



GLOBAL  
FINANCING  
FACILITY

**FASTR**  
FREQUENT ASSESSMENTS & HEALTH  
SYSTEMS TOOLS FOR RESILIENCE



# LEVERAGING RAPID-CYCLE PHONE SURVEYS FOR IMPROVED MEASUREMENT OF PRIMARY HEALTH CARE SERVICES

Insights from a health facility phone survey validation study in Western and Volta regions, Ghana

---

# GHANA

Western and Volta Regions  
Phone Survey Validation study



# Outline



## Introduction and objectives



## Methodology

- Study design and sample
- Statistical analysis
- Resource analysis



## Findings

- Key programmatic and policy indicators
- Resource comparisons
- Detailed validation metrics
- Key Takeaways
- Appendix



## Limitations

# Introduction

- Rapid-cycle phone surveys are a promising tool for complementing large-scale, traditional in-person health facility assessments (HFAs) by improving affordability and timely data use for decision-making.
- Emerging research indicates that phone surveys can serve as a valid and a reliable complement to in-person surveys, but more studies needed to contextualize findings.
- This study contributes to the evidence base by directly comparing two HFAs in the Ghana context: the FASTR phone-based HFA, and an in-person spot check survey.

## Objectives

- 1) **Compare reported availability of facility readiness tracer items between FASTR phone survey and in-person spot check facility survey.**
- 2) **Quantify the inter-mode reliability (precision) and criterion validity (accuracy) of phone-based FASTR survey for key facility readiness indicators.**
  - a. **To understand how inter-mode reliability (precision) and criterion validity (accuracy) of the phone-based FASTR survey compare when using the enumerator's physical verification.**
- 3) **Assess how the FASTR phone survey estimates compare with in-person spot check survey for key programmatic and policy-relevant facility readiness indicators.**
- 4) **Compare the time and resources required to implement the FASTR phone survey and the in-person spot check survey.**

# Methodology

## Study design & sample

- A cross-sectional sample of 64 facilities received both a phone-based (FASTR) health facility survey and an in-person spot-check verification survey in two regions of Ghana. Both surveys were conducted within a 2-week window of one another, with facilities receiving the phone survey first.
  - This analysis was restricted to a matched sample of n=63 facilities (there was 1 mismatched facility in the data).
- Sampled facilities included community-based health planning and services (CHPS), health centres, maternity homes, clinics, polyclinics and district hospitals.
- **Phone-based FASTR survey:** 30-45 minutes phone survey with facility in charge, conducted between 12 November 2025 – 19 December 2025.
  - Average response rate of 88%. Three call attempts were made on different days of the week.
  - FIC was the respondent in 97% of the surveys
- **In-person spot-check survey:** 55-65 minute in-person health facility survey with verification of tracer items with the facility administrator, held between 24 November 2025 – 22 December 2025.
  - Average response rate of 100%
  - FIC was the respondent in 65% of the surveys
- In 15 facilities (24%), the respondent of the phone and in person surveys differed.

# Methodology

## Statistical Analysis: Validation Metrics (Obj 1 and 2)

This analysis determined whether the phone survey measured what it was supposed to measure, when compared to the in-person survey as the gold standard.

### INTER-METHOD RELIABILITY = PRECISION

- **Percent agreement:**

Represents the overall proportion of responses that were identical between both phone and in-person surveys.

Values range between 0% (no agreement) and 100% (perfect agreement)

- **Prevalence adjusted and bias adjusted Kappa (PABAK):**

A variation of Cohen's Kappa that adjusts for skewed prevalence of certain responses and rater biases in the data; ensures that agreement measures are more stable and not overly affected by imbalances in response categories

Values range between -1 (perfect disagreement) and +1 (perfect agreement).

### CRITERION VALIDITY = ACCURACY\*

- **Sensitivity:**

Measures the phone survey's ability to correctly identify items that were truly available according to the in-person survey.

Values range between 0% (none of the items truly available were correctly identified as such on the phone) and 100% (all items truly available were correctly identified as such on the phone)

- **Specificity:**

Measures the phone survey's ability to correctly identify items that were truly unavailable according to the in-person survey.

Values range between 0% (none of the items truly unavailable were correctly identified as such on the phone) and 100% (all items truly unavailable were correctly identified as such on the phone)

*\* assumption: in-person spot check as "gold standard"*

# Methodology

## Statistical Analysis: Validation Metrics (Obj 2a)

**This supplementary analysis compared the phone survey with in-person enumerators' physical verification of select items, to understand how findings compare with the respondent's in-person self-report and whether divergence from the phone survey results primarily from respondent answers, or enumerator verification, or both.**

- For select facility accreditation, infrastructure, and breastfeeding readiness indicators, the in-person survey incorporated a physical verification of the item by the enumerator.
- In this supplementary analysis, we compared the phone survey response with the enumerator verification for these key indicators.
- The same validation metrics were used (% agreement, PABAK, sensitivity, specificity) to assess whether the phone survey measured what it was supposed to measure, but in this case, considering the enumerator verification is the gold standard.

# Methodology

## Statistical Analysis: Programmatic and Policy Indicators (Obj 3)

This analysis determined whether the phone survey provided comparable overall results to the in-person survey, which would result in similar policy and programmatic insights and similar next steps.

- **Domains:** infrastructure, equipment, diagnostics, personal protective equipment (PPE), vaccines, medicines.
- We computed the domain scores as the average availability of key tracer items, and we **compared mean domain scores** between the FASTR phone survey and the in-person spot check survey using t-tests.
- We also ranked tracer items within each domain (most vs. least available items) and **conducted rank correlations** across surveys:
  - Pearson correlation ( $r$ ): ranges between -1 (perfect negative relationship) and 1 (perfect positive relationship)
  - Spearman rank correlation ( $\rho$ ): ranges between -1 (perfect negative relationship) and 1 (perfect positive relationship)

# Methodology

## Cost Analysis: resources required for implementation (Obj 4)

This analysis compared the time and resources required to field the phone and in-person surveys.

### COST METRICS

- Total cost per completed survey
- Total cost of data collection per completed survey

### TIME METRICS

- Average interview duration for participants per completed survey
- Total # of days for data collection

### STAFFING METRICS

- Number of enumerators required
- Number of supervisors required
- Average enumerator duration per completed survey
- Average # of interviews completed per day

### OPERATIONAL METRICS

- Response rate
- Item nonresponse rate (missing answers)
- Survey incomplete rate
- Average # of attempts per completed survey

---

# **How did the FASTR phone survey perform on key programmatic and policy-relevant indicators?**

Obj 3. Preliminary findings

---

# Interpretation Guide

Comparison of FASTR phone survey vs in-person spot check survey

## Policy Question:

Key question being addressed, by facility readiness domain (infrastructure, equipment, medicines, etc.)

3 core questions for each domain:

- Average % of items available in facilities
- Most available items
- Least available items

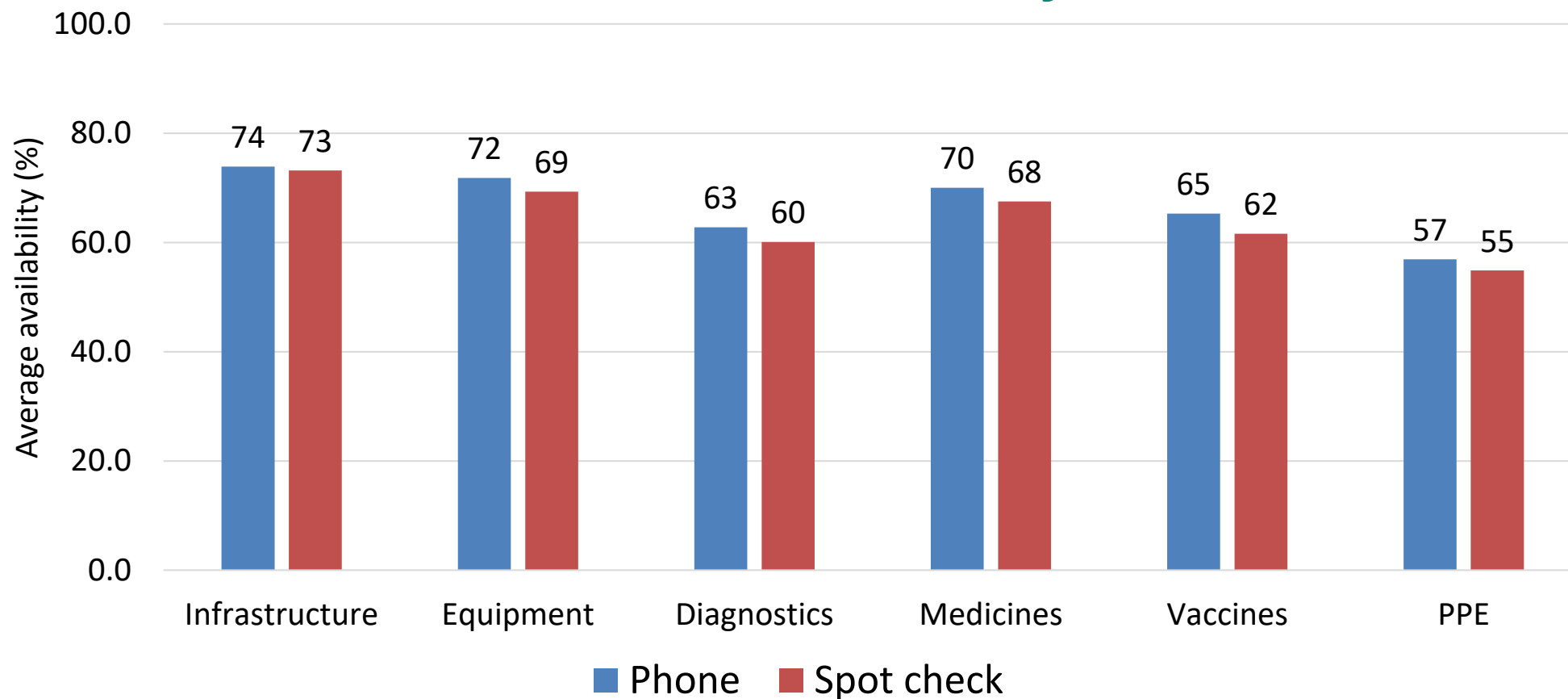
Key Policy Question	FASTR phone survey	Spot Check survey	Interpretation
<p><b>What is the average percentage of infrastructure tracer items available in facilities?</b></p> <p><b>12 tracer items included:</b> Facility connection to electric grid; guidelines/protocols for cleaning; improved source of water; improved staff toilet; improved private lockable toilet for staff, females only, males only, both males and females; handwashing with soap and water or alcohol sanitizer; safe method used for disposal of infectious medical waste; access to emergency transportation; room with auditory and visual privacy.</p>	<b>73.9%</b>	<b>73.2%</b>	<p><b>Very similar.</b></p> <p>Both indicate similar average availability of infrastructure items (12 tracers).</p> <p>The difference of +0.7 pp is not statistically significant (p-value=0.809).</p>
<p><b>What are the five most available infrastructure items in facilities?</b></p>	<ol style="list-style-type: none"> <li>1. Handwashing facility with water &amp; soap</li> <li>2. Electricity (grid connection)</li> <li>3. Improved water source</li> <li>4. Safe infectious waste disposal</li> <li>5. Improved staff toilet.</li> </ol>	<ol style="list-style-type: none"> <li>1. Improved water source</li> <li>2. Handwashing facility with water &amp; soap</li> <li>3. Electricity (grid connection)</li> <li>4. Improved female toilet (private/lockable)</li> <li>5. Improved staff toilet</li> </ol>	<p><b>Very similar.</b></p> <p>The top items identified are largely identical across surveys. Rank correlation is <b>moderate</b> (Pearson <math>r = 0.594</math>, Spearman <math>\rho = 0.594</math>, <math>N = 12</math> items).</p>
<p><b>What are the five least available infrastructure items in facilities?</b></p>	<ol style="list-style-type: none"> <li>1. Emergency transportation</li> <li>2. Improved female toilet (private/lockable)</li> <li>3. Improved male toilet (private/lockable)</li> <li>4. Improved M&amp;F toilet (private/lockable)</li> <li>5. Room with auditory &amp; visual privacy</li> </ol>	<ol style="list-style-type: none"> <li>1. Emergency transportation</li> <li>2. Room with auditory &amp; visual privacy</li> <li>3. Cleaning guidelines/protocols</li> <li>4. Safe infectious waste disposal</li> <li>5. Improved M&amp;F toilet (private/lockable)</li> </ol>	

