

Women's recall of iron-containing supplement adherence during pregnancy: Findings from a validation study in Nepal



BACKGROUND: NEED FOR VALID INDICATORS & METHODS TO ESTIMATE MMS/IFA ADHERENCE

Nutrition stakeholders need valid data on coverage of iron-containing supplements during pregnancy including 1) which supplements women receive (multiple micronutrient supplement, MMS vs. iron folic acid, IFA), 2) when in pregnancy supplementation is initiated, and 3) how many supplements women consume (adherence).

Household survey questions that ask postpartum women about the total number of IFA consumed during their last pregnancy are not well understood by respondents and do not elicit accurate responses.^{1,2} In recent years, there have been coordinated efforts to design better survey-based measures of MMS/IFA coverage and adherence.

This brief summarizes the results and conclusions of a recent study in Nepal that validated alternative questions and response options for measuring coverage and adherence to iron-containing supplements using household surveys. The findings are most relevant to survey implementers, measurement researchers, and policy and program stakeholders involved in designing monitoring and accountability systems.

KEY MESSAGES

- Pregnant women in Nepal accurately reported the number of MMS or IFA tablets taken in the previous ~1-month
- Asking about the number of tablets consumed had significantly higher validity than asking about the number missed across a given period
- Recently-delivered women could not accurately report about the total number of tablets consumed across pregnancy
- Phone interviews had lower validity than in person interviews, but may be sufficient for monitoring purposes
- In Nepal, large-scale surveys like the NDHS should stop asking recently-delivered women to recall the number of tablets across pregnancy and start asking currently pregnant women about tablets consumed in the previous month; this is a major shift in how adherence is measured that will require advocacy and technical support



FACE-TO-FACE INTERVIEW

Main study in 1 province
n = 1100–1200 women per visit



MOBILE PHONE INTERVIEW

Demo study across 6 provinces
n = 100–225 women per visit



QUESTIONS POSED AT STUDY VISITS

Currently Pregnant # pills consumed since enrollment
~30 day recall # pills missed since enrollment

Recently Delivered # pills consumed across pregnancy
~6 month recall # pills missed across pregnancy
pills usually consumed in a week during pregnancy

GOLD STANDARD

In-person pill count



Figure 1: Study Design

STUDY APPROACH

Parent studies: We nested our measurement validation study within the NAMASTE trial, an implementation research study conducted in Nepal in 2024-2025.³ In the main study, pregnant women in Lumbini province were randomized near the start of their second trimester to IFA in blister packs (standard of care), MMS in blister packs, or MMS in bottles. Women were followed to two months postpartum and data were collected at two home visits - 30 days after enrollment and within 30 days after birth. A demonstration study (“demo”) in six additional provinces only provided MMS packaged in bottles to pregnant women. Demo study visits followed the same schedule as the main study.

Validation study design (Figure 1): At the 30-day “currently pregnant” visit and the post-birth “recently-delivered” visit, we compared each woman’s recall of adherence in response to questions to her actual adherence estimated by pill count as a “gold standard”. Building from question formulations used in previous studies,^{4,5} women were randomized to either being asked about number of pills consumed or the number of pills missed. Recently-delivered women were also asked about their “usual” weekly consumption of pills across their pregnancy. In the main study, face-to-face interviews and pill counts were conducted during home visits. In the demo study we also used mobile phone interviews and in-person visit for pill count.

Analysis: We estimated sensitivity (proportion accurately classified as adherent), specificity (proportion accurately classified as non-adherent),

area under the receiver operating curve (AUC) (reflects overall accuracy of questions) and the inflation factor (population-level bias) for recalled adherence relative to the pill count. A threshold or cut-off point is needed to estimate sensitivity, specificity, and AUC. Based on a recent meta-analysis,⁶ we used two adherence thresholds: $\geq 75\%$ adherence or $\geq 90\%$ adherence.

FINDINGS

Which formulation of adherence questions could women answer most accurately?

