STATISTICS, OPTIMIZATION AND INFORMATION COMPUTING
Stat., Optim. Inf. Comput., Vol. **x**, Month 202x, pp 0–12.
Published online in International Academic Press (www.IAPress.org)

Sensitivity analysis for missing-not-at-random mechanisms in complex survey data using delta-adjusted multiple imputation

Mohyaldein Salih [1,*], Ali Satty [2], Henry Mwambi [1]

[1]*School of Mathematics, Statistics and Computer Science, University of KwaZulu-Natal, Private Bag X01 Scottsville 3209, Pietermaritzburg, South Africa*
[2]*Department of Mathematics, College of Science, Northern Border University, Arar, Saudi Arabia*

Abstract Missing data are common in complex surveys studies and can compromise statistical inference when assumptions about the missing data mechanism are violated. Although multiple imputation is widely used under the missing at random (MAR) assumption, MAR is inherently unverifiable and may be implausible in large population-based surveys. This study aimed to evaluate delta-adjustment sensitivity analysis as a principled approach for assessing the robustness of MAR-based inferences to missing not at random (MNAR) mechanisms in complex survey data. We assessed the performance of complete case analysis, multiple imputation under MAR and delta-adjusted pattern-mixture models in a survey-weighted logistic regression framework with data on 20,200 children from the 2022–2023 Yemen Multiple Indicator Cluster Survey. The assumption of missing completely at random (MCAR) was rejected. Multiple imputation yielded more efficient estimates than complete case analysis.Effect estimates remained stable within the prespecified delta range examined, and no tipping point was observed within these MNAR scenarios. These findings demonstrate the potential to routinely embed delta-adjustment sensitivity analysis within multiple imputation in order to enhance the transparency, believability and interpretability of analyses of incomplete complex survey data.

Keywords Missing data, Missingness assumptions, Sensitivity analysis, Delta adjustment, Complex survey data

AMS 2010 subject classifications 62D05, 62P10, 62F07

DOI: 10.19139/soic-2310-5070-3694

1. Introduction

Missing data are pervasive in complex surveys research and remain a major threat to valid statistical inference. Large observational studies and population-based surveys frequently experience incomplete data due to nonresponse, item skipping and survey design features, often affecting both outcomes and key covariates. When missing data are not appropriately addressed, parameter estimates may be biased and substantive conclusions distorted [1, 2]. As a result, statistical inference with incomplete data depends critically on assumptions about the underlying missing data mechanism. Within Rubin's framework, missingness is commonly classified as missing completely at random (MCAR), missing at random (MAR) or missing not at random (MNAR) [3]. While MCAR and MAR allow valid likelihood-based or multiple imputation (MI) analyses, MNAR mechanisms, where missingness depends on unobserved values, require additional, untestable assumptions and are not identifiable from the observed data alone [3, 4].

MI has become the predominant approach for handling missing data in surveys studies [5, 6]. By generating multiple plausible values for missing observations and combining results using Rubin's rules, MI appropriately

*Correspondence to: Mohyaldein Salih (Email: mohyyassin@gmail.com). School of Mathematics, Statistics and Computer Science, University of KwaZulu-Natal, Private Bag X01 Scottsville 3209, Pietermaritzburg, South Africa.

ISSN 2310-5070 (online) ISSN 2311-004X (print)
Copyright © 202x International Academic Press

reflects uncertainty due to missingness while retaining partially observed cases. Flexible implementations such as multiple imputation by chained equations (MICE) further extend MI by accommodating variables of different types through variable-specific imputation models [7]. Under MAR and correct model specification, MI yields consistent and efficient estimators and generally outperforms CCA, which is often inefficient and potentially biased [8, 9]. However, the validity of MI-based inference depends critically on the MAR assumption, which cannot be empirically verified using observed data alone [10]. In many survey-based settings, particularly large surveys and routine health data, missingness may plausibly depend on unobserved outcomes or exposures, raising concerns about bias under MNAR mechanisms [11].

Growing recognition of the unverifiable nature of MAR has led to increased emphasis on sensitivity analyses that assess the robustness of statistical inferences to alternative missing-data mechanisms [12, 13]. Rather than attempting to identify the true missingness mechanism, these approaches evaluate how much departure from MAR would be required to meaningfully change substantive conclusions and are now widely recommended as best practice [14, 15]. Among available methods, delta-adjusted pattern-mixture models are particularly appealing because they can be implemented within standard MI frameworks and allow transparent exploration of MNAR scenarios [16, 17, 18]. By varying a sensitivity parameter that represents systematic differences between observed and unobserved data, investigators can assess the stability of effect estimates and identify potential tipping points at which conclusions change [19, 20]. Despite these advantages, delta-adjustment remains underused in survey-based research, particularly in studies involving complex survey data, and empirical evaluations in real-world population-based settings remain limited [21, 22].

Large-scale population-based health surveys, such as the Multiple Indicator Cluster Survey (MICS), play a central role in research and public health decision-making. These surveys are widely used to estimate disease burden and identify determinants of health outcomes, but they frequently exhibit substantial and heterogeneous missing data due to nonresponse, survey design features, and respondent burden [23]. Because standard analytic approaches in these settings typically rely on MAR, uncertainty regarding the plausibility of this assumption raises important concerns about the validity of statistical inference. In this study, we evaluate delta-adjustment sensitivity analysis as a practical and transparent framework for examining departures from MAR in complex survey data. Using data from the 2022–2023 Yemen MICS as an empirical illustration, we compare complete case analysis (CCA), MI under MAR and delta-adjusted pattern-mixture models within a survey-weighted regression framework. The primary objective of this study is to methodologically investigate how increasingly severe MNAR assumptions influence parameter estimates, statistical inference and substantive conclusions, thereby strengthening the empirical basis for the routine incorporation of sensitivity analyses in complex surveys research with incomplete data.

