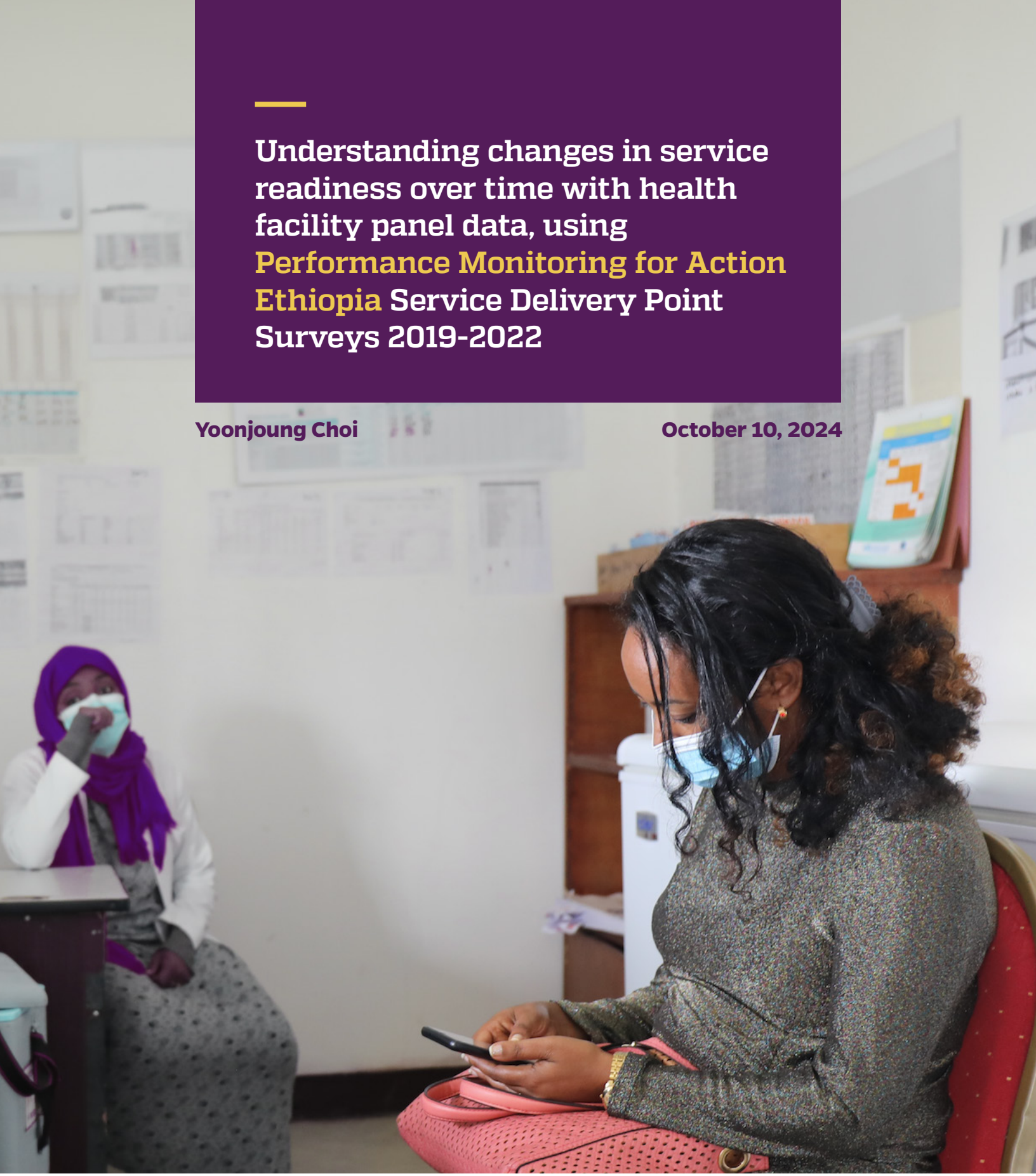
Understanding changes in service readiness over time with health facility panel data, using Performance Monitoring for Action Ethiopia Service Delivery Point Surveys 2019-2022

Yoonjoung Choi

October 10, 2024



JOHNS HOPKINS
BLOOMBERG SCHOOL
of PUBLIC HEALTH

PMA
PERFORMANCE MONITORING FOR ACTION

Suggested Citation:

Choi, Y., Desta, S., Silberg, C., Yihdego, M., Seme, A., Shiferaw,S., Zimmerman, L.A. (2024) Understanding changes in service readiness over time with health facility panel data, using Performance Monitoring for Action Ethiopia Service Delivery Point Surveys 2019-2022. Johns Hopkins Bloomberg School of Public Health and Addis Ababa University School of Public Health. Baltimore, MD and Addis Ababa, Ethiopia.

Table of Contents

List of Tables

List of Figures

List of Annex Tables

List of Annex Figures

ABSTRACT

With increasing emphasis on monitoring to improve the quality of care, it has become increasingly important to better understand the level of changes in measures of the quality of care. We assessed the magnitude of annual changes in select metrics of quality of care in Ethiopia over four years, between 2019 and 2022. We used panel data of health facilities that were constructed from PMA Ethiopia's SDP surveys. The overall readiness score for maternal and newborn health care was on average 73 out of a possible maximum of 100 and did not change over the four years. Six in ten facilities had an overall score of 70 or above, and there was no significant trend for facilities meeting the benchmark. In terms of family planning, eight out of ten facilities had a recommended range of contraceptive methods, and the level remained relatively stable over time. However, there was a declining trend in terms of a consistent three-month availability of the methods. Finally, regarding general readiness for infection prevention and control, the availability score for personal protection equipment (PPE) increased sharply between 2020 and 2021: by 36 points out of 100 for all PPE assessed in the survey and 61 points for four basic PPE items. The improvement, however, did not continue between 2021 and 2022. In conclusion, most of our measures did not change over the four years. Considering the COVID-19 pandemic, lack of deterioration may be a positive finding, reflecting the resilience of the health systems. Nevertheless, we also found elements of service readiness can change sharply even over a year as shown in the case of the PPE availability.

1. BACKGROUND

There has been an increasing emphasis on monitoring and improving the quality of care.[1] A large number of studies have examined how to measure quality of care in low-resource settings, for both process and experience of care[2–7], in addition to service readiness – the structural dimension of care.[8–12] However, recommendations on the frequency of monitoring have not been established clearly, although there are suggestions to assess it annually to use the information in annual health sector reviews.[13] Determining the optimal frequency needs to consider at least two factors. One is the quality and cost of data collection – since the quality metrics largely come from health facility assessments (HFA) which are typically outside routine health management information systems (HMIS). Another key factor is the expected changes in the measures, which have not been studied extensively. This is partially due to a lack of empirical data that can answer the question how quickly quality metrics may change. Repeated cross-sectional HFA surveys exist. Understandably, however, they do not have sufficient sample sizes to determine statistical significance in any time trends, since their objectives are to estimate levels at a time. Performance Monitoring for Action (PMA) Ethiopia's Service Delivery Points (SDP) surveys have unique longitudinal components that can be used to examine the magnitude of annual changes in the service readiness measures.

We aim to understand annual changes in select quality metrics in Ethiopia over four years, between 2019 and 2022, and to contribute to methodological recommendations for frequency of monitoring services readiness.
Three specific objectives are:

1. To assess annual variation in service readiness for maternal and newborn health (MNH) care services among facilities that offer MNH services.
2. To assess annual variation in contraceptive method availability among facilities that offer family planning (FP) services.
3. To assess annual variation in personal protection equipment (PPE) among all facilities.

2. METHODS

2.1 DATA: PMA ETHIOPIA AND SDP SURVEYS

Survey design

PMA Ethiopia is a project to generate a variety of reproductive, maternal, and newborn health (RMNH) information and knowledge from both cross-sectional and longitudinal population-based surveys. The cross-sectional surveys employ multistage stratified cluster sampling. Households are randomly selected in sampled clusters or enumeration areas (EAs), as in usual methods, but most EAs remain the same over time. The longitudinal women panel is nested in the cross-sectional surveys, and detailed sample methods are available elsewhere.[14]

The SDP surveys are another important component of PMA Ethiopia. Conducted annually, the surveys collect data on MNH and FP services as well as elements of general readiness to provide services. Questionnaires focus on inventory and service provision environment at the facility-level. The surveys are cross-sectional, and the survey sample is created each year to include SDPs (including both public and non-public health facilities, pharmacies, and drug stores) that are accessible to the population in EAs selected for the above population-based surveys – in terms of geographic proximity or health systems hierarchy in the country. For each EA, all public SDPs that serve the EA and up to 3 private facilities are selected. Because a majority of EAs remain over time, this design effectively creates a panel for many SDPs, especially higher-level health facilities such as hospitals, because their catchment areas can include multiple sampled EAs. Detailed SDP survey methods are described elsewhere.[15]

Survey implementation

Table 1 summarizes survey implementation. The SDP survey sample size was around 700-800 in 2019-2021 and 505 in 2022 when data collection was limited to four select regions. Over the four years, a total of 1233 SDPs were assessed at least once as part of the cross-sectional SDP surveys. Of those, 781 SDPs (63%) were assessed in two or more years between 2019-2022. About 4 in 5 hospitals and 7 in 10 lower-level health facilities (i.e., health centers, health posts, and clinics) were assessed multiple times over the four years **(Table 2)**. Excluding pharmacies and drug stores (which fall outside the scope of the current study), there were 993 health facilities (i.e., hospitals, health centers, health posts, and clinics), and 691 were assessed in multiple years between 2019 and 2022 – hereinafter referred to as panel facilities.

Table 1. PMA Ethiopia SDP surveys' time of data collection and sample size by year: 2019-2022

Survey year	Data collection	Number of EAs for household/ female surveys	Number of SDPs that completed the SDP interviews	Response rate (%)
2019	September - December 2019	265	800	97.9
2020	November - December 2020	231[1]	710	96.6
2021	November 2021 - January 2022	243[2]	744	96.2
2022	October 2022 - February 2023	162[3]	505	92.0[4]

[1] Data collection in Tigray was halted in November 2020, reducing the overall sample by all 34 Ems that were previously in the Tigray region. The region was included only in 2019.

[2] In 2021, a new region, Sidama, was formed from within SNNP. In addition to 8 EAs that had been in SNNP, 12 new EAs were included to provide data for the Sidama region..

[3] In 2022, SDP survey was conducted only in select four regions: Amhara, Oromia, SNNP, and Addis.

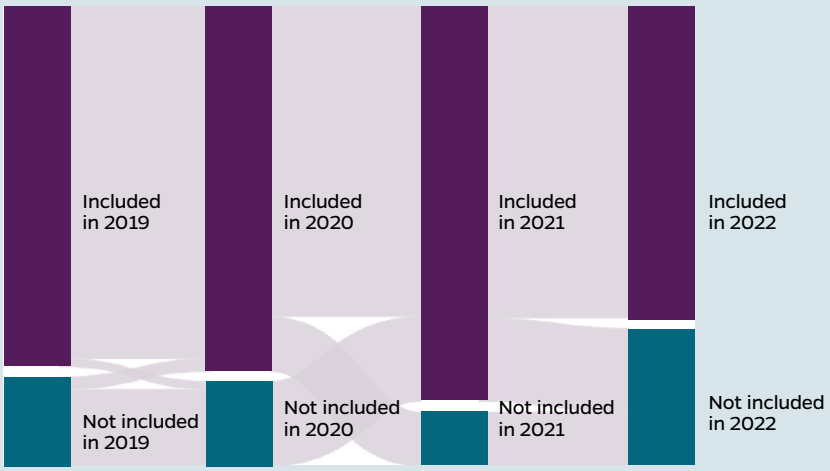
[4] In 2022, due to the ongoing security issues, a total of 23 facilities in the Oromia region could not reached, resulting a lower response rate.

Table 2. Number of SDPs that were assessed multiple times in PMA Ethiopia SDP surveys: by type

SDP type	Number of SDPs	Number of SDPs that were assessed in multiple years	Percent of SDPs that were assessed in multiple years
Hospital	189	154	81%
Health center, health post, clinic	804	537	67%
Pharmacy, drug store	240	90	38%
Total	1233	781	63%

As illustrated in Figure 1, not all panel facilities were assessed in all four years. This was due to primarily the change in regions in 2020 and 2022 (Table 1) as well as expected variation (especially among lower-level facilities) according to the SDP survey design. On average, the panel facilities were assessed 3.2 times, and there are 2206 facility-survey-year observations in total. In almost all panel facilities (99.2%), the time difference was one year between one observation and the immediate next one. Annex Table 1 shows further information on what years the panel facilities were assessed over the four years.

Figure 1.
Participation by survey year among the panel facilities (n=691)



Each bar height represents the number of facilities. The sum of two bars in a year is the total number of panel facilities. See an interactive version of the figure for further details of each flow between years.

All cross-sectional SDP survey data are available to the public at https://www.pmadata.org/countries/ethiopia. For the analysis, Addis Ababa University data management team linked SDP data sets between 2019 and 2022 and created a panel data set.

2.2 MEASURES: SERVICE READINESS

In each objective, service readiness measures were created to assess the structural domain of quality. The amount of available information in the PMA Ethiopia SDP survey varied by service (and thus objective). Also, the survey questionnaires evolved. Detailed measures are described below for each service.

2.2.1 MNH SERVICES

MNH is a priority topic in PMA Ethiopia. Therefore, the SDP surveys collected detailed information regarding MNH service readiness. We replicated an MNH service readiness measure from a recent in-depth study.[11] The study compared different items and approaches to create various summary indices for MNH service readiness, and we chose to focus on an index based on WHO technical elements. Using PMA Ethiopia 2019 SDP survey data, the study assessed levels and patterns of 79 items that are part of WHO technical elements (Column B in Table 3,) across four domains: Medicines & Commodities; Equipment, Supplies & Amenities; Staffing & Systems to Support Quality; and Performance of Signal Functions. The study excluded items with low variation and/or low discrimination in addition to items that are difficult to interpret, and created an MNH service readiness index, using 52 items for hospitals and 45 items for lower-level health facilities (Columns C-D in Table 3, Annex Table 2). In our study, we further assessed those items that could be measured in all four years of the survey. Thus, we determined to use 45 items for hospitals and 40 items for lower-level health facilities (Columns E-F in Table 3). Exploratory analyses using 2019 data showed summary scores – described shortly below – are comparable between using all items in the reference vs. using items measured across the four years (Annex Figure 1). For each item, facilities were classified into two: having the item observed (or reporting that they practice the item) vs. not. Annex Table 2 presents the list of individual items that were included in the reference index and the measure replicated in this study.

