**Communication in Physical Sciences, 2024, 12(1): 038-051**

# Assessing the Efficiencies of Calibration Ratio Estimators for Estimating Mean Weight of Babies in the Presence of Gestational Age Under Stratified Random Sampling

**Theophilus Obijuru Nelson and Emmanuel John Ekpenyong**

**Received: 12 September 2024/Accepted: 11 November 2024/First Published: 24 November 2024, doi:https://dx.doi.org/10.4314/cps.v12i1.5**

***Abstract:*** *This study compares the performance of various calibration ratio estimators in estimating the mean weight of newborn babies at the maternity ward of Federal Medical Centre (FMC), Umuahia, Abia State Nigeria, under stratified random sampling. Data were collected on maternal age, height, weight, and baby-related variables such as weight, gestational age and height. The maternal body mass index (BMI) was utilized as the stratification variable. The aim is to improve the accuracy of estimating mean weight of babies by incorporating gestational age as an auxiliary variable. An empirical study was carried out through population data sets obtained as to ascertain the efficiency and performance of various calibration ratio estimators considered in the study, and the results revealed that the estimator of Audu et al. (2020) outperformed the other estimators, and has proven to be consistent in all cases of sample size selection and the tuning parameter.*

***Keywords:*** *Calibration estimation, Efficiency, Auxiliary information, stratified random sampling, Gestational age.*

**Theophilus Obijuru Nelson**
Department of Statistics, Michael Okpara University of Agriculture, Umudike, Abia State Nigeria
**Email: nelson.theophilus@mouau.edu.ng**
**Orcid id: 0009-0008-7008-5560**

**Emmanuel John Ekpenyong**
Department of Statistics, Michael Okpara University of Agriculture, Umudike, Abia State Nigeria
**Email: ej.ekpenyong@mouau.edu.ng Orcid id: 0000-0003-4112-3712**

## 1.0 Introduction

Birth weight is a vital indicator of an infant's chances of survival, healthy growth and development. Good maternal health and nutrition, among other factors, significantly influence birth weight and subsequently impact child survival (Fall, 2009). A healthy and well-nourished mother is essential for a baby's optimal growth and development, whereas maternal malnutrition can lead to low birth weight and other complications. However, precise estimation of mean weight of babies is important for healthcare planning, policy-making and research (Kramer and Kakumar, 2001). Baby's birth weight is an essential indicator of neonatal health, and its estimation aids in revealing high risk groups (Alexander *et al.*, 1996). Gestational age is an important determinant of birth weight, as premature babies tend to weigh less than full-term babies. Also, other factors such as maternal body mass index (BMI) can influence birth weight (Catalano *et al.*, 2007). Body mass index is widely used to stratify individuals into weight strata or status categories (World Health Organization, 2016). Studies have shown that maternal body mass index have significant effect on birth weight (O' Tierney-Ginn and Loftin, 2017). Women with higher body mass index tend to have heavier babies, thereby increasing the risk of gestational diabetes, hypertension and cesarean delivery (Catalano and Shankar, 2012). Stratification is used to improve precision in estimation. By stratifying the population based on relevant variables, such as body mass index, estimators can account for differences within strata.

Calibration estimation, introduced by Deville and Särndal (1992) incorporates auxiliary information to improve the precision of the population estimates. It is a general technique of adjusting the original weights with the minimization of a given distance function based on a set of constraints under auxiliary information. Singh *et al.* (1998), extended the higher-level calibration approach in stratified random sampling, which is used to obtain optimum strata weights for improving the precision of the survey estimates of population mean. Singh (2003), Tracy *et al*. (2003), Rao *et al.* (2016), Nidhi *et al.* (2017), Sisodia *et al*. (2017), Ozgul (2018), Garg and Pachori (2019) and Alam *et al.* (2019) suggested calibration estimators for estimating the population mean in stratified sampling using different calibration constraints based on single auxiliary information. Clement (2015, 2017), Nidhi *et al*. (2018), Audu *et al.* (2020), Khare *et al*. (2022), and Muilli *et al*. (2022), among many others, have applied calibration estimation to ratio-type estimators in stratified sampling.

However, this paper seeks to compare the efficiency and performance of some calibration ratio estimators in estimating the mean weights of babies using gestational age as an auxiliary variable in stratified random sampling.

## 2. 0 Sampling procedure, Notations and Review of Existing Estimators

Consider the problem of estimating the population mean $\bar{Y}$ for a finite population of size N. Let a finite population $\Theta = \left(\Theta_1, \Theta_2, ..., \Theta_N\right)$, consist of N units. Let $\Upsilon = \left(\Upsilon_1, \Upsilon_2, ..., \Upsilon_n\right)$ be the set of sampled units under simple random sampling without replacement (SRSWOR). Let $y = 1, 2, ..., N$, be the study variable (Baby weight) and let $x = 1, 2, ..., N$, be the auxiliary variable (Gestational age) linearly related to *y*. Let the population be divided into $\Omega$ nonoverlapping homogeneous strata (stratified based maternal body mass index) and $n_d$ be the number of units drawn by simple random sampling without replacement (SRSWOR) from the $d^{th}$ stratum consisting of $N_d$ units,

$n = \sum_{d=1}^{\Omega} n_d$, $N = \sum_{d=1}^{\Omega} N_d$ give the total sample size and the population size. For the $d^{th}$ strata, let $\omega_d = N_d N^{-1}$ be the strata weights. Let the estimation of unknown population mean $\bar{Y}$ be of interest using the information from an auxiliary variable *X* and $\left(y_{di}, x_{di}\right)$ be the $i^{th}$ unit of the study variable and auxiliary variable respectively in the $\tau^{th}$ strata, where $i = 1, 2, .., n_d$ and $d = 1, 2, .., \Omega$.

$\bar{y}_d = n_d^{-1} \sum_{i=1}^{n_d} y_{di}$, $\bar{x}_d = n_d^{-1} \sum_{i=1}^{n_d} x_{di}$, $\bar{Y}_d = N_d^{-1} \sum_{i=1}^{N_d} Y_{di}$, $\bar{X}_d = N_d^{-1} \sum_{i=1}^{N_d} X_{di}$ are the sample and population means for the study variable and the auxiliary variable, respectively. However, it is assumed that the population mean $\bar{X} = \sum_{d=1}^{\Omega} \omega_d \bar{X}_d$ is completely known (Ozgul, 2018).

