# Consistency and Quality Check of Survey Data in India

Pushpanjali Swain, PhD,
National Institute of Health & Family Welfare,
Delhi, India

M. Hemanta Meitei , PhD,
Manipur University,
Manipur, India

Correspondence may be directed to:
Pushpanjali Swain, Department of Statistics and Demography,
National Institute of Health and Family Welfare,
Munirka, New Delhi, PIN 110067, India
Tel: 91-11-26165959 ext. 379,
E-mail: drpswain@yahoo.co.in

**Abstract**

**Objective:** Reliability of survey responses on topics such as utilization of health facilities by mother and child has long been a subject of concern. This paper explores consistency of responses from the same individuals over time on utilization of health services involving child delivery and child care.

**Methods:** A sub-sample survey was carried out by an independent monitoring agency in 13 states as a part of a larger Coverage Evaluation Survey of all states of India in 2009 by UNICEF, to recheck the responses to improve data quality. Our randomly chosen sub-sample consisted of 510 questionnaires regarding mothers and 497 regarding children. Differences of responses were noted and conveyed to field agencies to rectify recurring errors. Statistical analysis was conducted to find consistency of responses.

**Results:** Matching between the original and rechecked responses varied. Generally, however, the overall match was greater than 90%.

**Conclusion:** Findings suggest that response inconsistencies and the manner in which they are resolved are shown to have important implications for the overall estimate of indicators of utilization of health facilities. The monitoring exercise has, therefore, addressed the quality and consistency issue, which needs further consideration in large-scale surveys. Otherwise, it poses a validity threat to data quality and results on which national policy is framed.

## Introduction

The reliability of survey responses to questionnaires on topics such as utilization of health facilities by mother and child has long been a subject of concern. Evidence from developed countries suggests that the quality of responses appears to be as high as that found in studies of other topics. Studies on reliability of survey responses are far fewer in developing than in developed countries. Existing studies clearly indicate that poor reliability of responses is associated with poor design (Dare and Cleland 1994).

*Studies on reliability of survey responses are far fewer in developing than in developed countries*

In the first approach, a respondent is interviewed on two separate occasions, with the length of time between interviews ranging from a few hours to several days. The observations are evaluated for consistency of responses (e.g., contraceptive use, place of delivery, etc.) Studies of this nature have been carried out in the United States as well as in developing countries (Knodel and Piampiti 1977; Ryder 1979; Westoff et al. 1961). The clinic record approach is based on a comparison of an individual's interview response with some known program or clinic record (Stoeckel and Choudhury 1969).

While the pattern of inconsistency varies considerably from study to study, the most common discrepancy results from husbands reporting contraception use while their wives do not (Keonig et al. 1984; Coombs and Chang 1981). All these approaches are subject to inherent limitations as indicators of response validity. In spite of a large number of studies on response inconsistencies among a significant sub-sample of respondents, research on the underlying causes of inconsistencies appears lacking.

This paper deals with the reliability of information on a few aspects of reproductive health and immunization status of children aged 12 to 23 months. In the present study, we consider the reliability of data concerning pregnant mothers' antenatal care, deliveries, immunization and feeding practices for the child, as measured by the consistency between responses from the same individual by two different agencies, that is, field investigator and independent monitor. The duration between the two observations of response varied from a few hours to a few days during the survey period. We examine the level of response inconsistency and possible underlying factors, and we suggest an approach to minimize inconsistent responses. We subsequently consider the broader implications of reporting inconsistencies for both aggregate and state-level analyses of utilization of health facilities by women who delivered during the 12 months preceding the survey and mothers/caregivers of children aged 12 to 23 months.

The basic objective of UNICEF's 2009 Coverage Evaluation Survey (CES) (UNICEF 2009) was to assess routine immunization levels among children as well as maternal health services across all 28 states and seven union territories in India. Specifically, the CES assessed routine immunization and vitamin A coverage; antenatal care (ANC), delivery care and postnatal care (PNC) coverage; availability and use of iodized salt; initiation of

breastfeeding and colostrum feeding; accessibility and availability of immunization and maternal care services, and so forth.