2. Methods

2.1. Study population and variables

This study was conducted as a cross-sectional analysis using secondary data from the 2022–2023 Yemen MICS. The survey collected detailed information on children under five years of age, along with demographic, socioeconomic, maternal and environmental characteristics of their households. The study population comprised 20,200 children under five years of age. Comprehensive information on the survey design, methodological framework, sampling procedures, data availability, and key findings of the 2022–2023 Yemen MICS is provided in the official survey report issued by the Central Statistical Organization [24].

The primary outcome variable was childhood diarrhea, defined as whether a child experienced diarrhea during the two weeks preceding the survey. Independent variables included child sex, child age, breastfeeding status, vaccination status, health insurance coverage, mother's education level, area of residence, household wealth index, source of drinking water, sanitation facility, and the availability of a handwashing place. The extent of missing data varied across study variables. Child sex and area of residence were fully observed, whereas moderate levels of missingness were present for child age, health insurance coverage, and household wealth index. Breastfeeding and vaccination status showed substantial missingness and accounted for the highest proportions of incomplete

data. This marked heterogeneity in missingness across variables highlights the need for robust strategies to handle missing data during statistical analysis.

2.2. Missing data framework

Let $i = 1, 2, \ldots, n$ index individuals and let $j = 1, 2, \ldots, p$ index study variables. For each individual i, let $D_i = (Y_i, X_i)$ denote the complete data vector, where Y_i is a binary outcome indicating the occurrence of childhood diarrhea ($Y_i = 1$ if diarrhea occurred and $Y_i = 0$ otherwise), and let $X_i = (X_{i1}, X_{i2}, \ldots, X_{ip})'$ be a vector of covariates. Let $D = (D_1, D_2, \ldots, D_n)$ denote the complete data for the study population. Let $R = (R_1, R_2, \ldots, R_n)$ denote the collection of missingness indicators, where each R_i has the same dimension as D_i and indicates which components of D_i are observed. The complete data D can be partitioned into observed and missing components as $D = (D_{\text{obs}}, D_{\text{mis}})$.

Following Rubin [3], the joint distribution of the complete data and the missingness indicators is factorized as:

$$P(D, R \mid \gamma, \vartheta) = P(D \mid \gamma) P(R \mid D, \vartheta), \tag{1}$$

where γ denotes the parameters of the data model and ϑ denotes the parameters governing the missing data mechanism.

Under the MCAR assumption, the probability of missingness is independent of both observed and unobserved data:

$$P(R \mid D, \vartheta) = P(R \mid \vartheta). \tag{2}$$

Under the MAR assumption, the probability of missingness depends only on observed data:

$$P(R \mid D_{\text{obs}}, D_{\text{mis}}, \vartheta) = P(R \mid D_{\text{obs}}, \vartheta). \tag{3}$$

Under MNAR, the probability of missingness depends on the unobserved data even after conditioning on the observed data:

$$P(R \mid D_{\text{obs}}, D_{\text{mis}}, \vartheta) \neq P(R \mid D_{\text{obs}}, \vartheta). \tag{4}$$

2.3. Assessment of MCAR

To assess whether the missing data followed an MCAR mechanism, Little's MCAR test was applied [25]. This test evaluates the joint null hypothesis that the probability of missingness is independent of both observed and unobserved data. Formally, the null hypothesis is given by:

$$H_0 : P(R \mid D, \vartheta) = P(R \mid \vartheta), \tag{5}$$

where R denotes the missingness indicators, D the complete data and ϑ the parameters governing the missingness mechanism.

Little's MCAR test statistic follows a chi-square distribution and is defined as:

$$\chi^2 = \sum_{g=1}^{G} n_g (\bar{y}_g - \hat{\mu})' \hat{\Sigma}^{-1} (\bar{y}_g - \hat{\mu}), \tag{6}$$

where G is the number of distinct missingness patterns, n_g is the number of observations exhibiting missingness pattern g, $\bar{y}_g$ is the sample mean vector of the variables observed under pattern g, $\hat{\mu}$ is the estimated overall mean vector under the MCAR mechanism and $\hat{\Sigma}$ is the corresponding estimated covariance matrix. The quadratic form is evaluated over the subset of variables observed in each pattern.

The degrees of freedom for the test are given by:

$$df = \sum_{g=1}^{G} p_g - p, \tag{7}$$

where p_g denotes the number of variables observed in missingness pattern g, and p is the total number of variables in the complete data vector.

The null hypothesis is rejected when the observed χ^2 statistic exceeds the critical value from the chi-square distribution with the corresponding degrees of freedom at significance level α indicating a violation of the MCAR mechanism. Rejection of MCAR does not imply MAR, but motivates the use of MI as a baseline analytical approach, supplemented by sensitivity analyses to assess robustness to potential MNAR mechanisms [1].

2.4. Complete case analysis (CCA)

As an initial comparison, a CCA was performed by restricting the analysis to individuals with fully observed outcome and covariate data. Let

$$D_{\text{CC}} = \{i : R_{ij} = 1 \text{ for all } j = 0, 1, \ldots, p\}, \tag{8}$$

where R_{ij} denotes the missingness indicator for variable j of individual i, with $R_{ij} = 1$ if the value of variable j is observed and $R_{ij} = 0$ if the value is missing. Here, $j = 0$ indexes the outcome variable and $j = 1, \ldots, p$ index the covariates.

Given the complex survey design of the dataset, which involved stratification, multistage sampling and unequal sampling weights, survey-weighted logistic regression was used in place of standard logistic regression to obtain design consistent estimates and valid statistical inference [26]. The survey-weighted logistic regression model assumes the same mean structure as the standard logistic regression model and is specified as:

$$\text{logit}\{P(Y_i = 1 \mid X_i)\} = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \cdots + \beta_p X_{ip}, \tag{9}$$

where Y_i is the binary outcome indicating childhood diarrhea (1 = yes, 0 = no), $X_i = (X_{i1}, X_{i2}, \ldots, X_{ip})'$ is the vector of covariates for individual i and $\beta = (\beta_0, \beta_1, \ldots, \beta_p)'$ is the vector of regression coefficients.