All question formulations for adherence (number consumed, number missed, number usually taken in a week) had high levels of sensitivity ($>90\%$) using face-to-face interviews (Tables 1 and 2, Figure 2). However, specificity was consistently better when women were asked about the number of supplements consumed compared to when asked the number of supplements missed. The question about number taken in a usual week yielded extremely low specificity. Asking about the number consumed yielded the most accurate response in this population.

How did validity vary by recall period and pregnancy status?

Sensitivity was high and comparable for currently pregnant women asked about the last ~30 days and recently-delivered women asked about the last ~6 months; adherent women were correctly identified among both populations. Specificity was also high among currently-pregnant women asked about last ~30 days, suggesting overall accurate recall

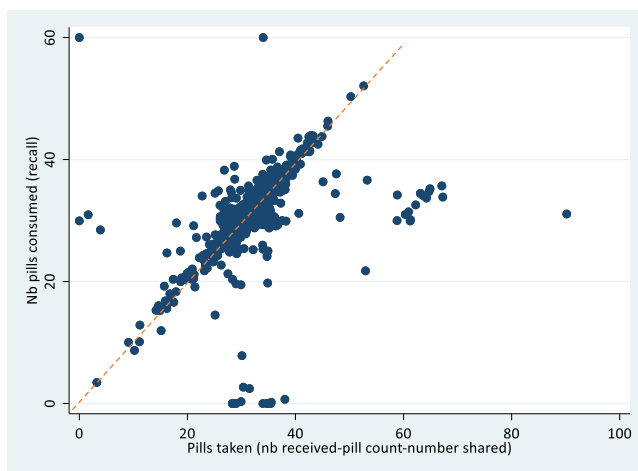


Figure 2: Number of pills recalled vs. number of pills taken in last ~30 days among currently pregnant women, in-person interview

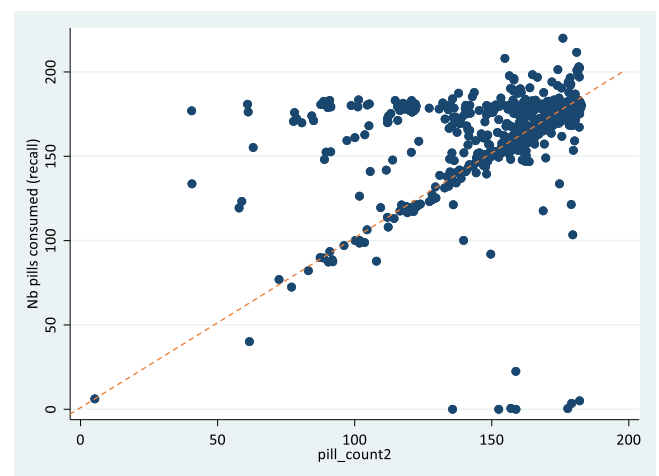


Figure 3: Number of pills recalled vs. number of pills taken across pregnancy among recently-delivered women, in-person interview