**Interpretation:**  
*Interpretation and implications of findings*

**P value for t-test comparing FASTR phone survey and in-person spot check domain averages.**  
*P < 0.05 indicates a significant difference between domain averages.*

**Pearson and Spearman rank correlations assess strength & direction of relationship between ranks, across surveys.**  
*Values range from -1 to +1.*

**Key takeaway: The FASTR phone survey produced average domain estimates and ranks very consistent with those from in-person spot check survey, resulting in near-identical policy and programmatic interpretations across the two surveys**



*T-test comparing phone and spot check surveys indicate no statistically significant differences in all domain means. Within each domain, rank correlations ranged from moderate (0.5) to very strong (>0.9)*

## INFRASTRUCTURE: Do the surveys have similar conclusions?

Key Policy Question	Phone survey	Spot Check survey	Interpretation
<p><b>What is the average percentage of infrastructure tracer items available in facilities?</b></p> <p><b>12 tracer items included:</b>                      Facility connection to electric grid; guidelines/protocols for cleaning; improved source of water; improved staff toilet; improved private lockable toilet for staff, females only, males only, both males and females; handwashing with soap and water or alcohol sanitizer; safe method used for disposal of infectious medical waste; access to emergency transportation; room with auditory and visual privacy.</p>	<p><b>73.9%</b></p>	<p><b>73.2%</b></p>	<p><b>Very similar.</b></p> <p>Both indicate good availability of infrastructure items (12 tracers).</p> <p>The difference of +0.7 pp is not statistically significant (p-value=0.809).</p>
<p><b>What are the five most available infrastructure items in facilities?</b></p>	<ol style="list-style-type: none"> <li>1. Handwashing facility with water &amp; soap</li> <li>2. Electricity (grid connection)</li> <li>3. Improved water source</li> <li>4. Safe infectious waste disposal</li> <li>5. Improved staff toilet.</li> </ol>	<ol style="list-style-type: none"> <li>1. Improved water source</li> <li>2. Handwashing facility with water &amp; soap</li> <li>3. Electricity (grid connection)</li> <li>4. Improved female toilet (private/lockable)</li> <li>5. Improved staff toilet</li> </ol>	<p><b>Similar.</b></p> <p>The top and bottom items identified are similar across surveys. Rank correlation is <b>moderate</b> (Pearson <math>r = 0.594</math>, Spearman <math>\rho = 0.594</math>, <math>N = 12</math> items).</p>
<p><b>What are the five least available infrastructure items in facilities?</b></p>	<ol style="list-style-type: none"> <li>1. Emergency transportation</li> <li>2. Improved female toilet (private/lockable)</li> <li>3. Improved male toilet (private/lockable)</li> <li>4. Improved M&amp;F toilet (private/lockable)</li> <li>5. Room with auditory &amp; visual privacy</li> </ol>	<ol style="list-style-type: none"> <li>1. Emergency transportation</li> <li>2. Room with auditory &amp; visual privacy</li> <li>3. Cleaning guidelines/protocols</li> <li>4. Safe infectious waste disposal</li> <li>5. Improved M&amp;F toilet (private/lockable)</li> </ol>	

## MEDICAL SUPPLIES: EQUIPMENT : Do the surveys have similar conclusions?

Key Policy Question	Phone survey	Spot Check survey	Interpretation
<p><b>What is the average percentage of equipment tracer items available in facilities?</b></p> <p><b>9 tracer items included:</b>                      thermometer; child/adult weighing scale; height measuring board; BP apparatus; oxygen delivery device; speculum; pap smear equipment; refrigerator with continuous temperature monitoring; cold box/vaccine carrier with ice packs.</p>	<p><b>71.8%</b></p>	<p><b>69.3%</b></p>	<p><b>Very similar.</b></p> <p>Both indicate generally good availability of equipment tracers (9 items). Difference of +2.5 pp is not statistically significant (p-value=0.248).</p>
<p><b>What are the five most available equipment items in facilities?</b></p>	<ol style="list-style-type: none"> <li>1. Child/adult weighing scale</li> <li>2. Blood pressure apparatus</li> <li>3. Thermometer</li> </ol>	<ol style="list-style-type: none"> <li>1. Child/adult weighing scale</li> <li>2. Blood pressure apparatus</li> <li>3. Thermometer</li> </ol>	<p><b>Identical.</b></p> <p>The items identified are identical. There is a <b>very strong correlation</b> for both values and rankings across the 9 equipment tracers (Pearson and Spearman = 0.98).</p>
<p><b>What are the five least available equipment items in facilities?</b></p>	<ol style="list-style-type: none"> <li>1. Pap smear collection equipment</li> <li>2. Oxygen delivery device</li> <li>3. Speculum</li> </ol>	<ol style="list-style-type: none"> <li>1. Pap smear collection equipment</li> <li>2. Oxygen delivery device</li> <li>3. Speculum</li> </ol>	

## MEDICAL SUPPLIES - DIAGNOSTICS : Do the surveys have similar conclusions?

Key Policy Question	Phone survey	Spot Check survey	Interpretation
<p><b>What is the average percentage of diagnostics tracer items available in facilities?</b></p> <p><b>7 tracer items included:</b>            Urine dipstick testing (protein/glucose/ketones); Malaria diagnostic testing; HIV rapid test; Urine pregnancy test; Syphilis diagnostic testing; Tuberculosis diagnostic test; Hepatitis B diagnostic test.</p>	<p><b>62.8%</b></p>	<p><b>60.1%</b></p>	<p><b>Very similar.</b></p> <p>Both indicate moderate availability of diagnostic items (7 tracers). Difference of +2.7 pp is not statistically significant (p=0.070).</p>
<p><b>What are the five most available diagnostics items in facilities?</b></p>	<ol style="list-style-type: none"> <li>1. Malaria diagnostic testing</li> <li>2. Urine pregnancy test</li> <li>3. HIV rapid test</li> </ol>	<ol style="list-style-type: none"> <li>1. Malaria diagnostic testing</li> <li>2. HIV rapid test</li> <li>3. Urine pregnancy test</li> </ol>	<p><b>Identical/very similar.</b></p> <p>The top rank is very similar, with minor difference between HIV and urine pregnancy tests.</p> <p>The bottom rank is identical.</p>
<p><b>What are the five least available diagnostics items in facilities?</b></p>	<ol style="list-style-type: none"> <li>1. Tuberculosis diagnostic test</li> <li>2. Hepatitis B diagnostic test</li> <li>3. Urine dipstick testing</li> </ol>	<ol style="list-style-type: none"> <li>1. Tuberculosis diagnostic test</li> <li>2. Hepatitis B diagnostic test</li> <li>3. Urine dipstick testing</li> </ol>	<p>There is a <b>very strong correlation</b> for both values and rankings (Pearson and Spearman &gt; 0.96, N = 7 items).</p>

## MEDICAL SUPPLIES – PPE: Do the surveys have similar conclusions?

Key Policy Question	Phone survey	Spot Check survey	Interpretation
<p><b>What is the average percentage of PPE tracer items available in facilities?</b></p> <p><b>8 tracer items included:</b> Respirator/mask (e.g., N95/FFP2); Surgical/medical mask; Latex/disposable gloves; Utility gloves; Wellington boots; Disposal gowns; Head cap; Protective goggles/face shield.</p>	<p><b>56.9%</b></p>	<p><b>54.9%</b></p>	<p><b>Very similar.</b></p> <p>Both indicate the lowest availability among supply domains (8 tracers). Difference of +2.0 pp is not statistically significant (p-value=0.333).</p>
<p><b>What are the five most available PPE items in facilities?</b></p>	<ol style="list-style-type: none"> <li>1. Latex/disposable gloves</li> <li>2. Surgical/medical mask</li> <li>3. Utility gloves</li> </ol>	<ol style="list-style-type: none"> <li>1. Latex/disposable gloves</li> <li>2. Wellington boots</li> <li>3. Surgical/medical mask</li> </ol>	<p><b>Very similar.</b></p> <p>Latex gloves rank #1 in both surveys. There is a <b>strong correlation</b> for both values and rankings (Pearson and Spearman = 0.857, N = 8 items).</p>
<p><b>What are the five least available PPE items in facilities?</b></p>	<ol style="list-style-type: none"> <li>1. Respirator/mask (N95/FFP2)</li> <li>2. Head cap</li> <li>3. Protective goggles/face shield</li> </ol>	<ol style="list-style-type: none"> <li>1. Respirator/mask (N95/FFP2)</li> <li>2. Protective goggles/face shield</li> <li>3. Head cap</li> </ol>	<p>The phone survey correctly identified PPE gaps.</p>

## MEDICAL SUPPLIES - VACCINES : Do the surveys have similar conclusions?

Key Policy Question	Phone survey	Spot Check survey	Interpretation
<p><b>What is the average percentage of vaccine tracer items available in facilities?</b></p> <p><b>11 tracer items included:</b> Measles &amp; rubella; Penta (DPT+HiB+HepB); BCG; Polio; PCV; HPV; Yellow fever; Rotavirus; RTSS (malaria); Men A; Tetanus toxoid.</p>	<p><b>65.3%</b></p>	<p><b>61.6%</b></p>	<p><b>Similar.</b> Both indicate moderate availability of vaccines (11 tracers). Difference of +3.7 pp is not statistically significant (p=0.056), though borderline.</p>
<p><b>What are the five most available vaccine items in facilities?</b></p>	<ol style="list-style-type: none"> <li>1. Polio vaccine</li> <li>2. Penta (DPT+HiB+HepB) vaccine</li> <li>3. BCG vaccine</li> </ol>	<ol style="list-style-type: none"> <li>1. PCV vaccine</li> <li>2. Polio vaccine</li> <li>3. Yellow fever vaccine</li> </ol>	<p><b>Moderately similar.</b> Polio appears in top 3 for both surveys.</p> <p>Rankings show more variability than other domains due to similar prevalence across many vaccines.</p>
<p><b>What are the five least available vaccine items in facilities?</b></p>	<ol style="list-style-type: none"> <li>1. HPV vaccine</li> <li>2. RTSS (malaria) vaccine</li> <li>3. Yellow fever vaccine</li> </ol>	<ol style="list-style-type: none"> <li>1. HPV vaccine</li> <li>2. RTSS (malaria) vaccine</li> <li>3. BCG vaccine</li> </ol>	<p>Rank correlation is <b>moderate</b> (Pearson and Spearman = 0.573, N = 11).</p> <p>Phone survey correctly identified HPV and RTSS (malaria) vaccines as the two least available.</p>

## MEDICAL SUPPLIES - MEDICINES: Do the surveys have similar conclusions?

Key Policy Question	Phone survey	Spot Check survey	Interpretation
<p><b>What is the average percentage of medicine tracer items available in facilities?</b></p> <p><b>13 tracer items included:</b> Amoxicillin; Oral rehydration salts (ORS); Zinc sulphate Oxytocin; Magnesium sulphate; Folic acid Iron; Chlorhexidine gel; Long-acting contraceptives (IUD/implant); Short-acting contraceptives (pills/condom); Any first-line anti-hypertensive; Any anti-malarials; Any anti-asthmatics.</p>	<b>70.0%</b>	<b>67.5%</b>	<p><b>Very similar.</b></p> <p>Both indicate good availability of medicine items (13 tracers). Difference of +2.5 pp is not statistically significant (<math>p=0.144</math>).</p>
<p><b>What are the five most available medicine items in facilities?</b></p>	<ol style="list-style-type: none"> <li>1. Any anti-malarials</li> <li>2. Folic acid</li> <li>3. Long-acting contraceptives (IUD/implant)</li> <li>4. Iron</li> <li>5. Short-acting contraceptives (pills/condom)</li> </ol>	<ol style="list-style-type: none"> <li>1. Any anti-malarials</li> <li>2. Folic acid</li> <li>3. Amoxicillin</li> <li>4. Short-acting contraceptives (pills/condom)</li> <li>5. Iron</li> </ol>	<p><b>Very similar.</b></p> <p>Phone survey identified top 2 most available items. All 5 least available items are the same across both surveys, though in slightly different order. The phone survey correctly identifies medicine supply gaps.</p>
<p><b>What are the five least available medicine items in facilities?</b></p>	<ol style="list-style-type: none"> <li>1. Any anti-asthmatics</li> <li>2. Chlorhexidine gel</li> <li>3. Magnesium sulphate</li> <li>4. Any first-line anti-hypertensive</li> <li>5. Zinc sulphate</li> </ol>	<ol style="list-style-type: none"> <li>1. Chlorhexidine gel</li> <li>2. Any anti-asthmatics</li> <li>3. Zinc sulphate</li> <li>4. Magnesium sulphate</li> <li>5. Any first-line anti-hypertensive</li> </ol>	<p>There is a <b>strong correlation</b> for both rankings (Pearson <math>r = 0.914</math>, Spearman <math>\rho = 0.867</math>, <math>N = 13</math> items).</p>

