 and at household level 0.3.

A multistage cluster sampling design with stratification was employed to generate the data. The true values of the fixed effects and random effects parameters for the simulation study were specified within a reasonable range of the estimates from the MPML fit of three level NHSP data. The number of simulations we performed was decided based on detailed literature review, and on a sample size calculation suggested by Burton *et al.* [17]. We generated 200 datasets for each scenario for three level data of the size that we have considered.

We equated $\text{ICC}_{(PSU)} = 0.05$ and $\text{ICC}_{(HH)} = 0.2$ for low ICC, and $\text{ICC}_{(PSU)} = 0.15$ and $\text{ICC}_{(HH)} = 0.3$ for high ICC and solve Equations (1) and (2) simultaneously to set the variances of PSU and household level random ef-

fects respectively. The resulting variances are $\sigma_{\text{PSU}}^2 = 0.2$ and $\sigma_{\text{HH}}^2 = 0.8$ for low ICC, and $\sigma_{\text{PSU}}^2 = 0.9$ and $\sigma_{\text{HH}}^2 = 1.8$ for high ICC. To achieve the specified prevalence of the outcome (for the different scenarios) we varied the values of the fixed effects parameters.

We set the fixed effects parameters for scenario 1 as $\beta_0 = -2.8$, $\beta_1 = 0.5$, $\beta_{21} = 0.7$, $\beta_{22} = 0.9$, $\beta_3 = 0.1$, $\beta_{41} = 0.7$, $\beta_{42} = 0.8$, $\beta_{43} = 1$ and the variances $\sigma_{\text{PSU}}^2 = 0.2$ and $\sigma_{\text{HH}}^2 = 0.8$. For scenario 2 we set the fixed effects parameters as $\beta_0 = -2.8$, $\beta_1 = 0.1$, $\beta_{21} = 0.09$, $\beta_{22} = 0.1$, $\beta_3 = 0.01$, $\beta_{41} = 0.1$, $\beta_{42} = 0.2$, $\beta_{43} = 0.3$ and the variance $\sigma_{\text{PSU}}^2 = 0.2$ and $\sigma_{\text{HH}}^2 = 0.8$. For scenario 3 we set the fixed effects parameters as $\beta_0 = -2.8$, $\beta_1 = 0.4$, $\beta_{21} = 0.6$, $\beta_{22} = 0.8$, $\beta_3 = 0.08$, $\beta_{41} = 0.6$, $\beta_{42} = 0.7$, $\beta_{43} = 1$ and the variances $\sigma_{\text{PSU}}^2 = 0.9$ and $\sigma_{\text{HH}}^2 = 1.8$. For scenario 4 we set the fixed effects parameters as $\beta_0 = -3.2$, $\beta_1 = 0.08$, $\beta_{21} = 0.07$, $\beta_{22} = 0.08$, $\beta_3 = 0.01$, $\beta_{41} = 0.09$, $\beta_{42} = 0.1$, $\beta_{43} = 0.2$ and the variance $\sigma_{\text{PSU}}^2 = 0.9$ and $\sigma_{\text{HH}}^2 = 1.8$.

To generate the outcome, a Bernoulli distribution with probability

$$p_{ijk} = \frac{\exp(\eta_{ijk})}{1 + \exp(\eta_{ijk})} \text{ was used. In order to incorporate unequal probability of se-}$$

lection the same weights were taken for simulated data as in the NHSP data. A computer program was prepared in SAS to simulate data for the four different scenarios. The weighted analysis was performed in GLLAMM of STATA to assess the performance of the MPML estimates.

To evaluate the performance of MPML we assessed the bias and coverage of the parameter estimates [18]. Let $\hat{\beta}_i$ represent the estimate of population parameter β where $i = 1 \dots B$ and B is the number of simulations performed. Then

$$\hat{\beta} - \beta \text{ indicates bias for parameter estimates } \hat{\beta}_i \text{, where } \hat{\beta} = \frac{\sum_{i=1}^B \hat{\beta}_i}{B} \text{ is the}$$

average of the estimates of interest over the B simulation runs. The accuracy of the standard error of a parameter estimate is assessed by computing the observed coverage of the 95% confidence intervals (CI) created by using the standard normal distribution [19]. Coverage is defined as the percentage of 95% CIs

$$\left[\hat{\beta}_i \pm z_{1-\frac{\alpha}{2}} SE(\hat{\beta}_i), i = 1, 2, \dots, B \right] \text{ that contain the true underlying value of } \beta.$$

The coverage should be approximately equal to the nominal coverage (95%).