Unlike standard logistic regression, parameter estimation is carried out using a design-based survey-weighted pseudo-likelihood, in which each observation contributes to the likelihood in proportion to its sampling weight. Specifically, the regression coefficients are obtained by maximizing:

$$\hat{\beta}_{\text{CC}} = \arg\max_{\beta} \sum_{i \in D_{\text{CC}}} w_i [Y_i \log(\pi_i) + (1 - Y_i) \log(1 - \pi_i)], \tag{10}$$

where w_i denotes the survey sampling weight for individual i, and $\pi_i = P(Y_i = 1 \mid X_i) = \text{expit}(X_i'\beta) = 1/(1 + \exp(-X_i'\beta))$. Variance estimation accounts for the survey design through appropriate incorporation of sampling weights, stratification and clustering, yielding design-consistent standard errors and confidence intervals.

While CCA yields unbiased estimates under the MCAR mechanism, it was included primarily as a benchmark analysis given its well-documented inefficiency and potential for bias when missingness depends on observed or unobserved data [1, 6].

2.5. Multiple imputation (MI) under MAR

To address missing data under the MAR assumption, MI was implemented following the framework originally proposed by Rubin [27]. MI proceeds in three steps: (i) missing values are imputed M times to generate M completed datasets; (ii) each completed dataset is analyzed separately using multivariable survey weighted logistic regression model; and (iii) results are combined across imputations to account for both within-imputation sampling variability and between-imputation variability due to missing data. In this study, the number of imputations was set to $M = 20$, consistent with recommendations for obtaining stable parameter estimates [12]. Imputation was performed under the MAR assumption using a fully conditional specification (FCS) approach, also known as MICE [7]. Let Y_{ij} denote the value of variable j for individual i, and let $Y_{i,-j}$ denote the set of all other variables for individual i excluding Y_{ij}. For binary variables with missing values, imputation was conducted using logistic regression:

$$P(Y_{ij} = 1 \mid Y_{i,-j}, \theta_j) = \frac{\exp(Y_{i,-j}'\theta_j)}{1 + \exp(Y_{i,-j}'\theta_j)}, \tag{11}$$

where θ_j denotes the parameter vector for the imputation model of variable j.

For a categorical variable Y_{ij} with K categories, let $Y_{i,-j}$ denote the set of all variables for individual i excluding variable j. The imputation model was:

$$P(Y_{ij} = k \mid Y_{i,-j}, \theta_{jk}) = \frac{\exp(Y'_{i,-j}\theta_{jk})}{\sum_{l=1}^{K} \exp(Y'_{i,-j}\theta_{jl})}, \tag{12}$$

for $k = 1, \ldots, K$, where K denotes the total number of categories and one category is designated as the reference (i.e., $\theta_{jk} = 0$).Survey sampling weights were included as auxiliary variables in the imputation models to improve compatibility between the imputation and analysis stages. Inclusion of the survey weights helps preserve population representativeness and reduce potential bias arising from the complex sampling design. The substantive analyses were subsequently conducted using survey-weighted logistic regression models incorporating sampling weights, stratification, and clustering.

For each imputed dataset $m = 1, \ldots, M$, the substantive survey-weighted logistic regression model was fitted:

$$\text{logit}\{P(Y_i^{(m)} = 1 \mid X_i^{(m)})\} = X_i^{(m)'}\beta^{(m)}, \tag{13}$$

yielding parameter estimates $\hat{\beta}^{(m)}$ and corresponding variance estimates $\text{Var}(\hat{\beta}^{(m)})$.

Estimates were combined across imputations using Rubin's rules. The pooled regression coefficient was computed as:

$$\bar{\beta}_{\text{MI}} = \frac{1}{M}\sum_{m=1}^{M} \hat{\beta}^{(m)}. \tag{14}$$

The within-imputation variance was estimated as:

$$\bar{W} = \frac{1}{M}\sum_{m=1}^{M} \text{Var}(\hat{\beta}^{(m)}), \tag{15}$$

while the between-imputation variance was computed as:

$$B = \frac{1}{M-1}\sum_{m=1}^{M} (\hat{\beta}^{(m)} - \bar{\beta}_{\text{MI}})(\hat{\beta}^{(m)} - \bar{\beta}_{\text{MI}})'. \tag{16}$$

The total variance of the pooled estimator was then obtained as:

$$T = \bar{W} + \left(1 + \frac{1}{M}\right) B. \tag{17}$$

The MI results were treated as the primary analysis under the MAR assumption and served as the baseline for subsequent sensitivity analyses [6, 27].The MICE algorithm was implemented using 20 imputations and 20 iterations per imputation chain. The choice of 20 imputations is consistent with contemporary recommendations advocating a larger number of imputations to improve the stability and efficiency of MI estimates [28]. Twenty iterations were used to promote adequate convergence and stable imputation chains, as convergence diagnostics rather than fixed iteration counts are generally recommended for assessing the adequacy of MICE algorithms [29]. These settings are commonly employed in recent epidemiological and health-related applications of MICE [30].

2.6. Delta-adjustment sensitivity analysis

To evaluate the robustness of the study findings to potential violations of the MAR assumption, delta-adjustment sensitivity analyses were conducted within a pattern-mixture modelling framework [29, 31]. Under MNAR mechanisms, the distribution of unobserved outcomes may systematically differ from that of observed outcomes, even after conditioning on observed covariates.

