



Survey implementation process and interviewer effects on skipping sequence of maternal and child health indicators from National Family Health Survey: An application of cross-classified multilevel model

Radhika Sharma^a, Laxmi Kant Dwivedi^{b,*}, Somnath Jana^a, Kajori Banerjee^c, Rakesh Mishra^d, Bidhubhusan Mahapatra^e, Damodar Sahu^f, S.K. Singh^b

^a International Institute for Population Sciences, Mumbai, India

^b Department of Survey Research & Data Analytics, International Institute for Population Sciences, Mumbai, India

^c SVKM's Narsee Monjee Institute of Management Studies (NMIMS), Mumbai, India

^d UNICEF, New Delhi, India

^e Population Council, New Delhi, India

^f National Institute of Medical Statistics, Indian Council of Medical Research, New Delhi, India

ARTICLE INFO

Keywords:

Survey design
Survey implementation
Interviewer effect
Cross-classified multilevel model
Level weights
Maternal and child health
Team level variation

ABSTRACT

Implementing a large-scale survey involves a string of intricate procedures exposed to numerous types of survey errors. Uniform and systematic training protocols, comprehensive survey manuals, and multilayer supervision during survey implementation help reduce survey errors, providing a consistent fieldwork environment that should not result in any variation in the quality of data collected across interviewers and teams. With this background, the present study attempts to delineate the effect of field investigator (FI) teams and survey implementation design on the selected outcomes. Data on four of the bigger Empowered Action Group (EAG) states of India, namely Uttar Pradesh, Madhya Pradesh, Bihar, and Rajasthan, were obtained from the fourth round of the National Family Health Survey (NFHS-4) for analysis. A fixed-effect binary logistic regression model was used to assess the effect of FI teams and survey implementation design on the selected outcomes. To study the variation in the outcome variables at the interviewer level, a cross-classified multilevel model was used. Since one interviewer had worked in more than one primary sampling unit (PSU) & district and did not follow a perfect hierarchical structure, the cross-classified multilevel model was deemed suitable. In addition, since NFHS-4 used a two-stage stratified sampling design, two-level weights were adjusted for the models to compute unbiased estimates. This study demonstrated the presence of interviewer-level variation in the selected outcomes at both inter- and intra-field agencies across the selected states. The interviewer-level intra-class correlation coefficient (ICC) for women who had not availed antenatal care (ANC) was the highest for eastern Madhya Pradesh (0.23) and central Uttar Pradesh (0.20). For 'immunisation card not seen', Rajasthan (0.16) and western Uttar Pradesh (0.13) had higher interviewer-level ICC. Interviewer-level variations were insignificant for women who gave birth at home across all regions of Uttar Pradesh. Eastern Madhya Pradesh, Rajasthan, and Bihar showed higher interviewer-level variation across the selected outcomes, underlining the critical role of agencies and skilled interviewers in different survey implementation designs. The analysis highlights non-uniform adherence to survey protocols, which implies that not all interviewers and agencies performed in a similar manner in the field. This study recommends a refined mechanism for field implementation and supervision, including focused training on the challenges faced by FIs, random vigilance, and morale building. In addition, examining

Abbreviations: FI, field investigator; EAG, empowered action group; PSU, primary sampling unit; ICC, intra-class correlation coefficient; FA, field agencies; NFHS, National Family Health Survey; CAPI, computer-assisted personal interviewing; SDGs, Sustainable Development Goals; ANC, antenatal care; UP, Uttar Pradesh; MP, Madhya Pradesh.

* Corresponding author.

E-mail addresses: radhikash279@gmail.com (R. Sharma), laxmikant@iipsindia.ac.in, laxmikdwivedi@gmail.com (L.K. Dwivedi), somnathjana443@gmail.com (S. Jana), kajori.b2012@gmail.com (K. Banerjee), rakeshgnu31@gmail.com (R. Mishra), bbmahapatra@popcouncil.org (B. Mahapatra), sahud@icmr.gov.in (D. Sahu), sk Singh31962@gmail.com (S.K. Singh).

<https://doi.org/10.1016/j.ssmph.2022.101252>

Received 22 December 2021; Received in revised form 28 August 2022; Accepted 2 October 2022

Available online 3 October 2022

2352-8273/© 2022 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

interviewer-level characteristics, field challenges, and field agency effects may also highlight the roots of interviewer-level variation in the data. However, based on the interviewer's performance in the field, the present study offers an intriguing insight into interviewer-level variations in the quality of data.

1. Introduction

Large-scale surveys are a substitute for the Census during the interim period as they are more economical and relatively less time consuming, and provide focused insights on specific characteristics of the larger population (Division, 1955; Gideon, 2012; Shryock, 1975; Srinivasan, 1998). In India, several nationally representative large-scale surveys, such as the National Family Health Survey (NFHS), National Sample Survey (NSS), Indian Human Development Survey (IHDS), Longitudinal Aging Survey (LASI), and Periodic Labor Force Survey (PLFS) have been vital facilitators for the formulation and assessment of policy-led programmatic interventions at various levels.

Besides tracking fertility, infant and child mortality, maternal and child health, and other health and family welfare indicators at the national and state levels, the fundamental focus of these surveys has been to provide high-quality data for use by researchers. Numerous measures have been adopted in these surveys over time to maintain high data quality standards. The implementation of a large-scale survey constitutes a series of complex challenges from survey planning to implementation to dissemination. As a strategic component of any survey design, survey implementation is prone to various types of errors, including sampling and non-sampling errors, which are inevitably present in every step of the survey (Verma & Le, 1996; Weisberg, 2008; Wolf, 2016).

Several researchers have highlighted inconsistencies in data quality in India in national and subnational assessments (Borkotoky & Unisa, 2014; Dandona, Pandey, & Dandona, 2016). However, unlike the NFHS (the Indian version of the Demographic Health Survey), none of these surveys have regular follow-up strategies to evaluate data quality issues (Srinivasan and Mishra, 2020). Since the inception of NFHS in India, numerous researchers have highlighted data quality issues to suggest improvements in survey design and implementation (Borkotoky & Unisa, 2014; Mari Bhat, 1995; Rajan, 2008; Retherford, 2001). Several studies have pointed out the underreporting and exaggeration of statistics for the same indicators obtained from different public data sources (Sandefur, 2014). In fact, discrepancies are also conspicuous in the estimates obtained from system data and external facility survey data (Phillips, 2019; Sharma et al., 2016). A majority of these attempts have highlighted inconsistencies in the reporting of age, fertility, mortality, timing of events, and reproductive and child health outcomes (Arnold, 1990; Bhat, 1995; Curtis, 1995; Dandona, Pandey, & Dandona, 2016; Gage, 1995; Namaste, 2018; Pullum, 2018; Schoumaker, 2014; Unisa, 2015). Many studies (Bogen, 1996; Johnson, 2009; Pullum, 2018; Sharma et al., 2016) have discerned the impact of length of questionnaire, positioning, time dimensions associated with specific questions, and biases associated with interviewer characteristics on data quality. The assessment of the first three rounds of NFHS in terms of questionnaire length and enquiries on socially restricted topics raised serious data quality concerns (James Ranjan, 2004; Ranjan and James 2004; Ranjan and James 2008).

In any populous and resource-deficient country, there are numerous challenges concerning the collection of high-quality data. Therefore, innovative strategies are required to collect quality data while maintaining time and cost effectiveness. A commitment to high-quality standards in large-scale surveys is manifested in the survey implementation strategies by field agencies (FAs) and compliance with protocols by field investigators (FIs) at the time of survey (Wolf, 2016). In a country like India with a vast spatial distribution, surveys are conducted in two phases to facilitate systematic monitoring and evaluation of both FAs and FIs (IIPS, 2015–16). However, the increasing demand for data at lower administrative units has amplified the risk of errors in survey implementation.