## SUMMARY: POLICY RELEVANCE

Domain	Phone survey mean (%)	Spot-check survey mean (%)	Difference (pp)	P-value	Spearman Rank Correlation
<b>Infrastructure</b>	73.9	73.2	+0.7	0.809	$\rho = 0.594$ (Moderate)
<b>Equipment</b>	71.8	69.3	+2.5	0.248	$\rho = 0.983$ (Very Strong)
<b>Diagnostics</b>	62.8	60.1	+2.7	0.070	$\rho = 0.964$ (Very Strong)
<b>Medicines</b>	70.0	67.5	+2.5	0.144	$\rho = 0.867$ (Strong)
<b>Vaccines</b>	65.3	61.6	+3.7	0.056	$\rho = 0.573$ (Moderate)
<b>PPE</b>	56.9	54.9	+2.0	0.333	$\rho = 0.857$ (Strong)

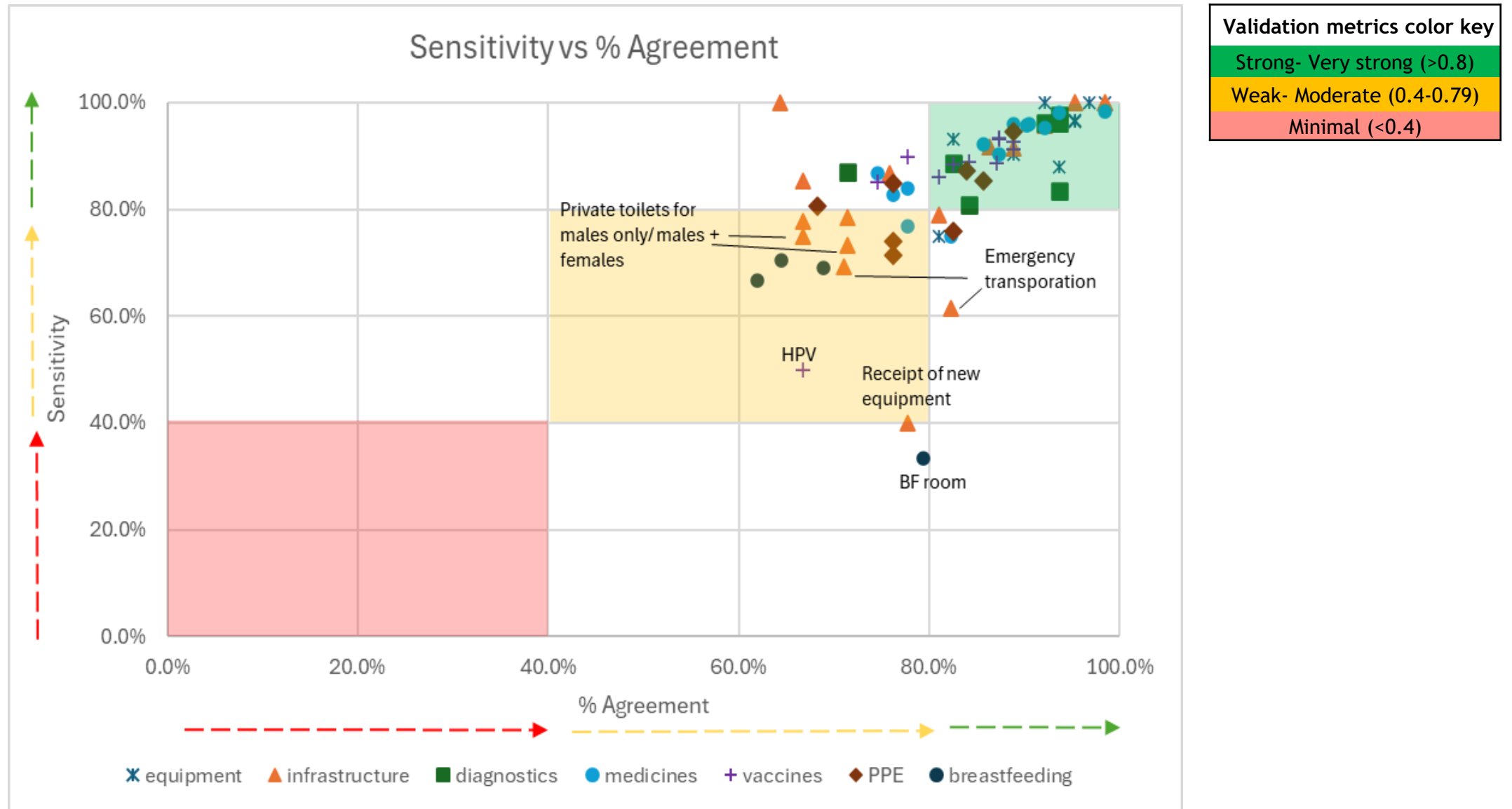
---

# Detailed Validation Metrics

Obj. 1 & 2. Preliminary findings

---

**Key takeaway: Most items performed strongly on the phone survey, some infrastructure, PPE, BF items performed moderately. The below figure summarizes these results across the 7 readiness domains and 2 validation metrics**



# Interpretation Guide

**Self-reported Survey results:** % of facilities reporting the item in the phone and in-person surveys, respectively (based on the validation sample of n=63 facilities)

**P value:** t test of difference in mean survey estimates

Validation metrics color key	
Very Strong (>0.9)	Color key legend* Categorizing validity and reliability results from Minimal (<0.4) to Very Strong (>0/9)
Strong (0.8-0.9)	
Moderate (0.6-0.79)	
Weak (0.4-0.59)	
Minimal (<0.4)	

Mapping	No	Indicator	Survey estimates			Precision metrics		Accuracy metrics		N
			FASTR	Spot Check	p value	% Agreement	PABAK	Sensitivity	Specificity	
Direct	1	Indicator 1	52.4%	42.9%	0.285	61.9%	0.24	66.7%	58.3%	63
Direct	2	Indicator 2	14.3%	19.0%	0.473	79.4%	0.71	33.3%	90.2%	63
Direct	3	Indicator 3	<b>53.2%</b>	<b>43.5%</b>	0.031	64.5%	0.29	70.4%	60.0%	62
Direct	4	Indicator 4	49.2%	47.5%	0.856	68.9%	0.38	69.0%	68.8%	61

**Mapping:** *Directly mapped* indicators have identical questions and response options in both surveys; *Indirectly mapped* indicators have different questions or response options across surveys.

**Bolding:** Bolded survey estimates indicate significant differences between FASTR and spot-check survey estimates (p<0.05)

**Agreement:** % of responses that were identical across both surveys

**PABAK:** An adjusted version of Cohen’s Kappa that accounts for prevalence and rater bias, offering a more stable measure of agreement. Most health facility indicators are based on high-volume items and generally have high availability. PABAK is only reported for *directly mapped* indicators.

**Sensitivity:** Proportion of actual positive cases (e.g., where the item was reported as available in-person) that are correctly identified as positive/available by the phone survey.

**Specificity:** Proportion of actual negative cases (e.g., where the item was not available in person) that are correctly identified as negative/unavailable by the phone survey

\*Validation metrics color key: Adapted from Landis & Koch (1977). Thresholds applied uniformly across percent agreement, PABAK, sensitivity, and specificity.

**KEY FINDINGS SUMMARIZED HERE**

# Important Interpretation notes (1/2)

- Close to 0% specificity and/or sensitivity values may result from indicator prevalences being very high (close to 100%), or very low (close to 0%). This is because there is a very small sample size for the calculation of these accuracy metrics.
- PABAK values are only reported for *directly mapped* indicators, for which the questions and response options were identical across the phone and in-person surveys.
- PABAK, sensitivity and specificity metrics are not available for indicators that had 100% prevalence in one of the surveys.
- Categorical indicators were re-categorized into binary indicators to assess sensitivity and specificity metrics: *i.e. Type of water source (categorical) → availability of an improved water source (binary).*

# Important Interpretation notes (2/2)

- Medical supplies item availability was collected differently in the phone survey (available/not available) and in the in-person survey (observed available unexpired, observed available expired, reported available not seen, not available, never available). This was done to provide further insights on how phone survey availability should be interpreted.

## FASTR phone survey

SUP_05A AA	Are the following vaccines available today anywhere in this facility?	[single select]
	<ul style="list-style-type: none"> <li>a. Measles and rubella vaccine</li> <li>b. Penta (DPT + HiB + HepB) vaccine</li> <li>c. BCG vaccine</li> <li>d. Polio vaccine</li> <li>e. PCV (pneumococcal vaccine)</li> <li>f. Human papilloma vaccine (HPV) injection</li> <li>g_gh. Yellow fever Vaccine</li> <li>h_gh. Rota vaccine</li> <li>i_gh. RTSS (malaria vaccine)</li> <li>j_gh. Men A</li> <li>k_gh. Tetanus toxoid vaccine</li> </ul>	<ul style="list-style-type: none"> <li>1 Yes</li> <li>2 No</li> <li>-99 Don't know</li> </ul>

## In-person spot check survey

SUP_05A AA	Are the following vaccines available in the health facility?	[Single Selection]
	<ul style="list-style-type: none"> <li>a. Measles and rubella vaccine</li> <li>b. Penta (DPT + HiB + HepB) vaccine</li> <li>c. BCG vaccine</li> <li>d. Polio vaccine</li> <li>e. PCV (pneumococcal vaccine)</li> <li>f. Human papilloma vaccine (HPV) injection</li> <li>g_gh. Yellow fever</li> <li>h_gh. Rota vaccine</li> <li>i_gh. RTSS (malaria vaccine)</li> </ul>	<ul style="list-style-type: none"> <li>1 Available, observed at least one not expired</li> <li>2 Available, observed but all expired</li> <li>3 Reported available but not seen</li> <li>4 Not available today</li> <li>5 Never available</li> </ul>

- For this analysis, an item was defined available in the in-person spot check if it was reported available (observed not expired, observed expired, or reported available but not seen).
- As a supplementary sensitivity analysis**, we replicated the analyses using a stricter definition (item available only if observed available and unexpired) where relevant; we also conducted additional analyses with enumerator verification of key infrastructure tracers. *Refer to synthesis slide 41 and appendix slides 40-50.*

# FACILITY ACCREDITATION

Validation metrics color key	
Very Strong (>0.9)	
Strong (0.8-0.9)	
Moderate (0.6-0.79)	
Weak (0.4-0.59)	
Minimal (<0.4)	

Mapping No	Indicator	Survey estimates			Precision		Accuracy		N
		Phone	In-person	p value	% Agreement	PABAK	Sensitivity	Specificity	
direct 1	Facility is accredited by Health Facility Regulatory Agency (HeFRA)	76.4%	60.0%	0.065	76.4%	0.59	93.9%	50.0%	55
direct 2	Facility is NHIA credentialed	85.7%	82.5%	0.626	96.8%	0.96	100.0%	81.8%	63
direct 3	Facility is a Network of Practice facility	77.0%	75.4%	0.832	85.2%	0.76	91.3%	66.7%	61

- **No statistically significant differences in self-reported estimates between surveys.**
- **Availability of NHIA credentials had very strong agreement, PABAK and sensitivity.**
- **Accreditation by HeFRA had moderate agreement, weak PABAK and specificity, suggesting responses diverged across surveys.**
  - **Enumerator verification indicates over-reporting of HeFRA accreditation in-person when compared to the enumerator’s verification, further explaining the weak/moderate agreement with the phone survey (*appendix slide 41*).**

# INFRASTRUCTURE: WATER, ELECTRICITY AND CLEANING PROTOCOLS

Validation metrics color key	
Very Strong (>0.9)	
Strong (0.8-0.9)	
Moderate (0.6-0.79)	
Weak (0.4-0.59)	
Minimal (<0.4)	