Under the MAR assumption, the imputation model for the binary outcome is specified as:

$$\text{logit}\{P(Y_i = 1 \mid X_i)\} = X_i'\theta, \tag{18}$$

where θ denotes the vector of imputation model parameters estimated under MAR.

To represent departures from MAR towards MNAR, a sensitivity parameter δ was introduced into the imputation model for individuals with missing outcome values. Specifically, for observations with $R_i = 0$, the delta-adjusted imputation model was defined as:

$$\text{logit}\{P(Y_i = 1 \mid X_i, R_i = 0)\} = X_i'\theta + \delta. \tag{19}$$

In complex survey data, the delta adjustment operates on the imputation model's conditional log-odds scale and is distinct from the survey weighting scheme. Survey weights account for the sampling design, whereas the sensitivity parameter δ represents an assumed departure from MAR among individuals with missing outcomes. Thus, δ modifies the distribution of missing values without altering the survey weights, and its impact is reflected in the final weighted estimates and uncertainty measures.

A sequence of delta values $\delta \in \{-1, -0.5, 0, 0.5, 1\}$ was examined, representing increasingly strong departures from MAR towards MNAR. This symmetric range around zero was selected in accordance with recommended practice for delta-adjustment sensitivity analyses [18, 20, 32].The selected δ values correspond to shifts on the log-odds scale and imply odds ratios ranging from approximately 0.37 $(\exp(-1))$ to 2.72 $(\exp(1))$ for individuals with missing outcomes relative to those predicted under MAR. These values were chosen to represent moderate-to-strong departures from MAR that remain plausible in large population-based surveys. We acknowledge that the selection of δ values is inherently subjective and that more extreme values could be considered. Therefore, the absence of a tipping point should be interpreted as evidence of robustness within the examined range rather than proof of robustness under all conceivable MNAR mechanisms. For each value of δ, missing outcome values were imputed using the delta-adjusted model specified above. Missing values in all other variables were imputed using standard FCS procedures under MAR. For each delta scenario, $M = 20$ completed datasets were generated.

For each imputed dataset $m = 1, 2, \ldots, M$ under a given delta value, the substantive survey-weighted logistic regression model was fitted:

$$\text{logit}\{P(Y_i^{(m,\delta)} = 1 \mid X_i^{(m,\delta)})\} = X_i^{(m,\delta)'}\beta^{(m,\delta)}, \tag{20}$$

yielding parameter estimates $\hat{\beta}^{(m,\delta)}$ and corresponding variance estimates $\text{Var}(\hat{\beta}^{(m,\delta)})$.

For each delta value, parameter estimates were combined across imputations using Rubin's rules. The pooled regression coefficient was computed as:

$$\bar{\beta}_{\text{MI}}(\delta) = \frac{1}{M}\sum_{m=1}^{M}\hat{\beta}^{(m,\delta)}. \tag{21}$$

The within-imputation variance was estimated as:

$$\bar{W}(\delta) = \frac{1}{M}\sum_{m=1}^{M}\text{Var}(\hat{\beta}^{(m,\delta)}), \tag{22}$$

while the between-imputation variance was computed as:

$$B(\delta) = \frac{1}{M-1}\sum_{m=1}^{M}(\hat{\beta}^{(m,\delta)} - \bar{\beta}_{\text{MI}}(\delta))(\hat{\beta}^{(m,\delta)} - \bar{\beta}_{\text{MI}}(\delta))'. \tag{23}$$

The total variance of the pooled estimator was then obtained as:

$$T(\delta) = \bar{W}(\delta) + \left(1 + \frac{1}{M}\right)B(\delta) \tag{24}$$

Changes in parameter estimates and inferential conclusions were examined across the set of delta values to assess sensitivity to departures from the MAR assumption. For each regression coefficient j, the sequence $\{\bar{\beta}_{\text{MI},j}(\delta) : \delta \in \{-1, -0.5, 0, 0.5, 1\}\}$, was evaluated to characterize robustness under increasingly severe departures from MAR towards MNAR.

A tipping-point analysis was conducted to identify a critical value δ^* at which substantive conclusions differed from those obtained under MAR ($\delta = 0$). A tipping point was defined as the smallest value of δ^* for which at least one of the following occurred: (i) Change in direction of association:

$$\text{sign}(\bar{\beta}_{\text{MI},j}(\delta^*)) \neq \text{sign}(\bar{\beta}_{\text{MI},j}(0)); \tag{25}$$

(ii) Change in statistical significance, indicated by the associated P-value crossing the threshold $\alpha = 0.05$; or (iii) Change in substantive interpretation of the effect.

2.7. Statistical analysis

All analyses were conducted using R software (version 4.5.2). Missing data mechanisms were initially assessed using Little's MCAR test. CCA was performed as a baseline analysis, accounting for the complex survey design through appropriate weighting, stratification and clustering. To address missing data under the MAR assumption, MI was applied and results were pooled using Rubin's rules.Variance estimation following MI accounted for the complex survey design through the incorporation of sampling weights, primary sampling units, and stratification variables within the survey-weighted regression framework. Sensitivity to departures from MAR was evaluated using delta-adjustment within a pattern-mixture framework to reflect plausible MNAR scenarios. Statistical inference was based on two-sided tests with a significance level of 0.05.

3. Results

3.1. Study variables

Table 1 summarizes the study variables, their categories and the extent of missing data. The binary outcome variable, diarrhea in the two weeks preceding the survey, contains 699 missing observations. Among the independent variables, child sex and area of residence are fully observed. Moderate levels of missingness are present for child age, health insurance coverage and household wealth index, whereas mother's education, source of drinking water and sanitation facility are nearly complete. In contrast, breastfeeding and vaccination status exhibit substantial missingness and account for the highest proportion of incomplete data.