The mean weight of babies is given as

$$\bar{Y} = \sum_{d=1}^{\Omega} \omega_d \bar{Y}_d \qquad (1)$$

By using the gestational age ( $X$ ) as auxiliary variable, which improves the estimate. Let

$$t^{ST} = \sum_{d=1}^{\Omega} \omega_d \bar{y}_d \quad (2)$$

be the stratified random sampling estimator of the mean weight of babies in (1), according to Cochran (1977).

The estimated variance of the estimator of the mean weight of babies is given by

$$\hat{V}ar\left(t^{ST}\right) = \sum_{d=1}^{\Omega} \omega_d^2 \psi_d s_{yd}^2 \quad (3)$$

The weights $\omega_d$ are modified or adjusted to obtain new calibration weights through minimization of the distance measure subject to some calibration constraints to enhance or improve the efficiency or precision of the estimator in (2) by applying a calibration approach.

The following notations and terminologies are defined below.

$\gamma_d = n_d N_d^{-1}$ : Sampling fraction,

$S_{yd}^2 = \left(N_d - 1\right)^{-1} \sum_{i=1}^{N_d} \left(Y_{di} - \bar{Y}_d\right)^2$ : Population variance for the $d^{th}$ stratum of the study variable,

$S_{xd}^2 = \left(N_d - 1\right)^{-1} \sum_{i=1}^{N_d} \left(X_{di} - \bar{X}_d\right)^2$ : Population variance for the $d^{th}$ stratum of the auxiliary variable,

$s_{yd}^2 = \left(n_d - 1\right)^{-1} \sum_{i=1}^{n_d} \left(y_{di} - \bar{y}_d\right)^2$ : sample variance for the $d^{th}$ stratum of the study variable,

$s_{xd}^2 = \left(n_d - 1\right)^{-1} \sum_{i=1}^{n_d} \left(x_{di} - \bar{x}_d\right)^2$ : sample variance for the $d^{th}$ stratum of the auxiliary variable,

$S_{yd}$ : Population standard deviation for the $d^{th}$ stratum of the of the study variable,

$S_{xd}$ : Population standard deviation for the $d^{th}$ stratum of the auxiliary variable,

$s_{yd}$ : sample standard deviation for the $d^{th}$ stratum of the of the study variable,

$s_{xd}$ : sample standard deviation for the $d^{th}$ stratum of the auxiliary variable,

$C_{yd} = S_{yd} \bar{Y}_d^{-1}$ : Population coefficient variation for the $d^{th}$ stratum of the study variable,

$C_{xd} = S_{xd} \bar{X}_d^{-1}$ : Population coefficient variation for the $d^{th}$ stratum of auxiliary variable,

$c_{yd} = s_{yd} \bar{y}_d^{-1}$ : Sample coefficient variation for the $d^{th}$ stratum of the study variable,

$c_{xd} = s_{xd} \bar{x}_d^{-1}$ : sample coefficient variation for the $d^{th}$ stratum of auxiliary variable,

$S_{xyd} = \left(N_d - 1\right)^{-1} \sum_{i=1}^{N_d} \left(X_{di} - \bar{X}_h\right)\left(Y_{di} - \bar{Y}_d\right)$ : Population covariance between auxiliary and study variables for the $d^{th}$ stratum,

$s_{xyd} = \left(n_d - 1\right)^{-1} \sum_{i=1}^{n_d} \left(x_{di} - \bar{x}_h\right)\left(y_{di} - \bar{y}_d\right)$ : Sample covariance between auxiliary and study variables for the $d^{th}$ stratum,

$\rho_{xyd} = S_{xyd}\left(S_{xd}S_{yd}\right)^{-1}$ : Population correlation coefficient between auxiliary and study variables for the $d^{th}$ stratum,

$r_{xyd} = s_{xyd}\left(s_{xd}s_{yd}\right)^{-1}$ : Sample correlation coefficient between auxiliary and study variables for the $d^{th}$ stratum,

$R_d = \bar{Y}_d \bar{X}_d^{-1}$ : Population ratio of the auxiliary variable to the study variable for the $d^{th}$ stratum,

$\hat{R}_d = \bar{y}_d \bar{x}_d^{-1}$ : sample ratio of the auxiliary variable to the study variable for the $d^{th}$ stratum and

$\psi_d = \left(1-\gamma_d\right)n_d^{-1}$ : Correction factor.

Under stratified random sampling using a single auxiliary variable, different calibration ratio estimators have been proposed. Some of the existing calibration ratio estimators are reviewed below.

Clement (2015) introduced calibration estimation to separate the ratio estimator for the population mean of the study variable $Y$ using the auxiliary variable $X$ in stratified sampling. The proposed estimator is defined as

$$t^C = \sum_{d=1}^{\Omega} \delta_d^C \hat{R}_d \bar{x}_d \tag{4}$$

where, $\delta_\tau^C$ are the new calibrated weights obtained by minimizing (5) for (6) and $\sigma_d = N_d n_d^{-1}$

$$G_c = \sum_{d=1}^{\Omega} \left(\delta_d^C \sigma_d^{-1} - 1\right)^2 \left(\Phi_d\right)^{-1} \tag{5}$$

$$\sum_{d=1}^{\Omega} \delta_d^C \bar{x}_d = \sum_{d=1}^{\Omega} \omega_d \bar{X}_d \tag{6}$$

The calibration weights $\delta_d^C$ was obtained as

$$\delta_d^C = \sigma_d + \left(\bar{X} - \sum_{d=1}^{\Omega} \delta_d^C \bar{x}_d\right)\left(\Phi_d \sigma_d \bar{x}_d\right)\left(\sum_{d=1}^{\Omega} \Phi_d \sigma_d \bar{x}_d^2\right)^{-1} \tag{7}$$

Thus, the proposed calibration approach separate ratio type estimator in stratified sampling is given as

$$t^C = \sum_{d=1}^{\Omega} \sigma_d \hat{R}_d \bar{x}_d + \left(\sum_{d=1}^{\Omega} \Phi_d \sigma_d \bar{x}_d \hat{R}_d^2\right)\left(\bar{X} - \sum_{\tau=1}^{\Omega} \sigma_d \bar{x}_d\right)\left(\sum_{d=1}^{\Omega} \Phi_d \sigma_d \bar{x}_d^2\right)^{-1} \tag{8}$$

Clement (2017) introduced a new improved separate ratio estimator for population mean in stratified random sampling using a calibration estimation approach. The proposed estimator is defined as

$$t^{CP} = \sum_{d=1}^{\Omega} \delta_d^{CP} \bar{y}_d \Lambda \tag{9}$$

where $\Lambda = \left[\lambda_d \bar{X}_d \bar{x}_d^{-1} + \left(1-\lambda_d\right)\bar{x}_d \bar{X}_d^{-1}\right]$ and $\delta_d^{CP}$ are the new calibrated weights obtained by minimizing (10) for equation 11, .

