



Level of Agreement of the Service Level Index (SLI) Tool with a Standard Service Provision Assessment Tool for Measurement of Antenatal Care Service Provision

Lorna Barungi Muhirwe¹ID*

¹Independent Researcher, Kampala, Uganda

Corresponding Author: **Lorna Barungi Muhirwe**

Address: Independent Researcher, Kampala, Uganda; Email: lornamuhirwe@gmail.com

Received date: 27 July 2023; **Accepted date:** 28 August 2023; **Published date:** 05 September 2023

Citation: Muhirwe LB. Level of Agreement of the Service Level Index (SLI) Tool with a Standard Service Provision Assessment Tool for Measurement of Antenatal Care Service Provision. J Health Care and Research. 2023 Sept 05;4(3):89-99.

Copyright © 2023 Muhirwe LB. This is an open-access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium provided the original work is properly cited.

Abstract

The generation of health service delivery data in middle- and low-income countries typically relies on health facility surveys and routine health monitoring data. Previous studies on antenatal care (ANC) service delivery have not employed measurement parameters that accurately and cost-effectively encompass various aspects of ANC service availability, content, and organization. This study aimed to assess the agreement level between the Service Level Index (SLI) tool, which measures ANC service provision, and a standard Service Provision Assessment tool. The SLI tool integrates pertinent sub-domains relevant to ANC service provision processes at the health facility level, employing selected key elements from these sub-domains as proxy measures. This approach minimizes the complexities associated with the time and effort required for assessing service delivery comprehensively.

The study examined the agreement between the Service Level Index tool and the ANC module of the MEASURE/DHS Service Provision Assessment tool. To accomplish this, the results obtained from the service level index measurement for each health facility were compared with the outcomes derived from the MEASURE/DHS Service Provision Assessment tool. Statistical analysis using the Bland-Altman method was employed to ascertain the significance of differences between measurements obtained from the Service Level Index tool and the Service Provision Assessment tool. The estimated mean difference (d) and standard deviation (sd) were 0.1 and 0.507, respectively.

The agreement level between the two tools (-1.19 to 1.46) indicated that, for 95% of observations, the scoring of ANC service provision conducted using the SLI tool deviated between 1.19% lower and 1.46% higher than the scores generated by the reference tool. The results revealed that the Service Level Index tool has the potential to serve as an alternative to the ANC module of the standard Service Provision Assessment tool for evaluating ANC services at the micro-level.

Keywords

Antenatal Care, Service Provision Assessment

Introduction

Efforts to measure the delivery of evidence-based, high-impact interventions aimed at reducing maternal mortality have primarily revolved around assessing the coverage of specific health services and interventions. The generation of health service delivery data in middle- and low-income countries typically relies on health facility surveys and routine monitoring data obtained from national health management information systems [1].

From a service provision perspective, scholars in the healthcare field commonly focus on two aspects of antenatal care (ANC) service delivery: content and quality. However, studies on ANC service delivery have yet to adopt an approach to measurement that comprehensively captures the fundamental aspects of service availability, content, and organization, with the goal of creating an integrated ANC service delivery index.

The ability to provide health services that effectively address the needs of the target population constitutes a crucial facet of health service accessibility [2]. As defined by both Donabedian and Dutton, this broadens the scope of healthcare accessibility to encompass how services and health resources either facilitate or impede the utilization of such services by the intended populations [2,3]. The measurement of service provision plays a pivotal role in ensuring that health managers possess a clear understanding of existing gaps and weaknesses. However, this measurement must be conducted using validated methods to ensure that the resulting data is deemed credible by all stakeholders, including policy makers [4]. Collecting service provision data at the health facility level offers the advantage of greater detail and relevance to the local contextual factors. This aspect is vital for steering evidence-based decision-making processes for local health managers.