NFHS meticulously trains the trainers and exercises strict supervision to monitor and evaluate the survey operations to ensure that the data is highly reliable however, a study highlighted the interviewer effect across two rounds of NFHS in the case of sensitive questions (Singh et al., 2022). Despite the rigorous training imparted to the investigators to keep the canvassing uniform, bias introduced by the investigators cannot be ignored. The characteristics of the interviewers, such as their age, education, religion, caste, and place of residence, influenced the quality of the data collected. One study noted a significant difference in the non-response and refusal rates based on interviewers' age (Pullum, 2018). Reportedly, older interviewers have low non-response and refusal rates compared to younger investigators. Similar findings have been observed with regard to interviewers' education level (Yang, 2008). A series of analytical studies (Adida, 2016; Liu, 2016; Wang, 2013) showed that the co-ethnicity of FIs and respondents positively impacted data quality. In addition, investigators with prior field experience show a decline in non-response and other inconsistencies during field investigation (Pullum, 2018). Numerous studies have argued that an interviewer's inexperience, inappropriate articulation of questions, inattentiveness, non-cooperation with interviewee, and inappropriate behaviour can seriously impede quality data in survey research (Anglewicz, 2009; Pullum, 2018). However, there is a paucity of literature on how field teams should be trained to perform complex field operations, particularly in the Indian context.

A considerable number of non-sampling errors were attributed to the way the questions in the questionnaire were designed. As part of the questionnaire design, a sequence of skipping is introduced in a questionnaire to navigate respondents to specific questions by eliminating sets of inapplicable questions (Allen, 2020; Manski & Molinari, 2008). These skips are regarded as an important caveat to reduce the burden of extra questions on the respondents and the interviewers (Amos, 2018). Simultaneously, they reduce the time and cost of the survey and enhance the accuracy of the reported data. However, there are several drawbacks to this approach. Skipping is often opted for as a strategic bypass to necessary qualifications, leading to the anomalous reporting of events (Bound, 2001; Mathews, 2012). A massive expansion of the NFHS questionnaire over time has created a noticeable increase in skipping of questions, thus increasing the possibility of erroneous responses. It is important to mention that the use of computer-assisted personal interviewing (CAPI) has facilitated programmed skipping in questionnaires (IIPS, 2015–16; IIPS & ICF, 2017). Thus, any discrepancies in data quality in a homogeneous and identical setting indicate severe incongruence in survey implementation in India. However, no study has attempted to explore this aspect in detail with regard to challenges in survey implementation, follow-up of survey protocols, and discrepancies in the performance of FAs and FIs in the field survey.

In the context described above, this study attempts to understand survey implementation strategies and identify the influence of FAs and FIs on data quality in India. However, due to the paucity of paradata on interviewers' characteristics, this study relied on the skipping behaviour of FIs associated with different FAs. This study examined the patterns of skipping in maternal and child health indicators within homogeneous rural clusters in the selected states. It also investigated the presence of inter- and intra-agency variations in data quality. We chose maternal and child health indicators to identify the pattern of inconsistencies in the country for two reasons. First, NFHS is an invaluable source of health information. As such, it is used for program-led policy formulations and for the assessment of the nation's commitment to the Sustainable Development Goals (SDGs), especially those pertaining to universal health coverage, including disaggregated targets of child nutrition and

mortality, maternal health, and gender-related goals (United Nations, 2015). Therefore, the quality of the data should be top-notch. Second, it is important to examine survey implementation strategies and the performance of FAs and FIs to ensure data quality in future surveys.

The present study hypothesised that with a uniform and systematic survey implementation procedure and a similar training protocol for fieldwork, there should be no difference in the quality of data collected across teams and interviewers. Based on this assumption, this study had two important aims. First, we attempted to assess the effect of the FI teams and the survey implementation design on the selected outcomes, namely births delivered in an institution (home versus health facility), visits for antenatal care (ANC) (no visit versus any visits), and the investigator seeing the immunisation card (no versus yes). Variations in responses with respect to these outcomes may indicate that the investigators made a deliberate attempt to reduce their workload by skipping certain questions. To achieve this objective, we adjusted the survey implementation design along with the PSU- and individual-level covariates as we hypothesised that after adjusting for these factors, the team-level variations related to a given outcome should not be significant. Our second aim was to explore variations in outcomes at the investigator level. For this, we used a cross-classified multilevel model to ascertain the impact of an interviewer on the same outcomes as mentioned above after adjusting for the PSU- and individual-level covariates. In summary, this study highlights the critical roles of survey implementation procedures and the interviewer in the data quality of a survey.

2. Materials and methods

2.1. Data

The NFHS is a nationally representative cross-sectional survey conducted under the guidance of the Ministry of Health and Family Welfare (MoHFW), Government of India. As the nodal agency, the International Institute for Population Sciences (IIPS), Mumbai, is responsible for collecting reliable unit-level data on various aspects of population, nutrition, maternal and child health, and family welfare in India (IIPS & ICF, 2017). Until 2016, four rounds of the survey had been carried out in India at different levels of disaggregation. In the fourth round (NFHS-4), nearly 0.6 million households, 2.8 million persons, and 0.7 million eligible women aged 15–49 years were covered across the nation using a stratified two-stage sampling procedure. With this round, NFHS, for the first time, reached down to the district level, covering all the states and union territories of India. It is important to mention that to cover such a large sample size, the survey implementation design of NFHS includes phase-wise collection of data and division of larger states into regions. These regions/states covering all parts of India were surveyed by 14 field agencies, constituting 789 teams (IIPS & ICF, 2017). Each team consisted of one field supervisor, three female interviewers, one male interviewer, two health investigators, and one driver.

2.1.1. Selection approach of the study sample

For the purpose of analysis, we selected four of the larger states from the Empowered Action Group (EAG): Uttar Pradesh (UP), Madhya Pradesh (MP), Bihar, and Rajasthan. The EAG states have relatively greater health and income inequalities than the other states. The NFHS-4 was conducted in two phases. In MP, Bihar, and eastern UP, the survey was conducted in Phase I, while western UP, central UP, and Rajasthan were covered in Phase II.

The populous state of UP was divided into three regions—western, central, and eastern—and the field operation was assigned to three different FAs. The Population Research Centre (PRC), Lucknow, covered an average of 45 PSUs and 1261 women per team in eastern UP, while the Development and Research Services (DRS) and the Goa Institute of Management (GIM) covered an average of 47 PSUs/1066 women and 40 PSUs/1150 women per team, respectively, in the central and western

UP. With an aggregate sample size of 97,661 women, each team covered an average of 44 PSUs, interviewing approximately 26 women per PSU in UP. The number of FIs assigned to each state was based on the number of districts and PSUs in the state and the sample size of the state (IIPS, 2015–16). On an average, 1.25 teams were assigned per district in India (IIPS, 2015–16).

As part of the survey implementation design, districts in each region were further divided into sets of 4–5 adjoining districts to ensure systematic accumulation of data and close monitoring of FAs and FIs. FI teams were assumed to complete the survey in a set and then move forward to the next assigned set of districts. Thus, in a particular set, several teams were assigned to different districts such that each team covered a number of districts belonging to different sets in a region. This systematic movement of teams in assigned districts from various sets formed a group consisting of districts from different sets covered by several teams (Fig. 1). The current study focuses only on the rural areas of the selected states to avoid sample size constraints and the heterogeneous movement of teams in urban areas.

2.1.2. Outcome variables

Skip sequencing is a widespread survey practice in which the response to an opening question is used to determine whether a respondent should be asked some of the subsequent questions (Dunn et al., 2015). The outcome variable in this study was defined as skipping of the opening questions, which may have led to bypassing detailed information on key indicators related to maternal and child health for a certain number of respondents. In rural clusters, it is believed that the pattern of skipping questions tends to be nearly identical because of the homogeneity of socioeconomic settings. Based on this assumption, three important maternal and child health indicators with lengthy follow-up sub-sections were considered in this study. These variables included ‘not availing antenatal care services’, ‘birth at home’, and ‘immunisation card not seen’. Women who had not received antenatal care services during pregnancy skipped approximately eight questions on antenatal care and a certain number of questions from the state module. Thus, with regard to skips, the outcome variable was coded as ‘1’ if antenatal care was not availed and as ‘0’ otherwise.