$$G_{cp} = \sum_{d=1}^{\Omega} \left(\delta_d^{CP} - \omega_d\right)^2 \left(\omega_d \Phi_d\right)^{-1} \tag{10}$$

$$\sum_{d=1}^{\Omega}\delta_d^{CP}s_{xd}^2 = V\left(\bar{x}_{st}\right) \tag{11}$$

The calibration weights $\delta_d^{CP}$ was obtained as

$$\delta_d^{CP} = \omega_d + \left(V\left(\bar{x}_{st}\right) - \sum_{d=1}^{\Omega}\omega_d s_{xd}^2\right)\left(\omega_d \Phi_d s_{xd}^2\right)\left(\sum_{d=1}^{\Omega}\omega_d \Phi_d \left(s_{xd}^2\right)^2\right)^{-1} \tag{12}$$

Thus, the proposed calibration estimator by Clement (2017) is given as

$$t^{CP} = \sum_{d=1}^{\Omega}\left[\omega_d + \left(V\left(\bar{x}_{st}\right) - \sum_{d=1}^{\Omega}\omega_d s_{xd}^2\right)\left(\omega_d \Phi_d s_{xd}^2\right)\left(\sum_{d=1}^{\Omega}\omega_d \Phi_d \left(s_{xd}^2\right)^2\right)^{-1}\right]\bar{y}_d \Lambda \tag{13}$$

where $V\left(\bar{x}_{st}\right) = \sum_{d=1}^{\Omega}\omega_d^2 \psi_d S_{xd}^2$ and $\hat{V}\left(\bar{x}_{st}\right) = \sum_{d=1}^{\Omega}\omega_d S_{xd}^2$

Nidhi *et al.* (2018) proposed a calibration-based ratio-type estimation of the mean in stratified sampling. The proposed estimator is given as

$$t^{ND} = \sum_{d=1}^{\Omega}\delta_d^{ND}\bar{y}_{ds} \tag{14}$$

where $\bar{y}_{ds} = \bar{y}_d\left(\frac{\bar{X}_d + C_{xd}}{\bar{x}_d + c_{xd}}\right)$ and $\delta_d^{ND}$ are the new calibrated weights obtained by minimizing (15) concerning (16)

$$G_{nd} = \sum_{d=1}^{\Omega}\left(\delta_d^{ND} - \omega_d\right)^2\left(\omega_d \Phi_d\right)^{-1} \tag{15}$$

$$\sum_{d=1}^{\Omega}\delta_d^{ND}\left(\bar{x}_d + c_{xd}\right) = \sum_{d=1}^{\Omega}\omega_d\left(\bar{X}_d + C_{xd}\right), \quad \sum_{d=1}^{\Omega}\delta_d^{ND} = 1 \tag{16}$$

The new calibration weights $\delta_d^{ND}$ was obtained as

$$\delta_d^{ND} = \omega_d + \left[\frac{\left[\omega_d \Phi_d\left(\bar{x}_d + c_{xd}\right)\right]\left(\sum_{d=1}^{\Omega}\omega_d \Phi_d\right) - \left(\omega_d \Phi_d\right)\left(\sum_{d=1}^{\Omega}\omega_d \Phi_d\left(\bar{x}_d + c_{xd}\right)\right)}{\left(\sum_{d=1}^{\Omega}\omega_d \Phi_d\left(\bar{x}_d + c_{xd}\right)^2\right)\left(\sum_{d=1}^{\Omega}\omega_d \Phi_d\right) - \left(\sum_{d=1}^{\Omega}\omega_d \Phi_d\left(\bar{x}_d + c_{xd}\right)\right)^2}\right]\left[\left(\sum_{d=1}^{\Omega}\omega_d\left(\bar{X}_d + c_{xd}\right)\right) - \left(\sum_{d=1}^{\Omega}\omega_d\left(\bar{x}_d + c_{xd}\right)\right)\right] \tag{17}$$

Hence, the proposed calibrated estimator is given as

$$t^{ND} = \sum_{d=1}^{\Omega}\omega_d \bar{y}_{ds} + \hat{\beta}_{nd}\left[\sum_{d=1}^{\Omega}\omega_d\left(\bar{X}_d + C_{xd}\right) - \sum_{d=1}^{\Omega}\omega_d\left(\bar{x}_d + c_{xd}\right)\right] \tag{18}$$

where, $$\hat{v}_{nd} = \frac{\left(\sum_{d=1}^{\Omega}\omega_d \Phi_d\right)\left(\sum_{h=1}^{\Omega}\omega_d \Phi_d \bar{y}_{ds}\left(\bar{x}_d + c_{xd}\right)\right) - \left(\sum_{d=1}^{\Omega}\omega_d \Phi_d \bar{y}_{ds}\right)\left(\sum_{d=1}^{\Omega}\omega_d \Phi_d\left(\bar{x}_d + c_{xd}\right)\right)}{\left(\sum_{d=1}^{\Omega}\omega_d \Phi_d\left(\bar{x}_d + c_{xd}\right)^2\right)\left(\sum_{d=1}^{\Omega}\omega_d \Phi_d\right) - \left(\sum_{d=1}^{\Omega}\omega_d \Phi_d\left(\bar{x}_d + c_{xd}\right)\right)^2}$$

Audu *et al.* (2020) developed a calibration ratio estimator expressed in (19) using the calibration constraints in (20)

$$t^{AU} = \sum_{d=1}^{\Omega}\delta_d^{AU}\bar{y}_d \tag{19}$$

$$\sum_{d=1}^{\Omega}\delta_d^{AU}\bar{x}_d = \sum_{d=1}^{\Omega}\omega_d^{*}\bar{X}_d, \quad \sum_{d=1}^{\Omega}\delta_d^{AU} = \sum_{d=1}^{\Omega}\omega_d^{*} \tag{20}$$

where, $\omega_d^* = \omega_d \bar{X}_d \bar{x}_d^{-1}$ and $\delta_d^{AU}$ are the new weights obtained by minimizing (21) with respect to (20).

$$G_{au} = \sum_{\tau=1}^{\Im} \left( \delta_\tau^{AU} - \eta_\tau^* \right)^2 \left( \eta_\tau^* \Phi_\tau \right)^{-1} \tag{21}$$