The conceptual framework underpinning this study is drawn from the World Health Organization's quality of care framework for maternal and newborn health [5]. As outlined by the World Health Organization, this framework places significant emphasis on the internal processes within health facilities and their connection

to both individual and facility-level outcomes [5]. Within the facility-level process domain of the framework, two key sub-domains come into focus: the provision of care and the experience of care [5]. The framework acknowledges the overarching significance of the availability of competent and motivated human resources, along with the essential physical resources that serve as prerequisites for delivering high-quality care in health facilities. These factors, in turn, profoundly influence the utilization of these services [5].

Limitations of existing tools for measuring ANC service provision include a focus on service inputs with limited measurement of the actual provision of care, as well as high costs, time requirements, and the need for technical assistance to administer the tools [6-8]. Additionally, current service provision assessment tools provide an incomplete picture of the ability of health facilities to deliver defined health service packages [9]. There is a need for the generation of data at the national, subnational, and facility levels using rapid, more cost-effective approaches for data collection and interpretation.

Health facility surveys typically concentrate on the domains of access, availability, patient-centered care, and organization and management [7]. Three tools have primarily been used for health facility surveys in these settings: The World Bank Service Delivery Indicators tool, the WHO Service Availability and Readiness Assessment (SARA) tool, and the WHO Service Provision Assessment (SPA) tool [7]. Of these, the SARA and SPA tools have been specifically useful in monitoring some aspects of maternal health service provision. According to Sheffel, Karp & Creanga, the use of the SARA tool to assess service provision generates data for the indicators comprising the WHO quality of care framework, but only to a limited extent. There is consensus that the main limitation of this tool remains the overwhelming focus on service inputs, to the detriment of the actual provision of care, as well as the experience of care indicators, both of which influence the quality of care in synergy with service inputs [6,7].

The WHO Service Provision Assessment (SPA) tool

can be used to generate data on service inputs such as financial resources, human resources, medicines, equipment, and health infrastructure, capturing all the core indicators already reflected in the SARA tool [10,11]. In addition, the SPA tool includes a module on the provision of care broken down into three sub-domains: evidence-based practices for routine care and the management of complications, actionable information systems and functional referral systems for assessing ANC and Family Planning services, and a module on the experience of care [6]. According to DHS, information on the provision and experience of care for ANC services using the Service Provision Assessment tool is collected using observations, exit interviews, and provider interviews.

Researchers point out that some of the limitations of the Service Provision Assessment tool include the high costs needed to collect information using this tool, the time needed to complete the assessment, and the need for technical assistance when administering the tool in a survey [8]. Nickerson et al conducted a systematic review of health facility assessment tools used in middle- and low-income countries. The authors found that the inconsistency of methodology as well as the exclusion of essential elements of service provision presents obstacles to the comparability of findings within and across countries and over time. Specifically, the authors highlighted the limitations of service provision modules, which provided an incomplete picture of a health facility's ability to deliver defined health service packages [9]. These limitations could potentially constrain the use of the tool at the subnational level for the purposes of collecting evidence to inform planning and programming.

For health managers at the facility and sub-national levels, health management information systems represent the most accessible source of data for decision-making and the assessment of service provision [12]. The primary data collection tools within the HMIS that generate service data include patient registers, client records, and periodic summaries. In a study conducted across 13 countries to assess maternal and newborn health content in the HMIS, Dwivedi et al. found that in Uganda, the integrated antenatal registers tracked ANC first and

fourth visits, thus generating data on enrollment into ANC and ANC dropout rates [12].

In conclusion, it can be stated that while standard and validated tools for health service provision assessment exist, simpler tools suitable for facility-level assessments in low- and middle-income countries have not been widely disseminated or rigorously compared to these standard tools as a means of increasing their use and acceptability at these levels of service delivery.

This study evaluated the level of agreement of a Service Level Index tool in comparison with a standard Service Provision Assessment tool. The Service Level Index tool integrates different sub-domains inherent in the process of service provision at the health facility level by using selected key elements of these sub-domains as proxy measures to minimize the complexities related to the time and effort needed to administer the tool.