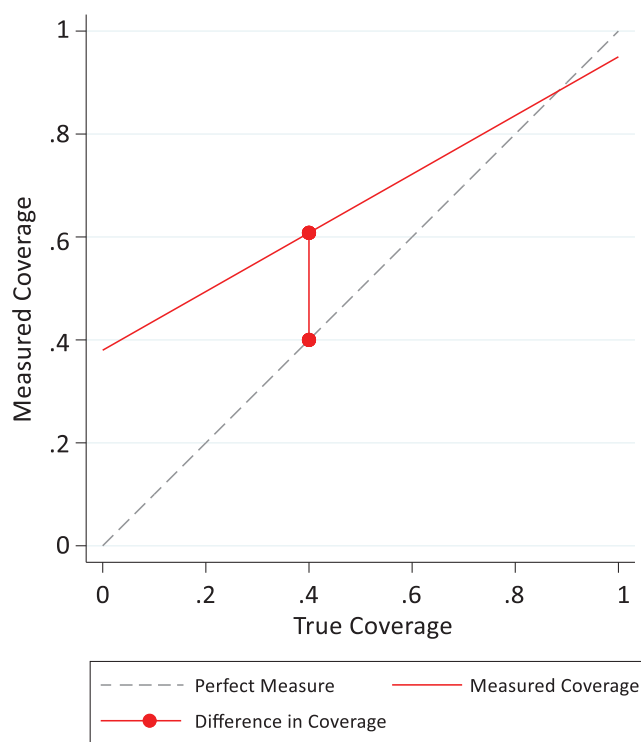


Figure 4: Measured (recalled) coverage vs. true coverage assuming 95% sensitivity and 62% specificity

of adherence (Table 1). However, specificity was significantly worse among recently-delivered women, suggesting over-reporting of adherence when asked about their entire pregnancy. This finding held true for every question formulation and modality. (Tables 1-3 and Figures 2-3)

Figure 4 shows how coverage of adherence measured via recall varies relative to true coverage when sensitivity of recall is 95% and specificity is 62% - as we found among recently delivered-women. If true adherence is 40%, asking recently-delivered would estimate 60% adherence, an approximately 20 percentage point over-estimate.

How did validity vary by modality?

Sensitivity was high and specificity was moderate for currently pregnant women asked about number consumed via mobile phone interview. Overall, validity of recalled adherence via mobile phone interview was somewhat lower relative to face-to-face interviews. However, the sample size for mobile phone interviews was small, leading to wide confidence intervals, particularly at the 75% adherence threshold. Our finding of moderate recall accuracy for mobile phone interviews should be treated with caution given small sample sizes.

IMPLICATIONS

Our findings suggest that, in Nepal, the indicators and questions used by household surveys like the Demographic and Health Survey (NDHS) to measure adherence to iron-containing supplements in pregnancy should be changed. Instead of asking recently-delivered women about pills consumed across pregnancy, currently pregnant women should be asked about the number of supplements consumed in the previous month. Receipt of iron-containing supplements should continue to be collected among recently-delivered women; previous research in Nepal shows that recently-delivered women can accurately report whether they received any iron-containing supplements.¹

Posing questions to currently pregnant women in the NDHS is likely feasible as the questionnaire includes questions to identify pregnant women. However, the sample size would be significantly smaller than for recently-delivered women and most women identified would be in the second or third trimester of pregnancy.

This change will require advocating for new indicators and methods with both survey and program stakeholders. It will have significant implications for monitoring trends across time as it would no longer be possible to directly report the proportion of women who receive at least 90 or 180 pills across their entire pregnancy. There are alternative ways to provide data on adherence across pregnancy. Guidance about how these data should be presented and interpreted will be needed.

Our results also suggest that, with caveats about our study sample size, mobile phone interviews with currently pregnant women may give a meaningful indication of adherence for program monitoring purposes. However, obtaining a list of mobile phone numbers of pregnant women may not be feasible in the programmatic context.

Validation studies for adherence questions should be replicated in other implementation settings using these and alternative question formulations, particularly in other regions and in settings with lower adherence. Questions on timing of supplement initiation should also be validated. Because our data were collected by interviewers during home visits, these results are not generalizable to collection of adherence data by health workers during ANC visits. Collection of adherence data during antenatal visits should be validated to assess social desirability bias, as well as the time required to routinely collect and report this data.

ANNEX

Table 1. Validity of recalled number of supplements consumed by in-person interview among currently pregnant and recently-delivered women

	N	Sn	95% CI	Sp	95% CI	AUC	95% CI	Prevalence (true)
Adherence at 30 days (currently pregnant women)								
≥90% adherence	1104	93.9%	(92.1-95.3%)	85.5%	(79.3-90.4%)	0.897	(0.869-0.924)	84%
≥75% adherence	1104	97.2%	(96.0-98.1%)	81.0%	(69.1-89.8%)	0.891	(0.842-0.940)	94%
Adherence over entire pregnancy (recently-delivered women)								
≥90% adherence	1167	95.4%	(96.0-98.4%)	62.3%	(57.4-67.0%)	0.798	(0.774-0.823)	66%
≥75% adherence	1167	98.9%	(98.0-99.4%)	34.4%	(26.1-43.6%)	0.666	(0.624-0.709)	90%