3.2. Test MCAR assumption (Little's test)

To evaluate whether missing data on childhood diarrhea followed an MCAR assumption, Little's MCAR test was applied using a chi-square statistic. The test produced a chi-square value of 14,097.8 with $P < 0.01$, providing strong evidence against the MCAR assumption. This finding indicates that the missingness was not MCAR and is likely related to observed data.

3.3. Complete case analysis (CCA) results

CCA indicated that vaccination status, area of residence and household wealth index were significantly associated with childhood diarrhea. Children from poorer households had markedly higher odds of diarrhea, whereas children living in rural areas exhibited lower odds compared with those residing in urban areas. Most demographic, maternal and environmental factors were not statistically significant, with wide confidence intervals reflecting reduced efficiency due to the exclusion of observations with missing data (Table 2).

Table 1. Study variables

Variable	Variable levels	Missing values
Outcome variable		
Diarrhea in last 2 weeks	Yes / No	699
Independent variables		
Child sex	Male / Female	0
Child age	0–5 / 6–11 / 12–23 / 24–35 / 36–47 / 48–59	639
Breastfeeding	Yes / No	8,536
Vaccination	Yes / No	13,822
Health insurance	Yes / No	661
Mother's education	Pre-primary or none / Primary / Lower secondary / Upper secondary / Higher	6
Area of residence	Rural / Urban	0
Household wealth index	Poorest / Poor / Middle / Rich / Richest	640
Source of drinking water	Unimproved / Improved	86
Sanitation	Unimproved / Improved	51
Handwashing place	Unobserved / Observed	327

Table 2. Complete case analysis (CCA) results

Variable	Levels	Estimate	SE	LCL	UCL	P-value
(Intercept)	–	−0.613	0.564	−1.718	0.493	0.28
Child sex	Male	0.025	0.075	−0.121	0.171	0.74
Child age	0–5	−0.823	0.511	−1.824	0.178	0.11
	6–11	−0.359	0.497	−1.332	0.615	0.47
	12–23	−0.306	0.505	−1.295	0.684	0.55
	24–35	−0.732	0.506	−1.724	0.261	0.15
Breastfeeding	Yes	0.113	0.125	−0.132	0.359	0.37
Vaccination	Yes	0.240	0.088	0.067	0.413	0.01
Health insurance	Yes	−0.184	0.924	−1.995	1.627	0.84
Mother's education	Pre-primary or none	−0.211	0.239	−0.679	0.257	0.38
	Primary	−0.226	0.243	−0.702	0.250	0.35
	Lower secondary	−0.005	0.252	−0.499	0.489	0.98
	Upper secondary	−0.001	0.248	−0.486	0.484	0.10
Area of residence	Rural	−0.417	0.133	−0.678	−0.155	< 0.01
Household wealth	Poorest	1.309	0.189	0.938	1.680	< 0.01
	Poor	1.239	0.168	0.909	1.568	< 0.01
	Middle	0.984	0.154	0.682	1.286	< 0.01
	Rich	0.443	0.143	0.163	0.722	< 0.01
Drinking water	Unimproved	0.005	0.108	−0.206	0.216	0.96
Sanitation	Unimproved	0.165	0.095	−0.022	0.351	0.08
Handwashing place	Unobserved	−0.034	0.112	−0.254	0.187	0.77

3.4. Multiple imputation (MI) results (Baseline analysis)

Compared to CCA, MI analysis resulted in additional significant covariates, particularly child sex and age. Effect estimates were generally consistent in direction and magnitude with those from the CCA but exhibited narrower confidence intervals, indicating improved precision through the inclusion of children with partially observed data.

The strong wealth gradient, the protective effect of rural residence and the positive association with vaccination status persisted under MI (Table 3).

Table 3. Multiple imputation (MI) results

Variable	Levels	Estimate	SE	LCL	UCL	P-value
(Intercept)	–	−1.701	0.163	−2.020	−1.383	< 0.01
Child sex	Male	0.133	0.039	0.057	0.210	< 0.01
Child age	0–5	0.243	0.084	0.077	0.409	< 0.01
	6–11	0.861	0.079	0.706	1.015	< 0.01
	12–23	0.880	0.071	0.741	1.019	< 0.01
	24–35	0.368	0.073	0.224	0.511	< 0.01
	36–47	0.163	0.075	0.016	0.310	0.03
Breastfeeding	Yes	−0.051	0.095	−0.237	0.136	0.59
Vaccination	Yes	0.154	0.096	−0.034	0.343	0.11
Health insurance	Yes	0.343	0.388	−0.418	1.103	0.38
Mother's education	Pre-primary or none	−0.032	0.121	−0.269	0.205	0.79
	Primary	0.069	0.125	−0.176	0.313	0.58
	Lower secondary	0.111	0.123	−0.131	0.353	0.37
	Upper secondary	0.090	0.122	−0.150	0.329	0.46
Area of residence	Rural	−0.409	0.080	−0.566	−0.251	< 0.01
Household wealth	Poorest	1.329	0.113	1.107	1.551	< 0.01
	Poor	1.230	0.102	1.031	1.429	< 0.01
	Middle	0.946	0.091	0.767	1.125	< 0.01
	Rich	0.666	0.085	0.500	0.832	< 0.01
Drinking water	Unimproved	−0.034	0.075	−0.182	0.113	0.65
Sanitation	Unimproved	0.187	0.061	0.068	0.306	< 0.01
Handwashing place	Unobserved	−0.111	0.084	−0.276	0.054	0.19