An acceptable criterion for the coverage is that the coverage should not fall outside of approximately two SEs of the nominal coverage probability (p) [17].

$$SE(\hat{p}) = \sqrt{\frac{p(1-p)}{B}} \text{ where } \hat{p} = 0.95 \text{ and with } B = 200 \text{ we get } SE(\hat{p}) = 0.015.$$

Hence according to within 2 SEs criteria (0.95 ± 0.03) the acceptable range for the coverage is 92% - 98%.

We also calculated the standardized bias that describes the bias as a percentage of the $SE(\hat{\beta})$. It can be calculated using the formula

$$\text{Standardized Bias} = \frac{\bar{\hat{\beta}} - \beta}{SE(\hat{\beta})}. \text{ Here } SE(\hat{\beta}) \text{ is the standard deviation of the esti-}$$

mate of interest of the overall simulation, and calculated as

$$SE(\hat{\beta}) = \sqrt{\left[\frac{1}{(B-1)} \right] \sum_{i=1}^B (\hat{\beta}_i - \hat{\beta})^2}$$

A standardized bias greater than 40% in either direction has adverse impact on the bias, coverage and efficiency [17].

The results of our simulation study indicate an acceptable bias for estimators of MPML for all scenarios. The performance of scaling method 1 is satisfactory (Table 2). In scenario 4 the variance components are showing relatively large bias compared to the fixed effects. The bias for PSU level variability for scenario 4 was largest for all parameter estimates among the four scenarios.

The results suggest that the effect of ICC is somewhat pronounced when the prevalence is low but it is not substantial, this could be due to the large sample size we have for the simulated data. The coverage (if not in the acceptable range) is close to the nominal value for few parameter estimates with mostly the coverage falling within two SEs of the nominal coverage probability (p). The standardized bias is less than 40% for all scenarios for scaling method 1. Hence we conclude that this procedure does not have an adverse impact on bias, coverage and efficiency of parameter estimates of fixed and random effects in the sample size ballpark we have used.

The simulation results show that the performance of scaling method 1 is reasonable for 3 level data with a binary outcome. In simulation results of scaling method 2 we also found acceptable bias for the fixed effects and random effects parameters for all scenarios (Table 3). In scenario 4 we did not observe any substantial increase or decrease in bias of fixed effect parameter estimates from

Table 2. The bias and rate of coverage for fixed effects parameter estimates and variance component estimates for four different scenarios using scaling method 1 (200 simulations for 3 level data).

| Scenario | Prevalence and ICC | Assessment | β_0 | β_1 | β_{21} | β_{22} | β_3 | β_{41} | β_{42} | β_{43} | σ_{HH}^2 | σ_{PSU}^2 |
|----------|--|-------------------|-----------|-----------|--------------|--------------|-----------|--------------|--------------|--------------|-----------------|------------------|
| 1 | 30% ¹ ³ ICC _{PSU} = 0.05 ³ ICC _{HH} = 0.2 | Bias | 0.030 | -0.010 | 0.010 | 0.010 | -0.020 | -0.040 | -0.020 | 0.007 | -0.010 | -0.040 |
| | | Coverage % | 88 | 95 | 92 | 97 | 90 | 94 | 92 | 94 | 93 | 94 |
| | | Standardized bias | 18 | -15 | 6 | 6 | -13 | -13 | -6 | 3 | -10 | -30 |
| 2 | 10% ² ³ ICC _{PSU} = 0.05 ³ ICC _{HH} = 0.2 | Bias | -0.010 | -0.010 | 0.010 | 0.030 | 0.020 | -0.004 | 0.030 | 0.040 | 0.020 | -0.060 |
| | | Coverage % | 95 | 94 | 93 | 91 | 92 | 92 | 93 | 94 | 95 | 92 |
| | | Standardized bias | -2 | -12 | 5 | 15 | 13 | -1 | 10 | 14 | 13 | -25 |
| 3 | 30% ¹ ⁴ ICC _{PSU} = 0.15 ⁴ ICC _{HH} = 0.3 | Bias | 0.020 | -0.010 | 0.020 | 0.020 | -0.030 | -0.050 | -0.040 | 0.00 | -0.060 | -0.070 |
| | | Coverage % | 94 | 97 | 94 | 93 | 94 | 95 | 93 | 95 | 93 | 93 |
| | | Standardized bias | 7 | -14 | 11.8 | 10 | -10 | -15 | -11 | 0 | -32 | -34 |
| 4 | 10% ² ⁴ ICC _{PSU} = 0.15 ⁴ ICC _{HH} = 0.3 | Bias | 0.030 | -0.010 | 0.010 | 0.020 | -0.040 | 0.040 | 0.010 | 0.040 | -0.050 | -0.100 |
| | | Coverage % | 90 | 95 | 92 | 95 | 92 | 94 | 94 | 92 | 93 | 91 |
| | | Standardized bias | 9 | -9 | 4 | 9 | -15 | 10 | 3 | 12 | -19 | -37 |