Mapping	No	Indicator	Survey estimates			Precision		Accuracy		N
			Phone	In-person	p value	% Agreement	PABAK	Sensitivity	Specificity	
direct	1	Facility has any electricity source (as a primary or secondary source)	96.8%	92.1%	0.243	95.2%	0.95	100.0%	40.0%	63
direct	2	Facility connected to electric grid (as a primary or secondary source)	93.7%	92.1%	0.729	98.4%	0.98	100.0%	80.0%	63
direct	3	Facility has an improved source of water*	88.9%	93.7%	0.344	88.9%	0.87	91.5%	50.0%	63
direct	4	Guidelines/protocols for cleaning facility (floors, counters, beds)	69.4%	61.3%	0.345	75.8%	0.56	86.8%	58.3%	62

\**Improved water sources: Piped, Ghana Water, public tap/standpipe, protected well/tube well/borehole, rainwater, bottled.*

- **Generally good availability of water/electricity with no statistically significant differences in self-reported estimates between surveys.**
- **Policy-relevant indicators (access to any electricity, connection to electric grid, access to an improved source of water) have strong to very strong %agreement, PABAK and sensitivity, although the survey responses diverged for the exact type of water/electricity source (categorical data not shown).**
- **Availability of cleaning guidelines/protocols performed moderate to weak overall; enumerator verification suggests significant over-reporting when self-reported in-person compared to the verification (*appendix slide 42*).**

# INFRASTRUCTURE: SANITATION & WASTE DISPOSAL

Validation metrics color key	
Very Strong (>0.9)	
Strong (0.8-0.9)	
Moderate (0.6-0.79)	
Weak (0.4-0.59)	
Minimal (<0.4)	

Mapping No	Indicator	Survey estimates			Precision		Accuracy		N
		Phone	In-person	p value	% Agreement	PABAK	Sensitivity	Specificity	
direct	1 Facility has an improved toilet within the facility used by staff*	74.6%	73.0%	0.839	92.1%	0.87	95.7%	82.4%	63
direct	2 Facility has at least one improved, functional, private and accessible toilet that can be locked designated for health facility staff only (among facilities that reported having an improved toilet)	81.8%	81.8%	1.000	86.4%	0.789	91.7%	62.5%	44
direct	3 Facility has at least one improved, functional, private and accessible toilet that can be locked designated for females only (among facilities that reported having an improved toilet for females)	71.4%	90.5%	0.116	81.0%	0.621	78.9%	100.0%	21
direct	4 Facility has at least one improved, functional, private and accessible toilet that can be locked designated for males only (among facilities that reported having an improved toilet for females)	71.4%	76.2%	0.726	66.7%	0.338	75.0%	40.0%	21
direct	5 Facility as at least one improved, functional, private and accessible toilet that can be locked designated for both males and females (among facilities that reported having an improved toilet for females)	61.9%	71.4%	0.513	71.4%	0.469	73.3%	66.7%	21
direct	6 Bin with lid available for disposal of used menstrual hygiene products in females only toilet (among facilities that reporting having a female-only toilet)	78.6%	42.9%	0.053	64.3%	0.334	100.0%	37.5%	14
direct	7 Handwashing facility with running water and soap, or alcohol-based hand sanitizer available	95.2%	93.7%	0.697	98.4%	0.982	100.0%	75.0%	63
direct	8 Safe method used for final disposal of infectious medical waste**	79.4%	65.1%	0.073	66.7%	0.443	85.4%	31.8%	63

\**Improved toilet: flush to sewer/septic, VIP latrine, pit latrine with slab, composting.*

\*\* *Safe waste disposal method: burn in incinerator with incinerator functional and with fuel, removed offsite stored in covered container or protected environment, dumped without burning in covered/protected pit, open burning in pit or protected ground, or offsite collection by government agency.*

- While responses diverged for the specific types of toilets/ waste disposal methods, there was very strong agreement and sensitivity for the availability of an improved staff toilet.
- Responses diverged for the availability of an improved, functional, private and accessible female only/male only toilet, and that of a bin with lid in female-only toilet (possibly influenced by small sample sizes), and for safe waste disposal method with over-reporting in phone survey.

# INFRASTRUCTURE: EMERGENCY TRANSPORTATION & PRIVACY

Validation metrics color key	
Very Strong (>0.9)	
Strong (0.8-0.9)	
Moderate (0.6-0.79)	
Weak (0.4-0.59)	
Minimal (<0.4)	

Mapping No	Indicator	Survey estimates			Precision		Accuracy		N
		Phone	In-person	p value	% Agreement	PABAK	Sensitivity	Specificity	
indirect	1 Facility has access to emergency transportation stationed at facility or that can be called for*	37.1%	21.0%	0.048	71.0%	-	69.2%	71.4%	62
indirect	2 Facility has access to emergency transportation stationed at facility or currently in use in working order with fuel and a driver 24 hours, or has access to a vehicle that can be called for	22.6%	21.0%	0.828	82.3%	-	61.5%	87.8%	62
direct	3 Room with auditory and visual privacy available for patient consultation	61.9%	58.7%	0.716	71.4%	0.452	78.4%	61.5%	63
direct	4 Facility received new infrastructure in past quarter	6.3%	1.6%	0.171	95.2%	0.948	100.0%	95.2%	63
direct	5 Facility received new equipment in past quarter	17.5%	23.8%	0.379	77.8%	0.670	40.0%	89.6%	63
direct	6 Facility received additional training for health care providers in past quarter	73.0%	71.4%	0.842	66.7%	0.443	77.8%	38.9%	63

\*defined as vehicle currently parked at health facility, usually parked but in use, or not parked but available on call

- Respondents significantly over-reported facilities' access to emergency transportation vehicle on the phone survey, compared to in-person (p<0.05): this indicator had moderate precision and accuracy.
- Availability of a room with auditory and visual privacy for patient consultations had moderate/weak precision and moderate accuracy: *possibly influenced by the complex nature of the question and response options.*
- Few facilities received of new infrastructure in past quarter; this indicator had very strong precision and accuracy, compared to that of equipment/training.
  - Enumerator verification suggests over-reporting of training received in-person compared to verification (*appendix slide 44*).

# MEDICAL SUPPLIES: EQUIPMENT

Validation metrics color key	
Very Strong (>0.9)	
Strong (0.8-0.9)	
Moderate (0.6-0.79)	
Weak (0.4-0.59)	
Minimal (<0.4)	

Mapping No	Indicator	Survey estimates			Precision		Accuracy		N
		Phone	In-person*	p value	% Agreement	PABAK	Sensitivity	Specificity	
indirect 1	Thermometer	88.9%	90.5%	0.770	95.2%	-	96.5%	83.3%	63
indirect 2	Child/adult weighing scale	93.7%	95.2%	0.697	95.2%	-	96.7%	66.7%	63
indirect 3	Height measuring board	76.2%	81.0%	0.515	88.9%	-	90.2%	83.3%	63
indirect 4	Blood pressure apparatus	93.7%	92.1%	0.729	98.4%	-	100.0%	80.0%	63
indirect 5	Oxygen delivery device	36.5%	39.7%	0.714	93.7%	-	88.0%	97.4%	63
indirect 6	Speculum	66.7%	58.7%	0.357	92.1%	-	100.0%	80.8%	63
indirect 7	Equipment to collect pap smear specimen**	25.4%	12.7%	0.070	81.0%	-	75.0%	81.8%	63
indirect 8	Refrigerator with continuous temperature monitoring	77.8%	69.8%	0.311	82.5%	-	93.2%	57.9%	63
indirect 9	Cold box or vaccine carrier with ice pack	87.3%	84.1%	0.611	96.8%	-	100.0%	80.0%	63

\* Defined as equipment item available if observed or reported available but not seen, and functional.

\*\* spatulas, endocervical brushes.

- No significant difference in reported availability of equipment items across surveys.
- Strong- Very strong agreement for all equipment items.
- Very strong sensitivity for most items except pap smear equipment and oxygen delivery device, indicating the phone survey correctly identified facilities where the equipment available.
- Strong specificity for most items, except refrigerator with continuous temperature monitoring.

# MEDICAL SUPPLIES: DIAGNOSTICS

Validation metrics color key	
Very Strong (>0.9)	
Strong (0.8-0.9)	
Moderate (0.6-0.79)	
Weak (0.4-0.59)	
Minimal (<0.4)	

Mapping	No	Indicator	Survey estimates			Precision		Accuracy		N
			Phone	In-person*	p value	% Agreement	PABAK	Sensitivity	Specificity	
indirect	1	Urine dipstick for protein, glucose or ketones	68.3%	65.1%	0.705	93.7%	-	97.6%	86.4%	63
indirect	2	Malaria diagnostic testing (rapid test or other)	82.5%	81.0%	0.818	92.1%	-	96.1%	75.0%	63
indirect	3	HIV rapid test	81.0%	81.0%	1.000	93.7%	-	96.1%	83.3%	63
indirect	4	Urine pregnancy test	82.5%	73.0%	0.199	71.4%	-	87.0%	29.4%	63
indirect	5	Syphilis diagnostic testing (rapid test or specific test or serology)	71.4%	69.8%	0.845	82.5%	-	88.6%	68.4%	63
indirect	6	Tuberculosis diagnostic test	12.7%	9.5%	0.571	93.7%	-	83.3%	94.7%	63
indirect	7	Hepatitis B diagnostic test	41.3%	41.3%	1.000	84.1%	-	80.8%	86.5%	63

\* Defined as commodity items available if observed and not expired, observed but expired, or reported available but not seen; equipment item available if observed or reported available but not seen.

- No significant difference in reported availability of diagnostic items across surveys.
- Most diagnostics had very strong agreement, except urine pregnancy tests (moderate), syphilis tests (strong) and hepatitis B test (strong).
- Strong – Very strong sensitivity for all items
- Minimal specificity for urine pregnancy tests, indicating the phone survey did not accurately identify facilities where the item was reported unavailable by the in-person spot check.

# MEDICAL SUPPLIES: PERSONAL PROTECTIVE EQUIPMENT (PPE)

Validation metrics color key	
Very Strong (>0.9)	
Strong (0.8-0.9)	
Moderate (0.6-0.79)	
Weak (0.4-0.59)	
Minimal (<0.4)	

Mapping	No	Indicator	Survey estimates			Precision		Accuracy		N
			Phone	In-person*	p value	% Agreement	PABAK	Sensitivity	Specificity	
indirect	1	Respirator/mask (e.g., N95/FFP2)	41.3%	36.5%	0.584	76.2%	-	73.9%	77.5%	63
indirect	2	Surgical or medical mask	66.7%	57.1%	0.271	68.3%	-	80.6%	51.9%	63
indirect	3	Latex or other disposable gloves	88.9%	87.3%	0.783	88.9%	-	94.5%	50.0%	63
indirect	4	Utility gloves	60.3%	52.4%	0.369	76.2%	-	84.8%	66.7%	63
indirect	5	Wellington boots	60.3%	65.1%	0.581	85.7%	-	85.4%	86.4%	63
indirect	6	Disposal gowns	53.2%	50.0%	0.719	83.9%	-	87.1%	80.6%	62
indirect	7	Head cap	41.3%	46.0%	0.590	82.5%	-	75.9%	88.2%	63
indirect	8	Protective goggles or face shield	42.9%	44.4%	0.857	76.2%	-	71.4%	80.0%	63

\* Defined item available if observed and not expired, observed but expired, or reported available but not seen.

- **No significant difference in reported availability of PPE items across surveys.**
- **Moderate to strong agreement and sensitivity.**
- **Weak specificity for surgical masks and latex gloves, with over-reported availability on the phone, though not statistically significantly.**

## MEDICAL SUPPLIES: VACCINES

Validation metrics color key	
Very Strong (>0.9)	
Strong (0.8-0.9)	
Moderate (0.6-0.79)	
Weak (0.4-0.59)	
Minimal (<0.4)	

Mapping	No	Indicator	Survey estimates			Precision		Accuracy		N
			Phone	In-person*	p value	% Agreement	PABAK	Sensitivity	Specificity	
indirect	1	Measles and rubella vaccine	69.8%	68.3%	0.847	82.5%	-	88.4%	70.0%	63
indirect	2	Penta (DPT + HiB + HepB) vaccine	73.0%	69.8%	0.693	87.3%	-	93.2%	73.7%	63
indirect	3	BCG vaccine	71.4%	61.9%	0.257	77.8%	-	89.7%	58.3%	63
indirect	4	Polio vaccine	74.6%	71.4%	0.688	87.3%	-	93.3%	72.2%	63
indirect	5	PCV (pneumococcal vaccine)	71.4%	73.0%	0.842	88.9%	-	91.3%	82.4%	63
indirect	6	Human papilloma vaccine (HPV) injection	<b>33.3%</b>	<b>16.7%</b>	<b>0.035</b>	66.7%	-	50.0%	70.0%	60
indirect	7	Yellow fever	67.7%	71.0%	0.697	87.1%	-	88.6%	83.3%	62
indirect	8	Rotavirus vaccine	71.4%	71.4%	1.000	84.1%	-	88.9%	72.2%	63
indirect	9	RTSS (malaria vaccine)	47.6%	42.9%	0.591	88.9%	-	92.6%	86.1%	63
indirect	10	Men A	68.3%	68.3%	1.000	81.0%	-	86.0%	70.0%	63
indirect	11	Tetanus toxoid vaccine	69.8%	63.5%	0.450	74.6%	-	85.0%	56.5%	63

\*Defined item available if observed and not expired, observed but expired, or reported available but not seen.