Table 3. Number of WHO technical items included in the MNH service readiness by domain: the reference and the current study

(A) Domain	(B) Number of items in WHO standards that were assessed in the reference	(C) Number of items included in the index for hospitals in the reference	(D) Number of items included in the index for lower-level facilities in the reference	(E) Number of items included in the index for hospitals in the reference and assessed consistently between 2019-2022	(F) Number of items included in the index for lower-level facilities in the reference and assessed consistently between 2019-2022
Medicines & Commodities	26	17	17	17	17
Equipment, Supplies & Amenities	25	15	11	13	10
Staffing & Systems to Support Quality	15	10	9	7	7
Performance of Signal Functions	13	10	8	8	6
Total	79	52	45	45	40

Reference: Stierman EK, Ahmed S, Shiferaw S, Zimmerman LA, Creanga AA. Measuring facility readiness to provide childbirth care: a comparison of indices using data from a health facility survey in Ethiopia. BMJ Glob Heal. 2021;6(10).

Next, summary scores were created first by domain and then overall. The number of items varied greatly by domain and subdomain (Table 4). For hospitals, it ranged from 7 in the Staffing & Systems to Support Quality domain to 17 in the Medicines & Commodities domain. For lower-level facilities, two subdomains were irrelevant, and the number of items ranged from 6 in Performance of Signal Functions to 17 in Medicines & Commodities. Previous studies examined different approaches to create summary measures for service readiness when the number of items is uneven across domains and subdomains.[8,9,11,16] It was reported different approaches – especially simple vs. weighted averages – produce fairly comparable distributions of summary scores as well as classification of facilities (i.e., relative ranking categories based on the scores).[9,11]

Our exploratory data analysis also confirmed a strong correlation in summary scores among different approaches (results not shown). Thus, we decided to use weighted averages for all scores, assuming that every domain (and subdomain within a domain) carries equal importance in providing quality MNH services. In each domain, we summed items within a subdomain, divided it by the total number of items in the subdomain, and then calculated a simple average across the subdomains. For the overall score, we again calculated a simple average across the four domains. The summary scores were multiplied by 100 to range from 0 (i.e., a facility had or practiced none of the items that were expected at the particular facility-level) to 100 (i.e., a facility had or practiced all of the items that were expected at the facility-level).

Table 4. Number of items for MNH service readiness used in the study by domain and subdomain

Domain and subdomain	Number of items applicable for hospitals		Number of items applicable for lower-level facilities	
	By domain	By subdomain	By domain	By subdomain
Medicines & Commodities	**17**		**17**	
Routine Newborn Care		4		4
Basic EmONC		10		10
Small and Sick Newborn Care		3		3
Equipment, Supplies & Amenities	**13**		**10**	
Routine Newborn Care		2		2
Basic EmONC		5		5
Small and Sick Newborn Care		3		0
Amenities – Routine Delivery		3		3
Staffing & Systems to Support Quality	**7**		**7**	
Information Systems & Quality Improvement Processes		4		4
Referral Systems		3		3
Performance of Signal Functions	**8**		**6**	
Basic EmONC		4		4
Comprehensive EmONC		2		0
Small and Sick Newborn Care		2		2

EmONC: Emergency obstetric and newborn care

Finally, to facilitate programmatic interpretation of the readiness scores, we created binary variables to classify facilities into two: facilities with a summary score at or above select thresholds vs. not. Summary scores range from 0 to 100, and we explored 60 and 70 as two potential benchmarks. We used the absolute thresholds – as opposed to relative percentiles within the study facilities. Relative percentiles are used often in research studies – for example, to understand an association between facility's relative characteristics and an outcome.8,16 Facility managers may also want to know their particular facility's standing compared to others.

However, for program managers at an aggregated level – such as regional or national-levels, ranking or categories within a group of facilities may not provide easily interpretable monitoring information. Of note, the two benchmarks were selected arbitrarily for exploratory purposes in this study. However, benchmark levels can be modified in future analyses and the target level can be determined at a level that correlates with better health utilization and outcomes. Or, attainability can be considered in the target setting, based on a baseline level or trends which will have direct programmatic use to understand levels, patterns, and trends of facilities meeting the benchmark.

The set of binary variables was created for each of the four domain-specific scores as well as the overall score. Further, an additional set of binary variables was created to identify facilities with summary scores above the benchmarks in all four domains.

2.2.2. FP SERVICES

The main data related to FP are limited to the availability of contraceptive methods. A total of 8 methods were assessed in the survey: IUD, implants, injectables, pills, emergency contraceptive pills (EC), male condoms, female condoms, and beads as a fertility awareness method. For each method, the survey asked about current availability and, if the method was reported to be available and verified by observation, stock-out in the past three months before the survey. Therefore, for each method, two types of binary measures were created: current availability (i.e., observed availability of a method, yes vs. no) and three-month availability (i.e., observed availability of a method without a history of stock-out of the method in the past three months, yes vs. no).

For a summary measure, we constructed and used two binary measures to indicate the availability of a recommended range of methods. The recommended FP service provision varies by facility type in Ethiopia,[17] and global monitoring guidelines also reflect a different number of methods that are expected to be available at varying types of health facilities.[18] Thus, based on facility types, we applied different definitions for 'a recommended range of methods': three or more types of methods for health posts and clinics, and five or more for hospitals and health centers. The binary measures were: 1) the current availability of any three/five or more types of methods and 2) the current availability of any three/five or more types of methods without stock-out history in the past three months.

2.2.3. IPC

Since the COVID-19 pandemic, there has been heightened demand to monitor infection prevention and control practices – especially the availability of personal protection equipment (PPE). Since 2020, the surveys collected data on the current availability of several PPE supplies and stock-outs in the past three months before the survey. For the analysis, we assessed 8 types of PPE: respirator mask (e.g., N95, FFP2, or equivalent), eye protection (i.e., goggles or face shield), gloves (i.e., examination gloves or surgical gloves), long-sleeved gown, waterproof apron, surgical scrubs, shoe covers, and alcohol/hand sanitizer. As with FP, for each PPE item, two types of binary measures were created: 1) current availability (i.e., observed availability of a PPE, yes vs. no) and 2) three-month availability (i.e., observed availability of a PPE without a history of stock-out of the PPE in the past three months, yes vs. no).

Then, as with MNH, we created summary scores by summing items, dividing the sum by the total number of PPE items, and then multiplying by 100 to range from 0 (i.e., having none of the PPE items assessed) to 100 (i.e., having all of the PPE items assessed). Two different lists of items were used to create the summary scores: all eight types of PPE assessed in the survey (hereinafter referred to as all PPE – although they are not the full list of PPE that may be required for advanced care) vs. a shorter list with only four types that are commonly used in outpatient care: respirator mask, eye protection, gloves, and alcohol/hand sanitizer (hereinafter referred

to as basic PPE). Unlike in MNH and FP, we did not apply different definitions per facility type. However, it is expected that higher-level facilities should have all or more types of PPE.

In addition, as with MNH metrics, we created binary variables to classify facilities into two: facilities with a summary score at or above vs. below select benchmarks (60 and 70 out of a possible total score of 100 – two different cutoffs for exploratory purposes).

Table 5 presents all summary measures used in the study by objective.

Table 5. Summary of service readiness measures by objective

Type of summary measure	Objective 1	Objective 2	Objective 3
Score measures			
Brief description of the outcome	MNH readiness scores, domain-specific and overall (a total of 45 items for hospitals, and 40 items for lower-level health facilities)	N/A	PPE availability score, all PPE and basic PPE (a total of 8 and 4 items, respectively)
Reference period	Current	N/A	Current vs. current without history of stock-out
Do definitions vary by facility type?	Yes	N/A	No
Range	0-100	N/A	0-100
Binary measures against benchmarks			
Brief description of the outcome	MNH readiness scores are at or above select thresholds: 60 and 70 out of possible maximum 100	Number of available contraceptive method types met benchmarks: 3 and more for health posts and clinics; and 5 and more for health centers and hospital	PPE availability scores are at or above select thresholds: 60 and 70 out of possible maximum 100
Reference period	Current	Current vs. three month	Current vs. three month
Do definitions vary by facility type?	Yes	Yes	No

2.3. STATISTICAL ANALYSIS

The analysis sample varied by study objective. For Objectives 1 and 2, analyses were restricted to facilities that reported offering relevant services (i.e., labor and delivery services for MNH and family planning services, respectively) in multiple years across the four survey years. For objective 3, the analysis sample included facilities that were assessed in multiple years across the three years, 2020-2022, when the survey collected information about the availability of PPE. We excluded pharmacies and drug stores from all analyses. Of the total 691 panel facilities in PMA Ethiopia SDP surveys 2019-2022 (i.e., facilities that were assessed in multiple years between 2019 and 2022), 56% and 93% were analyzed for objectives 1 and 2. For objective 3, 606 were analyzed. Table 6 summarizes further information about the analysis sample by objective.

Table 6. Summary of analysis sample by objective

Objectives	1	2	3
Analysis eligibility	Facilities that reported they offer MNH services in multiple years when they were assessed between 2019-2022	Facilities that reported they offer FP services in multiple years when they were assessed between 2019-2022	Facilities that were assessed in multiple years between 2020-2022
Number of facilities	385	645	606
Number of facility-year level observations	1280	2022	1566
Average number of observations per facility	3.3	3.1	2.6
Number of observation pairs that are immediate next assessments to each other	895	1377	960
Number of observation pairs that are immediate next assessments to each other and assessed one year apart	881	1339	960

The unit of analysis was a facility-survey-year observation in most analyses. In descriptive analyses about changes between two immediate assessments, the unit of analysis was a pair of observations between one year and the immediate next assessment. In most of the pairs, the two assessments were one year apart (last row in Table 6).

Under each objective, a set of descriptive and regression analyses was performed. A majority of our analysis was exploratory descriptive analysis to assess the overall level of the summary metrics and annual changes in them. Analysis was conducted by facility type: hospitals vs. lower-level facilities (i.e., health centres, health posts, and clinics). Then, we used bivariate and multivariate regression analyses to estimate the annual change in outcomes by year. In multivariate analyses, the estimate was adjusted for the facility type (hospitals vs. lower-level facilities), managing authority (public vs. non-public), and urban region (Addis Ababa vs. the rest). For the score outcomes, which are in fact proportions scaled to 0-100, we used a generalized linear model with a logit link, the binomial family, and clustered sandwich estimator to account for the panel data structure. For the binary outcomes, we used generalized estimation equation model with binomial distribution. P-value less than 0.05 was considered statistically significant. All statistical analyses were conducted using Stata SE 15, and R was used to create select data visualization.

Images of Empowerment

3. RESULTS

3.1. MNH SERVICE READINESS

3.1.1. Levels and changes in MNH service readiness scores

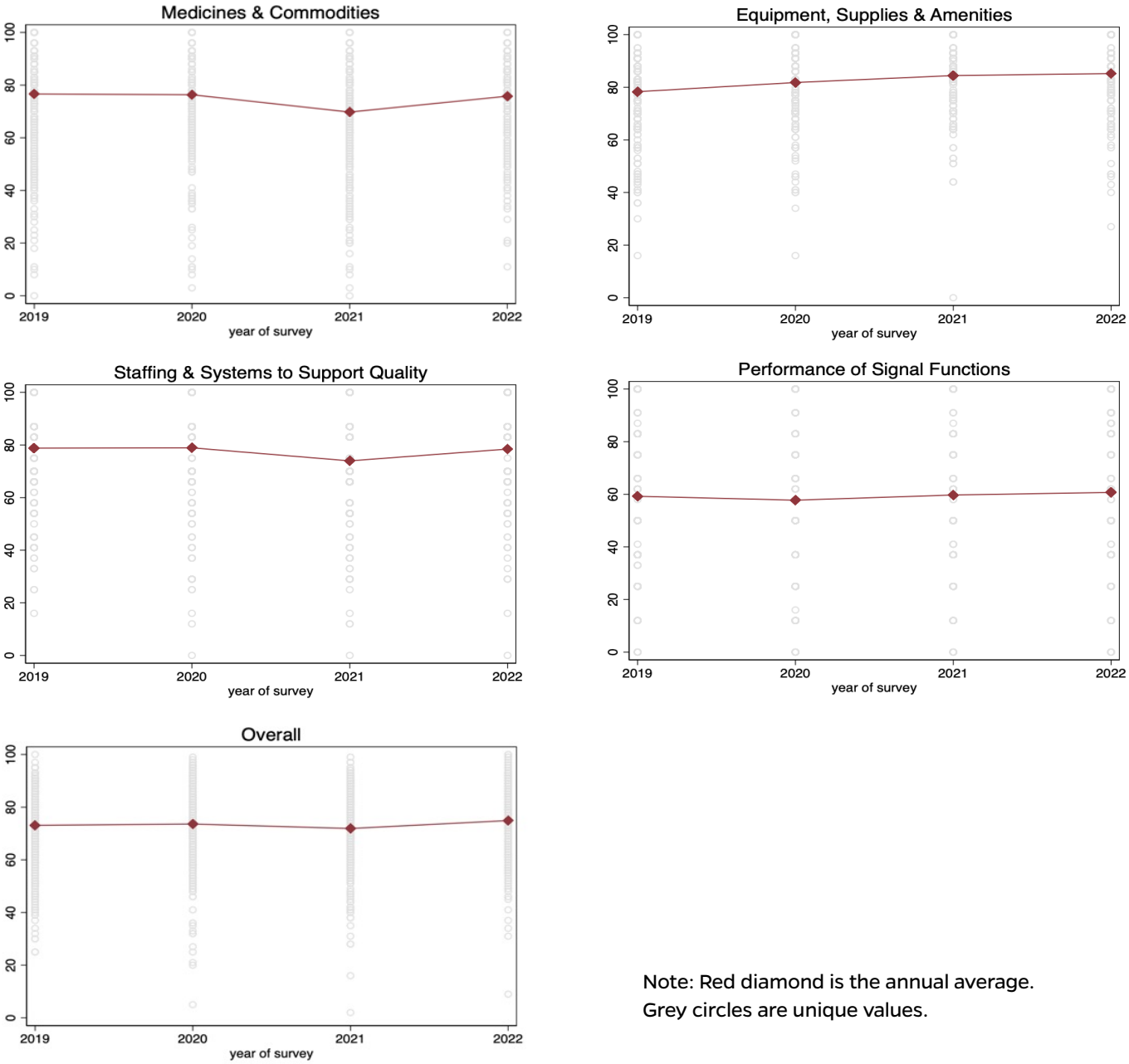
Figure 2 presents the distribution of the MNH readiness scores by domain as well as overall (i.e., the simple average across the four domains). The overall MNH services readiness score was 73 on average, out of a possible score of 100, and followed a relatively normal distribution. Domain-specific scores ranged from 59 for Performance of Signal Functions to 82 for Equipment, Supplies & Amenities. The average score was higher among hospitals – overall as well as by domain. The disparity was largest in the Performance of Signal Functions domain (Table 7). Figure 3 shows average domain-specific and overall scores by assessment year. Though slight, compared to other domains, the Equipment, Supplies & Amenities domain appeared to have a positive and steady trend.