The new calibration weights $\delta_d^{AU}$ was obtained as

$$\delta_d^{OA} = \omega_d^* + \frac{\left( \omega_d^* \Phi_d \bar{x}_d \right)\left( \sum_{d=1}^{\Omega} \omega_d^* \Phi_d \right) - \left( \omega_d^* \Phi_d \right)\left( \sum_{d=1}^{\Omega} \omega_d^* \Phi_d \bar{x}_d^2 \right)}{\left( \sum_{d=1}^{\Omega} \omega_d^* \Phi_d \bar{x}_d^2 \right)\left( \sum_{d=1}^{\Omega} \omega_d^* \Phi_d \right) - \left( \sum_{d=1}^{\Omega} \omega_d^* \Phi_d \bar{x}_d^2 \right)^2} \left( \sum_{d=1}^{\Omega} \omega_d^* \left( \bar{X}_d - \bar{x}_d \right) \right) \tag{22}$$

Thus, the proposed calibration estimator by Audu *et al.* (2020) is given as

$$t^{AU} = \sum_{d=1}^{\Omega} \omega_d \bar{X}_d \bar{x}_d^{-1} \left[ \bar{y}_d + \hat{v}_{au} \left( \bar{X}_d - \bar{x}_d \right) \right] \tag{23}$$

$$\hat{v}_{au} = \frac{\left( \sum_{d=1}^{\Omega} \omega_d \Phi_d \bar{X}_d \bar{x}_d^{-1} \right)\left( \sum_{d=1}^{\Omega} \omega_d \Phi_d \bar{X}_d \bar{y}_d \right) - \left( \sum_{d=1}^{\Omega} \omega_d \Phi_d \bar{X}_d \right)\left( \sum_{d=1}^{\Omega} \omega_d \Phi_d \bar{X}_d \bar{x}_d^{-1} \bar{y}_d \right)}{\left( \sum_{\tau=1}^{\Omega} \eta_\tau \Phi_\tau \bar{X}_\tau \bar{x}_\tau \right)\left( \sum_{d=1}^{\Omega} \omega_d \Phi_d \bar{X}_d \bar{x}_d^{-1} \right) - \left( \sum_{d=1}^{\Omega} \omega_d \Phi_d \bar{X}_d \bar{x}_d \right)^2}$$

where,

Khare *et al.* (2022) proposed a separate ratio estimator using a calibration approach for the population mean using stratified random sampling which is given by

$$t^{KH} = \sum_{d=1}^{\Omega} \delta_d^{KH} \bar{X}_d \bar{x}_d^{-1} \bar{y}_d \tag{24}$$

where $\delta_d^{KH}$ are the new calibrated weights obtained by minimizing (25) concerning (26)

$$G_{kh} = \sum_{d=1}^{\Omega} \left( \delta_d^{KH} - \omega_d \right)^2 \left( \omega_d \Phi_d \right)^{-1} \tag{25}$$

$$\sum_{d=1}^{\Omega} \delta_d^{KH} \bar{x}_d = \sum_{d=1}^{\Omega} \omega_d \bar{X}_d , \quad \sum_{d=1}^{\Omega} \delta_d^{KH} = 1 \tag{26}$$

The new calibration weights $\delta_d^{KH}$ was obtained as

$$\delta_d^{KH} = \omega_d + \frac{\left( \omega_d \Phi_d \bar{x}_d \right)\left( \sum_{d=1}^{\Omega} \omega_d \Phi_d \right) - \left( \omega_d \Phi_d \right)\left( \sum_{d=1}^{\Omega} \omega_d \Phi_d \bar{x}_d \right)}{\left( \sum_{d=1}^{\Omega} \omega_d \Phi_d \bar{x}_d^{\,2} \right)\left( \sum_{d=1}^{\Omega} \omega_d \Phi_d \right) - \left( \sum_{d=1}^{\Omega} \omega_d \Phi_d \bar{x}_d \right)^2} \left( \bar{X} - \sum_{d=1}^{\Omega} \omega_d \bar{x}_d \right) \tag{27}$$

Thus, the proposed calibration estimator is given as

$$t^{KH} = \sum_{d=1}^{\Omega} \omega_d \bar{X}_d \bar{x}_d^{-1} \bar{y}_d + \hat{v}_{kh} \left( \bar{X} - \sum_{d=1}^{\Omega} \omega_d \bar{x}_d \right) \tag{28}$$

where

$$\hat{v}_{kh} = \frac{\left( \sum_{\tau=1}^{\Im} \eta_\tau \Phi_\tau \right)\left( \sum_{\tau=1}^{\Im} \eta_\tau \Phi_\tau \bar{y}_\tau \bar{X}_\tau \right) - \left( \sum_{\tau=1}^{\Im} \eta_\tau \Phi_\tau \bar{X}_\tau \bar{x}_\tau^{-1} \bar{y}_\tau \right)\left( \sum_{\tau=1}^{\Im} \eta_\tau \Phi_\tau \bar{x}_\tau \right)}{\left( \sum_{\tau=1}^{\Im} \eta_\tau \Phi_\tau \right)\left( \sum_{\tau=1}^{\Im} \eta_\tau \Phi_\tau \bar{x}_\tau^2 \right) - \left( \sum_{\tau=1}^{\Im} \eta_\tau \Phi_\tau \bar{x}_\tau \right)^2}$$

## 3. 0 Results and Discussion

### *3.1 Empirical Study*

To compare the performance and efficiency of the calibration ratio estimators, real data on weights (kilogram), heights (metre) and ages (year or week) of mothers and babies from the maternity ward of Federal Medical Centre (FMC), Umuahia consisting of (N=430) individuals from February 2019 to August 2024 are considered. This dataset is further divided into four (4) categories (Strata) using the World Health Organization standard on body mass index (BMI) classification. The BIM of the mothers were used to stratify the data into underweight (BMI<18.5), normal (BMI=18.5-24.9), overweight (BMI=25-29.9) and Obese (BMI $\geq$30). Thus, the stratum sizes for these 4 strata are $N_1 = 19$, $N_2 = 314$, $N_3 = 87$ and $N_4 = 10$. In this study, the weight of babies is considered as a variable of interest (Y) utilizing gestational age as the auxiliary variable(X), having correlation ($\rho_{xy}$) 0.5121 between them. The true mean weight of babies is $\bar{Y} = 3.093488$. Further, samples of sizes *n* = 100 and 200 units are selected 50,000 times by simple random sampling without replacement (SRSWOR) from each stratum respectively using proportional allocation, that is,

$$n_d = nN_d N^{-1} = n\omega_d \text{, } d = 1,...,4 \qquad (30)$$

With known stratum means $\bar{X}_d$, the value of $\bar{Y}$ (assuming unknown) has been calculated using, a simple stratified estimator defined in Equation (2) and calibrated ratio estimators of Clement (2015), Clement (2017), Nidhi *et al.* (2018), Audu *et al.* (2020) and Khare *et al.* (2022) defined in Equations (8), (13), (18), (23) and (28) respectively. Thus, the performance of the estimators has been assessed using mean square error (MSE) defined in Equations (31) and (32) and percent relative efficiency (PRE) defined in Equation (33). An estimator with the least MSE and highest PRE is judged to have higher efficiency over other estimators.