- **Moderate vaccine availability overall.**
- **HPV vaccine availability lowest and significantly over-reported on the phone ( $p < 0.05$ ), and had moderate agreement, weak sensitivity.**
- **Agreement strong for most vaccines, and moderate for HPV, BCG and TT.**
- **BCG and TT vaccines had weak specificity: the phone survey did not identify all facilities that reported unavailable vaccines in the spot check survey.**

## MEDICAL SUPPLIES: MEDICINES

Validation metrics color key	
Very Strong (>0.9)	
Strong (0.8-0.9)	
Moderate (0.6-0.79)	
Weak (0.4-0.59)	
Minimal (<0.4)	

Mapping	No	Indicator	Survey estimates			Precision		Accuracy		N
			Phone	In-person*	p value	% Agreement	PABAK	Sensitivity	Specificity	
indirect	1	Amoxicillin	79.4%	82.5%	0.650	87.3%	-	90.4%	72.7%	63
indirect	2	Oral Rehydration Solutions (ORS)	78.7%	75.4%	0.667	90.2%	-	95.7%	73.3%	61
indirect	3	Zinc sulphates	60.3%	47.6%	0.153	74.6%	-	86.7%	63.6%	63
indirect	4	Oxytocin	66.7%	65.1%	0.851	92.1%	-	95.1%	86.4%	63
indirect	5	Magnesium sulphate	55.6%	49.2%	0.476	77.8%	-	83.9%	71.9%	63
indirect	6	Folic acid	87.3%	84.1%	0.611	93.7%	-	98.1%	70.0%	63
indirect	7	Iron	82.5%	79.4%	0.650	90.5%	-	96.0%	69.2%	63
indirect	8	Chlorhexidine gel	47.6%	36.5%	0.207	76.2%	-	82.6%	72.5%	63
indirect	9	Any long-acting reversible contraception methods (IUD/implant)	84.1%	79.4%	0.489	88.9%	-	96.0%	61.5%	63
indirect	10	Any short-acting reversible contraception methods (pills/condom)	82.5%	81.0%	0.818	85.7%	-	92.2%	58.3%	63
indirect	11	Any first-line anti-hypertensive	55.6%	61.9%	0.469	77.8%	-	76.9%	79.2%	63
indirect	12	Any anti-malarials	88.9%	90.5%	0.770	98.4%	-	98.2%	100.0%	63
indirect	13	Any anti-asthmatics	40.3%	45.2%	0.586	82.3%	-	75.0%	88.2%	62

\*Defined item available if observed and not expired, observed but expired, or reported available but not seen.

- **No significant difference in reported medicines availability across surveys.**
- **Moderate to very strong agreement and sensitivity for medicines, moderate specificity.**
- **Weak specificity for short-acting contraception methods.**

# BREASTFEEDING READINESS\*

Validation metrics color key	
Very Strong (>0.9)	Strong (0.8-0.9)
Moderate (0.6-0.79)	Weak (0.4-0.59)
Minimal (<0.4)	

Mapping	No	Indicator	Survey estimates			Precision		Accuracy		N
			Phone	In-person	p value	% Agreement	PABAK	Sensitivity	Specificity	
Direct	1	10 Steps of Breastfeeding posted in service delivery points	52.4%	42.9%	0.285	61.9%	0.24	66.7%	58.3%	63
Direct	2	Dedicated room for breastfeeding support to mothers	14.3%	19.0%	0.473	79.4%	0.71	33.3%	90.2%	63
Direct	3	Minimum ANC Breastfeeding Education Checklist	53.2%	43.5%	0.281	64.5%	0.29	70.4%	60.0%	62
Direct	4	Protocol/guideline for supporting breastfeeding in service delivery points	49.2%	47.5%	0.856	68.9%	0.38	69.0%	68.8%	61

\*not included in FASTER HFA generic tool

- <50% of facilities reported having these breastfeeding items available.
- No significant difference in mean reported availability of breastfeeding-relevant items across surveys.
- Moderate agreement and minimal PABAK for most items, indicating that responses diverged significantly across surveys within facilities;
- Minimal sensitivity for a dedicated breastfeeding support room: the phone survey did not identify all facilities where this room was indeed available; low prevalence (19%) leads to small sample size for sensitivity calculation.

## WORKFORCE & STAFFING

Mapping	No	Indicator	Survey estimates			N
			Phone	In-person*	p value	
direct	1	Total number of health care providers working in the facility	15.95	20.45	0.359	62
direct	2	Total number of medical doctors working in the facility**	1.60	0.60	0.306	62
direct	3	Total number of nursing personnel working in the facility (excludes volunteers)	8.82	14.55	0.126	62
direct	4	Total number of midwifery personnel working in the facility (excludes volunteers)	2.45	5.31	0.170	62

\* Spot check staffing reported based on numbers recorded in facility roster.

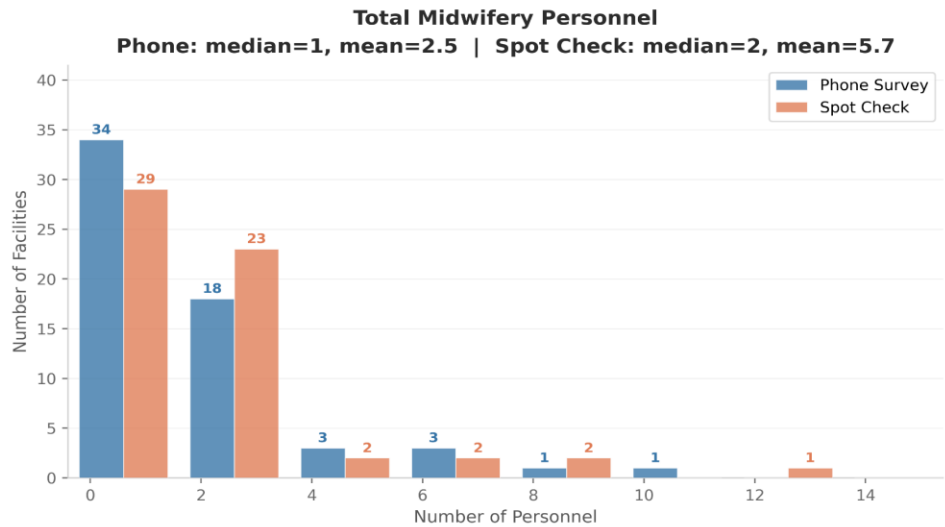
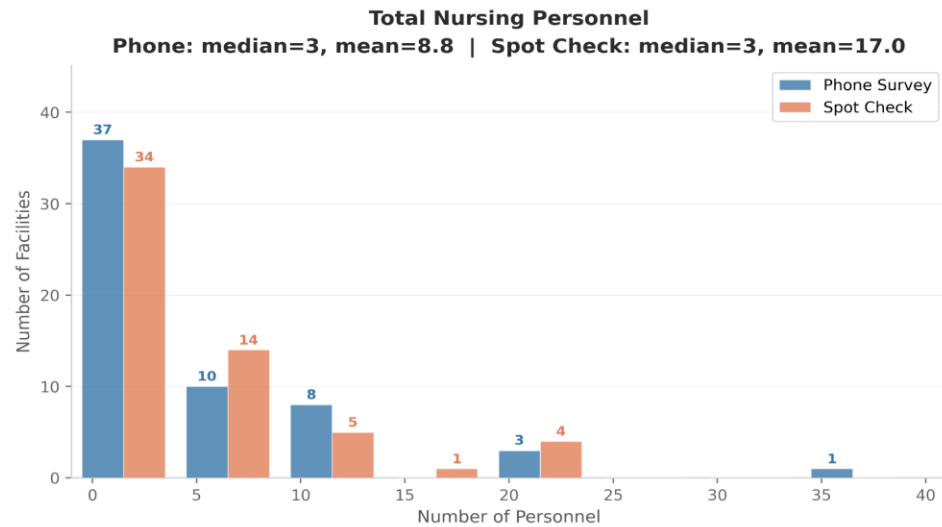
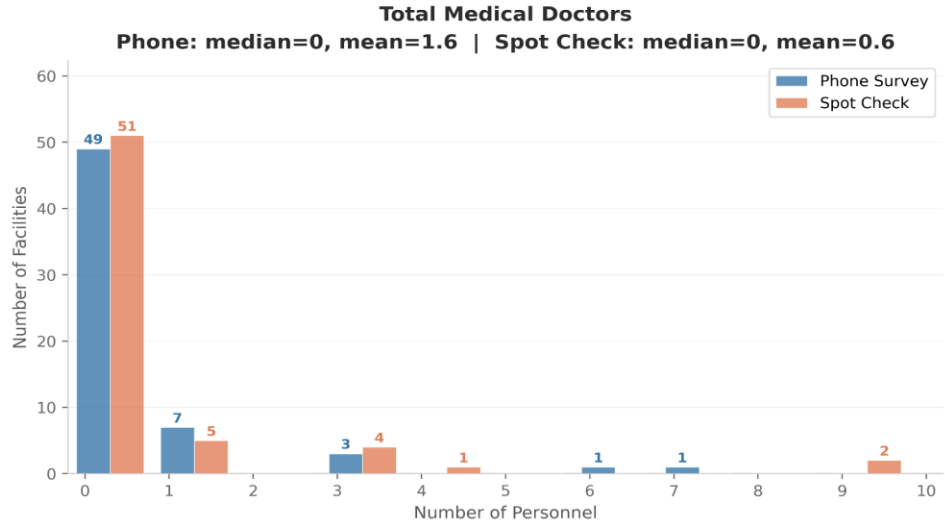
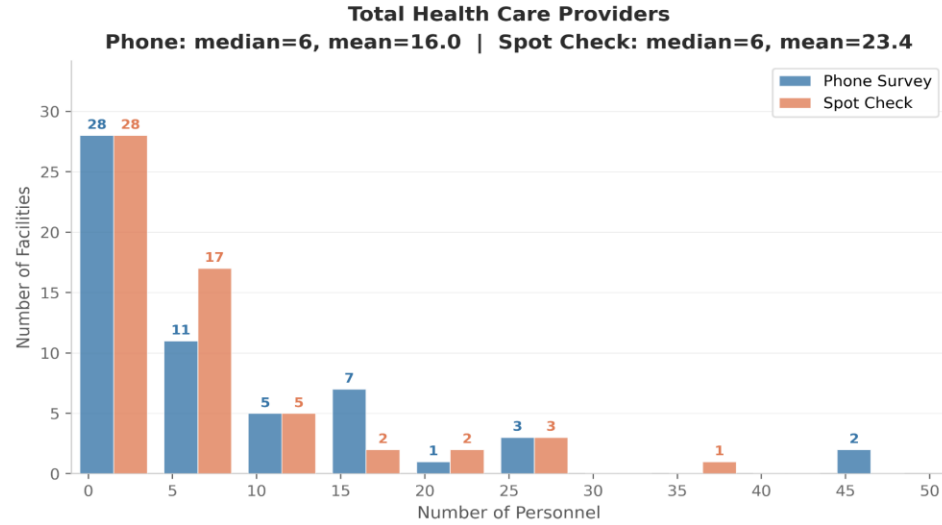
\*\* Medical doctors include generalists and specialist medical practitioners with a clinical doctoral degree

- **No significant differences in mean staffing numbers between FASTR phone survey and in-person spot check survey.**
- **Under-reporting of health care providers on the phone compared to in-person.**

# WORKFORCE & STAFFING:

## Distribution of responses between phone survey and in-person survey are consistent.

### Distribution of HR/Staffing Indicators: Phone Survey vs. In-Person Spot Check



---

# **How do the validation results compare with the enumerator's physical verification?**

Obj 2a. Preliminary findings synthesis

---

# Synthesis: Obj 2a

Comparing the phone survey responses with **enumerator's physical verification** of key infrastructure items provided the following insights: *(detailed results are in appendix)*

**1) Facility accreditation:** We found statistically significant differences ( $p < 0.05$ ) in phone vs verified availability of facility accreditations for HeFRA and NHIA, leading to minimal to moderate precision. These accreditations were also over-reported by the respondent in-person (i.e. 31.7% of facilities had verified HeFRA accreditation, vs 60% self-reported in person).