Similarly, 17 questions on institutional delivery care and costs were mandatory skips for women who had delivered at home. The immunisation card table entries were overlooked for children who had an immunisation card but which was not seen by the FIs. The ‘birth at home’ variable comprised women who delivered a child at home, including own home, parent’s home, or another home. The ‘immunisation card not seen’ variable was applied to women who had their child’s immunisation card but could not bring it to the interviewer on time, as a result of which the investigator might have avoided recording the data from the vaccination card. Immunisation information obtained from a vaccination card is considered more reliable than parental recall; however, due to the complexity of reading and recording the information on immunisation from the vaccination card, interviewers may do so without the card. Wagner (2019) showed substantial differences in immunisation estimates of diphtheria-tetanus-pertussis (DPT) across individuals with a vaccination card, without a vaccination card, and with a vaccination card not seen by the interviewer. This underlines the importance of recording immunisation information directly from the card and therefore the outcome variable ‘card not seen’ is a significant variable of study, which should not vary at the investigator level. Although the use of skip sequencing helps to pass over irrelevant questions for a particular respondent, reducing the respondent burden and the time and cost of the interview, it may also lead to data quality issues across the survey items for various reasons, thereby curtailing the informativeness of the survey.

2.2. Explanatory variables

The key explanatory variables were the field survey implementation

Group	Regional Sub Set	Set 1	Set 2	Set 3			
	Teams/Districts	Bara Banki	Fatehpur	Kanpur Nagar	Banda		
Group A	1	138	212	240	174		
	2	149	200	105	119		
	3	167	156	136	115		
	4	217	113	146	137		
	5	200	120	112	107		
Group B	Regional Sub Set	Set 1		Set 3	Set 4		
	Teams/Districts	Hardoi	Lucknow	Hamirpur	Auraiya	Jalaun	
Group B	6	181	81	263	266	121	
	7	179	69	164	180	118	
	8	213	167	209	262	185	
	9	149	51	29	63	153	
	10	113					
Group C	Regional Sub Set	Set 1	Set 2	Set 4	Set 5		
	Teams/Districts	Sitapur	Unnao	Kanpur Dehat	Jhansi	Lalitpur	Mahoba
Group C	11	206	216	180	134	110	168
	12	167	185	180	160	125	213
	13	149	125	153	137	120	118
	14	115	146	141	118	98	165
	15	192	204	200	118	80	33
Group D	Regional Sub Set	Set 1	Set 2	Set 3			
	Teams/Districts	Kheri	Rae Bareli	Chitrakoot			
Group D	16	218	196	259			
	17	271	249	244			
	18	175	290	201			
	19	262	271	244			

Fig. 1. Schematic representation of number of women interviewed per team across sets and districts in rural area of central region of Uttar Pradesh, NFHS-4, 2015-16.

teams and interviewers. Each team consisted of 7 people and a field supervisor. The field supervisor variable was considered a proxy for the team, and the interviewer identification variable was considered to represent the interviewer. These teams were exclusive to regions and states as different agencies were involved in data collection in different regions and states. The analysis was adjusted for the number of districts a particular team interviewed in a given region and for women’s demographic and socioeconomic characteristics, as follows:

- A) Individual-level characteristics: A fixed-effects model and a cross-classified multilevel model were adjusted for women’s age, educational status, religion, and caste and wealth status of the household.
- B) Village-level characteristics: At the village level, the proportion of literate women, poorest women, Hindu women, and scheduled caste (SC)/scheduled tribe (ST) women was adjusted in the analysis.

2.3. Statistical analysis

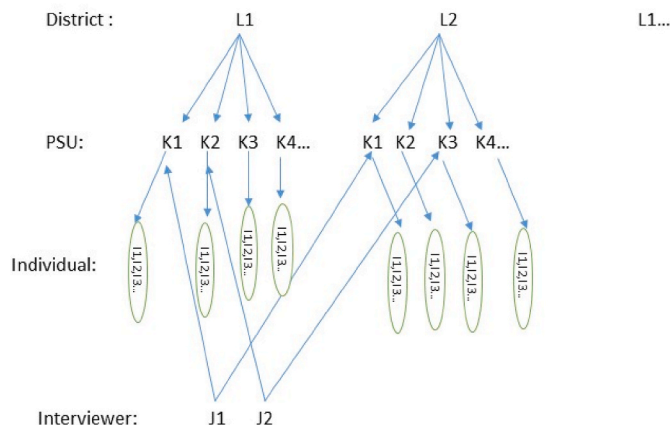
2.3.1. Fixed effects model

To fulfil the first objective, a fixed-effects binary logistic regression model was used after adjusting for the effects of teams, groups of districts, and selected women’s demographic, socioeconomic, and village-level variables. The outcome variables were binary in nature, whereby women who had not availed ANC, women who had given birth at home, and women whose immunisation card was not seen by the interviewer were marked as ‘1’ and as ‘0’ otherwise.

2.3.2. Cross-classified multilevel model

We employed a cross-classified random intercept multivariable multilevel (four-level) logit model to examine the interviewer variation in the selected outcomes. The model defined the outcome for the i^{th} respondent interviewed by the j^{th} investigator at the k^{th} PSU of the l^{th} district. The random intercept model was cross-classified because each interviewer had worked in more than one PSU&district and each PSU&district had more than one interviewer. Thus, interviewers with PSU&districts do not follow a perfect hierarchical structure as in the case of PSUs-districts. Therefore, cross-classified multilevel model was used to disentangle the non-hierarchical effects (Dunn et al., 2015).

A flowchart to clarify the hierarchical structure is given below:



Different socioeconomic and demographic variables at the individual and PSU levels were used as the independent variables. The model can be expressed as:

$$\log \left(\frac{\pi_{i(j,k)}}{1 - \pi_{i(j,k)}} \right) = \beta_0 + \beta_1^T x_{i(j,k)} + \beta_2^T x_{0k} + \varepsilon_{0j} + \varepsilon_{0k} + \varepsilon_{0l}$$

Here, $\pi_{i(j,k)}$ represents the probability that the i^{th} respondent interviewed by the j^{th} investigator in the k^{th} PSU of the l^{th} district takes the value of 1 for one of the categories of the dependent variables, as defined above. β_1^T is a vector containing the coefficients of individual-level socioeconomic and demographic characteristics, whereas β_2^T is a vector containing the coefficients of PSU-level covariates. ε_{0j} , ε_{0k} , and ε_{0l} represent the error terms at the interviewer, PSU, and district level, respectively.

$$\varepsilon_{0j} \sim N(0, \sigma_{0j}^2), \varepsilon_{0k} \sim N(0, \sigma_{0k}^2), \varepsilon_{0l} \sim N(0, \sigma_{0l}^2)$$

: σ^2 are variances at different levels

To estimate the cross-classified random intercept model, the Markov Chain Monte Carlo (MCMC) estimation process was adopted, and the model was executed using the runmlwin program in STATA 16.0. Level weights were applied to obtain precise estimates of the coefficients/parameters. In the absence of level weights, variance tends to be underestimated (Elkasabi et al., 2020).

We computed the intra-class correlation coefficient (ICC) to obtain

$$f = \frac{d_HH}{\frac{\text{Total number of census cluster for a particular stratum}}{\text{Total number of completed cluster for a particular stratum}} * \frac{\text{Average number of household per cluster for a particular stratum}}{\text{Number of household selected per cluster}}}$$

the share of total variance in the outcome variable at various levels. For example, the ICC at the district level was defined as the district-level variance relative to the sum of the variance at all levels.

ICC at the district level

$$= \frac{\text{District level variance}}{(\text{District level variance} + \text{PSU level variance} + \text{Investigator level variance} + \text{individual level variance})}$$

$$\text{ICC at the district level} = \left[\frac{\sigma_{0l}^2}{\sigma_{0l}^2 + \sigma_{0j}^2 + \sigma_{0k}^2 + 3.29} \right]$$

The ICC at the district level represents the correlation of the outcome between two different women from the same district who were interviewed by different interviewers in different PSUs.

2.3.3. Computation of sampling weight at different levels

Two-stage stratified sampling was used to collect rural and urban samples separately from each of the 640 districts in India. For villages larger than 300 households, two segments of size 100–150 households approximately were selected using probability proportional to size (PPS) sampling. Using the Census 2011 frame, villages were taken as the primary sampling units (PSUs) in rural areas, and census enumeration blocks (CEBs) in urban areas. A second-stage stratification in rural areas of each district was achieved based on village size (number of households), whereby three explicit strata were created. Implicit stratification was achieved by sorting the sampling frame according to the percentage of scheduled castes (SC)/scheduled tribes (ST) and female literacy rate.