$$MSE\left(\hat{\tau}_j\right) = \frac{1}{50000}\sum_{j=1}^{50000}\left(\hat{\tau}_j - \bar{Y}\right)^2 \qquad (31)$$

$$MSE\left(t^{ST}\right) = \frac{1}{50000}\sum_{j=1}^{50000}\left(t_j^{ST} - \bar{Y}\right)^2 \qquad (32)$$

$$PRE\left(\hat{\tau}_j\right) = \frac{MSE\left(t^{ST}\right)}{MSE\left(\hat{\tau}_j\right)} \times 100 \qquad (33)$$

where $\hat{\tau}_j = t^{ST},\ t^{C},\ t^{CP}, t^{ND},\ t^{AU}\ and\ t^{KH}$ are the estimators of Sample mean, Clement (2015), Clement (2017), Nidhi *et al.* (2018), Audu *et al.* (2020) and Khare *et al.* (2022) defined in Equations (7), (12), (17), (22) and (27) respectively.

The description of the population datasets is displayed on the various graphs below**.**

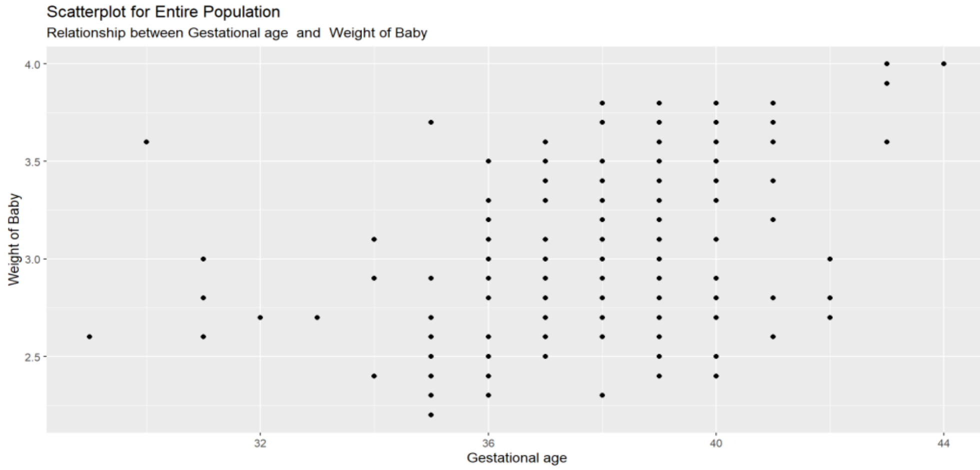


**Fig. 1: Scatterplot for the entire population**

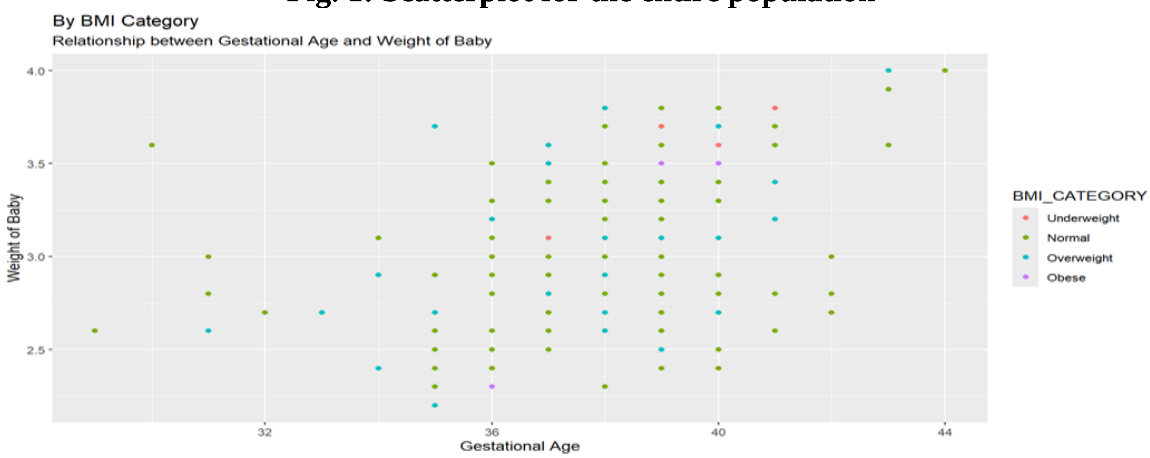


**Fig. 2: Scatterplot showing the population by BMI category**

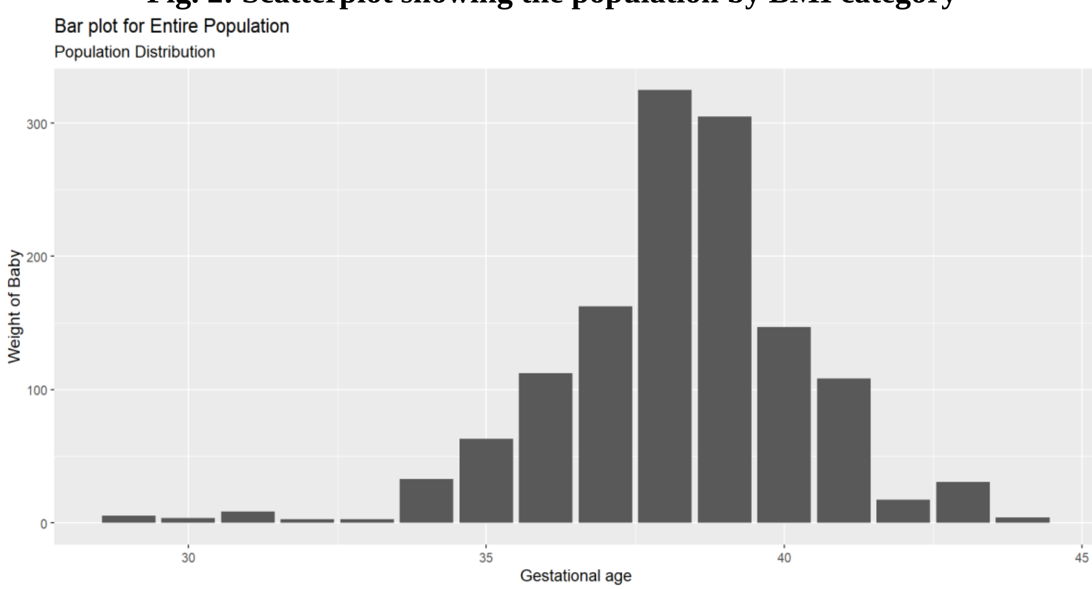


**Fig. 3: Bar plot showing the population distribution**

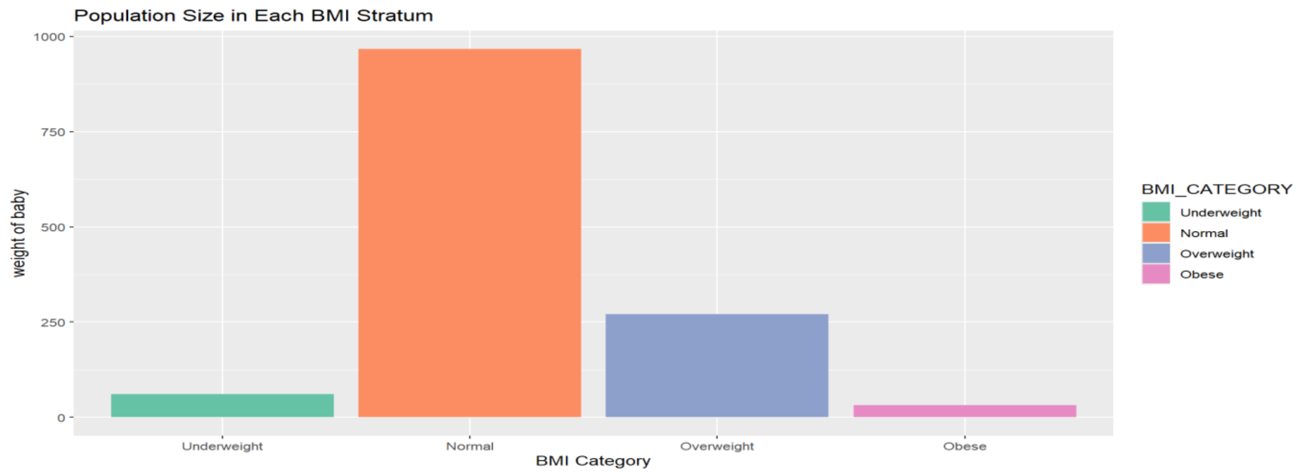


**Fig. 4: Baby's weight population size in each BMI category**

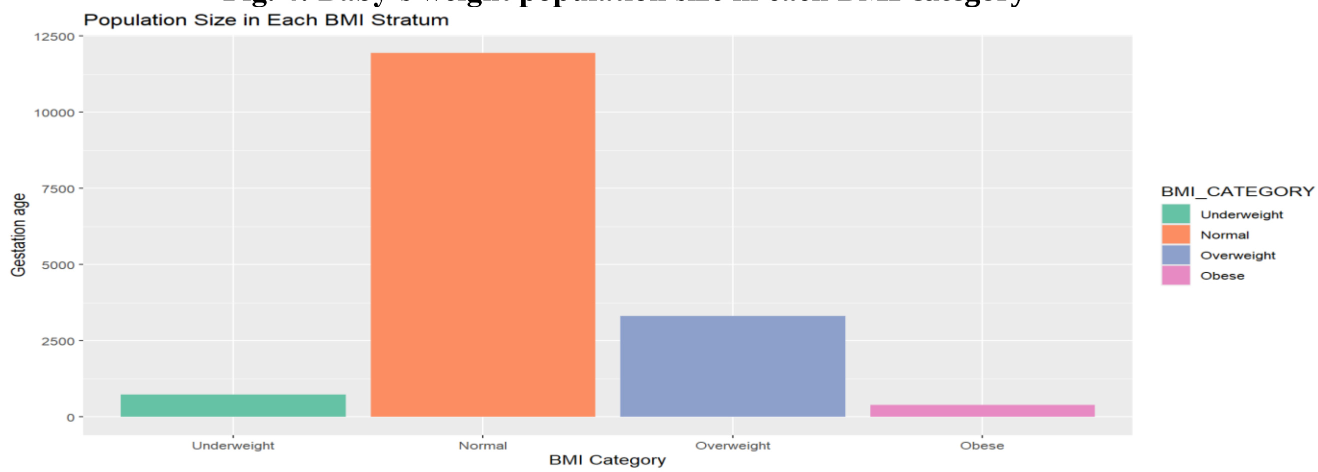


**Fig. 5: Gestational age population size in each BMI category**

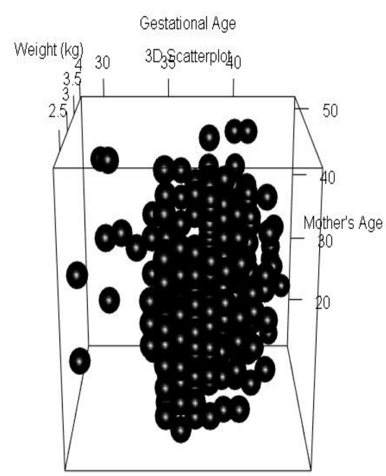


**Fig. 6: 3D scatterplot showing weight of babies, gestational age and age of mothers**

The original and calibrated weights of the estimators are presented in Table 1 below.