Figure 2. Distribution of domain-specific and overall MNH service readiness scores (0-100), pooled data across years

Table 7. Average domain-specific and overall MNH service readiness scores (0-100), pooled data across years: by facility type

Domain	Lower-level facilities (n=756)	Hospitals (n=524)	Total (n=1280)
Medicines & Commodities	67	84	74
Equipment, Supplies & Amenities	78	88	82
Staffing & Systems to Support Quality	73	83	77
Performance of Signal Functions	39	87	59
Overall	64	85	73

Figure 3. Average domain-specific and overall MNH service readiness scores (0-100): by year



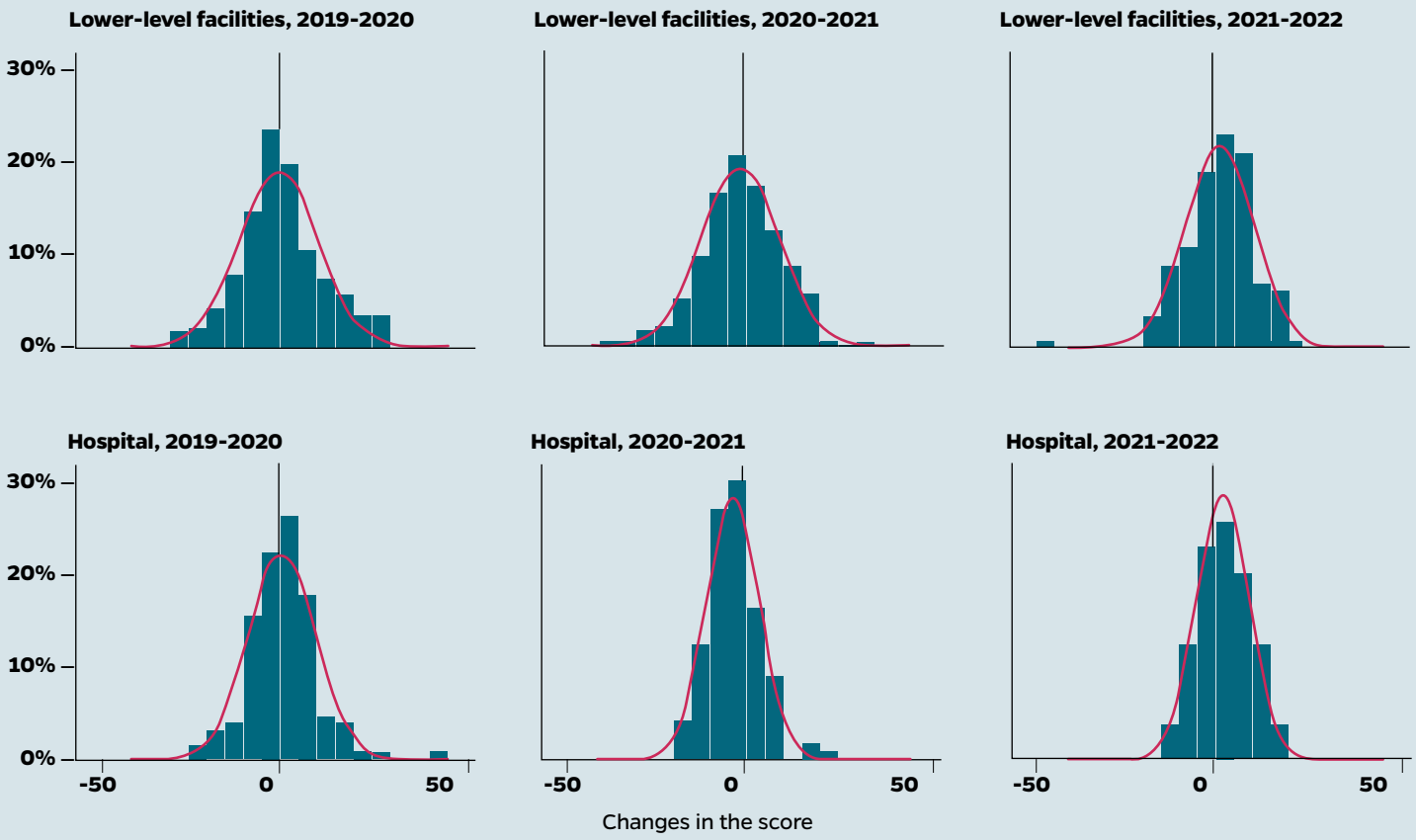
Note: Red diamond is the annual average.
Grey circles are unique values.

Between assessments within a facility, on average, the overall MNH service readiness score changed by only 0.5 points out of 100 per year. When examined by assessment years, the change distributions were roughly similar across all periods: 2019-2020, 2020-2021, and 2021-2022 (Figure 4). By domain, the change in the domain-specific score was on average small and similar across years (Table 8).

Table 8. Average absolute changes in domain-specific and overall MNH service readiness scores among observation pairs, pooled data across years: by facility type

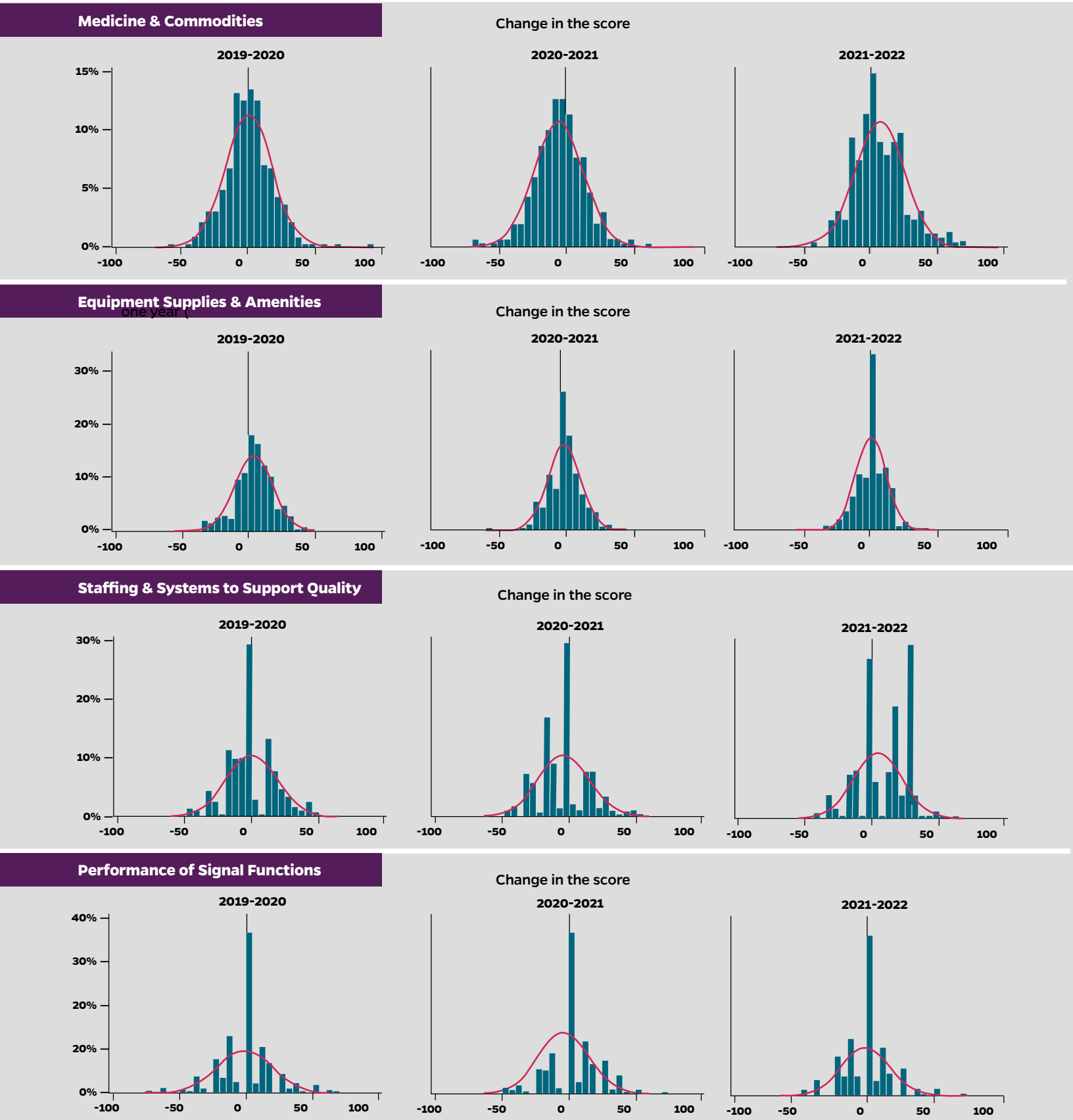
Domain	Lower-level facilities (n=524)	Hospitals (n=371)	Total (n=895)
Medicines & Commodities	0.4	-1.7	-0.5
Equipment, Supplies & Amenities	2.5	2.0	2.3
Staffing & Systems to Support Quality	-0.3	0.4	0.0
Performance of Signal Functions	-0.4	0.5	-0.1
Overall	0.5	0.3	0.5

Figure 4. Distribution of annual absolute changes in overall MNH service readiness scores among observation pairs: by facility type and assessment period



Note: Black vertical line refers to no change in the score. Red line shows normal distribution.

Figure 5. Distribution of annual absolute changes in domain-specific MNH service readiness scores among observation pairs: by assessment period



Note: Black vertical line refers to no change in the score. Red line shows normal distribution.
Figure presents only observation pairs that the time difference in assessment is one year (98% of all pairs).

Finally, we estimated the annual change in the MNH service readiness scores, using multivariate regression analyses with generalized linear model. For easier interpretation, we present marginal effects on the score in Table 9. The regression analysis results are consistent with descriptive analysis results. There was no statistically significant trend in the overall score, and the score for Medicines & Commodities declined by 1 point per year. However, the score for Equipment, Supplies and Amenities increased by 2 points each year steadily. There were no changes over time in the rest two domains: Staffing & Systems to Support Quality, and Performance of Signal Functions.

There were substantial differences in the scores by facility characteristics. In terms of the overall score, it was higher by 21 points among hospitals than among lower-level facilities. Public facilities had a higher overall score by 15 points than non-public facilities. Finally, the score was slightly higher (by 5 points) among facilities in Addis Ababa, compared to facilities outside the capital.

Table 9. Marginal effects on the MNH readiness score as a proportion among MNH panel facilities, based on multivariate regression analysis with generalized linear model

	Domain-specific				Overall
	Medicines & Commodities	Equipment, Supplies & Amenities	Staffing & Systems to Support Quality	Performance of Signal Functions	
Year	-0.01*** (-0.02 - -0.00)	0.02*** (0.02 - 0.03)	-0.01* (-0.01 - 0.00)	0.01 (-0.00 - 0.02)	0.00 (-0.00 - 0.01)
Facility type					
Lower-level facilities (reference)					
Hospital	0.16*** (0.14 - 0.19)	0.10*** (0.08 - 0.12)	0.10*** (0.07 - 0.13)	0.48*** (0.45 - 0.50)	0.21*** (0.19 - 0.23)
Managing authority					
Non-public (reference)					
Public	0.23*** (0.10 - 0.37)	0.01 (-0.05 - 0.07)	0.22*** (0.09 - 0.34)	0.15** (0.02—0.27)	0.15*** (0.05 - 0.24)
Region					
Regions outside Addis Ababa (reference)					
Addis Ababa	0.08*** (0.05 - 0.12)	0.08*** (0.07 - 0.10)	0.02 (-0.03 - 0.06)	0.04** (0.00 - 0.08)	0.05*** (0.04 - 0.07)
Observations	1,280	1,280	1,280	1,280	1,280

*** p<0.01, ** p<0.05, * p<0.1

3.1.2. LEVELS AND CHANGES IN MNH SERVICE READINESS AGAINST BENCHMARKS

Table 10 presents MNH service readiness levels against two different benchmarks that were selected for exploratory purposes, 70 and 60 out of 100. About six in ten observations had an overall score 70 or above out of 100. The level was much higher among hospitals (96%) than lower-level facilities (40%). With a benchmark score of 60, about eight in ten met the target level, and nearly all hospitals met the benchmark.

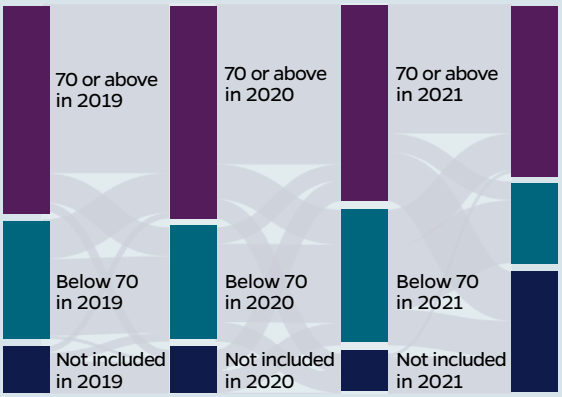
However, it was much less common for facilities to meet the benchmark in all four domains (29%, overall). Only 65% of hospitals and only 3% of lower-level facilities met that criterion. With a lower benchmark of 60, 82% of hospitals had or exceeded the target in all four domain-specific scores, compared to 9% among lower-level facilities.