Table 2. Validity of recalled number of supplements missed by in-person interview among currently pregnant and recently-delivered women

	N	Sn	95% CI	Sp	95% CI	AUC	95% CI	Prevalence (true)
Adherence at 30 days (currently pregnant women)								
≥90% adherence	1125	96.9%	(95.6-97.9%)	64.7%	(57.5-71.5%)	0.808	(0.774-0.843)	83%
≥75% adherence	1125	97.8%	(96.7-98.6%)	57.9%	(46-69.1%)	0.779	(0.722-0.835)	93%
Adherence over entire pregnancy (recently-delivered women)								
≥90% adherence	1218	97.0%	(95.6-98.1%)	21.4%	(17.6-25.5%)	0.592	(0.572-0.612)	64%
≥75% adherence	1218	99.7%	(99.2-99.9%)	14.8%	(9.3-21.9%)	0.573	(0.543-0.602)	89%

Table 3. Validity of recalled number of supplements consumed by mobile phone interview among currently pregnant and recently-delivered women

	N	Sn	95% CI	Sp	95% CI	AUC	95% CI	Prevalence (true)
Adherence at 30 days (currently pregnant women)								
≥90% adherence	109	92.6%	(85.3-97.0%)	66.7%	(38.4-88.2%)	0.796	(0.670-0.922)	86%
≥75% adherence	109	96.1%	(90.3-98.9%)	71.4%	(29.0-96.3%)	0.838	(0.656-1.000)	94%
Adherence over entire pregnancy (recently-delivered women)								
≥90% adherence	238	82.1%	(73.9-88.5%)	34.7%	(26.3-43.9%)	0.584	(0.529-0.639)	49%
≥75% adherence	238	92.7%	(88.2-95.8%)	12.1%	(3.4-28.2%)	0.524	(0.465-0.583)	86%

Table 4. Validity of recalled number of supplements missed by mobile phone interview among currently pregnant and recently-delivered women

	N	Sn	95% CI	Sp	95% CI	AUC	95% CI	Prevalence (true)
Adherence at 30 days (currently pregnant women)								
≥90% adherence	109	92.60%	(85.3-97.0%)	33.30%	(11.8-61.6%)	0.629	(0.503-0.756)	86%
≥75% adherence	109	97.10%	(91.6-99.4%)	28.60%	(3.7-71.0%)	0.628	(0.447-0.810)	94%
Adherence over entire pregnancy (recently-delivered women)								
≥90% adherence	238	95.70%	(90.3-98.6%)	5.80%	(2.4-11.6%)	0.508	(0.48-0.535)	49%
≥75% adherence	238	99.00%	(96.5-99.9%)	3.00%	(0.1-15.8%)	0.510	(0.48-0.540)	86%

Sn: Sensitivity; Sp: Specificity; AUC: Area Under the Receiver Operating Curve

Table 5. RCT sample characteristics by arm for 30-day follow-up visit and birth visit