In the first stage, within each rural stratum, villages were selected using probability proportional to size (PPS) sampling from the 2011 census frame. Within each urban sampling stratum, implicit stratification was achieved by sorting the sampling frames according to the percentage of the SC/ST population. Details of the sampling techniques and sample sizes are available in the NFHS-4 National Report (IIPS and ICF, 2017).

Level 1 and level 2 weight were computed by considering the survey design and adopting the following procedures. First, the total number of interviewed women in the 15–49 years age group was obtained from the state-specific NFHS dataset of ever-married women. Information related to the total number of women in the 15–49 years age group in the state census frame, the total number of households according to the census frame, and the total number of census clusters/PSUs in all strata (rural/urban area of each district) was obtained using the Census 2011 Primary Census Abstract (PCA) file (<https://censusindia.gov.in/census.websit e/>). It was not possible to obtain information related to each stratum of the rural area of each district; thus, rural/urban stratification was performed to compute level weights.

The de-normalized weight (d_HH) was calculated for a particular state as follows:

$$d_{HH} = \frac{sv005}{1000000} * \frac{\text{Total number of women in 15 - 49 age group in census frame}}{\text{Total number of women interviewed in 15 - 49 age group}}$$

Next, we approximated the level 2 weight based on the equal split method (Elkasabi et al., 2020)

Finally, the level 2 weight was generated by:

$$Wt2 = \frac{\text{Total number of census cluster for a particular stratum}}{\text{Total number of completed cluster for a particular stratum}} * f$$

The level 1 weight was calculated by:

$$Wt1 = \frac{d_HH}{Wt2}$$

Since this study was based on the state level, state weight (sv005) was used instead of the national weight (v005) while calculating the de-normalized weight. We took α as 0.5 to give optimal values to the weight (Elkasabi et al., 2020).

3. Results

3.1. Descriptive statistics of the sample

Table 2 presents the percentage of women who did not avail ANC services, the percentage of women who had delivered at home, and the percentage of women whose child's immunization card was not seen by investigators to record the data despite respondents having the card as an indicator of data quality outcome according to the selected

Table 1
Fieldwork implementation structure in the selected states in NFHS-4, 2015-16.

S. No.	States	Regions	Survey Agencies	No. Teams	Districts	No. of PSU	Sets	No. of Women
1	Uttar Pradesh	Western	PRC Lucknow	34	26	1531	5	42885
		Central	DRS	19	18	903	5	20266
		Eastern	GIM	30	27	1204	6	34510
2	Madhya Pradesh	Western	AMS	35	27	1421	6	38739
		Eastern	IIHMR	28	22	989	5	24064
3	Bihar	-	AMS	48	38	1677	8	45812
4	Rajasthan	-	IIHMR	41	31	1634	6	41965

Note: PRC- Population research center (PRC Lucknow was involved as FA in NFHS-I and NFHS-IV); DRS- Development and Research Services Pvt. Ltd., New Delhi (DRS was involved as FA in NFHS-III and NFHS-IV); GIM- Goa Institute of Management (GIM was for the first time involved as FA in NFHS-IV); AMS- Academy of Management Studies (AMS) (AMS was for the first time involved as FA in NFHS-IV); IIHMR- Indian Institute of Health Management Research (IIHMR) (IIHMR was involved as FA in NFHS-II, NFHS-III and NFHS-IV).

Table 2
Percentage distribution of skipping of maternal and child health information by socio-demographic characteristics of women in regions of rural Uttar Pradesh, NFHS-4, (2015–16).

Region	UP-C	UP-W	UP-E	UP-C	UP-W	UP-E	UP-C	UP-W	UP-E
Background Characteristic	ANC not availed			Birth at Home			Immunization card not seen		
Wealth Index Quintiles									
Poor	38.29	23.43	34.43	33.91	42.26	37.13	38.69	43.15	53.49
Middle	21.25	11.19	19.29	15.6	32.38	22.1	28.98	36.06	44.85
Rich	9.4	6.05	13.61	13.08	22.6	14.99	36.94	28.26	42.69
Education									
No education	44.56	23.37	43.62	40.88	42.77	44.2	40.49	41.65	59.04
Primary	31.83	16.14	27.25	30.31	40.74	31.04	38.43	38.02	50.15
Secondary	25.11	9.99	18.41	19.09	26.93	18.84	34.42	34.82	42.02
Higher	10.68	4.67	10.28	8.71	12.65	11.36	30.34	30.67	42.25
Women's Age									
15–24	25.11	11.87	22.36	23.01	28.73	22.35	31.87	35.23	43.05
25–34	34.33	16.52	30.5	29.24	37.93	31.82	38.22	38.48	51.24
35+	47.22	29.03	43.27	47.66	46.13	48.54	45	43.57	61.87
Religion									
Hindu	33.07	16.84	28.94	28.78	33.68	30.17	36.09	38.16	49.26
Others	38.45	16.05	38.44	39.98	42.41	41.12	46.52	36.5	57.25
Caste									
SC/ST	36.79	16.12	33.56	34.18	37.17	35.21	34.62	38.67	52.68
Others	31.74	16.83	28.94	27.54	35.24	30.32	38.9	37.56	49.25
Total	1291	997	2852	1688	3257	4382	1411	2010	4708

Note: UP-C= Uttar Pradesh Central region; UP-W= Uttar Pradesh Western region; UP-E= Uttar Pradesh Eastern region.

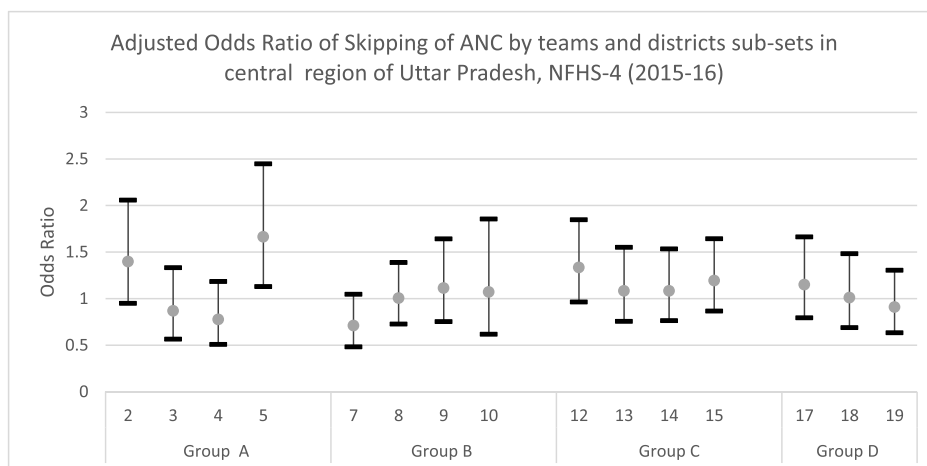
socioeconomic characteristics of women in the different regions of rural Uttar Pradesh. The estimates of the selected outcomes varied considerably across the different regions of Uttar Pradesh as well as by the selected covariates. The prevalence of the outcome variable decreased with an increase in the wealth quantile and level of education, whereas it increased with an increase in women's age and SC/ST population across the regions of Uttar Pradesh. The prevalence of the outcome variable varied marginally by religion and caste in the selected regions and states. Likewise, the prevalence of immunisation cards did not vary greatly in the eastern region of Uttar Pradesh as compared to the central and western regions of the state across the selected socioeconomic characteristics. For example, in western Uttar Pradesh, immunisation cards were not seen for 45 percent of women in the age group 35 years and above, as compared to 62 percent of women in the same age group in eastern Uttar Pradesh. Likewise, immunisation cards were not seen by the investigator for 40 percent and 59 percent of women with no education in central and eastern Uttar Pradesh, respectively. In contrast, immunisation cards were not seen for 30 percent and 42 percent of women with higher education in the same two regions of the state. The prevalence of 'immunisation card not seen' was higher among women belonging to the SC/ST category, except in the central region of the state. A similar pattern was observed in women who had not availed ANC services and those who had delivered at home across the regions of rural Uttar Pradesh. It was found that the prevalence of 'immunisation card not seen' was relatively higher than that of home deliveries and

non-use of ANC across the selected socioeconomic characteristics of women. It is important to mention that heterogeneity in the reporting of the outcome variables across different regions was also observed in all other selected states (Appendix Tables 1–3). These differences may be true in nature or an unexpected outcome of the investigator effect. Furthermore, an attempt was made to adjust the PSU and individual-level covariates to explore the contribution of teams and investigators to the reported differentials in the selected outcome variables.