**2) Availability of cleaning guidelines/protocols:** We found statistically significant differences between the phone survey and enumerator verification ( $p < 0.05$ ), and minimal PABAK. It was also over-reported by the respondent in-person (61.3%) when compared to the enumerator's verification (41.3%).

**3) Sanitation & waste disposal:** Phone survey results were fairly consistent with enumerator verification, except for the availability of a bin with a lid in female toilets where the responses were statistically significantly different (78.9% vs 31.6% respectively,  $p < 0.05$ ) and precision was weak/minimal. Some items were over-reported in-person by the respondent (handwashing facility), whereas others were under-reported (improved staff toilet) compared to enumerator verification.

**4) Access to emergency transportation** was generally over-reported by the respondent in-person, when compared to enumerator verification, with significant differences with the phone survey results ( $p < 0.05$ ).

**5) Enumerator verification in-person** led to a lower availability of all breastfeeding readiness items; 3 out of 4 indicators had minimal PABAK, and one had minimal sensitivity.

---

**Sensitivity analysis: applying a stricter definition to item availability in the in-person survey**

---

# Synthesis: Sensitivity analysis

As a supplementary sensitivity analysis, we replicated these analyses using a **stricter definition of supplies item availability: a supply item was considered available only if it was observed available and unexpired.**

**These analyses provided the following insights:** *(detailed results are in appendix)*

- 1) Using a strict item availability definition shows that the phone and in-person survey results were mostly consistent for equipment items, with strong agreement and sensitivity, except for the availability of pap smear specimen equipment which diverged significantly ( $p < 0.05$ ). This was not the case when we used the conservative item availability definition (slide 30), suggesting that such equipment item could have been malfunctioning/invalid yet reported available.
- 2) We found no significant difference in reported availability of diagnostic and PPE across phone and in-person survey when using a strict item availability definition; Findings were consistent with conservative item availability definition (slides 31 & 32).
- 3) The stricter definition of item availability leads to lower availability of vaccines in-person, and significant differences between phone and in-person surveys for Penta, BCG, Polio, HPV, Rotavirus and TT vaccines ( $p < 0.05$ ); Agreement remains strong for most vaccines, and sensitivity remains very strong except for HPV; this is aligned with analyses using a conservative item availability definition.
- 4) Applying a stricter item availability definition led to lower availability of all medicines in-person as compared to a conservative definition; There was a significant difference in the availability of Zinc sulphates as compared to the phone survey ( $p < 0.05$ ).

---

# Resource comparison

Obj. 4. Preliminary findings

---

# COST METRICS

Metric	Definition	Phone survey	In-Person survey	Difference
Total cost per completed survey	Total cost of data collection ÷ number of completed interviews	$\$4,420 \div 65 \text{ completes} =$  <u><b>\$68*</b></u> (international call from Nigeria)	$\$15,178.28 \div 64 \text{ completes} =$  <u><b>\$237.16</b></u>	The in-person interview is approximately 3.5 times more expensive than the phone interview.
Total cost of data collection	Total cost of completing the survey	<b>\$4,420</b>	<b>\$15,178.28</b>	Overall, the in-person data collection cost is approximately 3 times higher than the phone survey, reflecting the substantially greater resource requirements for field-based data collection.

\*To call the facilities from Nigeria, airtime costs \$0.72 per minute. Airtime costs to make domestic calls in Ghana are expected to be significantly cheaper:

- MTN-to-MTN calls are ~GHS 0.014 per minute (0.14 pesewas)
- MTN-to-other networks are ~GHS 0.075 per minute (0.75 pesewas)

# TIME METRICS

---

<b>Metric</b>	<b>Definition</b>	<b>Phone survey</b>	<b>In-Person survey</b>	<b>Difference</b>
Average interview duration (per completed survey)	Total interview time ÷ number of completed surveys	43 minutes	63 minutes	On average, the in-person survey took ~20 additional minutes.
Total number of days for data collection	Number of days required to complete data collection	10 days	15 days	The in-person survey took 5 additional days for data collection.

# STAFFING METRICS

Metric	Definition	Phone survey	In-Person survey	Difference
Number of enumerators required	Number of enumerators required to complete data collection	15	10	The phone survey had a larger interviewing team – this could contribute to the shorter data collection period. The larger interviewing team also cost significantly less, however.
Number of supervisors required	Number of supervisors required to complete data collection	3	4	There were more supervisors for the in-person survey.
Average # of interviews completed per day	Average number of interviews that were completed each day	6 interviews	2 interviews	The CATI interviewers were able to complete more interviews per day than the in-person teams.

# OPERATIONAL METRICS

Metric	Definition	Phone survey	In-Person survey	Difference
Response rate	Completed surveys ÷ total eligible respondents contacted	88%	100%	The response rate for the in-person survey was 100%. The higher response rate for the in-person survey reflects its sequencing after the phone survey and the ability to directly follow up with previously contacted facilities.
Average # of attempts per completed survey	Number of calls or visits needed to complete the survey	6 call attempts	1 visit per complete interview (initial visit)	Phone surveys required multiple call attempts to reach respondents and complete interviews, whereas in-person surveys were typically completed in a single visit. This is expected, as in-person data collection allows interviewers to directly access facilities and engage respondents, reducing the need for repeated contact attempts.

---

# Conclusions

---

# FIVE KEY TAKEAWAYS (1/5)

---

1. **At aggregate levels, the phone survey produced results that were consistent with the in-person spot check survey for policy-relevant indicators.**

**There were no statistically significant differences in mean domain scores, and the phone survey correctly identified the most/least available items with generally strong correlations.**

**Implication:** FASTR phone surveys can complement in-person surveys and adequately approximate facility readiness at aggregate levels for key infrastructure and medical supplies tracers. They can identify red flags to support timely decision-making, and can also identify strengths/weaknesses of facility readiness.

---

# FIVE KEY TAKEAWAYS (2/5)

---

**2. At a more granular level, the phone survey also produced point estimates that were comparable and not statistically significantly different from those reported in the spot check survey.**

- *Among all the infrastructure, supplies HR and breastfeeding items, only two indicators were significantly over-reported on the phone (access to an emergency vehicle stationed at the facility or that can be called for, availability of HPV vaccine).*
- *While other indicators (except HR) were over-reported on the phone, the difference was not statistically significant.*

**Implication:** FASTR phone surveys can adequately approximate item availability for most infrastructure and medical supplies, and can fairly accurately estimate facility readiness to inform actionable decisions.

---

# FIVE KEY TAKEAWAYS (3/5)

---

**3. Looking at the validity and reliability of the phone survey, the majority of medical equipment, diagnostics, medicines and vaccines had strong sensitivity and % agreement.**

**We found more variability in the performance of infrastructure items, PPE and breastfeeding readiness.**

**In general, sensitivity was higher than specificity. This means that the phone survey does a better job at correctly identifying facilities that truly had a tracer item available, rather than identifying facilities where the tracer item was truly unavailable.**

- *Moderate/Weak precision and accuracy: Access to emergency transportation, availability of improved, private, functional and accessible toilets that can be locked, safe waste disposal method, room with auditory and visual privacy, and availability of cleaning guidelines.*
  - *We found consistent evidence from other settings for the availability of an improved, accessible, private toilet and a room with auditory and visual privacy.*
- *While specific water/electricity sources and sanitation type responses diverged across surveys, programmatic-relevant indicators such as access to improved water/sanitation and connection to an electric grid performed well; consistent with other settings.*
- *Volatile items tend to be over-reported on the phone: PPE (surgical masks, latex gloves) short-acting contraception methods, etc.*

**Implication:** Most essential medical tracer items captured through the phone have adequate precision and accuracy, when compared to in-person spot checks. While the phone survey can be used to signal changes in facility readiness, it should not be used to identify specific facilities requiring action.

---

# FIVE KEY TAKEAWAYS (4/5)

---

**4. Simple indicators with single items perform better than composite indicators or multi-part questions. As the question complexity increases, % agreement in responses diverge across phone and in-person surveys.**

- *Availability of pap smear collection equipment, refrigerator with continuous temperature monitoring, private functional accessible and lockable toilets performed worse than “cold box or vaccine carrier with ice pack” an indicator derived from 3 simple items collected separately.*
- *Further analyses in other settings suggest that splitting a complex question into multiple parts helps to improve results. Example: private, accessible, functional toilet that can be locked: % agreement improved significantly when the “private” factor was removed.*

**Implication:** Complex/multi-part indicators collected by phone should be broken down into separate parts to improve precision and accuracy, or may need some adjustment factor to account for likely over-reporting by phone.

---

# FIVE KEY TAKEAWAYS (5/5)

---

## 5. Supplemental analyses with enumerator verification (Obj 2a) and applying stricter definitions of item availability highlight some over-reporting in the in-person spot check (self-report), but overall results are generally consistent (*appendix slides 42-52*)

- *In-person self-reported values for infrastructure items were over-reported when compared to enumerator verification: cleaning guidelines, bin with lid in female toilet, emergency transportation, receipt of health worker training etc., leading to more divergence with phone survey results and poorer precision. For ex: 61.3% of facilities reported the availability of cleaning guidelines during the in-person survey, whereas enumerator verification suggested it was 41.3%.*
- *As expected, applying a stricter definition of item availability (observed and unexpired) led to significant differences and lower agreement with the phone survey (where observations were not possible). Example: Penta, BCG, Polio, HPV, Rotavirus and Tetanus Toxoid vaccines and zinc sulphates were over-reported by the respondent on the phone. Further analyses indicated that these differences are primarily due to the item being reported available but not observed during the survey.*

**Implication:** Enumerator verification highlight that respondents may over-report even when surveyed in-person. Phone survey estimates may therefore represent an upper-bound of facility readiness and could need adjustment factors to account for possible over-reporting.

---

# LIMITATIONS

---

## **The questionnaires for the FASTR phone survey and in-person spot check survey were not identical:**

- Therefore, some expected level of variation was introduced in the results and can explain some of the divergence in responses across surveys as well as possible lower agreement for indirectly mapped questions.

## **The survey respondent was not always identical between surveys:**

- OIC was the respondent in 95% of the phone surveys vs. 64% of the in-person spot check surveys.
  - In 15 facilities, the respondents across surveys were different: this increases the likelihood of diverging responses and level of disagreement across surveys (in addition to difference in modality, we have a difference in respondent).
  - If the respondent was not the OIC, he/she may have been less familiar with the availability of key tracer items across services, influencing the responses.
-

# CATI: CHALLENGES AND LESSONS LEARNT

---

## **Challenges:**

- Difficulty reaching respondents: Incorrect contacts, switched-off phones, and refusals required multiple call attempts
- Network limitations: Connectivity issues delayed interviews; alternative platforms (e.g., WhatsApp) helped mitigate this
- Low initial awareness and trust concerns: Limited familiarity with FASTR and skepticism of international calls reduced participation
- Longer-than-expected timelines: Data collection was extended and achieved fewer completes (65 vs. 74 planned)
- District hospital constraints: While engagement was generally positive, interviews were often prolonged due to high patient volumes and the need to involve multiple staff not immediately available

## **Lessons Learnt:**

- Strategic call timing improves response rates by reaching OICs when they are available
- MoH/GHS engagement is essential—advance notification significantly increased participation
- Leveraging regional structures (e.g., health directorates) strengthens survey uptake

# CAPI: CHALLENGES AND LESSONS LEARNT

---

## **Challenges:**

- Scheduling constraints: Aligning visits with OIC availability led to delays and rescheduling
- Higher resource requirements: CAPI required substantially more personnel, time, and funding than CATI
- District hospital constraints: Challenges included difficult access to remote facilities and extended interviews due to the need to engage multiple staff (e.g., department leads) who were not always available

## **Lessons Learnt:**

Advance approval is critical: Even facilities that participated in CATI required additional approval for in-person visits

Early engagement improves access: Community sensitization and expectation management are key to facilitating entry

# THANK YOU!

Questions?



---

# Appendix

Additional analyses were performed: a) comparing FASTR phone survey reported supplies availability with the observed availability of valid/unexpired items, and b) with verification of tracer infrastructure items (Obj 2a).