1-Mid-range prevalence =30%. 2-Low prevalence =10%. 3-Low ICCs: ICC_{PSU} = 0.05; ICC_{HH} = 0.2. 4-High ICCs: ICC_{PSU} = 0.15; ICC_{HH} = 0.3.

Table 3. The bias and rate of coverage for fixed effects parameter estimates and variance component estimates for four different scenarios using scaling method 2 (200 simulations for 3 level data).

| Scenario | Prevalence and ICC | Assessment | β_0 | β_1 | β_{21} | β_{22} | β_3 | β_{41} | β_{42} | β_{43} | σ_{HH}^2 | σ_{PSU}^2 |
|----------|--|-------------------|-----------|-----------|--------------|--------------|-----------|--------------|--------------|--------------|-----------------|------------------|
| 1 | 30% ¹ ³ ICC _{PSU} = 0.05 ³ ICC _{HH} = 0.2 | Bias | -0.010 | -0.010 | 0.010 | 0.020 | -0.020 | -0.030 | 0.000 | 0.010 | 0.006 | -0.050 |
| | | Coverage % | 91 | 93 | 95 | 97 | 95 | 92 | 94 | 96 | 94 | 93 |
| | | Standardized bias | -4 | -15 | 7 | 14 | -14 | -11 | 0 | 4 | 5 | -25 |
| 2 | 10% ² ³ ICC _{PSU} = 0.05 ³ ICC _{HH} = 0.2 | Bias | -0.040 | -0.006 | 0.010 | 0.030 | -0.020 | -0.003 | 0.030 | 0.040 | 0.010 | -0.070 |
| | | Coverage % | 90 | 95 | 93 | 94 | 92 | 96 | 94 | 92 | 93 | 94 |
| | | Standardized bias | -14 | -6 | 6 | 16 | -13 | -1 | 11 | 15 | 7 | -34 |
| 3 | 30% ¹ ⁴ ICC _{PSU} = 0.15 ⁴ ICC _{HH} = 0.3 | Bias | -0.010 | -0.008 | 0.030 | 0.030 | -0.060 | -0.080 | 0.040 | -0.020 | -0.100 | -0.090 |
| | | Coverage % | 92 | 96 | 94 | 95 | 95 | 92 | 96 | 93 | 92 | 91 |
| | | Standardized bias | -4 | -11 | 18 | 17 | -21 | -25 | 11 | -4 | -41 | -35 |
| 4 | 10% ² ⁴ ICC _{PSU} = 0.15 ⁴ ICC _{HH} = 0.3 | Bias | 0.030 | 0.006 | 0.010 | 0.010 | -0.060 | -0.070 | 0.006 | 0.050 | -0.080 | -0.110 |
| | | Coverage % | 94 | 93 | 95 | 96 | 91 | 93 | 94 | 94 | 95 | 94 |
| | | Standardized bias | 9 | 5 | 5 | 4 | -18 | -18 | 2 | 15 | -34 | -39 |