Table 4. Delta-adjustment sensitivity analysis results

Variable	Levels	Coefficient estimates					Avg	SD	Sign	Significance	Tipping
		δ=−1	δ=−0.5	δ=0	δ=0.5	δ=1					
Child sex	Male	0.134	0.134	0.134	0.133	0.133	0.134	0.001	Same	Same	No
Child age	0–5	0.166	0.167	0.162	0.160	0.161	0.163	0.003	Same	Same	No
	6–11	0.373	0.375	0.374	0.374	0.373	0.374	0.001	Same	Same	No
	12–23	0.893	0.896	0.886	0.886	0.875	0.887	0.008	Same	Same	No
	24–35	0.863	0.865	0.859	0.857	0.852	0.859	0.005	Same	Same	No
	36–47	0.227	0.227	0.221	0.222	0.214	0.222	0.005	Same	Same	No
Breastfeeding	Yes	−0.006	−0.007	−0.010	−0.011	−0.009	−0.009	0.002	Same	Same	No
Vaccination	Yes	0.087	0.084	0.087	0.084	0.081	0.085	0.003	Same	Same	No
Health insurance	Yes	0.365	0.344	0.338	0.296	0.313	0.331	0.027	Same	Same	No
Mother's educ.	Pre-primary/none	0.104	0.092	0.096	0.085	0.089	0.093	0.007	Same	Same	No
	Primary	0.134	0.112	0.120	0.100	0.102	0.114	0.014	Same	Same	No
	Lower secondary	0.090	0.072	0.076	0.054	0.056	0.070	0.015	Same	Same	No
	Upper secondary	−0.004	−0.024	−0.028	−0.054	−0.057	−0.033	0.022	Same	Same	No
Area of residence	Rural	−0.409	−0.408	−0.413	−0.409	−0.407	−0.409	0.002	Same	Same	No
Household wealth	Poorest	0.654	0.651	0.656	0.655	0.647	0.653	0.004	Same	Same	No
	Poor	0.941	0.943	0.940	0.932	0.923	0.936	0.008	Same	Same	No
	Middle	1.227	1.226	1.228	1.219	1.197	1.219	0.013	Same	Same	No
	Rich	1.317	1.316	1.320	1.314	1.289	1.311	0.013	Same	Same	No
Drinking water	Unimproved	−0.031	−0.030	−0.035	−0.038	−0.037	−0.034	0.004	Same	Same	No
Sanitation	Unimproved	0.191	0.188	0.189	0.187	0.184	0.188	0.003	Same	Same	No
Handwashing	Unobserved	−0.110	−0.111	−0.117	−0.118	−0.116	−0.114	0.004	Same	Same	No

3.5. Delta-adjustment sensitivity analysis results

Table 4 presents the delta-adjustment sensitivity analysis results for diarrhea among children under five across a range of delta values $\delta = \{-1, -0.5, 0, 0.5, 1\}$ representing departures from the MAR assumption towards MNAR. Across all delta scenarios, the magnitude, direction and statistical significance of the coefficients' estimates remain the same and no tipping point was observed. These findings indicate that even under moderate to relatively strong MNAR assumptions, the estimated associations and inferential conclusions were robust.

4. Discussion

This study evaluated the plausibility of the MAR assumption relative to MNAR by applying delta-adjustment sensitivity analysis in the context of childhood diarrhea data with substantial missingness. Given the extent of incomplete data in both the outcome and covariates, reliance on CCA alone would likely have produced inefficient or potentially biased estimates [33]. Accordingly, a structured missing data framework was adopted that combined formal assessment of missingness, MI under MAR and sensitivity analyses, consistent with recommended best practice in complex surveys research [1, 6].

Little's MCAR test provided strong statistical evidence against the MCAR assumption, consistent with previous research showing that MCAR is rarely plausible in large observational health surveys [4, 34]. In such settings, missingness often depends on observed sociodemographic or health-related characteristics [35]. Rejection of MCAR therefore supported the use of MI under MAR as the primary analytical approach [36], followed by sensitivity analyses to assess robustness to departures towards MNAR mechanisms [37].

Comparisons between CCA and MI analyses highlighted the limitations of CCA in the presence of substantial missing data. Although CCA identified several key associations, confidence intervals were wide and statistical power was reduced, reflecting the well-documented inefficiency of CCA when missingness is related to observed covariates [9, 38]. In contrast, MI under MAR produced more precise estimates and identified additional associations with child sex and age, while preserving the direction and relative magnitude of effects observed in the CCA. These findings are consistent with both theoretical and empirical evidence showing that MI can recover information from partially observed cases when the MAR assumption is approximately satisfied [8, 22].

The delta-adjustment sensitivity analyses provided a systematic evaluation of robustness to plausible MNAR mechanisms. Across a broad range of delta values, effect estimates for the main predictors, particularly household wealth, area of residence, child age and sex, remained stable in direction, magnitude and statistical significance. No tipping point was identified at which substantive conclusions differed meaningfully from those obtained under the MAR-based MI analysis. Within established sensitivity-analysis frameworks, this level of stability indicates that the resulting inferences remained stable within the range of MNAR departures investigated [12, 18].

Some attenuation of the vaccination effect was observed across the MNAR scenarios. Although the direction of the association remained unchanged, the loss of statistical significance under certain sensitivity assumptions represents a meaningful reduction in evidential strength and suggests that conclusions regarding vaccination status are less robust than those for the major predictors. From a public health perspective, this distinction is important because it shifts the interpretation from evidence of association to a lack of strong statistical evidence under specific MNAR assumptions. One possible explanation is that vaccination status may be particularly vulnerable to recall bias, reporting inaccuracies, or social desirability effects, which could influence both response behaviour and the probability of missingness. Consequently, while the overall study conclusions remained stable, findings related to vaccination status should be interpreted with greater caution [10, 20].