Characteristics	30-day				Birth			
	Overall	IFA-blister	MMS-blister	MM-bottle	Overall	IFA-blister	MMS-blister	MM-bottle
	N=2474	N=799	N=843	N=832	N=2477	N=815	N=831	N=831
	% Mean (SD)	% Mean (SD)	% Mean (SD)	% Mean (SD)	% Mean (SD)	% Mean (SD)	% Mean (SD)	% Mean (SD)
Mean age (completed years)	24.5 (4.0)	24.5 (4.1)	24.6 (4.0)	24.3 (4.0)	24.4 (4.0)	24.5 (4.1)	24.6 (4.0)	24.3 (4.0)
Schooling (completed years)	6.6 (4.2)	6.6 (4.2)	6.1 (4.4)	7.0 (4.1)	6.6 (4.2)	6.7 (4.2)	6.1 (4.4)	7.0 (4.1)
First pregnancy	36.4	35.7	35.2	38.3	36.8	37.7	35.0	37.8
Number of previous pregnancies	1.9 (1.2)	1.8 (1.1)	2 (1.3)	1.8 (1.0)	1.9 (1.2)	1.8 (1.1)	2 (1.3)	1.8 (1.0)
Wealth quintile								
Bottom 20%	20.8	20.9	26.7	14.8	20.0	20.9	26.7	14.8
Below average 20%	23.4	23.3	23.5	23.6	20.0	23.3	23.5	23.6
Middle 20%	17.8	18.4	15.3	19.8	20.0	18.4	15.3	19.8
Above average 20%	22.2	20.2	21.8	24.5	20.5	20.2	21.8	24.5
Upper 20%	15.7	17.3	12.7	17.3	19.5	17.3	12.7	17.3

Table 6. Demo study sample characteristics for mobile phone and in-person interviews at 30 day follow-up visit and birth visit

Characteristics	30-day		Birth	
	Mobile	In-person	Mobile	In-person
	N=160	N=483	N=383	N=495
	% Mean (SD)	% Mean (SD)	% Mean (SD)	% Mean (SD)
Mean age (completed years)	24.8 (4.6)	24.7 (4.1)	24.7 (4.1)	24.8 (4.2)
Schooling (completed years)	9.3 (3.6)	9.1 (3.6)	9.2 (3.6)	9.2 (3.6)
First pregnancy	46.3	43.5	44.4	43.4
Number of previous pregnancies	1.4 (0.6)	1.5 (0.7)	1.4 (0.7)	1.4 (0.7)
Wealth quintile				
Bottom 20%	24.4	20.0	20.1	19.4
Below average 20%	24.4	20.7	21.4	19.6
Middle 20%	24.4	20.9	20.6	20.4
Above average 20%	15.6	19.7	15.9	20.2
Upper 20%	11.3	18.6	21.9	20.4