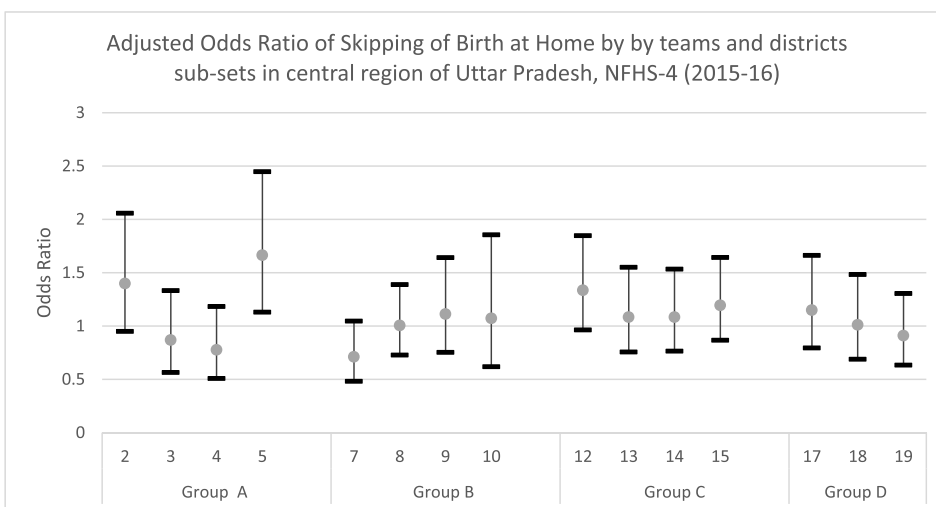
3.2. Team-level variations: Results from the fixed effects model

Fig. 2 presents the results of the fixed effects model by teams for women who did not avail ANC services, women who had delivered at home, and women whose child's immunisation cards were not seen by the investigators as an indicator of data quality outcome adjusted for the survey implementation process and for the PSU- and individual-level women covariates across various groups of teams in central UP.

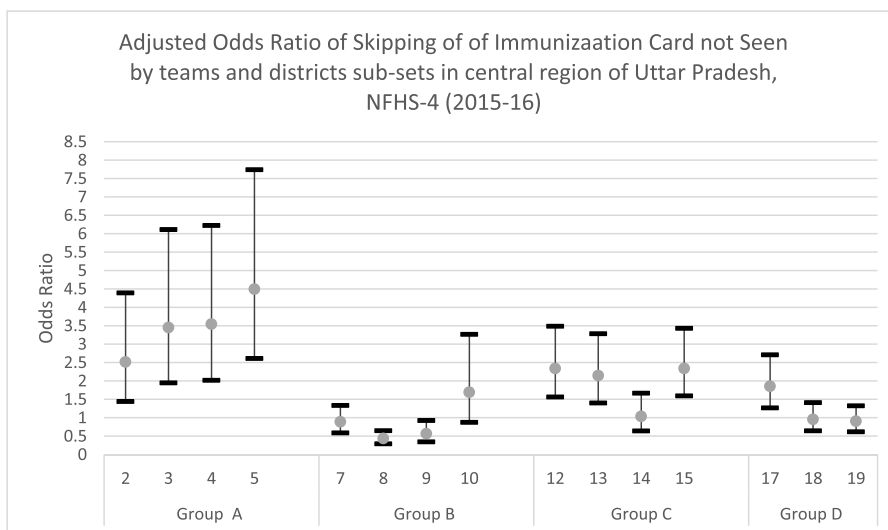
A notable difference was observed in the adjusted odds ratios of teams among various groups in the central region of Uttar Pradesh for the selected outcome variables. For example, all the teams that interviewed group A districts were 3–4 times odds more likely to skip immunisation card information than team 1 from the same group. Similarly, team 12 [Odds ratio (95% CI): 2.33(1.56–3.48)], 13 [2.15 (1.40–3.28)] and 15 [2.34(1.59–3.43)] had significantly higher odds of skipping immunisation card information than team 11 which



Note: Model is adjusted for PSU and Individual level covariates



Note: Model is adjusted for PSU and Individual level covariates



Note: Model is adjusted for PSU and Individual level covariates

Fig. 2. Adjusted odds ratio of skipping of selected maternal and child health information by interviewer teams and districts sub-sets in central region of Uttar Pradesh, NFHS-4 (2015–16).

Note: Model is adjusted for PSU and Individual level covariates.

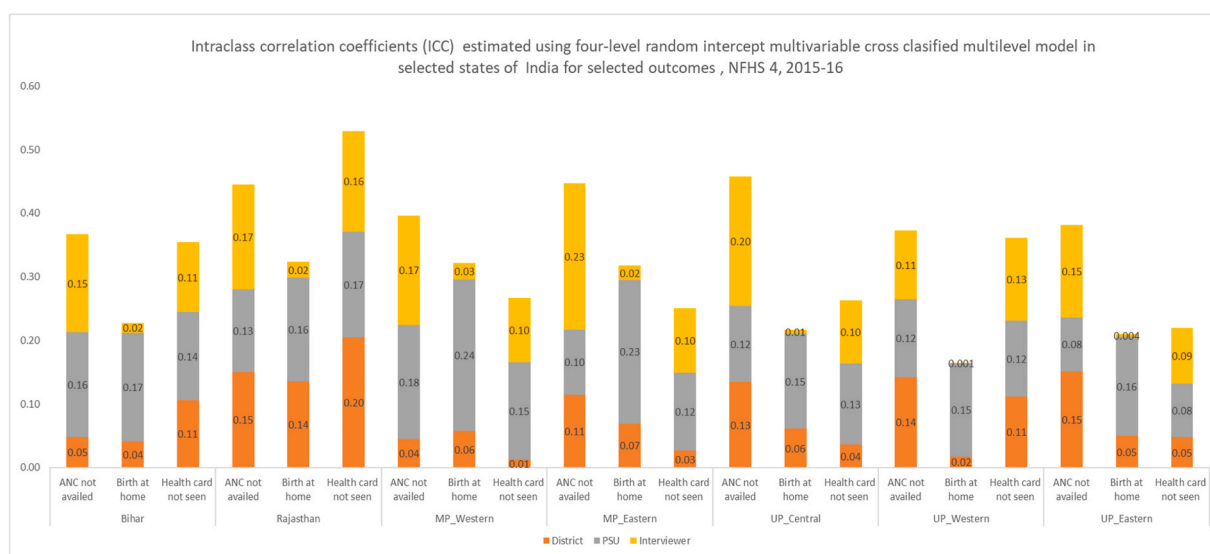
interviewed group C districts. However, among teams that interviewed group B districts, teams 8 [0.43(0.29–0.65)] and 9 [0.56(0.36–0.92)] had significantly lower odds of skipping immunisation card information than team 6 which interviewed the same group of districts. Almost all teams that interviewed groups A and D had significantly lower odds of interviewing women who had not received ANC. Interestingly, in group B, teams 7 [0.44(0.26–0.74)] and 8 [0.62(0.41–0.94)] were less likely to interview women who had not received ANC, whereas team 9 [2.08 (1.28–3.39)] from the same group was twice as likely to interview women who had not received ANC services. Team differentials were also observed among women who had delivered at home. Team 2 [1.39 (0.94–2.05)] and team 5 [1.66(1.13–2.44)] that interviewed group A districts had significantly higher odds of interviewing women who had delivered at home, whereas teams 3 and 4 had insignificantly lower odds of interviewing women who had delivered at home as compared to team 1 from group A. The results show that the different selected outcome variables varied greatly across the same group of teams that interviewed exclusive district groups. For example, among teams that interviewed group A districts, almost all teams had higher odds of skipping immunisation card information; however, the same group of teams had lower odds of interviewing women who had not received ANC (Fig. 2).