Research Methodology

Study Design and Setting:

This was a cross-sectional, quantitative study aimed at determining the level of agreement of the SLI tool when compared to the ANC module from the MEASURE/DHS SPA tool. The latter is a standardized health facility assessment tool that captures elements in the provision of care and experience of care domains for ANC services [6,13].

The study was conducted in a Private Not for Profit (PNFP) medical facility located in Iganga District, Uganda. Health workers interviewed included those directly involved in the management of the health facility. The data sources consisted of structured interviews with health facility managers.

Instrumentation and Measures:

Facility assessment scores regarding the service level factors of interest were gathered using both the Service Level Index tool and the Service Provision Assessment tool. The Service Provision Assessment methodology employs validated tools to gather information about the availability of various facility-based health services and the readiness level of health

facilities to provide these services [13]. When using the Service Provision tools, service availability and readiness are evaluated in the following four areas [13]:

- I. The availability of different health services.
- II. The extent to which facilities are prepared to offer various health services.
- III. The degree to which the service delivery process adheres to accepted standards of care.
- IV. The level of satisfaction that both clients and health providers experience within the service delivery environment.

The Service Provision Assessment employs an inventory questionnaire to collect information on the availability of health services at the time of the visit. Observation protocols for ANC, family planning, and child health are utilized to observe client-provider consultations. Client exit interviews are conducted to gather input from clients departing the facility, while health worker interviews are conducted to assess provider training, experience, and perceptions regarding working conditions. In this study, the ANC module of the SPA inventory questionnaire was administered, which also included the alignment of processes with standards of care. The client and health provider satisfaction modules of the standard SPA were not included in the assessment for the purposes of this study.

The Service Level Index tool was implemented at this health facility in conjunction with the ANC module of the SPA tool. This combined approach allowed for the gathering of information related to ANC service availability, ANC service content, and ANC service organization. The indicators incorporated within the Service Level Index tool are detailed in **Table-1**.

Statistical Analysis:

In order to assess the degree of agreement between the Service Level Index (SLI) tool and the SPA ANC module, the outcomes derived from the service level index measurement at the health facility were compared with the results obtained using the MEASURE/DHS Service Provision Assessment (SPA) tool. Bland-Altman analysis was employed to ascertain the statistical significance of differences between measurements obtained through the SLI tool and those acquired through the SPA tool [14,15]. The analysis centered on the comparison of paired measurements within the selected health facility using the two data instruments. Among the 30 assessment parameters that overlapped between the two tools, 12 fell under the service availability component, 17 were categorized as service content components, and 1 pertained to the service organization component.

All plots for the Bland-Altman method comparisons were generated using SPSS version 27.0. The correlation between scores from the SPA tool and those from the SLI tool was evaluated using linear regression. A graphical representation of the differences between the tools was created in accordance with the methodology outlined by J. M. Bland and D. G. Altman [15]. Values are presented as means and standard deviations (SD). Additionally, regression analysis to detect any proportional bias was also conducted using SPSS 27.0.

Results and Discussion

The overall scores and observation-specific scores obtained using the two data collection tools for assessing ANC service provision in the same health facility are presented in **Table-2** and **Table-3**, respectively. Notably, higher scores were observed when utilizing the Service Level Index tool (SLI)

Table-1: Key Parameters of ANC Service Provision in the Service Level Index Tool

ANC service availability	Availability of trained Health Provider
	Availability of Medicines and medical supplies
	Availability of Medical Equipment
ANC service content	Range of recommended services provided
ANC service organization	Timing of services
	Integration with other services

Table-2: Overall Scores Obtained by Using the Two Data Collection Tools to Assess ANC Service Provision in the Same Health Facility

Observation	SLI Tool (S1)	SPA Tool (S2)	S1 – S2	Mean (S1+S2)/2
ANC Service Availability	14	12	2	13
ANC Service Content	20	17	3	18.5
ANC Service Organization	2	3	-1	6.5
Overall Score	36	32		