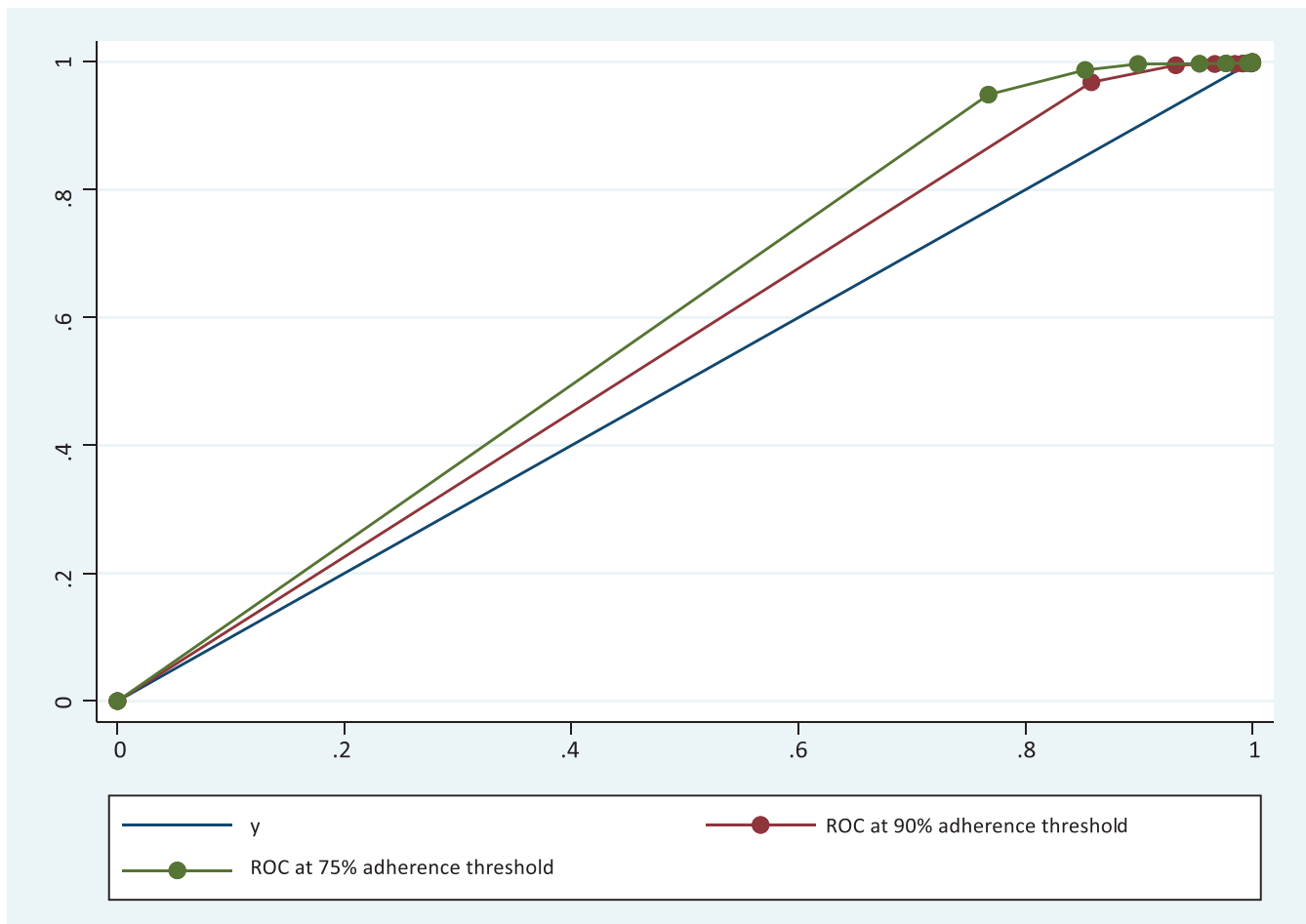


Figure 1: Receiver operating curve for usual number of days per week question

REFERENCES

1. Bryce E, Munos M, Lama TP, Khatry SK, LeClerq S, Katz J. Validation of Maternal Report of Receipt of Iron-Folic Acid Supplementation during Antenatal Care in Rural Southern Nepal. *J Nutr.* Jan 11 2022;152(1):310-318. doi:10.1093/jn/nxab336
2. Thorne-Lyman AL, Lama TP, Heidkamp RA, et al. Cognitive testing of questions about antenatal care and nutrition interventions in southern Nepal. *Soc Sci Med.* Oct 2022;311:115318. doi:10.1016/j.socscimed.2022.115318
3. Cunningham K, Poudel S, Hoang MA, et al. Non-inferiority Assessment of Maternal Adherence to Supplements, a Trial on the Effects of Multiple Micronutrient Supplementation (NAMASTE-MMS) in Nepal: study protocol. *Trials.* Nov 13 2025;26(1):502. doi:10.1186/s13063-025-09031-1
4. Horino M, Al-Khatib L, Habash R, et al. Validation Study of Adherence to Daily Multiple Micronutrient Supplementation by Palestine Refugee Women Attending United nations Relief and Works Agency for Palestine Refugees in the Near East Health Clinics in Jordan. *Curr Dev Nutr.* Dec 2025;9(12):107588. doi:10.1016/j.cdnut.2025.107588
5. Walton S, Wood E, Berhane HY, et al. Formative Research in Ethiopia and Bangladesh: Strengthening Household Survey Measures of Maternal Micronutrient Supplementation in Low- and Middle-Income Countries. *J Glob Health (In Press).* 2026;
6. Smith ER, Gomes F, Adu-Afarwuah S, et al. Contribution of Maternal Adherence to the Effect of Multiple Micronutrient Supplementation During Pregnancy: A Systematic Review and Individual Participant Data Meta-analysis. *Adv Nutr.* Jul 2025;16(7):100455. doi:10.1016/j.advnut.2025.100455