The skipping pattern of the selected outcome variables varied across the groups of districts, and there was notable team-level variation in the different regions of the selected states (Appendix Fig. 1). For example, teams from group C in western Madhya Pradesh exhibited a vast variation in the case of women who had not availed ANC and women who had given birth at home, whereas variation in the estimates could be observed for antenatal care information across teams of different groups in the eastern region of Madhya Pradesh. In addition, the analysis shows dissimilarities across teams in the state of Bihar, largely in the absence of ANC and immunisation card not seen in various groups, namely A, B, C, and F (Appendix Fig. 1). Teams belonging to groups D, E, and F in Rajasthan showed significant variations in the estimates of the selected outcome variables (Appendix Fig. 1). Thus, the results clearly highlight the presence of variations in the estimates of the selected maternal and child health variables across groups of teams in the selected states.

3.3. Investigator-level variations: results from the cross-classified multilevel model

Fig. 3 presents the intra-class correlation coefficient (ICC) estimates for the selected outcomes at the district, PSU-village, and interviewer levels using the four-level random intercept multivariable cross-classified model for the selected states of India. The outcome variables varied marginally at the district level across regions and states, with the exception of Rajasthan. At the district level, a major variation was observed for women who had not availed ANC, followed by skipping of immunisation card information. However, at the PSU level, the variation was mainly observed for women who had delivered at home, especially in both regions of Madhya Pradesh. Variations in the outcome variables at the district and PSU levels may be attributed to a combination of diverse latent factors. However, with uniform training and field implementation protocols, after adjusting for selected socioeconomic and demographic covariates at the PSU and individual levels, the variation at the interviewer level should become statistically insignificant.

Based on the above argument, this study presents mixed results, whereby substantial variation in the interviewer effect may be observed for some of the selected outcomes, whereas for the other outcomes, the variation is not statistically significant across the selected states. For the women who did not avail ANC, the interviewer-level ICC varied considerably from the lowest of 0.11 in western Uttar Pradesh to 0.20 in central Uttar Pradesh, to and the highest of 0.23 in eastern Madhya Pradesh (Fig. 3). After controlling for an array of socioeconomic background characteristics, a fair degree of disproportion in the outcome variables across the regions may be attributed to non-uniformity in the survey implementation protocols, which may be further ascribed to several reasons, including lack of motivation to work, geographical and logistical constraints, lack of moral liability, lack of sincerity in work, and delay in salary payment. These are identified as critical impediments to the quality of the data. In addition, for the majority of the selected states, the districts and PSUs contributed less to the variance partition than did the interviewers. The interviewer-level ICC ranged from 0.001 to 0.03 for women who delivered at home, and was found to be the lowest compared to the other selected outcomes. Likewise, for ‘immunization card not seen’, the interviewer-level ICC ranged from 0.09 to 0.16 (Fig. 3). Among the selected states, Rajasthan (0.16) and western Uttar Pradesh (0.13) had a higher interviewer-level ICC,



Note: Model is adjusted for PSU and individual level covariates

Fig. 3. Intra-class correlation coefficient (ICC) estimated using four-level random intercept multivariable cross classified multilevel model in selected states of India for selected outcome, NFHS 4, 2015-16.

Note: Model is adjusted for PSU and Individual level covariates.

Table 3

Estimated Variance from **Cross-classified Multilevel Model** for skipping of maternal and child health information at district, PSU and Interviewer level in selected states of India for selected outcomes, NFHS-4, 2015-16.

State	Levels	ANC not availed		Birth at home		Health card not seen	
		Variance	95% Credible interval	Variance	95% Credible interval	Variance	95% Credible interval
Bihar	District	0.252	0.136, 0.437	0.175	0.097, 0.296	0.541	0.314, 0.906
	PSU	0.854	0.752, 0.962	0.729	0.634, 0.826	0.706	0.584, 0.830
	Interviewer	0.806	0.612, 1.045	0.065	0.038, 0.098	0.563	0.412, 0.748
Rajasthan	District	0.894	0.488, 1.508	0.659	0.366, 1.129	1.432	0.800, 2.455
	PSU	0.768	0.512, 1.054	0.794	0.638, 0.954	1.163	0.928, 1.434
	Interviewer	0.978	0.668, 1.363	0.120	0.058, 0.195	1.103	0.751, 1.579
Western MP	District	0.244	0.106, 0.472	0.280	0.124, 0.547	0.054	0.007, 0.137
	PSU	0.982	0.766, 1.252	1.157	0.943, 1.402	0.689	0.550, 0.840
	Interviewer	0.933	0.646, 1.32	0.124	0.062, 0.203	0.458	0.291, 0.662
Eastern MP	District	0.683	0.315, 1.361	0.336	0.151, 0.649	0.117	0.036, 0.262
	PSU	0.608	0.424, 0.851	1.087	0.875, 1.332	0.538	0.385, 0.715
	Interviewer	1.371	0.940, 1.962	0.108	0.045, 0.197	0.444	0.295, 0.647
Central UP	District	0.816	0.352, 1.771	0.257	0.108, 0.546	0.161	0.034, 0.409
	PSU	0.724	0.461, 0.992	0.627	0.476, 0.797	0.571	0.412, 0.763
	Interviewer	1.235	0.786, 1.879	0.024	0.001, 0.072	0.442	0.249, 0.715
Western UP	District	0.75	0.369, 1.397	0.07	0.026, 0.145	0.58	0.297, 1.070
	PSU	0.64	0.478, 0.819	0.58	0.476, 0.692	0.61	0.460, 0.782
	Interviewer	0.57	0.377, 0.814	0.0001	0.000, 0.010	0.67	0.454, 0.973
Eastern UP	District	0.806	0.416, 1.474	0.207	0.098, 0.391	0.201	0.093, 0.385
	PSU	0.450	0.335, 0.584	0.653	0.548, 0.769	0.357	0.269, 0.451
	Interviewer	0.772	0.529, 1.098	0.015	0.001, 0.038	0.369	0.241, 0.537

Note: Model is adjusted for selected PSU and Individual level covariates; UP- Uttar Pradesh; MP-Madhya Pradesh

whereas eastern Madhya Pradesh (0.10), Central Uttar Pradesh (0.10), and eastern Uttar Pradesh (0.09) had a lower interviewer-level ICC for 'immunization card not seen'. Interestingly, the findings highlight that the estimated interviewer effect varied largely by outcomes across the selected regions and states (Table 3). For example, the interviewer-level variance for women who gave birth at home was not statistically significant for all the regions of Uttar Pradesh; however, in the same regions of Uttar Pradesh, the interviewer-level factors contributed significantly to the case of women who had not availed ANC (Table 3). The interviewer effect was significantly higher for 'immunization card not seen' and non-use of ANC than for women delivered at home (Table 3). These findings highlight the discordance at the interviewer level, which is a critical impediment to the quality of survey data. This calls for an urgent need to examine interviewer- and agency-level characteristics with regard to the appropriate implementation of survey operations.

4. Discussion and conclusions

The National Family Health Survey in India provides an array of information on maternal and child health indicators, which holds great value for strategic policies and program management. However, an important concern with the estimates from large-scale surveys such as the NFHS pertains to non-sampling errors, which include non-response, measurement, interviewer, adjustment, and data processing errors that may occur at any stage of survey design.

This study explored the interviewer effect in explaining the variation among selected maternal and child health indicators. To explore this effect in teams of differentials in the reporting and capture of the

selected outcome variables, the teams and districts were restructured into mutually exclusive groups (Fig. 1). In each group, the districts were arranged in chronological order of team movement for the purpose of data collection. Therefore, a group consisted of exclusive teams and districts; that is, a particular group of teams covered a number of districts chronologically (Fig. 1). Because only the rural parts of the districts were selected, a homogenous environment within a group of districts should have provided considerably uniform results across teams and interviewers. Additionally, the use of comprehensive survey manuals, training protocols, and multilevel supervision should have helped standardise survey implementation procedures. Under such conditions, the interviewer variance should not vary significantly across the outcome variables.