The nationwide CES survey was carried out by a private agency appointed by UNICEF. Data collection was monitored and checked independently by a premier national-level institute, the National Institute of Health and Family Welfare (NIHFW), New Delhi, in collaboration with population research centres located in various regions of the country. UNICEF designated NIHFW as the independent monitoring agency (IMA) to monitor survey activities carried out by the private agency. Survey monitoring by independent agencies was felt to minimize errors due to non-adherence of protocols during investigator training and during data collection, thus making it easier to correct errors and compensate for them at the time of data tabulation and analysis. The idea of survey monitoring by an independent agency was to minimize non-sampling errors that could creep in for reasons such as training of investigators and supervision by field monitors, which would not be amenable for correction at the later stage of analysis and tabulation. More specifically, monitoring was designed to (i) assess whether work (training of field investigators and data collection in the field) was performed according to the methodology; (ii) assess whether the private firm was adhering to time lines; (iii) provide technical support by identifying systemic errors made by field investigators during fieldwork – any misinterpretation of questions that may lead to non-sampling errors and others that would have adverse effects on the survey and its results – and reporting them to the agency's supervisors/UNICEF; and (iv) ensure immediate reporting to the agency's supervisors/UNICEF of grave inadequacies for immediate action (such as repeating the part of the survey).

The IMA's observations provided an opportunity for onsite correction during data collection in the field and feedback to the private agency where the inadequacy was found, as well as improvement in overall field investigator performance regardless of where mistakes were detected. For the purpose of monitoring fieldwork, the NIHFW, in consultation with Ministry of Health and Family Welfare (MOHFW) and UNICEF, developed a monitoring protocol and also engaged staff from the selected population research centres (PRCs).

*The IMA's observations provided an opportunity for onsite correction during data collection in the field and feedback to the private agency where the inadequacy was found.*

To analyze and interpret the underlying causes of inconsistencies in responses, it was necessary to understand the process of data collection by the private agency's field investigators and the strategy adopted by the IMA to monitor data collection. The household interview was undertaken by trained field investigators engaged by the private agency, using interview questionnaires prepared by UNICEF. The IMA collected the completed questionnaires and re-interviewed randomly selected respondents, using the monitoring protocol (a checklist) prepared separately. This enabled the IMA to check the completeness and consistency of the original interview.

## Methodology

A sub-sample survey was carried out by the IMA in 13 randomly selected states, broadly covering six regions – north, south, east, west, central and northeast – of the country. This survey was part of a nationwide CES conducted in all 35 states and union territories of India in 2009 by UNICEF, to monitor and improve the quality of data. Field

investigators from a private agency collected data at the field level from face-to-face interviews in selected primary sampling units (PSUs) in each state and union territory. PSUs are generally villages in rural areas and wards in urban areas. The selected 13 major states were: Uttar Pradesh, Rajasthan, Bihar, Madhya Pradesh, Chhattisgarh, Orissa, West Bengal, Maharashtra, Gujarat, Karnataka, Andhra Pradesh, Manipur and Arunachal Pradesh. Responses to two sets of questionnaires administered by private agency investigators were rechecked for (i) mothers who had delivered in the last 12 months (ii) the mother or guardian of children 12 to 23 months old.

In each of the selected states, 10% of PSUs were monitored, and in each PSU, responses from at least four households where mother had delivered in the last 12 months and four households with children 12 to 23 months old were rechecked for consistency of responses. Finally, 128 PSUs were selected for rechecking from all 13 selected states. The sub-sample survey consisted of a randomly chosen sample of 510 administered interview questionnaires of mothers and 497 such questionnaires of children by field investigators. The format developed for monitoring the survey is given in Appendix 1 (monitoring format of mothers is appended). Two monitoring strategies were adopted: (i) concurrent evaluation and (ii) back-check (the term used in this survey). Responses given by the individuals of sampled questionnaires were rechecked by monitors from the IMA. These individuals used checklists to check almost half of completed questionnaires concurrently, after the field investigators had completed the interviews. Concurrent monitoring was helpful for rectifying errors, as both monitors and field investigators were present. Similarly, monitors back-checked responses of the selected households a few days after the survey. Differences were noted and conveyed to the field agencies for rectification. Field investigators received reorientation, and instructions were passed on to senior levels of the private agency as well as to UNICEF. Whenever inconsistencies arose between field investigators and the monitoring agency, the latter prevailed because of their greater fieldwork experience. This approach resulted in improved data quality in terms of coverage, completeness and correctness. It should be noted that monitors assessed all field investigators at the beginning of the survey so that any inconsistencies in responses would be detected and minimized through the monitor's intervention. The monitoring plan was prepared by the IMA based on the field investigators' movement plan. Each monitor was assigned a group of investigators for assessment in the field of a particular state. Field teams comprised three female investigators and a supervisor. Field investigators were generally responsible for gathering information from respondents, who were generally females, to maintain sensitivity to the gender issue. Female respondents are generally comfortable with female investigators. The field supervisors' role was to oversee investigators' activities. Each field team was assessed more than once by the monitor. Statistical analysis of matching paired sub-samples was carried out to assess the level of consistency and reliability of responses.