1-Mid-range prevalence =30%. 2-Low prevalence =10%. 3-Low ICCs: ICC_{PSU} = 0.05; ICC_{HH} = 0.2. 4-High ICCs: ICC_{PSU} = 0.15; ICC_{HH} = 0.

scenarios 1, 2 and 3 but relatively greater bias for random components estimates were observed as compared to scenarios 1 and 2. These biases are not alarming but in comparison to scenarios 1 and 2 they are somewhat large. The coverage for 95% Wald CIs of the simulated model for scaling method 2 is close to the nominal value for fixed and random effect parameter estimates. Mostly coverage is falling within 2 SEs of the nominal coverage probability (p). The standardized bias is less than 40% for all scenarios for scaling method 2 except for a variance component at household level of scenario 3 (41%). However, the latter is a rather small deviation.

5. Discussion

Parameter estimates of the 3-level logistic regression model obtained from un-weighted and weighted (scaling methods 1 and 2) analysis of NHSP data converged to similar results with few exceptions. Relatively small ICC, and a large number of PSUs and households are probably the key components of consistency in the results of the weighted and un-weighted analyses in the present application.

Our simulation results showed that the performance of the scaled weighted estimators is satisfactory for scaling methods (1 and 2) for 3-level data with a binary outcome in all scenarios that we considered. The results of the analysis of health care utilization from the NHSP data are consistent with the results of the simulation study with regard to agreement between the two scaling methods.

When substantial divergence is found in inferential conclusion, it suggests

that sampling weights are highly informative (the design weights are correlated to the response) and also the ICC may be larger [1]. Informative weights might be considered an important reason for the difference in the results of weighted and un-weighted analyses. We correlated the design weights with the outcome using t-test for two independent samples to check informativeness at each level in the NHSP data. The weights did not differ much at both levels (PSU level and household level) with respect to the binary response variable (p-value 0.04 and 0.045 respectively).

In our analysis of NHSP the standard errors for both the scaling methods are larger than those for the un-weighted analysis. The larger standard errors under the two scaled weighted analyses (methods 1 and 2) showed marginal divergence in inferential conclusion compared to un-weighted analysis.

Simulation studies [1] and [8] have recommended scaling method 1 for estimating between cluster variance. Our simulation results also show less bias for between cluster variance for MPML estimates for scaling method 1 for three-level binary data. So we refer to the results of scaling method 1 for reporting the variance of the NHSP data. The performance of MPML is affected by the intra-class correlation [20]. As we observed in the multilevel analysis of the NHSP data, the ICCs are relatively small for both the un-weighted and weighted analysis. This could be a reason for agreement in the results of the un-weighted and scaled-weighted methods.

We are unaware of any simulation study reported to date that has assessed the impact of different prevalence of the binary outcome, and ICCs at different clustering levels for three-level complex survey data on the bias of estimators of MPML using the two scaling methods. The results of our simulation will provide guidelines for future surveys with regard to obtaining unbiased estimates of MPML weighted analysis for three-level data with a binary outcome.

A simulation study assessed the influence of the prevalence of the outcome, and that of ICC for different sample sizes on the estimates of maximum marginal likelihood for two-level data. However, this study did not involve a complex survey design with weights [21]. Large bias was observed for low prevalence (10%) of the outcome when the size of the cluster was small, that is 5. Similarly, large bias was found for different values of ICCs with small sample size. The influence of ICC and prevalence seemed to be more pronounced when sample sizes were small. However in our simulation this is not the case probably due to the large sample size. In addition, Moineddin *et al.* [21] did not vary prevalence of the outcome and ICC simultaneously, whereas we varied these for 3rd level and 2nd level clusters simultaneously. Though we did not find any particular pattern in the bias with respect to varying prevalence of the outcome and ICCs, the bias of the variance component increased slightly when the ICC increased for both scaling methods. The sample weight scaling methods considered in this paper can be biased when the number of clusters is small. We can explore this further in future research by varying sample sizes at different levels for the various scenarios of our simulation. We can conclude that both scaling methods 1 and 2 are

effective for complex survey designs for the sample size, cluster size, ICC and range of binary outcome prevalence we have considered.