**Table 1: Adjusted and original (actual) calibrated weights of the calibrated ratio estimators with different $\Phi_d$**

| | | | | | | |
|---|---|---|---|---|---|---|
| | | | $\Phi_d = 1$ | | | |
| S/N | $\omega_d$ | $\delta_d^{C}$ | $\delta_d^{CP}$ | $\delta_d^{ND}$ | $\delta_d^{AU}$ | $\delta_d^{KH}$ |
| 1. | 0.0442 | 0.1787 | 0.0385 | 0.1047 | 0.0398 | 0.0394 |
| 2. | 0.7302 | 0.2633 | 0.2950 | 0.8928 | 0.7150 | 0.7178 |
| 3. | 0.2023 | 0.3983 | 0.0759 | 0.2481 | 0.2240 | 0.2227 |
| 4. | 0.0233 | 0.1575 | 0.0216 | 0.0620 | 0.0205 | 0.0202 |
| Sum | 1.0000 | 0.9978 | 0.4310 | 1.3076 | 0.9993 | 1.0000 |
| | | | $\Phi_d = \frac{1}{\bar{x}_d}$ | | | |
| 1. | 0.0442 | 0.2548 | 0.0387 | 0.0442 | 0.0399 | 0.1983 |
| 2. | 0.7302 | 0.2307 | 0.2983 | 0.7302 | 0.7146 | 3.3420 |
| 3. | 0.2023 | 0.2333 | 0.0726 | 0.2023 | 0.2242 | 0.9501 |
| 4. | 0.0233 | 0.2682 | 0.0217 | 0.0233 | 0.0206 | 0.1039 |
| Sum | 1.0000 | 0.9871 | 0.4314 | 1.0000 | 0.9993 | 4.5943 |

The true mean weight of babies is $\bar{Y} = 3.093488$. Thus, the estimates of the mean weight of babies are presented in the table below.

**Table 2: Estimates of the mean weight of babies**

| | n=100 | | n=200 | |
|---|---|---|---|---|
| | Estimates | | Estimates | |
| Estimators | $\Phi_d = 1$ | $\Phi_d = \frac{1}{\bar{x}_d}$ | $\Phi_d = 1$ | $\Phi_d = \frac{1}{\bar{x}_d}$ |
| Sample mean $t^{ST}$ | 3.134940 | 3.134940 | 3.107407 | 3.107407 |
| Clement (2015) $t^{C}$ | 3.166629 | 3.193889 | 3.103378 | 3.102655 |
| Clement (2017) $t^{CP}$ | 1.360358 | 1.361930 | 1.494110 | 1.494530 |
| Nidhi *et al.* (2018) $t^{ND}$ | 4.120742 | 3.134940 | 3.440735 | 3.107407 |
| Audu *et al.* (2020) $t^{AU}$ | 3.128341 | 3.128368 | 3.093877 | 3.093963 |
| Khare *et al.* (2022) $t^{KH}$ | 3.130485 | 14.39853 | 3.099879 | 289.3698 |

The calculated mean square errors (MSEs) and percent relative efficiencies (PREs) of the various estimators are presented in the Table 3 below.

**Table 3: MSE and PRE of calibration ratio estimators with different sample sizes**

| n=100 | | | | |
|---|---|---|---|---|
| Estimators | $\Phi_d = 1$ | | $\Phi_d = \frac{1}{\bar{x}_d}$ | |
| | MSE | PRE | MSE | PRE |
| Sample mean $t^{ST}$ | 0.0017182 | 100.00000 | 0.0017182 | 100.00000 |
| Clement (2015) $t^{C}$ | 0.0053495 | 32.119660 | 0.0100803 | 17.045640 |
| Clement (2017) $t^{CP}$ | 3.0037410 | 0.0572040 | 2.9982930 | 0.0573080 |
| Nidhi *et al.* (2018) $t^{ND}$ | 1.0552510 | 0.1628280 | 0.0017182 | 100.00080 |
| Audu *et al.* (2020) $t^{AU}$ | 0.0012147 | 141.45390 | 0.0012166 | 141.23290 |
| Khare *et al.* (2022) $t^{KH}$ | 0.0013688 | 125.53260 | 127.80390 | 0.0013440 |
| n=200 | | | | |
| Sample mean $t^{ST}$ | 0.0001937 | 100.00000 | 0.0001937 | 100.00000 |
| Clement (2015) $t^{C}$ | 0.0000978 | 198.08890 | 0.0000840 | 230.54550 |
| Clement (2017) $t^{CP}$ | 2.5580110 | 0.0075730 | 2.5566670 | 0.0075770 |
| Nidhi *et al.* (2018) $t^{ND}$ | 0.1205805 | 0.1606570 | 0.0001937 | 100.00010 |
| Audu *et al.* (2020) $t^{AU}$ | 0.0000002 | 128405.30 | 0.0000002 | 86144.800 |
| Khare *et al.* (2022) $t^{KH}$ | 0.0000408 | 474.29800 | 81954.140 | 0.0000002 |

### *3.2 Discussion of Results*

Fig. 1 shows the relationship between the weight of babies and the gestational age of the entire population. The points indicate individual data points. Fig. 2 shows the relationship between the weight of babies and gestational age by body mass index categories. Fig. 3 compares the weight of babies across different gestational ages of the entire population. Higher bars indicate a higher baby's weight. The chart helps to identify gestational age associated with higher or lower birth weights. Figs. 4 and 5 compare the weight of babies and gestational age across different categories respectively. Fig. 6 shows relationships between weight of babies, gestational age and age of mothers.

Table 1 presents the original weights and calibrated weights of the existing estimators considered with different $\Phi_d$. The results revealed that some calibrated weights did not satisfy the desired adjustments to approximate the original weights. The average sum of adjusted calibrated weights of Clement (2017) for both values of $\Phi_d$, Nidhi *et al.* (2018) for $\Phi_d = 1$ and Khare *et al.* (2022) for $\Phi_d = \frac{1}{\bar{x}_d}$ did not approximate to the average sum of the actual calibrated weights. Thus, this could be the reason for the estimators' poor performance in terms of efficiency and inconsistency. However, calibrated weights of the Clement (2015) and Audu *et al.* (2020) estimators have satisfied the condition of adjusting for the actual or original weights.

Table 2 shows the mean weight of babies estimates of various estimators. The results revealed that the estimator of Clement (2017), $t^{CP}$ underestimated the mean weight of babies, Nidhi *et al.* (2018), $t^{ND}$ estimator also underestimated the mean weight of babies at $\Phi_d = 1$ and Khare *et al*. (2022), $t^{KH}$ overestimated the mean weight of babies at $\Phi_d = \frac{1}{\bar{x}_d}$. The stratified sample mean, ($t^{ST}$) and Clement (2015), $t^{C}$ estimators estimated the mean weight of babies very close the true mean value, though the estimator of Audu *et al.* (2020), $t^{AU}$ gave the results that are much closer to the actual mean weight of babies.