Dependent variables in the analysis are a dichotomous measure assuming the values 1 (the responses agree) and 2 (the response disagree) between the monitor's and field investigator's response. Though we resolved some inconsistency in the field and corrected the actual data, we did not rectify the investigator error in our monitoring checklist. Therefore, the percentage of mismatch was higher in our dataset. Data analysis was done using SPSS. Tables were prepared for each matched response across the sample states by all selected sample questions. We examined the results of each question's response matching. Monitors also asked respondents about the average time field investigators took to complete the questionnaire. Differences in monitors' gender were also investigated for

possible gaps in the agreement of responses. Further analysis was done using variables such as type of residence and duration of time to complete the interview.

## Results

### Matching of Responses Based on Mother's Questions

Table 1 shows the matched responses of mothers regarding care during the last pregnancy, place of delivery, postnatal care, breastfeeding practices, children's illness, using facilities from government programs and intake of iodized salt. State-wise matching of responses was undertaken to further understand the level of matching (see Table 1 at www.longwoods.com/content/23793).

One set of response matching was related to the question on testing salt used by the household. Responses obtained by field investigators and monitors were compared. The level of similarity ranged from 73% in Orissa to 100% in Jharkhand, Manipur and Uttar Pradesh. Overall, 89% of responses in all sampled states were consistently matched. Barring Arunachal Pradesh (77%), Madhya Pradesh (76%) and Karnataka (78%), the rate of matching responses was about 95%. Therefore, the question on consumption of iodized salt was not perfectly matched, although it was matched to a high degree. The reason for inconsistency of responses was primarily dependent on whether field investigators treated a village or ward as a homogeneous group or not. It was observed that in a few states, investigators tested salt in just a few households. If test outcomes were similar, they skipped the test for other sampled households. However, data were reliable and valid despite this inconsistency.

The matched response for the question "did the mother receive any supplementary nutrition from the Anganwadi centre during last pregnancy?" was 96%. One hundred percent matching responses were found in Andhra Pradesh, Arunachal Pradesh, Manipur and Uttar Pradesh, whereas the lowest rates were found in Madhya Pradesh (87%) and Gujarat (89%). Remaining states performed well above 90%.

In all sampled states, the question on "place of last delivery" yielded consistent response matches in 96% of cases. Analysis of responses to the "type of delivery" question found very little difference between the original and rechecked survey data. Except for Madhya Pradesh, Karnataka and Arunachal Pradesh, the matched response rate was above 95%. The typology for place of last delivery was a long list. The response mismatch might be associated with characteristics of investigators, monitors and respondents. Moreover, validity of responses concerning government facility, private facility or home did not deviate. However, validity issues might arise with respect to exact place of delivery within government facilities, or within private facilities; where there were various categories of facilities within government or privates.

"Was the last delivery normal or cesarean or assisted?" was a straightforward question. As expected, the overall reliability of responses was about 99%. For the question "who conducted the delivery," overall, the matched response rate was 93%, ranging from 85% in Arunachal Pradesh to 100% in Jharkhand and Orissa. Matched responses to the question on postnatal checkup was comparatively low (91%) for all sampled states; lowest was Rajasthan (67%), followed by Madhya Pradesh (81%) and Karnataka (84%). Almost all responses on continuing breastfeeding were matched (99.4%), except for Rajasthan, Uttar Pradesh and Madhya Pradesh, each at 98%.

The monitoring team also matched the information on "any government assistance during pregnant women's delivery." This question was thought important for identifying funds utilized under the national rural health mission i.e., how effectively the Janani Surakya Yojana program motivated expecting mothers to deliver in the institutions (Table 1). Responses were found congruent in 97% of cases. Matched responses were lowest in

Arunachal Pradesh (85%), while exceeding 93% in other states. In response to the question "whether the Accredited Social Health Activist (ASHA) accompanied the delivering women to reach to the health facility," overall, 94% of the responses were correctly matched. However, in Jharkhand (81%) and Karnataka (84%) matching responses were lower.