Delta-adjustment represents only one class of MNAR sensitivity-analysis methods. Alternative approaches include selection models, controlled MI, reference-based imputation strategies and the Not-at-Random Fully Conditional Specification (NARFCS) procedure [12, 19, 32]. Pattern-mixture approaches such as delta-adjustment are attractive because they are relatively straightforward to implement and interpret within standard multiple-imputation workflows. However, they assume that the MNAR departure can be represented by a constant shift applied to all missing observations. In practice, the magnitude of departure from MAR may vary across individuals or subgroups, making this assumption potentially unrealistic. Consequently, future studies could compare multiple

sensitivity-analysis frameworks to evaluate the consistency of conclusions across alternative MNAR specifications. Overall, evidence from Little's test, the consistency between CCA and MI estimates and the robustness observed in the delta-adjustment sensitivity analyses support the plausibility of MAR as a working assumption for these data. Although MNAR mechanisms cannot be definitively ruled out, the results indicate that departures from MAR would need to be implausibly large to meaningfully alter substantive conclusions. These findings reinforce current methodological guidance advocating the use of MI under MAR, complemented by formal sensitivity analyses, as a principled and transparent approach for handling missing data in large-scale complex surveys [29, 39].

5. Conclusion

Within the range of MNAR scenarios examined, statistical inferences obtained under the MAR assumption remained stable and no tipping point was observed. The results support current methodological guidance recommending MI under MAR as a primary analytic strategy, accompanied by formal sensitivity analyses to assess robustness to departures from this assumption. The routine incorporation of MNAR sensitivity analyses, such as delta-adjustment, can enhance transparency and strengthen confidence in inferences derived from incomplete data in population-based studies.

Strengths and Limitations

A key strength of this study is the explicit evaluation of the missing data mechanism through the integration of formal testing, MI, and delta-adjustment sensitivity analysis within a unified framework, extending beyond common practice where MI under MAR is often applied without robustness assessment [15]. The use of complementary analytic approaches enabled transparent comparisons of efficiency and stability, while sensitivity analyses demonstrated that the primary findings were not driven solely by the MAR assumption. Additional strengths include the large sample size and the inclusion of diverse demographic, socioeconomic and environmental covariates, supporting precision and generalizability.

Several limitations should be acknowledged. As with all missing data analyses, the true missingness mechanism is not identifiable from observed data alone, and delta-adjustment relies on assumptions about MNAR departures that cannot be empirically verified [37].An additional limitation is that the delta-adjustment sensitivity analysis was applied only to the outcome variable, whereas missing covariate values were imputed under the MAR assumption. In particular, vaccination status had substantial missingness and may also be subject to MNAR mechanisms. Therefore, the analysis does not account for possible departures from MAR in the covariates. Future studies could extend the framework by incorporating sensitivity analyses for key incomplete covariates to assess the impact of alternative MNAR assumptions on the study conclusions. Substantial missingness in some covariates may have reduced precision despite MI [36], and the cross-sectional design precludes causal interpretation [40].

Supplementary Material

Supplementary material can be found online at the journal website.

Authors Contributions

MS: Conceptualization, Methodology, Formal analysis, Writing - original draft. AS: Conceptualization, Writing - review and editing, Supervision. HM: Writing - review and editing, Supervision, Project administration.

Funding

The authors declare that no funding was received for this study.

Data Availability

The data used in this study are publicly available through the MICS program upon request via https://mics.unicef.org/surveys.

Ethical Approval

Informed consent was obtained from all participants during the original data collection of the 2022–2023 Yemen MICS. Ethical approval was secured at the time of the survey. As this study involved secondary analysis of anonymized data, no additional ethical approval was required.

Conflict of Interest

The authors declare no conflicts of interest.

REFERENCES

1. J.A.C. Sterne, I.R. White, J.B. Carlin, et al., *Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls*, BMJ, vol. 338, b2393, 2009.
2. A.B. Pedersen, E.M. Mikkelsen, D. Cronin-Fenton, et al., *Missing data and multiple imputation in clinical epidemiological research*, Clinical Epidemiology, vol. 9, pp. 157–166, 2017.
3. D.B. Rubin, *Inference and missing data*, Biometrika, vol. 63, pp. 581–592, 1976.
4. R.J.A. Little, and D.B. Rubin, *Statistical analysis with missing data*, 3rd ed., Hoboken, NJ: Wiley, 2019.
5. O. Harel, E.M. Mitchell, N.J. Perkins, et al., *Multiple imputation for incomplete data in epidemiologic studies*, American Journal of Epidemiology, vol. 187, pp. 576–584, 2018.
6. D.B. Rubin, *Multiple imputation for nonresponse in surveys*, New York, NY: Wiley, 1987.
7. S. van Buuren, and K. Groothuis-Oudshoorn, *mice: multivariate imputation by chained equations in R*, Journal of Statistical Software, vol. 45, pp. 1–67, 2011.
8. I.R. White, P. Royston, and A.M. Wood, *Multiple imputation using chained equations: issues and guidance for practice*, Statistics in Medicine, vol. 30, pp. 377–399, 2011.
9. P.C. Austin, I.R. White, D.S. Lee, and S. van Buuren, *Missing data in clinical research: a tutorial on multiple imputation*, Canadian Journal of Cardiology, vol. 37, pp. 1322–1331, 2021.
10. S.R. Seaman, J. Galati, D. Jackson, and J.B. Carlin, *What is meant by "missing at random"*, Statistical Science, vol. 28, pp. 257–268, 2013.
11. R.A. Hughes, J. Heron, J.A.C. Sterne, and K. Tilling, *Accounting for missing data in statistical analyses: multiple imputation is not always the answer*, International Journal of Epidemiology, vol. 48, pp. 1294–1304, 2019.
12. J.R. Carpenter, and M.G. Kenward, *Multiple imputation and its application*, Chichester, UK: Wiley, 2013.
13. M.J. Daniels, and J.W. Hogan, *Missing data in longitudinal studies: strategies for Bayesian modeling and sensitivity analysis*, Boca Raton, FL: Chapman & Hall/CRC, 2008.
14. G. Molenberghs, G. Fitzmaurice, M.G. Kenward, A. Tsiatis, and G. Verbeke, eds., *Handbook of missing data methodology*, Boca Raton, FL: Chapman & Hall/CRC, 2014.
15. J.R. Carpenter, and M. Smuk, *Missing data: a statistical framework for practice*, Biometrical Journal, vol. 63, pp. 915–947, 2021.
16. R.J. Little, *Pattern-mixture models for multivariate incomplete data*, Journal of the American Statistical Association, vol. 88, pp. 125–134, 1993.
17. R.J. Little, *A class of pattern-mixture models for normal incomplete data*, Biometrika, vol. 81, pp. 471–483, 1994.
18. F.P. Leacy, S. Floyd, T.A. Yates, and I.R. White, *Analyses of sensitivity to the missing-at-random assumption using multiple imputation with delta adjustment: application to a tuberculosis/HIV prevalence survey with incomplete HIV-status data*, American Journal of Epidemiology, vol. 185, pp. 304–315, 2017.
19. D.M. Tompsett, F. Leacy, M. Moreno-Betancur, J. Heron, and I.R. White, *On the use of the not-at-random fully conditional specification (NARFCS) procedure in practice*, Statistics in Medicine, vol. 37, pp. 2338–2353, 2018.
20. P.H. Rezvan, K.J. Lee, and J.A. Simpson, *Sensitivity analysis within a multiple imputation framework using delta-adjustment: application to a longitudinal study*, Longitudinal and Life Course Studies, vol. 9, pp. 259–278, 2018.