Table 10. Percent of observations with MNH service readiness scores above a benchmark, pooled data across years: by facility type

	Lower-level facilities (n=756)	Hospitals (n=524)	Total (n=1280)
MNH readiness score 70 or above out of 100			
Medicines & Commodities	54	86	67
Equipment, Supplies & Amenities	79	96	86
Staffing & Systems to Support Quality	62	86	72
Performance of Signal Functions	7	86	40
Overall	40	96	63
In all domains	3	65	29
MNH readiness score 60 or above out of 100			
Medicines & Commodities	67	95	79
Equipment, Supplies & Amenities	90	99	94
Staffing & Systems to Support Quality	74	90	81
Performance of Signal Functions	14	94	47
Overall	71	98	82
In all domains	9	82	39

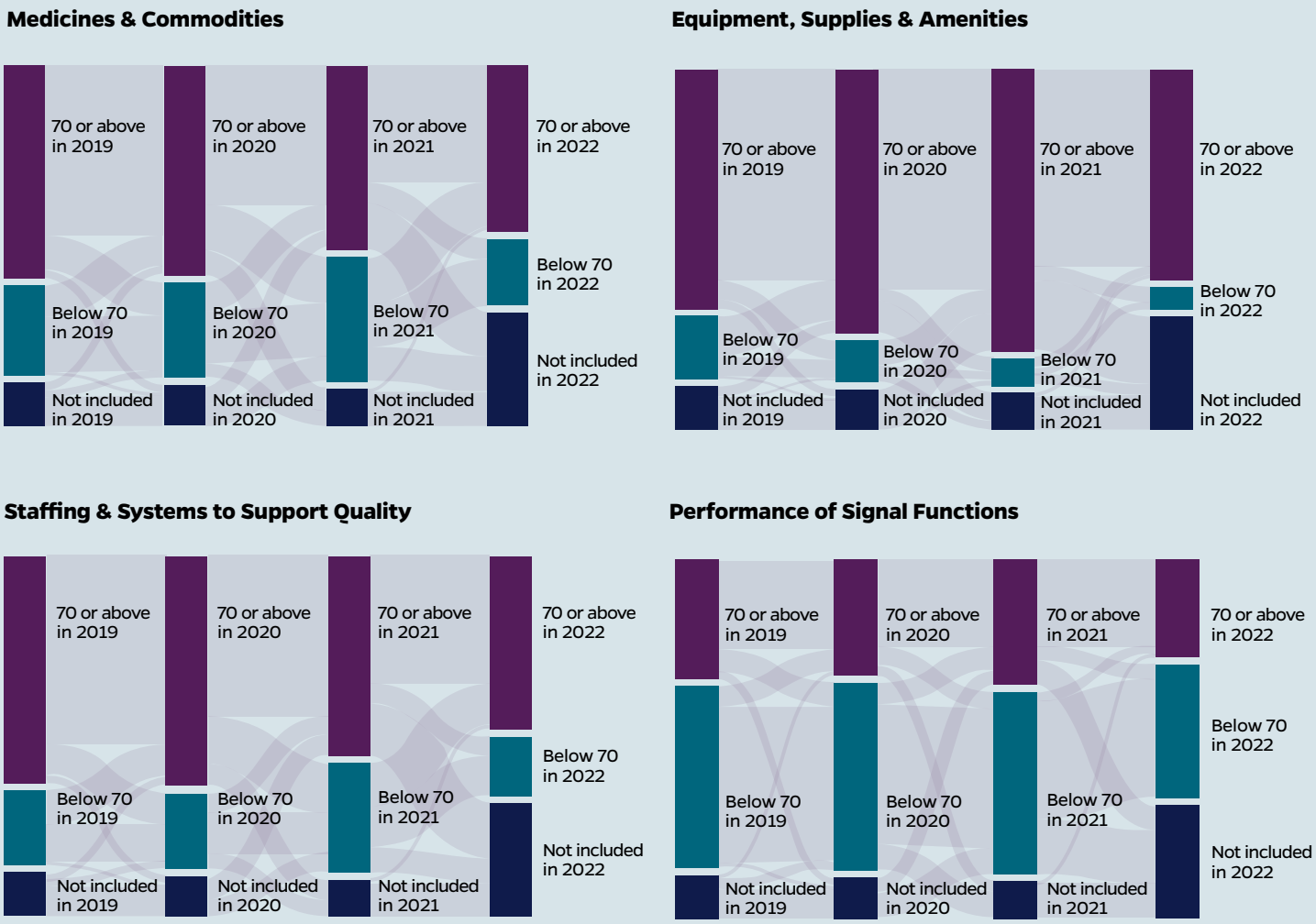
Figure 6 and Figure 7 show changes in the binary measures – overall and domain-specific, respectively – among all facilities that are included in Objective 1. Purple bars represent facilities with an MNH readiness score 70 or above. The sum of the three purple, teal, and blue bars in each year is the total number of panel facilities included in the analysis (n=385). For the overall MNH readiness score, a majority of facilities tended to stay in their group over time (Figure 6). For example, 85% of facilities that had the overall score 70 or above in 2019 remained in the same group in 2020 (175/175+30); 15% had a score 70 or below in 2020 (30/175+30), moving from the green bar in 2019 to the orange bar in 2020. See an interactive version of the figures for further details of each flow between years.

Figure 6. Changes in overall MNH readiness score against a benchmark of 70 out of 100 by year among the panel facilities for MNH services (n=385 facilities)



Each bar height represents the number of facilities. The sum of all bars in a year is the total number of panel facilities.

Figure 7. Changes in domain-specific MNH readiness score against a benchmark of 70 out of 100 by year among the panel facilities for MNH services (n=385 facilities)



Each bar height represents the number of facilities. The sum of all bars in a year is the total number of panel facilities.

We estimated the annual change in having the overall MNH score at or above benchmarks, using multivariate regression analyses with generalized estimation equation model. Table 11 presents marginal effects on the percent of MNH panel facilities with the overall score at or above benchmarks. There was no significant trends with either threshold, consistent with descriptive analysis results.

Like the results on the summary scores (Table 9), however, there were substantial differences in meeting the benchmark by facility characteristics. Using 70 or above as a target, the percent of hospitals meeting the benchmark was higher by 58% points, compared to that among lower-level facilities. It was higher by 46% points, among public facilities than non-public facilities. A higher percent of facilities in Addis Ababa also met the benchmark, compared to facilities in regions outside the capital. When using the lower benchmark of 60 or above, the gap by facility type and by region was reduced by about half, except when comparing public and non-public facilities which stayed roughly similar.

Table 11. Marginal effects on the proportion of MNH panel facilities with their MNH service readiness score at or above a benchmark, based on multivariate regression analysis with generalized estimation equation

Model

	70 or above out of 100	60 or above out of 100
Year	0.00	0.01
	(-0.02 - 0.02)	(-0.00 - 0.02)
Facility type		
Lower-level facilities (reference)		
Hospitals	0.58***	0.27***
	(0.53 - 0.63)	(0.23 - 0.32)
Managing authority		
Non-public (reference)		
Public	0.46***	0.42***
	(0.21 - 0.71)	(0.16 - 0.67)
Region		
Regions outside Addis Ababa (reference)		
Addis Ababa	0.18***	0.07***
	(0.12 - 0.25)	(0.04 - 0.10)
Observations	1,280	1,280
Number of facilities	385	385

*** p<0.01, ** p<0.05, * p<0.1

3.2. CONTRACEPTIVE METHODS AVAILABILITY FOR FP SERVICES

3.2.1. Levels and changes in contraceptive methods availability against benchmarks

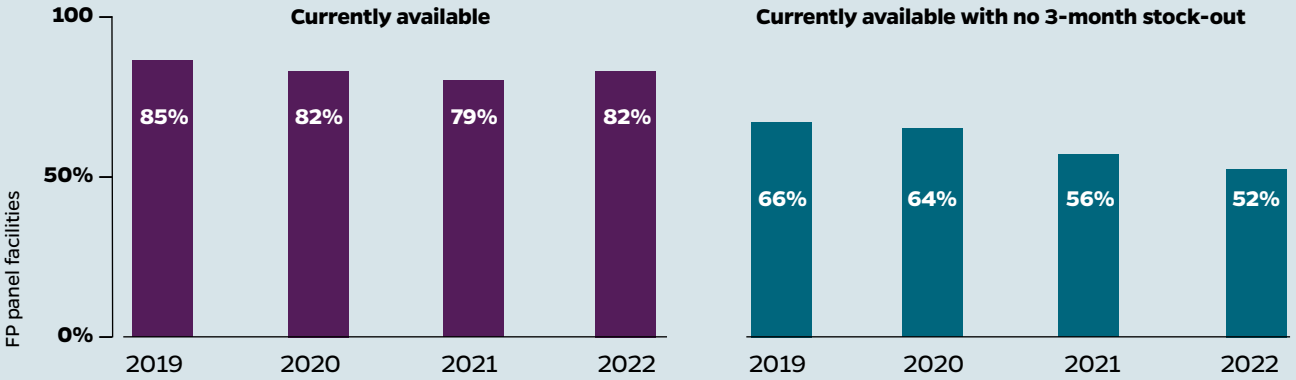
Overall, 82% of observations had a recommended range of methods (3 or more types for health posts and clinics, and 5 or more for hospitals and health centers), with the availability higher among hospitals and health centers (86%) than in health posts and clinics (75%). Availability decreased when the definition included no history of stock-out in the past three months: 62% among hospitals and health centers and 56% among health posts and clinics (Table 12). Over time, on average, the current availability remained relatively stable (left panel in Figure 8) but the current availability without stockout history decreased somewhat steadily (right panel in Figure 8).

Images of Empowerment

Table 12. Percent of observations with a recommended range of methods by reference period, pooled data across years: by facility type

	Health posts & clinics (n=764)	Hospitals & health centers (n=1258)	Total (n=2022)
Currently available	75	86	82
Currently available with no history of stock-out in the past three months	56	62	60

Figure 8. Percent of observations with a recommended range of methods (3 or more types for health posts and clinics, and 5 or more for hospitals and health centers) by reference period and year



Focusing on facility-level changes, Figure 9 shows changes in the availability among all facilities that are included for Objective 2 (n=645). Green bars represent facilities that have a recommended range of methods. As in objective 1, availability appears to be relatively stable – i.e., most facilities that met the benchmark in one year tended to do so again in the subsequent year, as long as they are assessed again in the survey. See an interactive version of the figures for further details of each flow between years.

Figure 9. Changes in the current availability of a recommended range of methods by year among the panel facilities for FP services (n=645 facilities)

We estimated the annual change in the availability, using multivariate regression analyses with a generalized estimation equation model. Table 13 presents marginal effects on the availability (i.e., the proportion of FP panel facilities with a recommended range of methods). The current availability did not change over time, but the percent of facilities without a history of 3-month stock-out decreased by 5% points per year, consistent with the descriptive analysis results (Figure 8). The current availability was higher in hospitals and health centers by 6% points than in health posts and clinics, but there was no significant difference by facility type in terms of the availability without a history of 3-month stock-out. Both availability measures were higher among public facilities than non-public, by about 10% points. By region, there was little difference in the current availability between facilities in Addis Ababa and their counterparts. The availability without a history of stock-out, however, was higher in Addis Ababa by about 10% points.

Table 13. Marginal effects on the proportion of FP panel facilities with a recommended range of methods, based on multivariate regression analysis with generalized estimation equation model

	Currently available	Currently available without history of 3-month stock-out
Year	-0.01	-0.05***
	(-0.02 - 0.00)	(-0.07 - -0.03)
Facility type		
Health posts and clinics (reference)		
Hospitals and health centers	0.06**	0.01
	(0.01 - 0.11)	(-0.05 - 0.07)
Managing authority		
Non-public (reference)		
Public	0.11***	0.13***
	(0.03 - 0.18)	(0.04 - 0.22)
Region		
Regions outside Addis Ababa (reference)		
Addis Ababa	0.01	0.13***
	(-0.06 - 0.07)	(0.05 - 0.21)
Observations	2,022	2,022
Number of facilities	645	645

*** p<0.01, ** p<0.05, * p<0.1

3.3. PPE COMMODITY AVAILABILITY FOR IPC

3.3.1. Levels and changes in PPE availability scores

PPE availability was relatively low. On average, the current availability score was 47 out of 100 for the full list of PPE and 57 for the basic list. For each set of PPE, the scores were lower by about 10 points when the definition of availability was restricted to include no history of stock-out in the past three months (Table 14). Figure 10 and Figure 11 show distributions of the four scores in the pooled data.

Table 14. Average PPE availability scores for all and basic PPE (0-100), pooled data across years: by facility type

	Health posts & clinics (n=643)	Health centers (n=530)	Hospitals (n=393)	Total (n=1566)
All PPE				
Currently available	30	51	71	47
Currently available with no history of stock-out in the past three months	24	43	62	40
Basic PPE				
Currently available	47	61	69	57
Currently available with no history of stock-out in the past three months	36	49	58	46

Figure 10. Distribution of PPE availability scores for all PPE (0-100), pooled data across years: by reference time

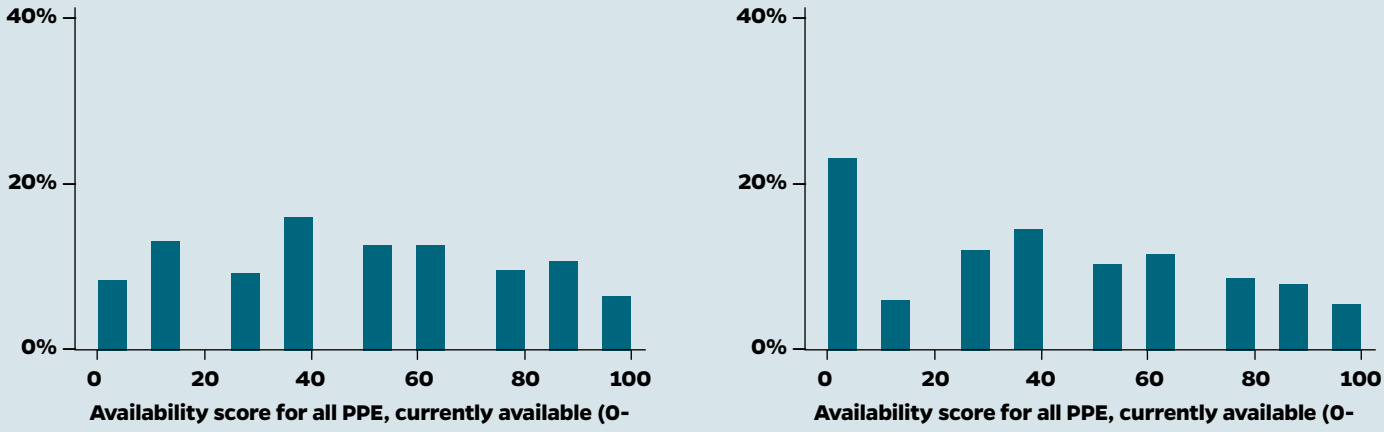
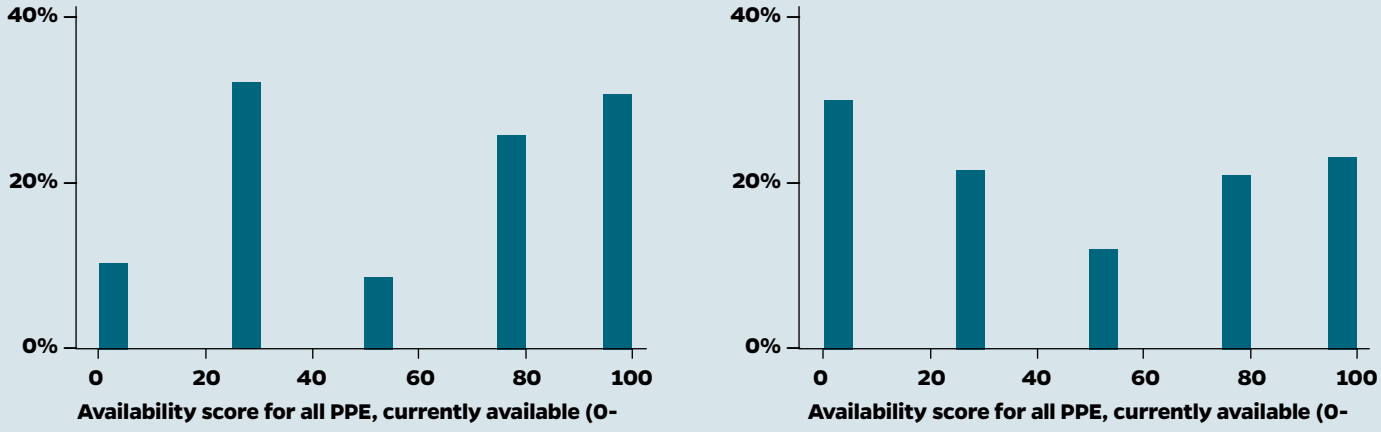
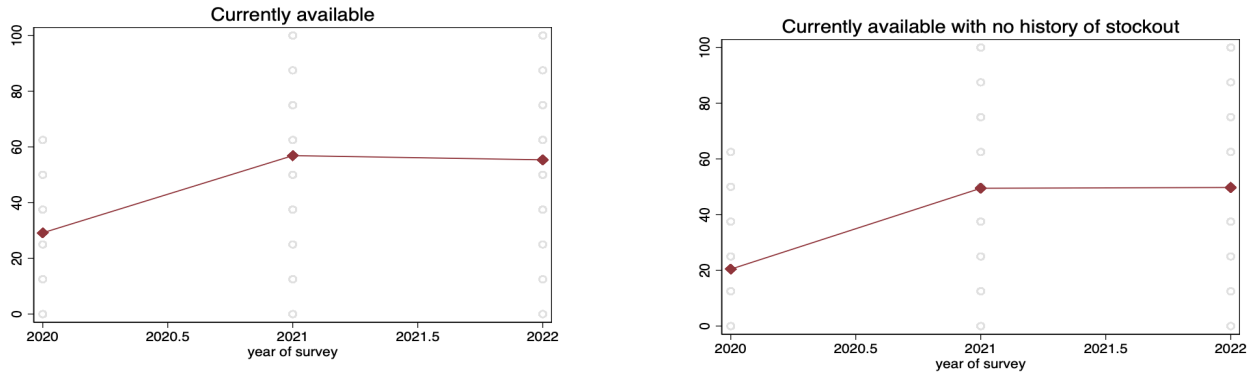