---

## FACILITY ACCREDITATION (with enumerator verification)

Validation metrics color key	
Very Strong (>0.9)	
Strong (0.8-0.9)	
Moderate (0.6-0.79)	
Weak (0.4-0.59)	
Minimal (<0.4)	

Mapping	No Indicator	Reported vs Verified	Survey estimates			Precision		Accuracy		N	
			Phone	In-person	p value	% Agreement	PABAK	Sensitivity	Specificity		
direct	1	Facility is accredited by Health Facility Regulatory Agency	Self-reported	76.4%	60.0%	0.065	76.4%	0.59	93.9%	50.0%	55
			Verified	<b>76.7%</b>	<b>31.7%</b>	<b>&lt;0.001</b>	51.7%	0.04	94.7%	31.7%	60
direct	2	Facility is NHIA credentialed	Self-reported	85.7%	82.5%	0.626	96.8%	0.96	100.0%	81.8%	63
			Verified	<b>85.7%</b>	<b>63.5%</b>	<b>0.004</b>	77.8%	0.64	100.0%	39.1%	63
direct	3	Facility is a Network of Practice facility	Self-reported	77.0%	75.4%	0.832	85.2%	0.76	91.3%	66.7%	61

*Notes: Enumerator verification entailed physically checking the credentials documentation in-person; no verification was done on the phone. Sample size for verification metrics are based on a paired sample across phone survey and verification responses.*

- The in-person enumerator verification led to a lower availability of facility accreditation.
- Statistically significant differences ( $p < 0.05$ ) in phone reported vs verified accreditations for HeFRA and NHIA, leading to minimal-moderate precision.
- These accreditations were over-reported by the respondent in-person, compared to the enumerator's verification.

# INFRASTRUCTURE: WATER, ELECTRICITY AND CLEANING PROTOCOLS (with enumerator verification)

Validation metrics color key	
Very Strong (>0.9)	
Strong (0.8-0.9)	
Moderate (0.6-0.79)	
Weak (0.4-0.59)	
Minimal (<0.4)	

Mapping	No Indicator	Reported vs Verified	Survey estimates			Precision		Accuracy		N	
			Phone	In-person	p value	% Agreement	PABAK	Sensitivity	Specificity		
direct	1	Facility has any electricity source (primary or secondary)	Self-reported	96.8%	92.1%	0.243	95.2%	0.95	100.0%	40.0%	63
direct	2	Facility connected to electric grid (primary or secondary source)	Self-reported	93.7%	92.1%	0.729	98.4%	0.98	100.0%	80.0%	63
direct	3	Facility has an improved source of water*	Self-reported	88.9%	93.7%	0.344	88.9%	0.87	91.5%	50.0%	63
direct	4	Guidelines/protocols for cleaning facility (floors, counters, beds)	Self-reported	69.4%	61.3%	0.345	75.8%	0.56	86.8%	58.3%	62
			Verified	<b>69.8%</b>	<b>41.3%</b>	<b>0.001</b>	65.1%	<b>0.31</b>	92.3%	45.9%	63

\*Improved water sources: Piped, Ghana Water, public tap/standpipe, protected well/tube well/borehole, rainwater, bottled

Note: No verification done on the phone; Sample size for verification metrics are based on a paired sample across phone survey and verification responses.

- Availability of guidelines/protocols for cleaning was significantly different between the phone survey and the enumerator verification ( $p < 0.05$ ), and had minimal PABAK. This difference was not statistically significant when comparing phone vs in-person self-reported values.
- The respondents over-reported this item when self-reported in-person (61.3%) as compared to the enumerator verification (41.3%).

# INFRASTRUCTURE: SANITATION & WASTE DISPOSAL

## (with enumerator verification)

Validation metrics color key	
Very Strong (>0.9)	
Strong (0.8-0.9)	
Moderate (0.6-0.79)	
Weak (0.4-0.59)	
Minimal (<0.4)	

Mapping	No Indicator	Reported vs Verified	Survey estimates			Precision		Accuracy		N	
			Phone	In-person	p value	% Agreement	PABAK	Sensitivity	Specificity		
direct	1	Facility has an improved toilet within the facility used by staff*	Self-reported	74.6%	73.0%	0.839	92.1%	0.87	95.7%	82.4%	63
			Verified	74.6%	76.2%	0.836	95.2%	0.924	95.8%	93.3%	63
direct	2	Facility has at least one improved, functional, private and accessible toilet that can be locked designated for health facility staff only (among facilities that reported having an improved toilet)	Self-reported	81.8%	81.8%	1.000	86.4%	0.789	91.7%	62.5%	44
			Verified	81.8%	86.4%	0.560	81.8%	0.729	86.8%	50.0%	44
direct	3	Facility has at least one improved, functional, private and accessible toilet that can be locked designated for females only (among facilities that reported having an improved toilet for females)	Self-reported	71.4%	90.5%	0.116	81.0%	0.621	78.9%	100.0%	21
			Verified	71.4%	85.7%	0.259	76.2%	0.527	77.8%	66.7%	21
direct	4	Facility has at least one improved, functional, private and accessible toilet that can be locked designated for males only (among facilities that reported having an improved toilet for females)	Self-reported	71.4%	76.2%	0.726	66.7%	0.338	75.0%	40.0%	21
direct	5	Facility as at least one improved, functional, private and accessible toilet that can be locked designated for both males and females (among facilities that reported having an improved toilet for females)	Self-reported	61.9%	71.4%	0.513	71.4%	0.469	73.3%	66.7%	21
direct	6	Bin with lid available for disposal of used menstrual hygiene products in females only toilet (among facilities that reporting having a female-only toilet)	Self-reported	78.6%	42.9%	0.053	64.3%	0.334	100.0%	37.5%	14
			Verified	78.9%	31.6%	0.003	52.6%	0.173	100.0%	30.8%	19
direct	7	Handwashing facility with running water and soap, or alcohol-based hand sanitizer available	Self-reported	95.2%	93.7%	0.697	98.4%	0.982	100.0%	75.0%	63
			Verified	95.2%	90.5%	0.299	95.2%	0.945098	100.0%	50.0%	63
direct	8	Safe method used for final disposal of infectious medical waste*	Self-reported	79.4%	65.1%	0.073	66.7%	0.443	85.4%	31.8%	63
			Verified	79.4%	68.3%	0.156	66.7%	0.457	83.7%	30.0%	63

\**Improved toilet: flush to sewer/septic, VIP latrine, pit latrine with slab, composting.*

\*\* *Safe waste disposal method: burn in incinerator with incinerator functional and with fuel, removed offsite stored in covered container or protected environment, dumped without burning in covered/protected pit, open burning in pit or protected ground, or offsite collection by government agency.*

*Note: No verification on the phone ; Sample size for verification metrics are based on a paired sample across phone survey and verification responses.*

- **Most phone survey results were consistent with enumerator verification, except for the availability of a bin with a lid in female toilets where the responses were statistically significantly different (78.9% vs 31.6% respectively, p<0.05) and precision was weak/minimal.**
- **Some indicators were over-reported in-person by the respondent (#3,6,7), whereas others were under-reported (#1,2) compared to enumerator verification**

# INFRASTRUCTURE: EMERGENCY TRANSPORTATION & PRIVACY

## (with enumerator verification)

Validation metrics color key	
Very Strong (>0.9)	
Strong (0.8-0.9)	
Moderate (0.6-0.79)	
Weak (0.4-0.59)	
Minimal (<0.4)	

Mapping	No Indicator	Reported vs Verified	Survey estimates			Precision		Accuracy		N	
			Phone	In-person	p value	% Agreement	PABAK	Sensitivity	Specificity		
indirect	1	Facility has access to emergency transportation stationed at facility or that can be called for* (defined as vehicle currently parked at health facility, usually parked but in use, or not parked but available on call)	Self-reported	37.1%	21.0%	0.048	71.0%	-	69.2%	71.4%	62
			Verified	36.1%	14.8%	0.007	68.9%	-	66.7%	69.2%	61
indirect	2	Facility has access to emergency transportation stationed at facility or currently in use in working order with fuel and a driver 24 hours, or has access to a vehicle that can be called for	Self-reported	22.6%	21.0%	0.828	82.3%	-	61.5%	87.8%	62
			Verified	21.4%	7.1%	0.031	85.7%	-	100.0%	84.6%	56
direct	3	Room with auditory and visual privacy available for patient consultation	Self-reported	61.9%	58.7%	0.716	71.4%	0.452	78.4%	61.5%	63
			Verified	61.9%	57.1%	0.586	73.0%	0.479	80.6%	63.0%	63
direct	4	Facility received new infrastructure in past quarter	Self-reported	6.3%	1.6%	0.171	95.2%	0.948	100.0%	95.2%	63
			Verified	6.3%	3.2%	0.403	90.5%	0.895	0.0%	93.4%	63
direct	5	Facility received new equipment in past quarter	Self-reported	17.5%	23.8%	0.379	77.8%	0.670	40.0%	89.6%	63
			Verified	17.5%	25.4%	0.278	76.2%	0.641	37.5%	89.4%	63
direct	6	Facility received additional training for health care providers in past quarter	Self-reported	73.0%	71.4%	0.842	66.7%	0.443	77.8%	38.9%	63
			Verified	73.0%	52.4%	0.017	63.5%	0.314	84.8%	40.0%	63

- Access to emergency transportation was generally over-reported by the respondent in-person, when compared to enumerator verification, with further divergences with the phone survey results (p<0.05).
- Results from the self-reported and verification were consistent for an available room with auditory and visual privacy.
- Receipt of health care provider training was also statistically significantly different between the phone survey and enumerator verification, and was over-reported by the respondent in-person (71.4% self-report vs 52.4% verified in-person)

## BREASTFEEDING READINESS (with enumerator verification)

Validation metrics color key	
Very Strong (>0.9)	
Strong (0.8-0.9)	
Moderate (0.6-0.79)	
Weak (0.4-0.59)	
Minimal (<0.4)	

Mapping	No	Indicator	Reported vs Verified	Survey estimates			Precision		Accuracy		N
				Phone	In-person*	p value	% Agreement	PABAK	Sensitivity	Specificity	
Direct	1	10 Steps of Breastfeeding posted in service delivery points	Self-reported	52.4%	42.9%	0.285	61.9%	0.24	66.7%	58.3%	63
			Verified		38.1%	0.107	66.7%	0.34	75.0%	61.5%	63
Direct	2	Dedicated room for breastfeeding support to mothers	Self-reported	14.3%	19.0%	0.473	79.4%	0.71	33.3%	90.2%	63
			Verified		17.5%	0.626	81.0%	0.74	36.4%	90.4%	63
Direct	3	Minimum ANC Breastfeeding Education Checklist	Self-reported	53.2%	43.5%	0.281	64.5%	0.29	70.4%	60.0%	62
			Verified		38.7%	0.105	59.7%	0.20	66.7%	55.3%	62
Direct	4	Protocol/guideline for supporting breastfeeding in service delivery points	Self-reported	49.2%	47.5%	0.856	68.9%	0.38	69.0%	68.8%	61
			Verified		41.0%	0.363	68.9%	0.38	72.0%	66.7%	61

*Note: Enumerator verification involved verifying the following in-person: hat the 10 Steps of breastfeeding were posted in various service delivery points, verifying the dedicated room for breastfeeding support, verifying the minimum ANC breastfeeding education checklist, and verifying the protocol/guideline for supporting breastfeeding was in the service delivery points.*

*No verification was done on the phone.*

- **Enumerator verification in-person led to a lower availability of all breastfeeding items; for some indicators, this leads to stronger agreement with the phone survey (10 steps of breastfeeding, dedicated breastfeeding support room), suggesting the in-person response was over-reported. For other indicators (ANC checklist), agreement is reduced, suggesting the phone survey response was over-reported.**
- **PABAK is minimal for most indicators, and sensitivity is minimal for the availability of a breastfeeding room.**

## MEDICAL SUPPLIES: EQUIPMENT (with strict item availability definition)

Validation metrics color key	
Very Strong (>0.9)	Dark Green
Strong (0.8-0.9)	Light Green
Moderate (0.6-0.79)	Light Yellow-Green
Weak (0.4-0.59)	Yellow
Minimal (<0.4)	Red