Contextual phenomena are highly important for public health research. To measure the contextual aspect of health care utilization we used multilevel measures of variance and intra-cluster correlation in addition to fixed effects at different levels of clustering. Our analysis indicates that in rural areas more developed communities seek significantly more health care than lesser developed communities. However, in urban areas there was no significant difference among communities with different levels of development. Communities with low development index are illiterate and economically deprived. In Pakistan health care centers of good quality are generally located in urban areas, hence it seems that subjects from communities with low development index in rural areas do not have the resources and/or awareness to travel to urban areas for seeking health care. In addition, our analysis indicates marked differences with regard to health care utilization among the four provinces of Pakistan; Punjab being the most developed in this regard and Baluchistan being the least developed. Underdevelopment of Baluchistan in general is a phenomenon that the present Government and politicians of Pakistan are advocating to address. Moreover, our analysis indicates that the socio-economic status of a household is associated with health care utilization; households belonging to higher socio-economic status seek significantly more health care. The people of Pakistan need to make progress in many areas including health behavior, and health care practices, and improve economic and social policy to improve the nation's overall health.

In our analysis of NHSP we found that social and economic factors have created differences in the distribution of health utilization among different individuals, households and geographical areas. The epidemiological vision of multilevel analysis must therefore focus on measures of health variation that provide information about the distribution and determinants of health status within multiple contexts.

The investigation of variance components (or ICCs) in multilevel models for health care research provide additional insight than odds ratios and regression coefficients that are the traditional measures of association. If accurate information on contextual effects is provided to health policy makers, their decisions or policies could be very efficient in reducing health inequalities.

Data analysis is critical to formation of policy and program in the health sector, yet in low and middle income countries like Pakistan, there is little use of statistical information [22]. The policy implication of this study for Pakistan lies primarily in building confidence in the quality of data produced in the country and its subsequent analysis. Policy makers in Pakistan frequently make decisions without data (national and provincial surveys being weak). This study gives advocates a stronger position in relation to decision makers in the government, as they marshal data to promote their policies for reform. As we discussed earlier well designed economic policies, development of under developed geographical

regions, the launch of mass health education and health promotion campaigns could be helpful to address the issues of unawareness about health care, its availability and accessibility to the Pakistani population.

Acknowledgements

We express our heartiest gratitude to the Aga Khan University (AKU) for funding this study through a “Faculty Development Award” (FDA). We thank Dr Andrew Titman (Lecturer, Lancaster University) for helping in the procedure for the simulation. We thank Dr Nida Zahid (Instructor, Aga Khan University) for references of the manuscript.

Conflict of Interest

The authors have declared no conflict of interest.

References

- [1] Carle, A.C. (2009) Fitting Multilevel Models in Complex Survey Data with Design Weights: Recommendations. *BMC Medical Research Methodology*, **9**, 49.
- [2] Goldstein, H. (2003) Multilevel Statistical Models. 3rd Edition.
- [3] Rabe-Hesketh, S. and Skrondal, A. (2006) Multilevel Modelling of Complex Survey Data. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, **169**, 805-827. <https://doi.org/10.1111/j.1467-985X.2006.00426.x>
- [4] Grilli, L. and Pratesi, M. (2004) Weighted Estimation in Multilevel Ordinal and Binary Models in the Presence of Informative Sampling Designs. *Survey Methodology*, **30**, 93-104.
- [5] Chambers, R.L. and Skinner, C.J. (2003) Analysis of Survey Data. John Wiley & Sons. <https://doi.org/10.1002/0470867205>
- [6] Longford, N.T. (1996) Model-Based Variance Estimation in Surveys with Stratified Clustered Designs. *Australian Journal of Statistics*, **38**, 333-352. <https://doi.org/10.1111/j.1467-842X.1996.tb00687.x>
- [7] Pfeffermann, D., Skinner, C.J., Holmes, D.J., Goldstein, H. and Rasbash, J. (1998) Weighting for Unequal Selection Probabilities in Multilevel Models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **60**, 23-40. <https://doi.org/10.1111/1467-9868.00106>
- [8] Asparouhov, T. (2006) General Multi-Level Modeling with Sampling Weights. *Communications in Statistics—Theory and Methods*, **35**, 439-460. <https://doi.org/10.1080/03610920500476598>
- [9] Stapleton, L.M. (2002) The Incorporation of Sample Weights into Multilevel Structural Equation Models. *Structural Equation Modeling*, **9**, 475-502. https://doi.org/10.1207/S15328007SEM0904_2
- [10] Jenkins, F. (2008) Multilevel Analysis with Informative Weights. *Proceedings of the Joint Statistical Meeting*, ASA Section on Survey Research Methods, 2225-2233.
- [11] Potthoff, R.F., Woodbury, M.A. and Manton, K.G. (1992) Equivalent Sample Size and Equivalent Degrees of Freedom Refinements for Inference Using Survey Weights under Superpopulation Models. *Journal of the American Statistical Association*, **87**, 383-396. <https://doi.org/10.1080/01621459.1992.10475218>
- [12] Mallick, M.D. (1992) Sample Design for the National Health Survey of Pakistan. *Pakistan Journal of Medical Sciences*, **31**, 289-290.