Figure 11. Distribution of PPE availability scores for basic PPE (0-100), pooled data across years: by reference time



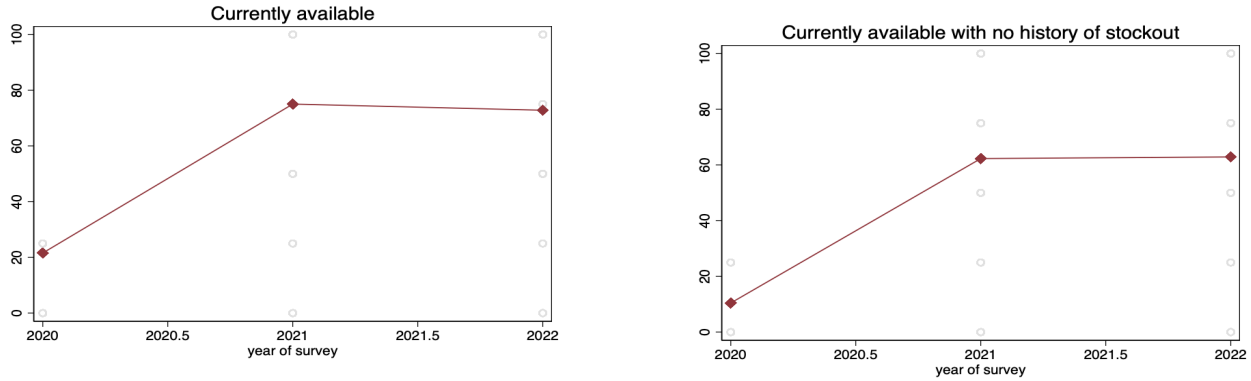
In both sets of PPE, the score improved greatly over time – especially between 2020 and 2021 (Figure 12 and Figure 13). On average among all observation pairs, the currently available score improved by 13 points for all PPE and 25 points for basic PPE (Table 15). Hospitals and health centers had much larger gains than health posts and clinics. Again, the changes were concentrated between 2020 and 2021 (Left panels in Figure 14 and Figure 15). The availability with no history of stock-out had similar distributions of the annual changes (Annex Figure 5).

Figure 12. Distribution of PPE availability scores for all PPE (0-100): by reference time and year



Note: Red diamond is the annual average. Grey circles are unique values.

Figure 13. Distribution of PPE availability scores for basic PPE (0-100): by reference time and year

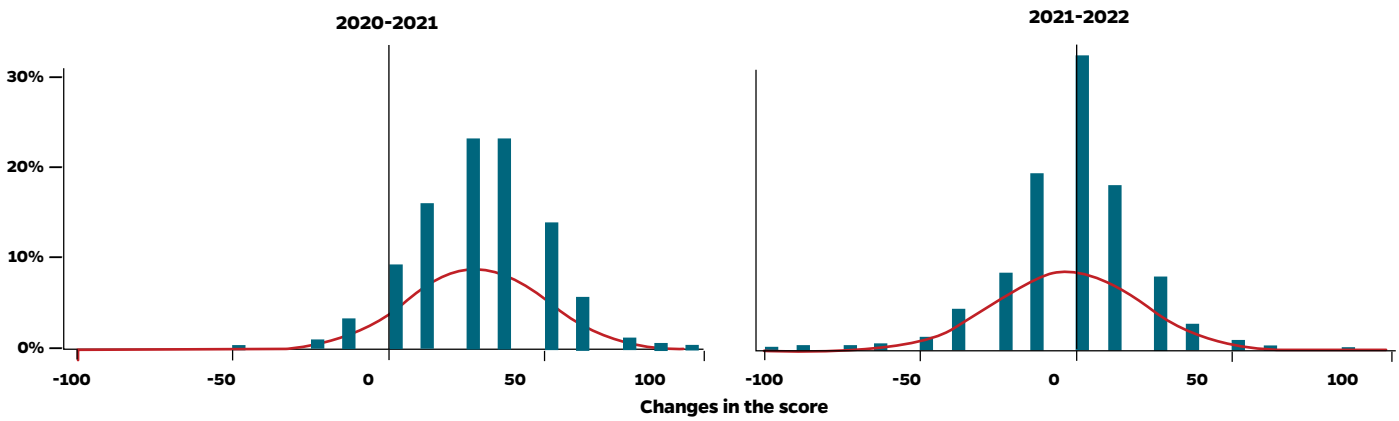


Note: Red diamond is the annual average. Grey circles are unique values.

Table 15. Average absolute changes in PPE availability scores for all and basic PPE among observation pairs, pooled data across years: by facility type

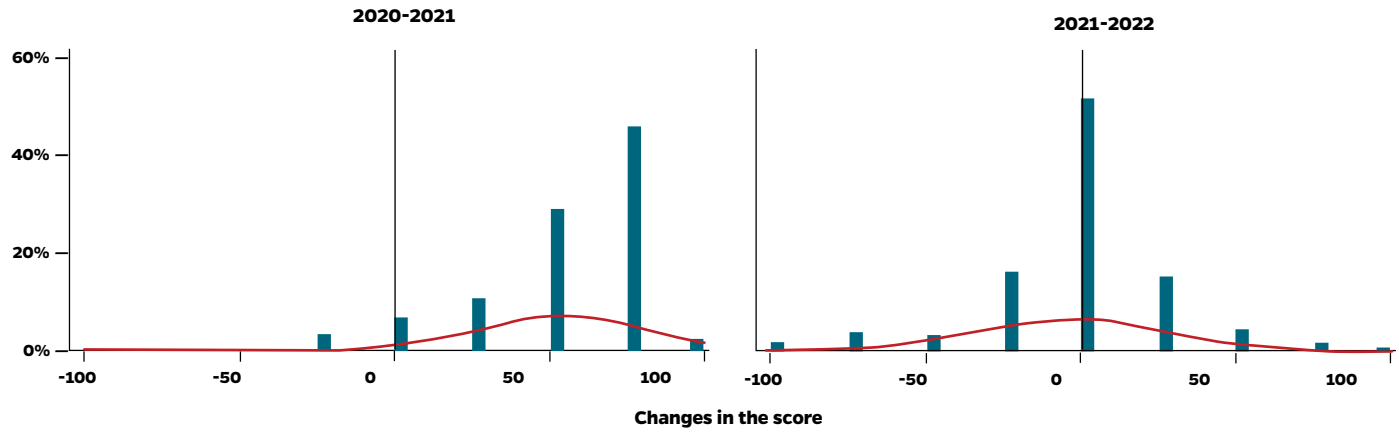
	Health posts and clinics (n=382)	Health centers (n=330)	Hospitals (n=248)	Total (n=960)
All PPE				
Currently available	5	19	18	13
Currently available with no history of stock-out in the past three months	6	20	22	15
Basic PPE				
Currently available	12	32	38	26
Currently available with no history of stock-out in the past three months	14	33	38	27

Figure 14. Distribution of annual absolute changes in availability scores for all PPE currently available: by observation period



Note: Black vertical line refers to no change in the score. Red line shows normal distribution.

Figure 15. Distribution of annual absolute changes in availability scores for basic PPE currently available: by observation period



Note: Black vertical line refers to no change in the score. Red line shows normal distribution.

Finally, we estimated the annual change in the PPE availability scores, using multivariate regression analyses with generalized linear model. Again, we present marginal effects on the score in Table 16. The current availability score increased by 36 points and 61 points between 2020 and 2021, for all and basic PPE, respectively. However, there were no further additional annual changes after 2021. A similar pattern was observed for the availability score without a history of 3-month stock-out.

In all availability measures, there was a gradual increase as the facility level became higher: health posts and clinics vs. health centers vs. hospitals. For example, the availability scores for all PPE among hospitals and among health centers were higher than the score among health posts and clinics by 51 points (confidence interval: 48 – 53 points) and 31 points (confidence interval: 28 – 34 points), respectively. The availability was higher in facilities in Addis Ababa, compared to those in regions outside the capital.

Table 16. Marginal effects on the PPE availability score as a proportion among panel facilities between 2020-2022, based on multivariate regression analysis with generalized linear model

	All PPE		Basic PPE	
	Currently available	Currently available without history of 3-month stock-out	Currently available	Currently available without history of 3-month stock-out
Year	-0.02*	-0.00	-0.04*	0.00
	(-0.05 - 0.00)	(-0.03 - 0.02)	(-0.07 - 0.00)	(-0.04 - 0.04)
Years in group				
2020				
2021-2022	0.36***	0.34***	0.61***	0.56***
	(0.33 - 0.39)	(0.31 - 0.37)	(0.57 - 0.65)	(0.52 - 0.59)
Facility type				
Health posts and clinics (reference)				
Health centers	0.31***	0.33***	0.27***	0.30***
	(0.28 - 0.34)	(0.29 - 0.37)	(0.23 - 0.31)	(0.24 - 0.35)
Hospitals	0.51***	0.53***	0.36***	0.43***
	(0.48 - 0.53)	(0.50 - 0.57)	(0.32 - 0.39)	(0.38 - 0.47)
Managing authority				
Non-public (reference)				
Public	-0.12***	-0.15***	-0.10***	-0.15***
	(-0.16 - -0.07)	(-0.21 - -0.10)	(-0.15 - -0.05)	(-0.21 - -0.09)
Region				
Non-urban regions (reference)				
Addis Ababa	0.25***	0.28***	0.18***	0.23***
	(0.21 - 0.28)	(0.23 - 0.33)	(0.14 - 0.22)	(0.16 - 0.30)
Observations	1,566	1,566	1,566	1,566

*** p<0.01, ** p<0.05, * p<0.1

3.3.2. LEVELS AND CHANGES IN PPE AVAILABILITY AGAINST BENCHMARKS

Table 17 presents PPE availability levels against two benchmarks, 70 and 60 out of a maximum score of 100. When the eight types of PPE were considered (i.e., all PPE), availability was low. About three in 10 observations had the current availability score of 70 or above out of 100. The level was much higher among hospitals (57%) than lower-level facilities: 29% among health centers and 7% among health posts and clinics. With a benchmark score of 60, overall, 40% met the benchmark, with a still significant disparity between hospitals (76%) and lower-level facilities: 45% among health centers and 13% among health posts and clinics.

When only the four types of PPE were assessed (i.e., basic PPE), the availability was higher but still remained low in terms of absolute value. About half of the observations met the benchmark of 70. The percent of hospitals meeting the benchmark was higher at 63%, but the difference across facility types declined, with 56% among health centers and 42% among health posts and clinics. Because only four items were assessed, there was no difference between the percent of the Basic PPE availability score exceeding 70 and 60 – see Figure 11.

Table 17. Percent of observations with all and basic PPE availability scores above a benchmark by reference period, pooled data across years: by facility type

	Health posts and clinics (n=643)	Health centers (n=530)	Hospitals (n=393)	Total (n=1566)
PPE availability scores 70 or above out of 100				
Eight types of PPE				
Currently available	7	29	57	27
Currently available with no history of stock-out in the past three months	6	22	47	22
Basic PPE: mask, eye protection, gloves, and alcohol hand sanitizer				
Currently available	42	56	63	52
Currently available with no history of stock-out in the past three months	31	42	52	40
PPE availability scores 60 or above out of 100				
Eight types of PPE				
Currently available	13	45	76	40
Currently available with no history of stock-out in the past three months	11	38	63	33
Basic PPE: mask, eye protection, gloves, and alcohol hand sanitizer				
Currently available	42	56	63	52
Currently available with no history of stock-out in the past three months	31	42	52	40

Figure 16 and Figure 17 show changes in the binary measures among all facilities that are included in Objective 3. They reflected the substantial facility-level changes in the PPE availability scores between 2020 and 2021 (Figure 14 and Figure 15). Green bars represent facilities with an availability score 70 or above. For both all PPE (Figure 16) and basic PPE (Figure 17), no facility met the benchmark in 2020. However, a majority of facilities met the benchmark in 2021 and the pattern continued in 2022. See an interactive version of the figures for further details of each flow between years.

Figure 16. Changes in all PPE availability score against a benchmark of 70 out of 100 by year among the panel facilities in 2020-2022 (n=606 facilities)

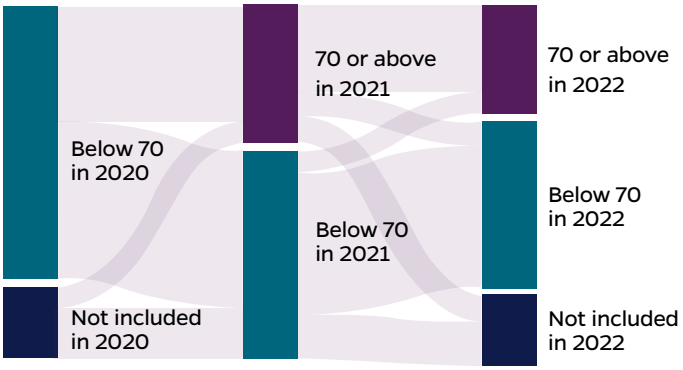
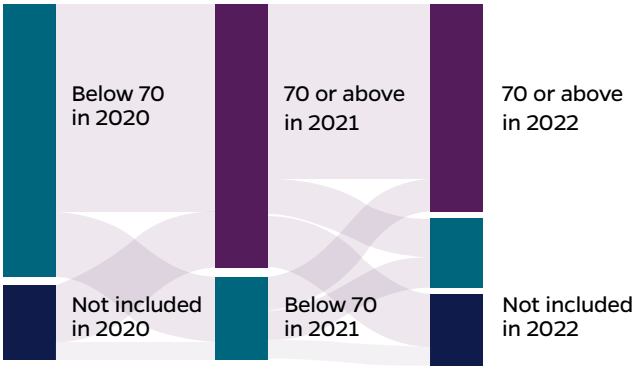


Figure 17. Changes in basic PPE availability score against a benchmark of 70 out of 100 by year among the panel facilities in 2020-2022 (n=606 facilities)



Each bar height represents the number of facilities. The sum of all bars in a year is the total number of panel facilities.