However, the present analysis clearly demonstrates the presence of interviewer-level variations in the selected outcomes at both the inter- and intra-field agency levels across the selected states. Eastern Madhya Pradesh, Rajasthan, and Bihar showed a higher interviewer level of variation across selected outcomes, underlining the critical role of agencies and skilled interviewers in different survey implementation designs. This highlights the nonuniform adherence to survey protocols, meaning that not all interviewers and agencies performed in a similar manner in the field. The study strongly recommends further examination of the interviewer field and agency effects, especially for the lengthier sections of the household questionnaire and in states with a larger sample size in order to modify the current survey implementation design and adopt new stringent measures to minimise non-sampling errors.

Although NFHS follows an array of measures to standardise survey implementation protocols and minimise non-sampling errors, there may

be numerous latent factors contributing to interviewer-level variance. Training is of utmost importance to follow a standard approach for data collection (Roy and Pandey, 2008). Highly motivated and well-trained interviewers are crucial for collecting data efficiently and maintaining survey quality (DHSM3, 2009). Training sessions include a detailed discussion on all questions, methods, and procedures of data collection, mock interviews, and field practice, followed by training assessment tests and retraining if needed (IIPS, 2015–16). Such detailed training equips interviewers with the appropriate skills and confidence for adequate field operations. However, individual interviewer factors, such as indolent behaviour, missing out on training sessions, minimal involvement in training, poor communication skills, and stubborn nature, may negatively influence survey quality and create variability in the outcomes at the interviewer level. In addition, the sample size of the NFHS-4 increased by approximately five times. This increase in the workload of the field interviewers is associated with longer field work and, therefore, fatigue, which may potentially result in an erroneous collection of data or an intentional negative response to the opening question to avoid follow-up questions due to time constraints.

Individual motivation to perform better in a challenging environment and the seriousness and sincerity towards the assigned task also contribute to the effort made to gather quality information from respondents. For example, at times, the immunisation card is not readily available with a mother and the interviewer must wait until the mother finds the card. Under such circumstances, the interviewer may potentially skip filling in the immunisation information through the card to meet the survey completion targets. Prior field experience with data collection may also contribute to better, or sometimes even poor, reporting on the selected outcome variables within a homogeneous rural setting through two pathways. Arguably, prior field experience may improve an interviewer's understanding of the questionnaire, resulting in better performance. By contrast, with the successive movement of interviewers across districts, their field experience and repetitive exposure to the questionnaire may result in intentional manipulation. This study confirms both types of clustering in NFHS fieldwork.

Other factors, such as inappropriate articulation of questions, inattentiveness, lack of knowledge of the local language, lack or loss of communication between FIs and respondents, non-cooperation from respondents, prior experience of survey implementation, and pressure to complete survey targets in an unfamiliar area may also result in interviewer-level variation in survey estimates (Johnson, 2009; Anglewicz, 2009; Pullum, 2018). More qualitative research is needed to obtain better insights into the factors that contribute to the interviewer effect in large-scale surveys.

This study highlighted that the interviewer-level variance in NFHS-4 estimates was disproportionate across the selected regions and states at the level of field agencies. This implies that not all field agencies followed standardised field procedures. Such findings are of great importance as they underline the need to strengthen field operations and supervision. Apart from planning and training activities, supervision and field operation management are essential for minimising non-sampling errors in large-scale surveys (Roy and Pandey, 2008). In addition to a two-level supervision, it would be helpful to have spot checks, field check tables, frequent field visits by nodal agencies, random vigilance, recording parts of an interview, and revised training to reduce inter-agency interviewer-level variation.

The findings shed light on the extent to which the magnitude of the interviewer effect varies across the selected outcomes. Women who had not availed ANC and whose children's immunisation card was not seen by the interviewer had the highest interviewer-level variance. However, interestingly, among women who had given birth at home, interviewer-level variance was not significant in Uttar Pradesh. More distinguished interviewer effects were elucidated in the sections of the questionnaire which had a long series of follow-up questions, such as 'ANC not availed'

and 'immunisation card not seen'. These findings are similar to those reported by (Leone et al., 2021).

Existing studies on paradata analysis show significant effects of interviewer characteristics on the quality of data recorded during a field survey (Pullum, 2018). Interviewer background characteristics, use of a translator during an interview in case of language mismatch, number of visits to a household, number of days the interviewer spent in the cluster, and duration of the interview during fieldwork (beginning, middle, or end of the fieldwork period) impact the data quality. These factors could explain the existing team-level variations in field operations in NFHS. A comprehensive collection and examination of para data analysis could provide more insight into the quality of data, especially on how lengthier sections of the questionnaire are dealt with by FIs from different FAs. However, the paucity of para data and its alignment with the unit-level data of NFHS restricts such empirical investigations. Nevertheless, there is considerable scope for analysing para data in India.

In general, skipping critical questions requires examining the interviewer effects on data quality. However, based on interviewers' performance in the field, the present study offers an intriguing insight into interviewer-level variations in data quality. Multivariate analysis revealed a notable effect of interviewer bias on data quality. The non-uniform clustering effects in the pattern of reporting are specific to certain sections of the questionnaire and to the interviewer conducting the survey operations. The study also highlights that interviewer-level variance is higher than PSU-level variance except for the outcome 'women who have given birth at home'. This study recommends a refined mechanism to improve data quality. Providing training with a particular focus on the issues and challenges faced by FIs during field operations may profoundly affect the quality of the data collected. Repeated in-depth surveys of interviewers could add a new dimension to the overall data quality and field operations. This would enable nodal agencies to study interviewer characteristics that are susceptible to erroneous reporting. Random vigilance of field agencies and interviewers and regular field monitoring during the survey operations could work as a vital strategy to check the real-time data quality and resolve critical impediments to efficient survey implementation. By and large, it would help reduce the chances of interviewer-level variations in survey implementation. Most importantly, interviewers and teams should be sensitised to their crucial role in the country's policy formulation and the success of national programs.

Ethics statement

The authors used publicly available deidentified data for this analysis.

Authorship contributions

Category 1

Conception and design of study: LKD, RS and KB

Acquisition of data: RS, SJ, RM, KB and LKD

Analysis and/or interpretation of data: LKD, RS, SJ, RM and KB

Category 2

Drafting the manuscript: RS, LKD and RM

Revising the manuscript critically for important intellectual content: LKD, RS, SJ and KB

Category 3

Approval of the version of the manuscript to be published (the names of all authors must be listed):

Radhika Sharma (RS), Laxmi Kant Dwivedi (LKD), Somnath Jana (SJ), Kajori Banerjee (KB), Rakesh Mishra (RM), Bidhubhusan Mahapatra (BM), Damodar Sahu (DS) and S. K. Singh (SKS)

Declaration of competing interest

The authors declare that they have no conflict of interest.

Data availability

<https://dhsprogram.com/data/available-datasets.cfm>

Acknowledgments

This paper was written as part of DataQi project of the Population Council funded by Bill & Melinda Gates Foundation (grant # OPP1194597). The views expressed herein are those of the authors and do not necessarily reflect the official policy or position of the Bill & Melinda Gates Foundation and/or Population Council. The authors would like to acknowledge the support from National Data Quality Forum (NDQF) in providing feedback to initial draft of the paper. Also, authors wish to acknowledge the valuable contribution of anonymous reviewers for their insightful comments and suggestions which have helped improve this research work tremendously.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.ssmph.2022.101252>.