- [13] Hadden, W.C., Pappas, G. and Khan, A.Q. (2003) Social Stratification, Development and Health in Pakistan: An Empirical Exploration of Relationships in Population-Based National Health Examination Survey Data. *Social Science & Medicine*, **57**, 1863-1874. [https://doi.org/10.1016/S0277-9536\(03\)00052-2](https://doi.org/10.1016/S0277-9536(03)00052-2)
- [14] Skrondal, R.-H. (2008) Multilevel and Longitudinal Modeling Using Stata. STATA Press, College Station.
- [15] Fatmi, Z., Hadden, W.C., Razzak, J.A., Qureshi, H.I., Hyder, A.A. and Pappas, G. (2007) Incidence, Patterns and Severity of Reported Unintentional Injuries in Pakistan for Persons Five Years and Older: Results of the National Health Survey of Pakistan 1990-94. *BMC Public Health*, **7**, 152. <https://doi.org/10.1186/1471-2458-7-152>
- [16] Pappas, G., Akhtar, T., Gergen, P.J., Hadden, W.C. and Khan, A.Q. (2001) Health Status of the Pakistani Population: A Health Profile and Comparison with the United States. *American Journal of Public Health*, **91**, 93. <https://doi.org/10.2105/AJPH.91.1.93>
- [17] Burton, A., Altman, D.G., Royston, P. and Holder, R.L. (2006) The Design of Simulation Studies in Medical Statistics. *Statistics in Medicine*, **25**, 4279-4292. <https://doi.org/10.1002/sim.2673>
- [18] Collins, L.M., Schafer, J.L. and Kam, C.-M. (2001) A Comparison of Inclusive and Restrictive Strategies in Modern Missing Data Procedures. *Psychological Methods*, **6**, 330. <https://doi.org/10.1037/1082-989X.6.4.330>
- [19] Snijders, T.A.B. and Bosker, R.J. (1999) Multilevel Analysis. An Introduction to Basic and Advanced Multilevel Modeling. Sage, London.
- [20] Kovacevic, M.S. and Rai, S.N. (2003) A Pseudo Maximum Likelihood Approach to Multilevel Modelling of Survey Data. *Communications in Statistics-Theory and Methods*, **32**, 103-121. <https://doi.org/10.1081/STA-120017802>
- [21] Moineddin, R., Matheson, F.I. and Glazier, R.H. (2007) A Simulation Study of Sample Size for Multilevel Logistic Regression Models. *BMC Medical Research Methodology*, **7**, 34. <https://doi.org/10.1186/1471-2288-7-34>
- [22] AbouZahr, C. and Boerma, T. (2005) Health Information Systems: The Foundations of Public Health. *Bulletin of the World Health Organization*, **83**, 578-583.



Submit or recommend next manuscript to SCIRP and we will provide best service for you:

Accepting pre-submission inquiries through Email, Facebook, LinkedIn, Twitter, etc.

A wide selection of journals (inclusive of 9 subjects, more than 200 journals)

Providing 24-hour high-quality service

User-friendly online submission system

Fair and swift peer-review system

Efficient typesetting and proofreading procedure

Display of the result of downloads and visits, as well as the number of cited articles

Maximum dissemination of your research work

Submit your manuscript at: <http://papersubmission.scirp.org/>

Or contact ojepi@scirp.org