We were not able to conduct regression analyses with a dichotomous variable to distinguish year 2020 vs. years 2021-2022, due to perfect collinearity – as no facility met benchmarks in year 2020.

DISCUSSION

We assessed the magnitude of annual changes in select metrics of quality of care in Ethiopia over four years, between 2019 and 2022. Unique longitudinal components from PMA Ethiopia's SDP surveys provided a panel dataset for analysis, which is a representative sample of health facilities that are accessible to the population – geographically and administratively. Understanding the level of changes is important, as there is increasing emphasis on monitoring to improve the quality of care. Our study was motivated by a question about the optimal frequency to monitor service readiness, given limited human and financial resources in many settings. Of note, the data were collected before and throughout the COVID-19 pandemic, which posed immense challenges to the health systems throughout the country.

For MNH, the overall readiness score was on average 73 out of a possible maximum of 100 and did not change over the four years. In terms of meeting a select benchmark, six in ten facilities had an overall score of 70 or above. There was no significant trend for facilities meeting either benchmark, 70 or 60 out of 100. By domain, the score for Equipment, Supplies & Amenities Domain was relatively high – on average 82 out of 100 – and increased slightly by 2 points per year. On the other hand, the score for Medicines and Commodities was 74 on average but showed a small decline of 1 point per year. Given the large number of items included in the domain, we further explored changes in the availability of individual medicines. The results are presented by facility type in Annex Figure 6. As expected, there were more variations in trends among the 17 medicines. However, within the same group by facility type, there were relatively consistent patterns that certain medicines tend to be available more than others across years.

In terms of FP, eight out of ten facilities had a recommended range of contraceptive methods, and the level remained relatively stable over time. By individual method, the current availability also has been similar over time (Annex Figure 7). In other words, not only the number of methods but also the type of methods remained similar over the four years. However, when a history of recent stock-out was examined additionally, the level was lower: six out of ten facilities had the methods consistently for three months. Further, it declined over the four years – by 5% points per year. PMA Ethiopia relied on self-reported recall for the history of stock-out, while the current availability was verified by observation. Therefore, there might have been reporting errors, although we do not have reasons to speculate that the errors might have increased over time.

Regarding PPE, the availability score increased sharply between 2020 and 2021: by 36 points out of 100 for all PPE assessed in the survey and 61 points for four basic PPE items. This was likely thanks to emergency responses to the COVID-19 pandemic. The improvement, however, did not continue between 2021 and 2022, even though there is still room to improve. In 2022, among health posts and clinics, about half had only two of the

Images of Empowerment

four basic PPE items (results not shown). It is unclear if the health systems can continue improving the availability of basic PPE at lower facilities as emergency resources for the pandemic decrease.

Our mixed findings can be interpreted in a few ways. First, it is rather disappointing to see no significant improvement in most of our measures over the four years. At the individual facility level, we also found that typically there was no dynamic change in their service readiness. However, considering the COVID-19 pandemic and interruptions in the supply chains, maintaining a level of service readiness may well be a positive finding and reflect the resilience of the health systems. Utilization of key health services remained relatively stable in Ethiopia at the national level,[19] although regional disparities were reported.[20] At the same time, we found that elements of service readiness can change sharply even over a year – availability of PPE, specifically. This was most likely due to domestic and international emergency resources to respond to the global pandemic but does show that with concerted effort, significant change over a short period is possible.

Determining optimal frequency for routine monitoring assumes that there are: ongoing and/or emergency interventions (e.g., programs to procure and distribute PPE during the pandemic); and, therefore, reasons to believe the quality of care improves (e.g., increase in the PPE availability between 2020 and 2021) or remains at an optimal level. Unless there are programmatic investments that are expected to create rapid and/or substantial changes over time, annual – and even more frequent – routine monitoring of service readiness may not necessarily be efficient. A potentially more pragmatic way to assess quality is to strengthen capacity for rapid assessment, without necessarily recommending and committing to frequent routine monitoring.

In terms of methods, we explored the potential use of a benchmark. The amount of score change may or may not be meaningful or interpretable for program managers. Though the thresholds were selected arbitrarily in the study, we demonstrated how using an absolute benchmark can be useful for the audience. In addition, we also had mixed findings in terms of the impact of including a history of stock-out on trends of availability measures. For contraceptive availability, trends of current availability were stable, whereas there were decreasing trends of consistent availability without a history of stock-out. Trends of PPE availability, however, did not differ by the reference period. It is a data collection burden to ask about the history of stock-out for individual items in a survey. Further evidence would be helpful to determine if the value of the stock-out data outweighs the cost.

In conclusion, we examined the magnitude of annual changes in select metrics of quality of care in Ethiopia over four years, between 2019 and 2022. PMA Ethiopia's SDP surveys provided unique panel data. Most of our measures did not change significantly or meaningfully over the four years. Considering the COVID-19 pandemic, lack of deterioration may be a positive finding. Nevertheless, we also found elements of service readiness can change sharply even over a year as shown in the case of the PPE availability.

REFERENCES

1. World Health Organization; *The Network for Improving Quality of Care for Maternal, Newborn and Child Health: Evolution, Implementation and Progress. 2017-2020 Report*.; 2021. https://iris.who.int/bitstream/handle/10665/343370/9789240023741-eng.pdf?sequence=1

2. Afulani PA, Diamond-Smith N, Golub G, Sudhinaraset M. Development of a tool to measure person-centered maternity care in developing settings: Validation in a rural and urban Kenyan population. *Reprod Health*. 2017;14(1). doi:10.1186/s12978-017-0381-7

3. Sudhinaraset M, Afulani PA, Diamond-Smith N, Golub G, Srivastava A. Development of a Person-Centered Family Planning Scale in India and Kenya. *Stud Fam Plann*. 2018;49(3). doi:10.1111/sifp.12069

4. Holt K, Zavala I, Quintero X, Hessler D, Langer A. Development and Validation of the Client-Reported Quality of Contraceptive Counseling Scale to Measure Quality and Fulfillment of Rights in Family Planning Programs. *Stud Fam Plann*. 2019;50(2):137-158. doi:10.1111/sifp.12092

5. Jain A, Aruldas K, Mozumdar A, Tobey E, Acharya R. Validation of Two Quality of Care Measures: Results from a Longitudinal Study of Reversible Contraceptive Users in India. *Stud Fam Plann*. 2019;50(2). doi:10.1111/sifp.12093

6. Montagu D, Landrian A, Kumar V, et al. Patient-experience during delivery in public health facilities in Uttar Pradesh, India. *Health Policy Plan*. 2019;34(8):574-581. doi:10.1093/heapol/czz067

7. Karp C, OlaOlorun FM, Guiella G, et al. Validation and Predictive Utility of a Person-Centered Quality of Contraceptive Counseling (QCC-10) Scale in Sub-Saharan Africa: A Multicountry Study of Family Planning Clients and a New Indicator for Measuring High-Quality, Rights-Based Care. *Stud Fam Plann*. 2023;54(1). doi:10.1111/sifp.12229

8. Wang W, Do M, Hembling J, Ametepi P. Assessing the quality of care in family planning, antenatal, and sick child services at health facilities in Kenya, Namibia, and Senegal . *DHS Anal Stud No 44* . Published online 2014. http://dhsprogram.com/pubs/pdf/AS44/AS44.pdf

9. Sheffel A, Zeger S, Heidkamp R, Munos MK. Development of summary indices of antenatal care service quality in Haiti, Malawi and Tanzania. *BMJ Open*. 2019;9(12). doi:10.1136/bmjopen-2019-032558

10. Brizuela V, Leslie HH, Sharma J, Langer A, Tunçalp Ö. Measuring quality of care for all women and newborns: how do we know if we are doing it right? A review of facility assessment tools. *Lancet Glob Heal*. 2019;7(5). doi:10.1016/S2214-109X(19)30033-6

11. Stierman EK, Ahmed S, Shiferaw S, Zimmerman LA, Creanga AA. Measuring facility readiness to provide childbirth care: a comparison of indices using data from a health facility survey in Ethiopia. *BMJ Glob Heal*. 2021;6(10). doi:10.1136/bmjgh-2021-006698

12. Qian J, Shiferaw S, Seme A, et al. Data for local decision-making, not a mere reporting requirement: Development of an index to measure facility-level use of HMIS data. *J Glob Heal Reports*. 2023;7. doi:10.29392/001c.75141

13 O'Neill K, Takane M, Sheffel A, Abou-Zahr C, Boerma T. Monitoring service delivery for universal health coverage: the Service Availability and Readiness Assessment. *Bull World Health Organ*. 2013;91(12):923-931. doi:10.2471/BLT.12.116798

14. Zimmerman L, Desta S, Yihdego M, et al. Protocol for PMA-Ethiopia: A new data source for cross-sectional and longitudinal data of reproductive, maternal, and newborn health. *Gates Open Res*. 2020;4. doi:10.12688/gatesopenres.13161.1

15. Choi Y, Safi S, Nobili J, PMA2020 Principal Investigators Group, Radloff S. *Levels, Trends, and Patterns of Contraceptive Method Availability: Comparative Analyses in Eight Sub-Saharan African Countries, Performance Monitoring and Accountability 2020 Methodological Report No. 5.*; 2019.

16. Assaf S, Wang W, Mallick L. Quality of care in family planning services at health facilities in Senegal . *DHS Anal Stud No 55* . Published online 2015. http://dhsprogram.com/pubs/pdf/AS55/AS55.pdf

17. Federal Democratic Republic of Ethiopia M of H. *National Guideline for Family Planning Services In Ethiopia, 3rd Edition*.; 2020. https://scorecard.prb.org/wp-content/uploads/2022/03/National-Guideline-for-Family-Planning-Services-In-Ethiopia-2020.pdf

18. FP2030. FP2030 Measurement Framework. https://wordpress.fp2030.org/wp-content/uploads/2023/10/Framework-06-03-2022-EXTERNAL.pdf

19. Arsenault C, Gage A, Kim MK, et al. COVID-19 and resilience of healthcare systems in ten countries. *Nat Med*. 2022;28(6). doi:10.1038/s41591-022-01750-1

20. Mebratie AD, Nega A, Gage A, Mariam DH, Eshetu MK, Arsenault C. Effect of the COVID-19 pandemic on health service utilization across regions of Ethiopia: An interrupted time series analysis of health information system data from 2019–2020. *PLOS Glob Public Heal*. 2022;2(9). doi:10.1371/journal.pgph.0000843

ANNEXES

Annex Table 1. Number of SDPs and health facilities by participation in PMA Ethiopia SDP surveys

Number of survey years that the facility was assessed	Survey years that the facility was assessed	Number of SDPs	Number of health facilities[1]
Once			
	2019	219	153
	2020	142	80
	2021	90	68
	2022	1	1
Multiple times			
	2019, 2020	98	83
	2019, 2020, 2021	121	118
	2019, 2020, 2021, 2022	355	342
	2019, 2021	7	2
	2019, 2021, 2022	15	10
	2020, 2021	7	7
	2020, 2021, 2022	12	12
	2021, 2022	166	117
Total		**1233**	**993**

[1] Of the total 993 health facilities, only those that were assessed multiple times were included in analyses. Analysis sample differed by objective. See Table 6 for details.

Annex Table 2. List of WHO technical elements included in the MNH readiness reference index and assessed consistently between 2019-2022

Note: This table is adapted using Table S2 in a study by Steirman *et al.* (2021). The study describes detailed reasons for an item excluded from the reference index.

Items by domain and subdomain	(B) Items included in the reference index for hospitals	(C) Items included in the reference index for lower-level facilities	(D) Items included in the reference index for hospitals and assessed consistently between 2019-2022	(E) Items included in the reference index for lower-level facilities and assessed consistently between 2019-2022
Medicines & Commodities Domain				
Routine Delivery				
Urine dipsticks				
Routine Newborn Care				
Chlorohexidine gel in facility*	Yes	Yes	Yes	Yes
Injectable vitamin K in facility*	Yes	Yes	Yes	Yes
Tetracycline ointment in facility*	Yes	Yes	Yes	Yes
BCG vaccine	Yes	Yes	Yes	Yes
Oral polio vaccine				
Basic EmONC				
Magnesium sulphate in facility*	Yes	Yes	Yes	Yes
Calcium gluconate in facility*	Yes	Yes	Yes	Yes
At least one antihypertensive (hydralazine, nifedipine, or methyldopa)*	Yes	Yes	Yes	Yes
Injectable diazepam in facility*	Yes	Yes	Yes	Yes
Oxytocin in facility*	Yes	Yes	Yes	Yes
Misoprostol in facility*	Yes	Yes	Yes	Yes
Injectable ergometrine/methergine in facility*	Yes	Yes	Yes	Yes
Injectable ampicillin	Yes	Yes	Yes	Yes
Injectable gentamicin	Yes	Yes	Yes	Yes
Injectable metronidazole	Yes	Yes	Yes	Yes
IV solution in facility and infusion set (cannula, needle, and syringe)*	Yes	Yes	Yes	Yes
Comprehensive EmONC				
Blood transfusion available at all times facility open				
Access to blood bank either inside or outside facility				
Lignocaine/lidocaine				
Small and Sick Newborn Care				
Dexamethasone / betamethasone*	Yes	Yes	Yes	Yes
Nevirapine (NVP)*	Yes	Yes	Yes	Yes
Maternal ARV regimen				
HIV rapid test				
Benzathine benzylpenicillin (for prevention of congenital syphilis)	Yes	Yes	Yes	Yes
Syphilis testing (VDRL)				