Table 3 shows the mean squared errors (MSEs) and percent relative efficiencies (PREs) of the various calibration ratio estimators considered in this work. The results revealed that Audu *et al.* (2020), $t^{AU}$ estimator have lesser MSEs and higher PREs than the stratified sample mean, $t^{ST}$, Clement (2015), $t^{C}$, Clement(2017), $t^{CP}$, Nidhi *et al.* (2018), $t^{ND}$, and Khare *et al*. (2022), $t^{KH}$ estimators for all cases of $\Phi_d$ and sample sizes. However, from the results discussed, it can be deduced that Audu *et al.* (2020), $t^{AU}$ estimator has demonstrated consistency in all the tuning parameter ($\Phi_d$) values and in all sample sizes. The estimator $t^{AU}$ has the least MSE and the highest PRE for all values of the tuning parameter, indicating its highest performance in terms of efficiency.

## 4. Conclusion and Recommendation

In the present study, we have compared the performance of various calibration ratio estimators. An empirical study has been carried out through population data sets obtained as to ascertain the efficiency and performance of various calibration ratio estimators considered in the study, and the results reveals that Audu *et al.* (2020), ($t^{AU}$)has the least MSE and the highest PRE for all values of the tuning parameter and sample sizes, indicating its highest performance in terms of efficiency and has proven to be consistent as the sample size increases. However, the findings offer valuable insights on the practical suitability of these estimators for accurately estimating the mean weight of babies. Thus, it is recommended that $t^{AU}$ an estimator should be adopted by healthcare practitioners for the baby's mean weight estimation.

## 5.0 References


Alam, S., Singh S. and Shabbir, J. (2019). New Methodology of Calibration in Stratified Random Sampling. *Survey Research Methods Section*, 3, pp.90-95.

Alexander, G. R., Himes, J. H., Kaufman, R. b., Mor, J. & Kogan, M. (1996). Use of birth weight percentile standards in perinatal epidemiology. *American Journal of Epidemiology*, 143, 7, pp. 665-671.

Audu, A., Singh, R. V. K., Muhammed S., Ishaq, O. & Zakari, Y. (2020). *On the Efficiency of Calibration Ratio Estimators of Population mean.* Proceeding of Royal Statistics Society Nigeria Local Group, pp. 247-261.

Catalano, P. M. & Shankar, K. (2012). Obesity and pregnancy: A review of risks and recommendations. *American Journal of Clinical Nutrition*, 95, 5, pp. 1105-1113.

Catalano, P. M., Presley, L., Minium, J. & Hauguel-de Mouzon, S. (2007). Impact of maternal obesity on fetal growth and neonatal outcomes. *Journal of Endocrinology and Metabolism,* 92, 11, pp. 4322-4326

Clement, E. P. (2015). Calibration Approach Separate Ratio Estimator for Population Mean in Stratified Sampling. *International Journal of Modern Mathematical Sciences*, 13, 4, pp. (4): 377-384.

Clement, E. P. (2017). A New Ratio Estimator of Mean in Survey Sampling by Calibration Estimation. *Elixir International Journal*, 106, pp. 46461-46465.

Cochran, W. G., (1977). *Sampling Technique.* 3rd edition. Wiley Eastern Limited, New York.

Deville, J. C., & Sarndal, C. E. (1992). Calibration Estimators in Survey Sampling. *Journal of the American Statistical Association*, 87, 418, pp. 376-382.

Fall, C. (2009). Maternal Nutrition: Effects on Health in next gestation. *Indian Journal of Medicine*, 130, pp. 593-599.

Garg, N., & Pachori, M. (2019). Use of coefficient of variation in calibration estimation of Population mean in stratified sampling. *Communications in Statistics - Theory and Methods,* 49, 23, pp. 842-5852.

Khare, B. B., Singh, S. &Mishra, M. (2022). Separate Ratio Estimator Using Calibration Approach for the Population Mean in Stratified Random Sampling. *Asian Journal of Probability and Statistics*, 20, 3, pp. 64-73.

Kramer, M. S. & Kakumar, R. (2001). *Birth weight and fetal growth*. In A. A. Fanaroff and R. J. Martin (Eds.), Neonatal - perinatal medicine (7th ed., pp. 43-55). Mosby

Muili, J. O., Singh, R. V. K., Onwuka, G. I. & Audu, A. (2022). Calibration of Stratified Random Sampling with Combined Ratio Estimators. *Oriental Journal of Physical Sciences*, 7, 1, pp. 0-36.

Nidhi, B. V. S., Sisodia, S., Singh, P. & Kumar, S. (2018). Calibration Ratio based type estimation of mean in stratified sampling. *International Journal of Agricultural Statistical Science, 14, 1, pp. 433-440.*

Nidhi, B. V. S., Sisodia, S., Singh, S. K., & Singh (2017). Calibration approach estimation of the mean in stratified sampling stratified double sampling. *Communication in Statistics-Theory and Methods* 46, 10, pp. 4932-4942.

O'Tierney- Ginn, P. F. & Loftin, C. W. (2017). Contributions of Maternal BMI and gestational weight gain to birth. *Journal of Clinical Endocrinology and Metabolism*, 102, 11, pp. 3978- 3986.

Ozgul, N. (2018). New calibration estimator in stratified Sampling. *Journal of Statistical Computation and Simulation,* 88, 13, pp. 2561-2572.

Rao, D., Tekabu, T., & Khan, M. (2016). New calibration estimators in stratified sampling. *International Journal of Computer Science and Engineering* (APWC on CSE), 3rd Asia Pacific World Congress, Nadi, Fijion, 5, 6, pp. 66-70.

Singh, S. (2003). *Advanced Sampling Theory with Applications: How Michael "Selected" Amy,* Volume 2. The Netherlands: Kluwer Academic Publishers.

Singh, S., Horn, S. & Yu, F. (1998). Estimation of variance of the general regression Estimator: Higher level calibration approach. *survey methodology* 48, 1, pp. 41-50.

Sisodia, B. V. S., Singh, S., & Singh, S. K. (2017). Calibration approach estimation of the Mean in stratified sampling and stratified double sampling. *Communications in Statistics–Theory and Methods* 46, 10, pp. 4932-4942.

Tracy, D. S., Singh, S., & Arnab, R. (2003). Note on calibration in stratified and double Sampling. *Survey Methodology* 29, 1, pp. 99-104.

World Health Organization. (2016). *Body mass index (BMI).*

**Compliance with Ethical Standards Declaration**

**Ethical Approval**

Not Applicable

**Competing interests**

The authors declare that they have no known competing financial interests

**Funding**

The author declared no source of external funding

**Availability of data and materials**
Data would be made available on request.

**Author's Contribution**
Both authors contributed to the writing of the manuscript, under the supervision of EJE