Responding to the question on the amount of money received by delivered mothers, 490 of 510 responses matched correctly, although in Gujarat only 84% of responses matched. On the question of availability of a separate bed for delivery in the hospital/health facility, 97% of responses matched, while in Arunachal Pradesh and Manipur only 88% matched.

On the "prevalence of diarrhea in last two weeks" the lowest match was observed in Bihar, Gujarat and Rajasthan. In aggregate, nearly 97% of responses across the states matched. Similarly, for "prevalence of cough in last two weeks" consistent matching responses were only 93%, while 100% matching was attained in Andhra Pradesh, Arunachal Pradesh and Jharkhand.

For responses on the nearest place for getting the health services, the overall matching rate was 89%. Correctly matched responses varied from 76% in Madhya Pradesh to 98% in Maharashtra.

Among the 14 questions monitors posed to mothers for the re-check, responses matched in more than 95% of cases across eight questions. Whether the difference in responses was due to "type of residence" or "length of interview by field investigators to ask all questions," the analysis was conducted using chi-square tests.

Table 2 depicts the percentage of matching responses by type of residence, sex of monitor and duration of interview by investigator. Agreement of responses was lower in rural than in urban areas for all questions put to the mothers. Significant differences were found for "place of last delivery," "accompanied by accredited social health activist (ASHA)" and "had separate bed for institutional deliveries." The higher proportion of mismatches in rural areas compared to urban areas was most probably due to prevailing local languages or dialects. Field investigators may have not have been well versed in them (see Table 2 at www.longwoods.com/content/23793).

*The higher proportion of mismatches in rural compared to urban areas was most probably due to prevailing local languages or dialects.*

Regarding the role of gender in matching information, a higher proportion of matching was observed among female monitors in most questions on delivery-related cases, child illness and nearest facilities; this finding did not extend to "salt test" or "nutrition supplementation from AWW" questions. This higher proportion did not reach statistical significance, except for the question "What is the nearest place where you get antenatal services by skilled provider?"

During the monitoring visit, the respondent was asked, "How much time did the investigators take to ask all questions?" A higher proportion of matching responses was found where interviews lasted 30 to 45 minutes over those lasting less than 30 minutes. The highest proportion of consistent matching was found in interviews to understand the child's illness or mother's breastfeeding practices that lasted more than 45 minutes.

## Matching of Responses Based on Children's Questions

In response to questions on children's health answered by the mother or guardian, a greater number of questions matched the responses collected by field investigators and those of independent monitoring agencies compared to the matched responses on the questions of mothers. Therefore, information sought for children was more accurate and

better matched than that sought for mothers, because parents and guardians must be more vigilant about their children's health than own. Nevertheless, a few questions about children were not perfectly matched (see Table 3 at www.longwoods.com/content/23793).

The response on the salt testing was matched 91% across the states, which is close to the matching response secured from mothers (89%, Table 1). On the question to confirm child's age, the match rate of investigator to independent monitor responses was 98%.

It may be noted that only 91% of responses to the question on whether the immunization card was available matched perfectly. Among the valid cases of BCG (bacillus Calmette-Guérin) vaccination, 95% of responses matched correctly. Lowest matching was observed in Karnataka (80%), followed by Gujarat (85%). Reliability of matching on the source of vaccination for children was 94%; lowest matching was observed in Karnataka (75%) and Uttar Pradesh (88%).

The question on conducting routine immunization in local areas was matched up to 90%. In Arunachal Pradesh, the rate was only 61%, and in five states – Karnataka, Madhya Pradesh, Manipur, Rajasthan, and Uttar Pradesh – the rate was below 90%. Of nine questions about child immunization, five responses were matched above 95%. The concentration of mismatch between investigators and monitoring agents was a random chance, but it clearly showed the difficulties field investigators experienced in what to ask or the way they communicated with respondents.

## Discussion and Conclusion

There are many possible sources of response inconsistency in health surveys. For example, inconsistencies may be partly attributable to the survey itself, since lack of clarity of questions may contribute to erroneous responses. Response inconsistencies may also be a function of the data collection process to some extent, through factors such as errors in recording or transferring information. In this investigation, we were primarily concerned with the main source of response inconsistency – misreporting, both deliberate and unintentional, by respondents themselves.