Mapping	No	Indicator	Survey estimates			Precision		Accuracy		N
			Phone	In-person (strict)*	p value	% Agreement	PABAK	Sensitivity	Specificity	
indirect	1	Thermometer	88.9%	90.5%	0.770	95.2%	-	96.5%	83.3%	63
indirect	2	Child/adult weighing scale	93.7%	93.7%	1.000	93.7%	-	96.6%	50.0%	63
indirect	3	Height measuring board	76.2%	76.2%	1.000	87.3%	-	91.7%	73.3%	63
indirect	4	Blood pressure apparatus	93.7%	90.5%	0.510	96.8%	-	100.0%	66.7%	63
indirect	5	Oxygen delivery device	36.5%	34.9%	0.853	92.1%	-	90.9%	92.7%	63
indirect	6	Speculum	66.7%	50.8%	0.070	84.1%	-	100.0%	67.7%	63
indirect	7	Equipment to collect pap smear specimen**	<b>25.4%</b>	<b>9.5%</b>	<b>0.019</b>	77.8%	-	66.7%	78.9%	63
indirect	8	Refrigerator with continuous temperature monitoring	77.8%	69.8%	0.311	82.5%	-	93.2%	57.9%	63
indirect	9	Cold box or vaccine carrier with ice pack	87.3%	82.5%	0.455	95.2%	-	100.0%	72.7%	63

\* *Strict item availability defined as equipment item available if observed and functional.*

\*\* *spatulas, endocervical brushes.*

- Using a strict item availability definition shows that the phone and in-person survey results were mostly consistent and had strong %agreement and sensitivity.
- Availability of pap smear specimen equipment diverged significantly ( $p < 0.05$ ), whereas this was not the case when we used the conservative item availability definition

## MEDICAL SUPPLIES: DIAGNOSTICS (with strict item availability definition)

Validation metrics color key	
Very Strong (>0.9)	
Strong (0.8-0.9)	
Moderate (0.6-0.79)	
Weak (0.4-0.59)	
Minimal (<0.4)	

Mapping	No	Indicator	Survey estimates			Precision		Accuracy		N
			Phone	In-person (strict)*	p value	% Agreement	PABAK	Sensitivity	Specificity	
indirect	1	Urine dipstick for protein, glucose or ketones	68.3%	61.9%	0.455	90.5%	-	97.4%	79.2%	63
indirect	2	Malaria diagnostic testing (rapid test or other)	82.5%	79.4%	0.650	90.5%	-	96.0%	69.2%	63
indirect	3	HIV rapid test	81.0%	77.8%	0.660	90.5%	-	95.9%	71.4%	63
indirect	4	Urine pregnancy test	82.5%	69.8%	0.094	71.4%	-	88.6%	31.6%	63
indirect	5	Syphilis diagnostic testing (rapid test or specific test or serology)	71.4%	65.1%	0.444	81.0%	-	90.2%	63.6%	63
indirect	6	Tuberculosis diagnostic test	12.7%	7.9%	0.380	95.2%	-	100.0%	94.8%	63
indirect	7	Hepatitis B diagnostic test	41.3%	38.1%	0.716	81.0%	-	79.2%	82.1%	63

\* *Strict item availability defined as commodity items available if observed and not expired,; equipment item available if observed.*

- **No significant difference in reported availability of diagnostic items across phone and in-person survey when using a strict item availability definition.**
- **Findings very consistent with conservative item availability definition (slide 31).**

# MEDICAL SUPPLIES: PERSONAL PROTECTIVE EQUIPMENT (with strict item availability definition)

Validation metrics color key	
Very Strong (>0.9)	Green
Strong (0.8-0.9)	Light Green
Moderate (0.6-0.79)	Light Yellow
Weak (0.4-0.59)	Yellow
Minimal (<0.4)	Red

Mapping	No	Indicator	Survey estimates			Precision		Accuracy		N
			Phone	In-person (strict)*	p value	% Agreement	PABAK	Sensitivity	Specificity	
indirect	1	Respirator/mask (e.g., N95/FFP2)	41.3%	31.7%	0.267	77.8%	-	80.0%	76.7%	63
indirect	2	Surgical or medical masks	66.7%	50.8%	0.070	61.9%	-	78.1%	45.2%	63
indirect	3	Latex or other disposable gloves	88.9%	82.5%	0.309	84.1%	-	94.2%	36.4%	63
indirect	4	Utility gloves	60.3%	44.4%	0.074	68.3%	-	82.1%	57.1%	63
indirect	5	Wellington boots	60.3%	57.1%	0.717	84.1%	-	88.9%	77.8%	63
indirect	6	Disposal gowns	53.2%	45.2%	0.369	79.0%	-	85.7%	73.5%	62
indirect	7	Head cap	41.3%	41.3%	1.000	77.8%	-	73.1%	81.1%	63
indirect	8	Protective goggles or face shield	42.9%	34.9%	0.361	76.2%	-	77.3%	75.6%	63

\* *Strict item availability defined as item available if observed and not expired.*

- **No significant difference in reported availability of PPE items across phone and in-person survey when using a strict item availability definition.**
- **Findings consistent with conservative item availability definition (slide 32).**

## MEDICAL SUPPLIES: VACCINES (with strict item availability definition)

Validation metrics color key	
Very Strong (>0.9)	
Strong (0.8-0.9)	
Moderate (0.6-0.79)	
Weak (0.4-0.59)	
Minimal (<0.4)	

Mapping	No	Indicator	Survey estimates			Precision		Accuracy		N
			Phone	In-person (strict)*	p value	% Agreement	PABAK	Sensitivity	Specificity	
indirect	1	Measles and rubella vaccine	69.8%	55.6%	0.097	82.5%	-	97.1%	64.3%	63
indirect	2	Penta (DPT + HiB + HepB) vaccine	<b>73.0%</b>	<b>55.6%</b>	<b>0.041</b>	82.5%	-	100.0%	60.7%	63
indirect	3	BCG vaccine	<b>71.4%</b>	<b>47.6%</b>	<b>0.006</b>	73.0%	-	96.7%	51.5%	63
indirect	4	Polio vaccine	<b>74.6%</b>	<b>55.6%</b>	<b>0.025</b>	81.0%	-	100.0%	57.1%	63
indirect	5	PCV (pneumococcal vaccine)	71.4%	55.6%	0.064	81.0%	-	97.1%	60.7%	63
indirect	6	Human papilloma vaccine (HPV) injection	<b>33.3%</b>	<b>11.7%</b>	<b>0.004</b>	68.3%	-	57.1%	69.8%	60
indirect	7	Yellow fever	67.7%	54.8%	0.140	80.6%	-	94.1%	64.3%	62
indirect	8	Rotavirus vaccine	<b>71.4%</b>	<b>54.0%</b>	<b>0.043</b>	79.4%	-	97.1%	58.6%	63
indirect	9	RTSS (malaria vaccine)	47.6%	36.5%	0.207	88.9%	-	100.0%	82.5%	63
indirect	10	Men A	68.3%	55.6%	0.142	77.8%	-	91.4%	60.7%	63
indirect	11	Tetanus toxoid vaccine	<b>69.8%</b>	<b>52.4%</b>	<b>0.044</b>	76.2%	-	93.9%	56.7%	63

\* *Strict item availability defined as Item available if observed and not expired.*

- The stricter definition of item availability leads to significant differences between phone and in-person surveys for Penta, BCG, Polio, HPV, Rotavirus and TT vaccines ( $p < 0.05$ );
- Agreement remains strong for most vaccines, and sensitivity remains very strong except for HPV; this is aligned with analyses using a conservative item availability definition.

## MEDICAL SUPPLIES: MEDICINES (with strict item availability definition)

Validation metrics color key	
Very Strong (>0.9)	
Strong (0.8-0.9)	
Moderate (0.6-0.79)	
Weak (0.4-0.59)	
Minimal (<0.4)	

Mapping	No	Indicator	Survey estimates			Precision		Accuracy		N
			Phone	In-person (strict)*	p value	% Agreement	PABAK	Sensitivity	Specificity	
indirect	1	Amoxicillin	79.4%	81.0%	0.823	85.7%	-	90.2%	66.7%	63
indirect	2	Oral Rehydration Solutions (ORS)	78.7%	72.1%	0.400	86.9%	-	95.5%	64.7%	61
indirect	3	Zinc sulphates	<b>60.3%</b>	<b>42.9%</b>	<b>0.050</b>	73.0%	-	88.9%	61.1%	63
indirect	4	Oxytocin	66.7%	63.5%	0.709	90.5%	-	95.0%	82.6%	63
indirect	5	Magnesium sulphate	55.6%	44.4%	0.212	79.4%	-	89.3%	71.4%	63
indirect	6	Folic acid	87.3%	82.5%	0.455	92.1%	-	98.1%	63.6%	63
indirect	7	Iron	82.5%	76.2%	0.379	90.5%	-	97.9%	66.7%	63
indirect	8	Chlorhexidine gel	47.6%	33.3%	0.102	73.0%	-	81.0%	69.0%	63
indirect	9	Any long-acting reversible contraception methods (IUD/implant)	84.1%	71.4%	0.086	87.3%	-	100.0%	55.6%	63
indirect	10	Any short-acting reversible contraception methods (pills/condom)	82.5%	74.6%	0.278	85.7%	-	95.7%	56.3%	63
indirect	11	Any first-line anti-hypertensive	55.6%	60.3%	0.588	76.2%	-	76.3%	76.0%	63
indirect	12	Any anti-malarials	88.9%	88.9%	1.000	96.8%	-	98.2%	85.7%	63
indirect	13	Any anti-asthmatics	40.3%	40.3%	1.000	83.9%	-	80.0%	86.5%	62

\* *Strict item availability defined as Item available if observed and not expired.*

- Applying a stricter item availability definition led to lower availability of all medicines in-person as compared to a conservative definition;
- There was a significant difference in the availability of Zinc sulphates as compared to the phone survey ( $p < 0.05$ ).
- Agreement and sensitivity remain moderate to very strong; weak specificity for long and short acting contraception methods.

# Detailed description of validation metrics



# Percentage Agreement and PABAK – Measures of Reliability

## % Percentage Agreement

What proportion of facilities gave the same answer on both surveys?

Formula:  $(a + d) \div N$

a = both Yes | d = both No

## PABAK — Bias-Adjusted Agreement

Corrects for how available the item is (prevalence) and any directional bias.

Formula:  $PABAK = 2 \times (\% \text{ Agreement}) - 1$

Range: -1 (total disagreement)  $\rightarrow$  0 (chance)  $\rightarrow$  +1 (perfect)

PABAK is particularly useful when item prevalence is very high or very low.

### The 2x2 table— building block of all validity and reliability metrics

	In-person: YES	In-person: NO	Total
Phone: YES	a ✓ Both Yes (True Positive)	b ✗ Phone Yes, Spot No (False Positive)	a+b
Phone: NO	c ✗ Phone No, Spot Yes (False Negative)	d ✓ Both No (True Negative)	c+d
Total	a+c	b+d	N

% Agreement =  $(a+d) \div N$

PABAK =  $2 \times (a+d) \div N - 1$

# Sensitivity and Specificity — Measures of Validity

We treat the in-person survey as the truth or “Gold Standard”.

Sensitivity and specificity tell us whether the phone survey can accurately identify items that were truly available and unavailable in-person.

## Sensitivity — Does Phone Identify True Positives?

Of facilities that TRULY have the item available (in-person), what proportion did phone survey correctly identify available?

Formula:  $a \div (a + c)$

Low sensitivity = under-reporting: phone says No when the truth is Yes.

## Specificity — Does Phone Identify True Negatives?

Of facilities that truly do NOT have the item, what share did phone correctly identify as No?

Formula:  $d \div (b + d)$

Low specificity = over-reporting: phone says Yes when the truth is No.

### The 2x2 table— building block of all validity and reliability metrics

	In-person: YES	In-person: NO	Total
Phone: YES	a ✓ Both Yes (True Positive)	b ✗ Phone Yes, Spot No (False Positive)	a+b
Phone: NO	c ✗ Phone No, Spot Yes (False Negative)	d ✓ Both No (True Negative)	c+d
Total	a+c	b+d	N

Sensitivity =  $a \div (a+c)$   
Specificity =  $d \div (b+d)$

# Sensitivity and Specificity — Measures of Validity (cont'd)

## How to read these two metrics together

### High Sens + High Spec

Phone accurately mirrors truth in both directions — it is valid.

### High Sens + Low Spec

Phone over-reports. Readiness is inflated — looks better than reality

### Low Sens + High Spec

Phone under-reports. Gaps (unavailable items) are caught, but strengths (available items) are missed.

### Low Sens + Low Spec

Phone disagrees in both directions — does not correctly identify gaps nor strengths. Do not use phone survey alone.