References

- Adida, C. L., Ferree, K. E., Posner, D. N., & Robinson, A. L. (2016). Who's asking? Interviewer coethnicity effects in African survey data. *Comparative Political Studies*, 49(12), 1630–1660.
- Allen, C. K., Fleuret, J., & Ahmed, J. (2020). *Data quality in demographic and health surveys that used long and short questionnaires*. Rockville, Maryland, USA: Retrieved from.
- Amos, M. (2018). Interviewer effects on patterns of nonresponse: Evaluating the impact on the reasons for contraceptive nonuse in the Indonesia and the Philippines. *DHS. Demographic Research*, 39, 415–430.
- Anglewicz, P. a. j., Obare, F., Watkins, S. C., & Kohler, H.-P. (2009). The Malawi Diffusion and Ideational Change Project 2004-2006: Data collection, data quality and analysis of attrition. *Demographic Research*, 20(21), 503–540.
- Arnold, F. (1990). *Assessment of the quality of birth history data in the demographic and health surveys*. Maryland, USA: Retrieved from.
- Bhat, P. (1995). Age misreporting and its impact on Adult mortality estimates in South Asia. *Demography India*, 24(1), 59–80.
- Bogen, K. (1996). *The effect of questionnaire length on response rates: A review of the literature*. Retrieved from.
- Borkotoky, K., & Unisa, S. (2014). Indicators to examine quality of large scale survey data: An example through district level household and facility survey. *PLoS One*, 9(3), Article e90113. <https://doi.org/10.1371/journal.pone.0090113>
- Bound, J., Brown, C., & Mathiowetz, N. (2001). *Measurement error in survey data* (Vol. 5). Elsevier.
- Curtis, S. (1995). *Assessment of the quality of data used for direct estimation of infant and child mortality in DHS-II surveys*. Maryland, USA: Retrieved from Calverton.
- Dandona, R., Pandey, A., & Dandona, L. (2016). A review of national health surveys in India. *Bulletin of the World Health Organization*, 94(4), 286–296A. <https://doi.org/10.2471/BLT.15.158493>
- Division, U. N. D. o. E. a. S. A. P. (1955). *Methods of appraisal of quality of basic data for population estimates*. New York: Retrieved from.
- Dunn, E. C., Milliren, C. E., Evans, C. R., Subramanian, S. V., & Richmond, T. K. (2015). Disentangling the relative influence of schools and neighborhoods on adolescents' risk for depressive symptoms. *American Journal of Public Health*, 105(4), 732–740.
- Elkasabi, M., Ren, R., & Pullum, T. W. (2020). *Multilevel modeling using DHS surveys: A framework to approximate level-weights*. Rockville, Maryland, USA: ICF. DHS Methodological Reports No. 27 <https://www.dhsprogram.com/pubs/pdf/MR27/MR27.pdf>.
- Gage, A. (1995). *An assessment of the quality of data on age at first union, first birth, and first sexual intercourse for phase II of the demographic and health surveys program*. Maryland, USA: Retrieved from Calverton.
- Gideon, L. (2012). *Handbook of survey methodology for the social sciences*. New York: Springer.
- (IIPS) , I. I. f. P. S. (2015-16). *Request for proposal (RFP) from field agencies to implement national family health survey (NFHS)-4*. Retrieved from Mumbai: http://rchiips.org/NFHS/NFHS4/pdf/NFHS4%20FA%20BID_1.pdf.
- International Institute for Population Sciences (IIPS) and ICF. (2017). *National family health survey (NFHS-4), 2015-16: India*. Mumbai: IIPS. <http://rchiips.org/nfhs/NFHS-4Reports/India.pdf>.
- James, K. S., & Rajan, S. I. (2004). Respondents and quality of survey data. *Economic and Political Weekly*, 39(7), 659–663.
- Johnson, K., & Kreuter, F. (2009). *Fieldwork-related factors and data quality in the demographic and health surveys program*. Retrieved from.
- Leone, T., Sochas, L., & Ernestina Coast. (2021). Depends who's asking: Interviewer effects in demographic and health surveys Abortion data. *Demography*, 58(1), 31–50.
- Liu, M. (2016). The effect of the perception of an interviewer's race on survey responses in a sample of Asian Americans. *Asian American Journal of Psychology*, 7(3), 167.
- Manski, C. F., & Molinari, F. (2008). Skip sequencing: A decision problem in questionnaire design. *Annals of Applied Statistics*, 2(1), 264–285. <https://doi.org/10.1214/07-aoas134>
- Mari Bhat, P. N. (1995). On the quality of birth history data collected in National Family Health Survey. *Demography India*, 24, 245–258.
- Mathews, P., Sear, R., Coast, E., & Iacovou, M. (2012). *Do preceding questions influence the reporting of childbearing intentions in social surveys?* San Francisco: Paper presented at the Population Association of America Annual Meeting.
- Namaste, S., Benedict, R. K., & Henry, M. (2018). *Enhancing nutrition data quality in the DHS program*. Rockville, Maryland, USA: Retrieved from.
- Nations, United. (2015). *Transforming our world: The 2030 agenda for sustainable development, preamble*.
- Phillips, B. S., Singhal, S., Mishra, S., Kajal, F., Cotter, S. Y., & Sudhinaraset, M. (2019). Evaluating concordance between government administrative data and externally collected data among high-volume government health facilities in Uttar Pradesh, India. *Global Health Action*, 12(1).
- Pullum, T. W., Christina, J., Nizam, K., & Sarah, S. (2018). *The effect of interviewer characteristics on data quality in DHS surveys*. Rockville, Maryland, USA: Retrieved from.
- Rajan, S. I., & James, K. S. (2004). Second national family health survey: Emerging issues. *Economic and Political Weekly*, 39(7), 647–651.
- Rajan, S. I., & James, K. S. (2008). Third national family health survey in India: Issues, problems and prospects. *Economic and Political Weekly*, 33–38.
- Retherford, R. D., & Mishra, V. K. (2001). *An evaluation of recent estimates of fertility trends in India*. Retrieved from.
- Roy, T. K., & Pandey, A. (2008). National family health survey in India: Survey design and related issues. *Demography India*, 37(Supplement), 13–41.
- Sandefur, J., & Glassman, A. (2014). *The political economy of bad data: Evidence from African survey & administrative studies-working paper*.
- Schoumaker, B. (2014). *Quality and consistency of DHS fertility estimates 1990 to 2012*. Rockville, Maryland, USA: Retrieved from.
- Sharma, A., Rana, S. K., Prinja, S., & Kumar, R. (2016). Quality of health management information system for maternal & child health care in Haryana state, India. *PLoS One*, 11(2), Article e0148449. <https://doi.org/10.1371/journal.pone.0148449>
- Shryock, H. S., & Siegel, J. S. (1975). *The methods and materials of demography* (Vol. 102). USA: Department of Commerce, Bureau of the Census.
- Singh, A., Kumar, K., & Arnold, F. (2022). How interviewers affect responses to sensitive questions on the justification for wife beating, the refusal to have conjugal Sex, and domestic violence in India. *Studies in Family Planning*, 53(2), 259–279. <https://doi.org/10.1111/sifp.12193>.
- Srinivasan, K. (1998). *Basic demographic techniques and applications*. New Delhi: Sage Publications.
- Srinivasan, K., & Mishra, R. (2020). *Quality of data in NFHS-4 compared to earlier rounds* (Vol. 6, pp. 40–45). Economic & Political Weekly. February 2020.
- Unisa, S., Dwivedi, L. K., Reshmi, R. S., & Kumar, K. (2015). Missing age of population in Indian censuses during the last 50 Years: Prevalence, trends and differentials. *Social Science Spectrum*, 1(1), 8–19.
- Verma, V., & Le, T. (1996). An analysis of sampling errors for the Demographic Health Surveys. *International Statistical Review*, 64(3), 265–294. Retrieved from <https://www.ncbi.nlm.nih.gov/pubmed/12179593>.
- Wagner, A. L. (2019). The use and significance of vaccination cards. *Human Vaccines & Immunotherapeutics*, 15(12), 2844–2846.
- Wang, K., Kott, P., Moore, A., & Of, R. T. I. (2013). *Assessing the relationship between interviewer effects and NSDUH data quality*. Retrieved from.
- Weisberg, H. F. (2008). *The methodological strengths and weaknesses of survey research the SAGE handbook of public opinion research*.
- Wolf, C., Schneider, S. L., Behr, D., & Joye, D. (2016). Harmonizing survey questions between cultures and over time. In S. A. G. E. The (Ed.), *Handbook of survey methodology* (pp. 502–524). Los Angeles: Sage.
- Yang, M. L., & Yu, R. R. (2008). The interviewer effect when there is an education gap with the respondent: Evidence from a survey on biotechnology in Taiwan. *Social Science Research*, 37, 1321–1331. <https://censusindia.gov.in/census.website/>.