The independent monitoring agency responsible for the UNICEF 2009 CES was concerned about the response inconsistency problem. UNICEF was concerned about data quality, and that is why the concept of using a monitoring agency arose. Indeed, all the IMA officials received the same training, along with the key trainers of the private agency. Subsequently, key trainers from the private agency trained the field investigators, and the IMA monitor observed the training. The independent monitoring officials were present during investigator training to ensure resolution of all doubts among investigators. Monitors observed mock interviews as well as interviews in the field, to ensure investigator proficiency. This process implied that data collected by field investigators and monitoring agency must be the same. Yet despite these processes, there were many discrepancies in responses collected by field investigators and independent monitoring officials. These discrepancies arose for the following reasons, observed in the field situation. Firstly, field investigators were educated up to a university graduation level, whereas monitors had a much higher level of education (post graduation and above), knowledge regarding survey work and number of years of experience. Secondly, the fieldwork itself was tedious, and investigators may have been fatigued. (Although this issue is not recorded and reflected much in the literature, it was observed during monitoring.) Thirdly, the level of field investigator motivation may have reflected the amount of remuneration they received (which was beyond the domain of the monitoring exercise).

Findings suggest that the interview context is an important factor in reporting agreement.

In other words, the same individual was interviewed twice highlighted agreement or disagreement in responses. Moreover, consistency checks could have been done without rechecking the responses again in the field but by glancing through the questionnaires filled out by the investigators, as is done in many surveys. In this process, non-sampling errors could not have been reduced much.

The main thrust of the paper is to highlight the existence of the discrepancies. It is very hard to find the solutions to this generic problem. One of the ways to resolve this lies with the field
agencies. They should limit the number of interviews undertaken by an investigator and should ensure investigators receive adequate remuneration. Therefore, it is strongly suggested that field agencies use a reward-penalty mechanism to motivate field investigators. As discussed earlier, response inconsistencies were resolved in the field during (concurrent/back-check) rechecking of responses. Since field investigator teams were monitored at PSUs and their movements tracked until the end of the survey, errors were detected early and resolved by identifying the type of errors, communicating them to the field agencies, conducting reorientation training for investigators if required, and tracking the team to ensure that similar errors would not be repeated. In this manner, mismatched responses were resolved to some extent.

How corrections and inconsistencies were treated and resolved by the field agency after we left the monitoring visits with our monitoring mechanism might be the limitation to the whole exercise. The monitoring process facilitated a sense of the reliability and to some extent quality of the data. Generally, monitoring had important implications for the overall estimate of important indicators for the utilization of health facilities in the country. The exercise addressed the fundamental issue of reliability and data quality, which needs further consideration in large-scale surveys. Otherwise, it will pose validity threats to data quality. If reliability of data is to be ensured, adoption of a separate monitoring process is desirable in such kind of surveys (given the cost consideration) and must be built in. The current survey could serve as the introduction of a mechanism to improve quality of data, so that subsequent surveys could adopt similar mechanisms. The implication of discrepancies at the aggregate level was well known at UNICEF, and therefore concurrent evaluation and back-checks by monitoring agencies were built in. Findings from this study strongly suggest the need for a strong and layered supervision of the data collection process to ensure data quality and consistency of survey data.

## References

Coombs, L.C. and M.C. Chang. 1981. "Do Husbands and Wives Agree? Fertility Attitude and Later Behavior." *Population and Environment* 4(2): 109–127.

Dare, O.O. and J.C. Cleland. 1994. "Reliability and Validity of Survey Data on Sexual Behavior." *Health Transition Review* 4(Suppl): 93–110.

Knodel, J. and S. Piampiti. 1977. "Response Reliability in a Longitudinal Survey in Thailand. Studies in Family Planning" 8(3): 55–66.

Koenig, M.A., G.B. Simmons and B.D. Misra. 1984. "Husband-Wife Inconstancies in Contraceptive Use Responses." *Population Studies* 38(2): 281–98.

Ryder, N.B. 1979. "Consistency of Reporting Fertility Planning Status." *Studies in Family Planning* 10(4): 115–28.

Stoeckel, J. and M.A. Choudhury. 1969. "Pakistan: Response Validity in a KAP Survey." S*tudies in Family Planning* 1(47): 5–9.

Westhoff, C.F., R.G. Potter Jr. and P.C. Sagi. 1961. "Some Estimates of the Reliability of Survey Data on Family Planning." *Population Studies* 15(1): 52–69.

UNICEF. 2009. Coverage Evaluation Survey 2009. *National Facts Sheet*. UNICEF, India.