Items by domain and subdomain	(B) Items included in the reference index for hospitals	(C) Items included in the reference index for lower-level facilities	(D) Items included in the reference index for hospitals and assessed consistently between 2019-2022	(E) Items included in the reference index for lower-level facilities and assessed consistently between 2019-2022
Equipment, Supplies & Amenities Domain				
Routine Delivery				
Thermometer				
Blood pressure apparatus				
Fetal stethoscope and/or fetal scope				
Alcohol hand scrub and/or soap and running water for staff handwashing	Yes	Yes	No	No
Sterile gloves				
Waste receptable with lid and plastic liner				
Sharps container				
Already mixed decontaminating solution (0.5% chlorine)				
Routine Newborn Care				
Sterile cord ties/clamp and scissors/blades (either in sealed delivery kit or separate)	Yes	Yes	Yes	Yes
Infant weight scale	Yes	Yes	Yes	Yes
Basic EmONC				
Obstetric forceps and/or functioning electrical vacuum extractor	Yes	Yes	Yes	Yes
Dilatation and curettage (D&C) kit and/or manual vacuum aspiration (MVA) and cannula	Yes	Yes	Yes	Yes
Functional suction apparatus for use with catheter and/or manual suction device for fluid extraction	Yes	Yes	Yes	Yes
Self-inflating bag and newborn masks (size 0 and size 1) for resuscitation	Yes	Yes	Yes	Yes
Resuscitation table/trolley with light source	Yes	Yes	Yes	Yes
Comprehensive EmONC				
Functional sterilizing equipment	Yes	No	No	No
Small and Sick Newborn Care				
Functional incubator	Yes	No	Yes	No
Oxygen supply	Yes	No	Yes	No
Pulse oximeter	Yes	No	Yes	No
Amenities – Routine Delivery				
Delivery bed				
Private delivery room or visual privacy ensured in delivery area	Yes	Yes	Yes	Yes
Newborn corner	Yes	Yes	Yes	Yes
Water outlet on site				
Electricity available at all times when facility was open in last 7 days and/or back-up energy source	Yes	Yes	Yes	Yes
Functional toilet available for patient use				

Items by domain and subdomain	(B) Items included in the reference index for hospitals	(C) Items included in the reference index for lower-level facilities	(D) Items included in the reference index for hospitals and assessed consistently between 2019-2022	(E) Items included in the reference index for lower-level facilities and assessed consistently between 2019-2022
Staffing & Systems to Support Quality Domain				
Routine Delivery Care				
Skilled birth attendant present at facility or on call 24 hours	Yes	Yes	No	No
Ratio of skilled health personnel to delivery volume meets or exceeds desired scenario in FIGO statement	Yes	Yes	No	No
Management protocols on selected obstetric topics (FMOH, 2010)				
Routine Newborn Care				
Baby-friendly guidelines				
Comprehensive EmONC				
Skilled provider available 24h to provide C-section				
At least one staff trained in anesthesia	Yes	No	No	No
Small and Sick Newborn Care				
National Comprehensive and Integrated Prevention of Mother-to-Child Transmission of HIV Guideline (FMOH, 2015)				
Information Systems & Quality Improvement Processes				
Functional mechanism for reporting data on maternal deaths to the Maternal and Perinatal Death Surveillance Response	Yes	Yes	Yes	Yes
Conducts maternal death reviews	Yes	Yes	Yes	Yes
Produces monthly reports for the HMIS and receives feedback that includes recommendations for action	Yes	Yes	Yes	Yes
Has a performance monitoring team that meets at least quarterly	Yes	Yes	Yes	Yes
Referral Systems				
Access to a functional ambulance/car on-site for emergency transportation	Yes	Yes	Yes	Yes
Access to phone or radio system at all times				
Printed referral form observed	Yes	Yes	Yes	Yes
Functional system for recording and sharing outcomes of cases referred in and out	Yes	Yes	Yes	Yes

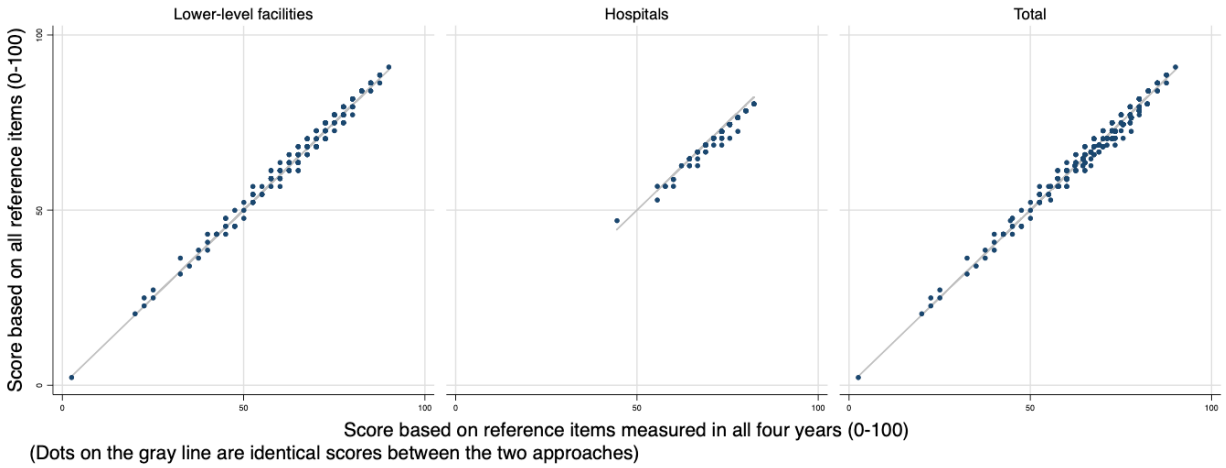
Items by domain and subdomain	(B) Items included in the reference index for hospitals	(C) Items included in the reference index for lower-level facilities	(D) Items included in the reference index for hospitals and assessed consistently between 2019-2022	(E) Items included in the reference index for lower-level facilities and assessed consistently between 2019-2022
Performance of Signal Functions Domain				
Routine Delivery				
Used partographs to monitor labor				
Routine Newborn Care				
Skin-to-skin care				
Assist mother to breastfeed				
Basic EmONC				
Provided parenteral anticonvulsants to manage high blood pressure in pregnancy	Yes	Yes	No	No
Provided parenteral or oral uterotonics to prevent or treat PPH	Yes	Yes	No	No
Performed manual removal of placenta	Yes	Yes	Yes	Yes
Provided instrument/ assisted deliveries	Yes	Yes	Yes	Yes
Provided parenteral antibiotics for infections related to pregnancy, abortion, labor or delivery	Yes	Yes	Yes	Yes
Performed neonatal resuscitation	Yes	Yes	Yes	Yes
Comprehensive EmONC				
Performed blood transfusions for maternity care	Yes	No	Yes	No
Performed cesarean deliveries	Yes	No	Yes	No
Small and Sick Newborn Care				
Provided antenatal corticosteroids for fetal lung maturation	Yes	Yes	Yes	Yes
Neonatal intensive care provided at facility	Yes	Yes	Yes	Yes

EmONC: Emergency obstetric and newborn care

* In 2020-2022, the indicated medicines were observed either in and outside the delivery room. In 2019, however, they were observed both in and outside the delivery room. For analysis, if the medicine was observed in either place, it was considered to be available in 2019.

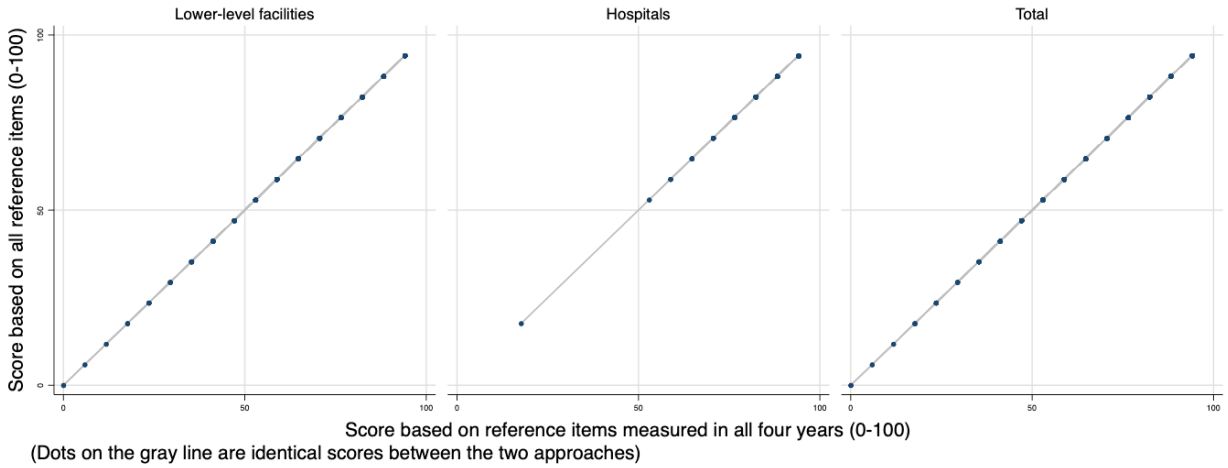
Annex Figure 1. Among facilities that were assessed for MNH services in 2019 and assessed at least once more between 2020 and 2022, scatterplot of MNH readiness scores created based on two approaches, using all reference items vs. reference items measured in all four years: lower-level facilities (n=204), hospitals (n=139), and all facilities (n=343)

Overall Score



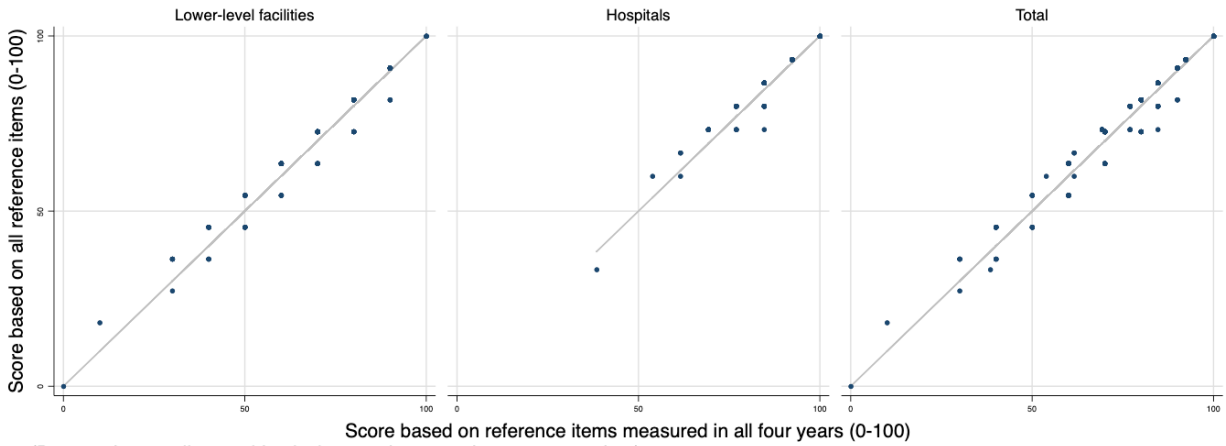
(Dots on the gray line are identical scores between the two approaches)

Domain-Specific Score: Medicines & Commodities



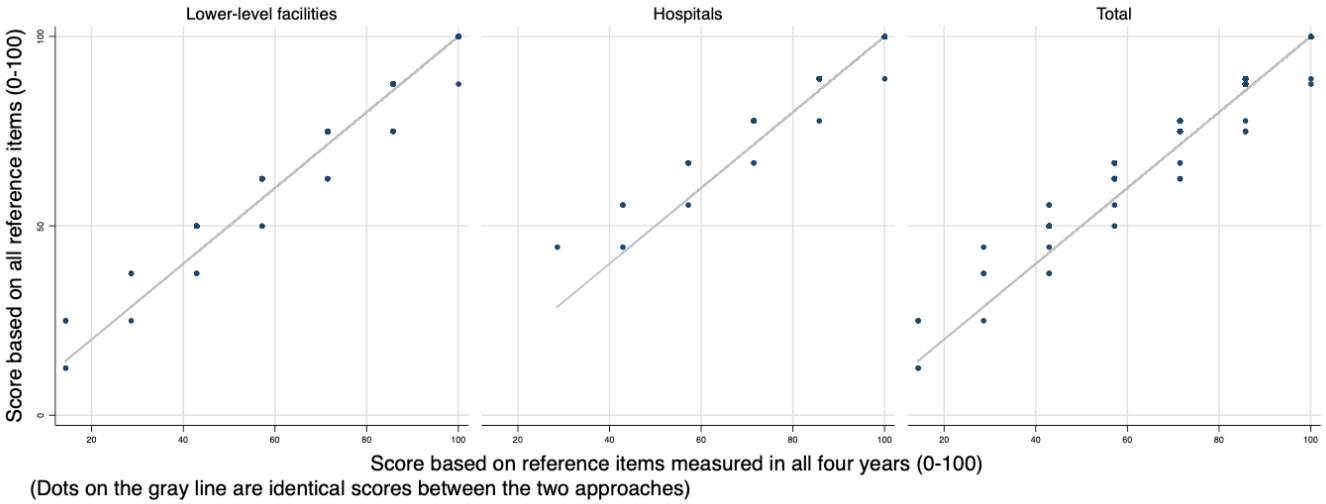
(Dots on the gray line are identical scores between the two approaches)

Domain-Specific Score: Equipment, Supplies & Amenities



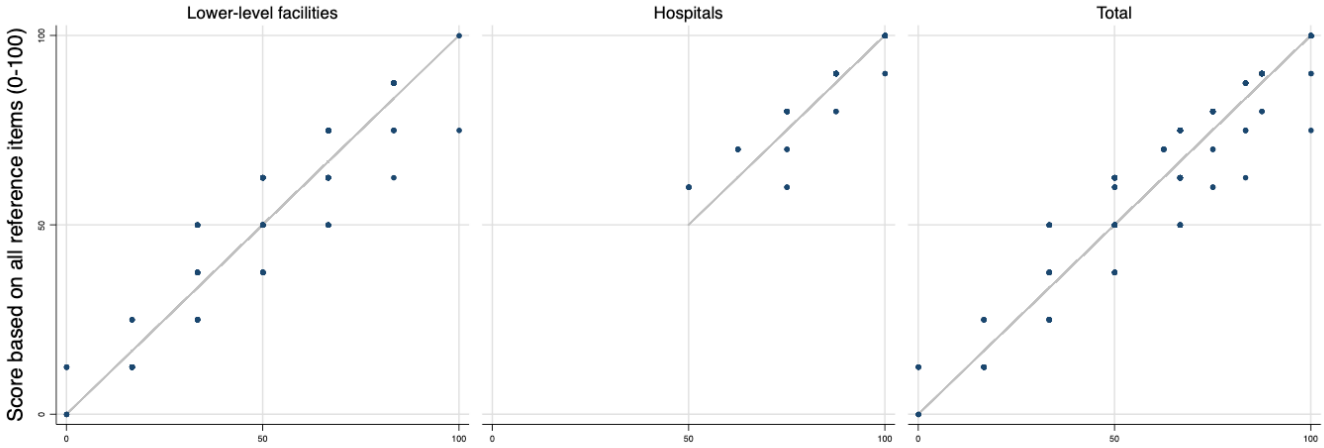
(Dots on the gray line are identical scores between the two approaches)