21. J.C. Jakobsen, C. Gluud, J. Wetterslev, and P. Winkel, *When and how should multiple imputation be used for handling missing data in randomised clinical trials: a practical guide with flowcharts*, BMC Medical Research Methodology, vol. 17, 162, 2017.
22. J. Zhang, S.G. Dashti, J.B. Carlin, K.J. Lee, and M. Moreno-Betancur, *Recoverability and estimation of causal effects under typical multivariable missingness mechanisms*, Biometrical Journal, vol. 66, e202200326, 2024.
23. S. Khan, and A. Hancioglu, *Multiple Indicator Cluster Surveys: delivering robust data on children and women across the globe*, Studies in Family Planning, vol. 50, pp. 279–286, 2019.
24. Central Statistical Organization (CSO), United Nations Children's Fund (UNICEF), *Yemen Multiple Indicator Cluster Survey 2022–2023: survey findings report*, Sana'a, Yemen: CSO and UNICEF, 2023.
25. R.J.A. Little, *A test of missing completely at random for multivariate data with missing values*, Journal of the American Statistical Association, vol. 83, pp. 1198–1202, 1988.
26. D.W. Hosmer, and S. Lemeshow, *Applied logistic regression*, 2nd ed., New York, NY: Wiley, 2000.
27. D.B. Rubin, *Multiple imputations in sample surveys: a phenomenological Bayesian approach to nonresponse*, in Proceedings of the Survey Research Methods Section, American Statistical Association, Vol. 1, Alexandria, VA: American Statistical Association, pp. 20–34, 1978.
28. P. T. von Hippel, *How many imputations do you need? A two-stage calculation using a quadratic rule*, Sociological Methods & Research, 49(3), 699–718, 2020.
29. S. van Buuren, *Flexible imputation of missing data*, 2nd ed., Boca Raton, FL: Chapman & Hall/CRC, 2018.
30. E. Costantini, K. M. Lang, K. Sijtsma, and T. Reeskens, *Solving the many-variables problem in MICE with principal component regression*, Behavior Research Methods, 56, 1715–1737, 2024.
31. V. Liublinska, and D.B. Rubin, *Sensitivity analysis for a partially missing binary outcome in a two-arm randomized clinical trial*, Statistics in Medicine, vol. 33, pp. 4170–4185, 2014.
32. S. Cro, T.P. Morris, M.G. Kenward, and J.R. Carpenter, *Sensitivity analysis for clinical trials with missing continuous outcome data using controlled multiple imputation: a practical guide*, Statistics in Medicine, vol. 39, pp. 2815–2842, 2020.
33. E. Kawabata, D. Major-Smith, G.L. Clayton, et al., *Accounting for bias due to outcome data missing not at random: comparison and illustration of two approaches to probabilistic bias analysis: a simulation study*, BMC Medical Research Methodology, vol. 24, 278, 2024.
34. K.J. Lee, J.B. Carlin, J.A. Simpson, and M. Moreno-Betancur, *Framework for the treatment and reporting of missing data in observational studies*, Journal of Clinical Epidemiology, vol. 134, pp. 79–88, 2021.
35. C. Rioux, and T.D. Little, *Missing data treatments in intervention studies: what was, what is, and what should be*, International Journal of Behavioral Development, vol. 45, pp. 51–58, 2019.
36. P. Madley-Dowd, E. Curnow, R.A. Hughes, and K. Tilling, *Analyses using multiple imputation need to consider missing data in auxiliary variables*, American Journal of Epidemiology, vol. 193, pp. 1721–1729, 2025.
37. S. Gachau, M. Quartagno, E.N. Njagi, et al., *Handling missing data in modelling quality of clinician-prescribed routine care: sensitivity analysis of departure from missing at random assumption*, Statistical Methods in Medical Research, vol. 29, pp. 3076–3092, 2020.
38. T.R. Sullivan, A.B. Salter, P. Ryan, and K.J. Lee, *Approaches to addressing missing values, measurement error, and confounding in epidemiologic studies*, Journal of Clinical Epidemiology, vol. 131, pp. 89–100, 2021.
39. J. Weberpals, S.R. Raman, P.A. Shaw, et al., *A principled approach to characterize and analyze partially observed confounder data from electronic health records*, Clinical Epidemiology, vol. 16, pp. 235–249, 2024.
40. M.A. Hernán, and J.M. Robins, *Causal inference: what if*, Boca Raton, FL: Chapman & Hall/CRC, 2020.