Domain-Specific Score: Staffing & Systems to Support Quality



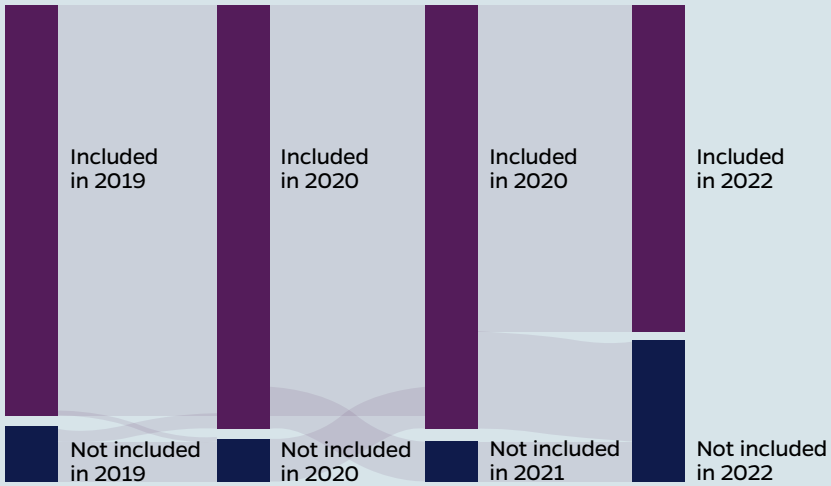
(Dots on the gray line are identical scores between the two approaches)

Domain-Specific Score: Performance of Signal Functions



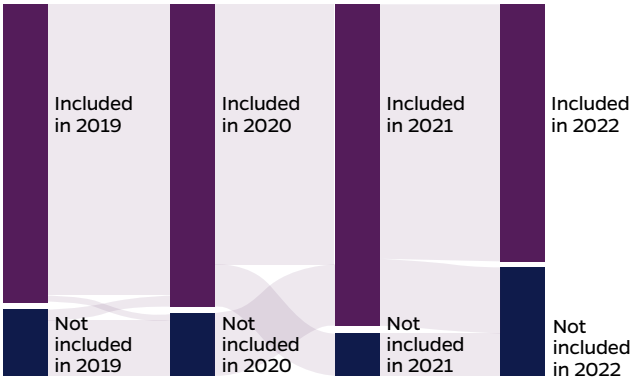
(Dots on the gray line are identical scores between the two approaches)

Annex Figure 2. Participation by survey year among the panel facilities included in analyses for objective 1 (n=385)



Each bar height represents the number of facilities. The sum of all bars in a year is the total number of panel facilities for objective 1.

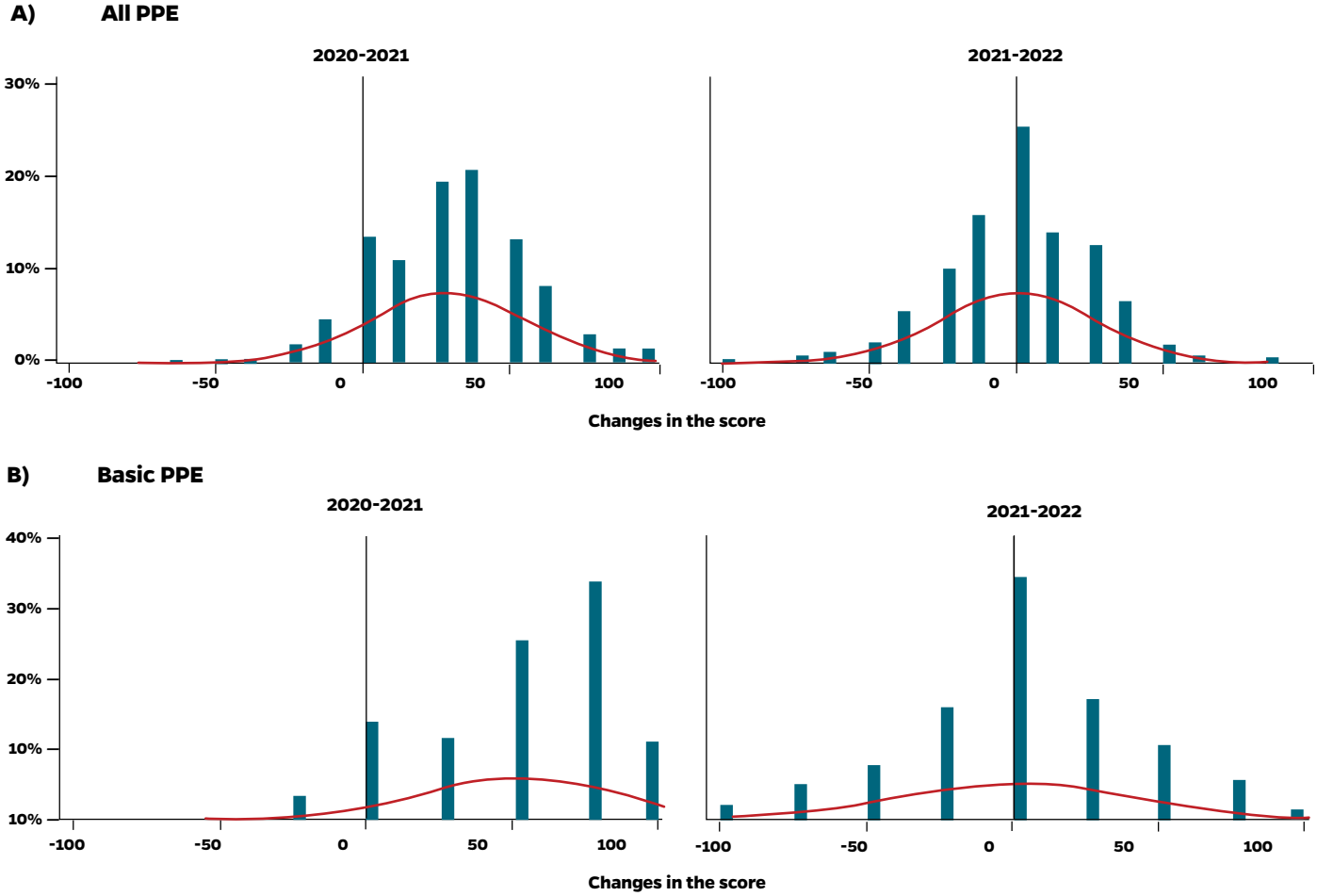
Annex Figure 3. Participation by survey year among the panel facilities included in analyses for objective 2 (n=645)



Annex Figure 4. Participation by survey year among the panel facilities included in analyses for objective 3 (n=606)

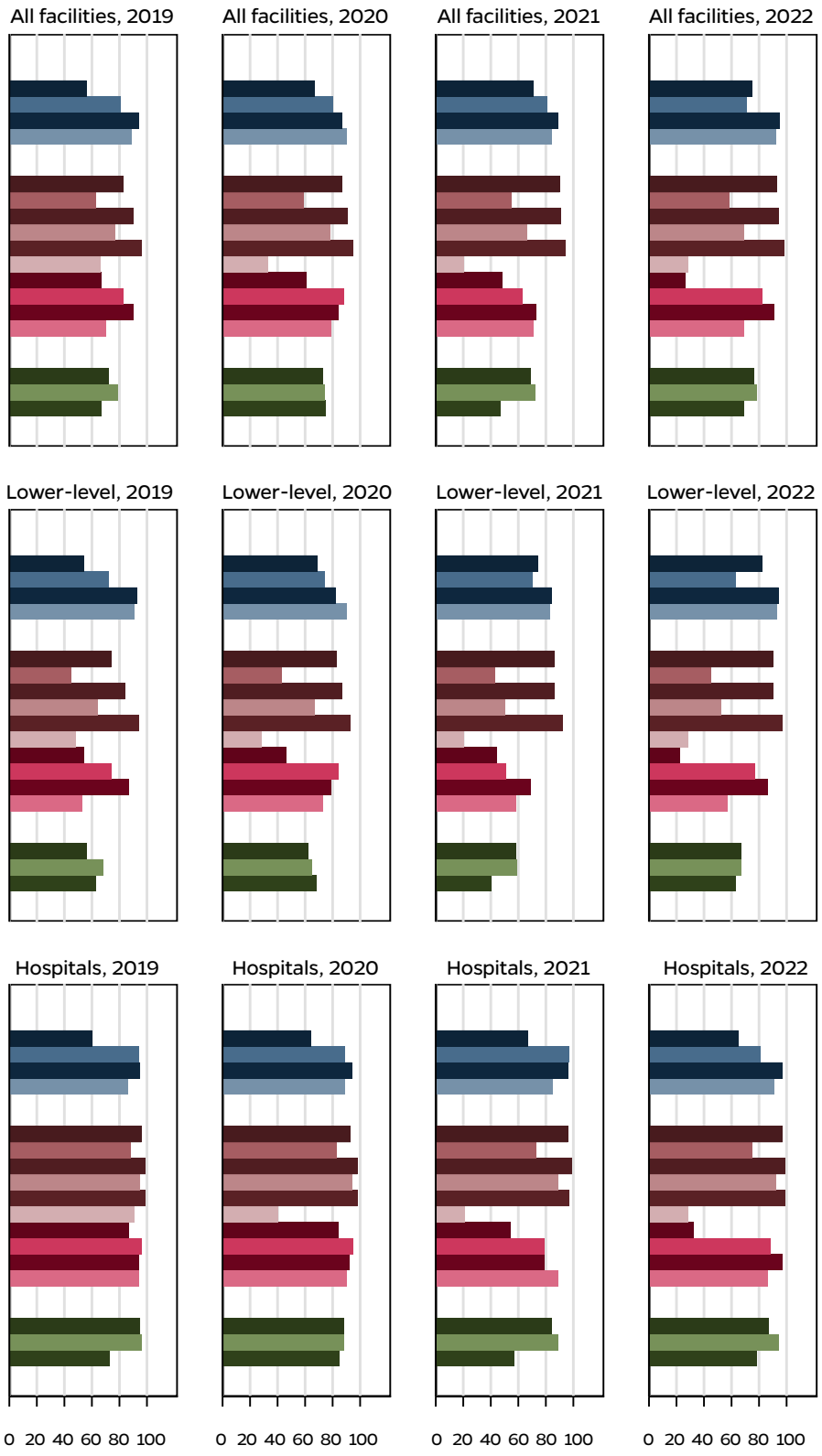
Each bar height represents the number of facilities. The sum of all bars in a year is the total number of panel facilities for objective 2.

Annex Figure 5. Distribution of annual absolute changes in availability scores for all and basic PPE currently available with no history of stock-out in the past three months: by observation period



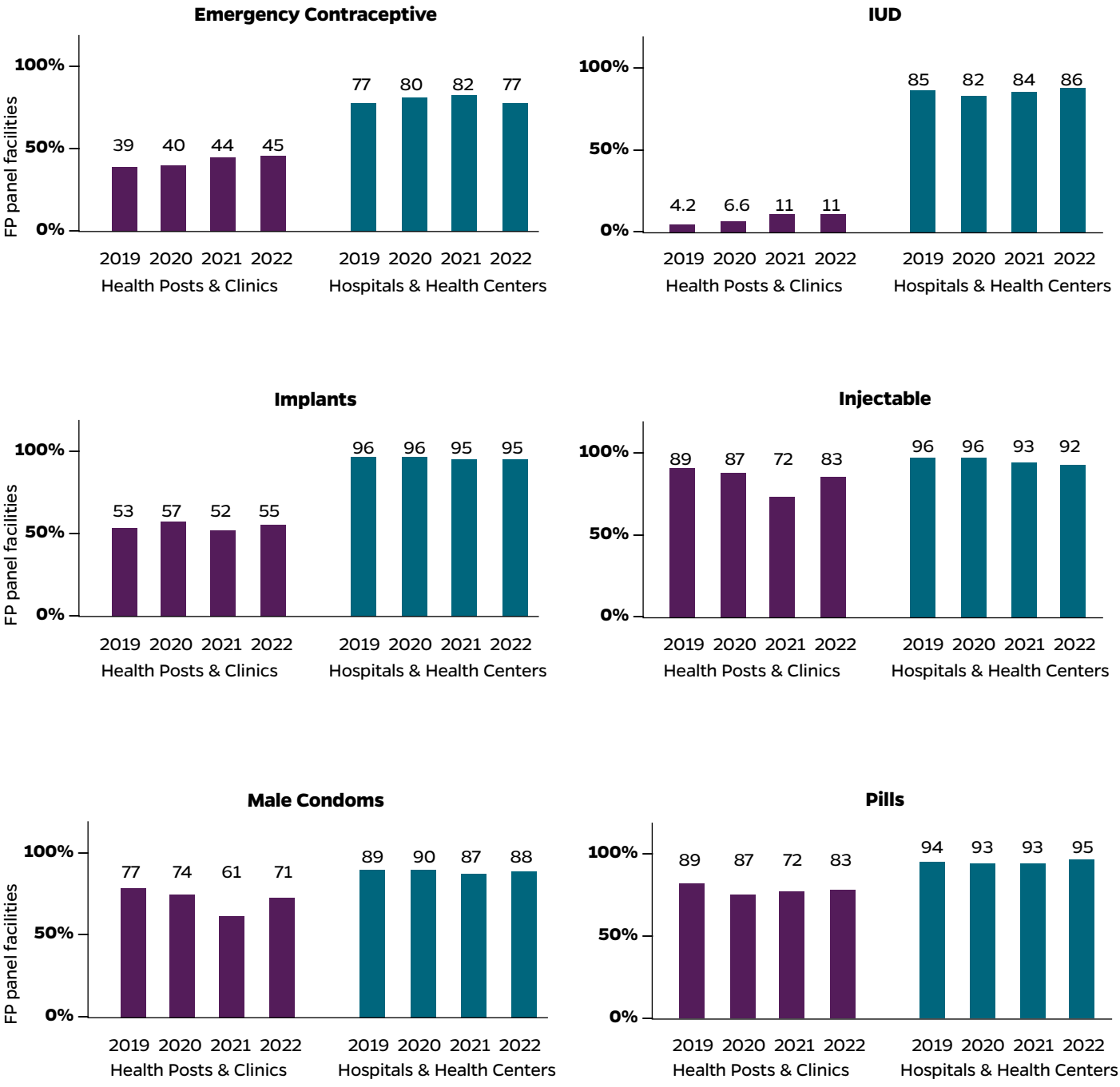
Note: Black vertical line refers to no change in the score. Red line shows normal distribution.

Annex Figure 6. Percent of MNH panel facilities with each of the 17 medicines included in the readiness score by year and facility type



BEmONC: Basic Emergency obstetric and newborn care; PNC: Postnatal Newborn Care; SSNC: Small and sick newborn care

Annex Figure 7. Percent of FP panel facilities with each of six individual contraceptive methods by year and facility type



Note: Female condoms and beads were available in less than 2% of facilities in any given year and excluded from the figure.

JOHNS HOPKINS
BLOOMBERG SCHOOL
of PUBLIC HEALTH

PMA
PERFORMANCE MONITORING FOR ACTION