Table-3: Bland Altman Analysis Comparing the Service Provision Assessment Tool (ANC Module) and the Service Level Index Tool

Observation	SLI	SPA	Mean	Diff
Adult stethoscope	3	1	2	2
Blood pressure apparatus (may be digital or manual sphygmomanometer with stethoscope)	1	1	1	0
Fetal stethoscope (pinard or electronic)	1	1	1	0
Light source (lamp or hand torch, with power source)	1	1	1	0
Fetoscope	1	1	1	0
Weighing Scale	1	1	1	0
Syphilis rapid tests	1	1	1	0
HIV rapid tests	1	1	1	0
Urosticks (protein)	1	1	1	0
Tetanus vaccine (prevention care)	1	1	1	0
Iron and folate (anti-anemia)	1	1	1	0
Sulfadoxine-pyrimethamine (fansidar) (anti-malarial)	1	1	1	0
Iron Supplementation	1	1	1	0
Folic Supplementation	1	1	1	0
Tetanus Toxoid Vaccination	1	1	1	0
Intermittent preventive treatment for malaria	1	1	1	0
HIV Rapid Diagnostic Test	1	1	1	0
Urine Protein Testing	2	1	2	1
Urine Glucose Testing	2	1	2	1
Rapid Test for Syphilis	1	1	1	0
Diagnosis and Treatment of STIs	2	1	2	1
Counseling on recommended minimum number of ANC visits for each pregnancy	1	1	1	0
Counseling on Birth Preparedness	1	1	1	0
Family Planning Counseling	1	1	1	0
HIV Counseling and Testing	1	1	1	0
Counseling on use of insecticide treated nets to prevent mosquito bites and malaria	1	1	1	0
Nutrition Counseling	1	1	1	0
Counseling about breastfeeding	1	1	1	0
Counseling on newborn care	1	1	1	0
How many days per week does this facility offer Antenatal care services?	2	3	2.5	-1

compared to the ANC module of the Service Provision Assessment tool (SPA). Sample statistics and the one-sample T-test outcomes are displayed in **Table-4** and **Table-5**. The mean difference (d) and standard deviation (sd) were calculated as 0.1 and 0.507, respectively. This finding suggests that, on average, the Service Level Index tool scores 0.1 higher than the ANC module of the Service Provision Assessment tool.

the 30 corresponding observations measured using the two data collection tools. The limits of agreement are noted as -0.9 and 1.1 (lower and upper LOA) respectively. With a range spanning from -0.9 to 1.1, these limits of agreement imply that the SLI tool's scores could be up to 1.1 higher or 0.9 lower than those derived from the ANC module of the SPA tool.

The results from the regression analysis aimed at detecting proportional bias are outlined in **Table-6**. The standardized beta-coefficient is -0.205 (p-value = 0.276), indicating the absence of proportional bias. **Fig-1** and **Fig-2** illustrates the Bland-Altman plots for

Out of the total observation scores (n=30), the majority exhibited a difference of zero (25 observations). For 3 observations, a difference of 1 was noted, and for the remaining 2 observations, differences of -1 and -2 were recorded.

Table-4: One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
Diff	30	0.13	0.507	0.093

Table-5: One Sample T-test

	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
Diff	1.439	29	0.161	0.133	-0.06	0.32

Table-6: Regression Analysis to Detect Proportional Bias

Model Summary						
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate		
1	.205a	0.042	0.008	0.505		
a. Predictors: (Constant), Mean						
ANOVA a						
Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	0.315	1	0.315	1.233	.276b
	Residual	7.152	28	0.255		
	Total	7.467	29			
a. Dependent Variable: Diff						
b. Predictors: (Constant), Mean						
Coefficients a						
Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	0.263	0.149		1.766	0.088
	Mean	-0.103	0.092	-0.205	-1.11	0.276
a. Dependent Variable: Diff						

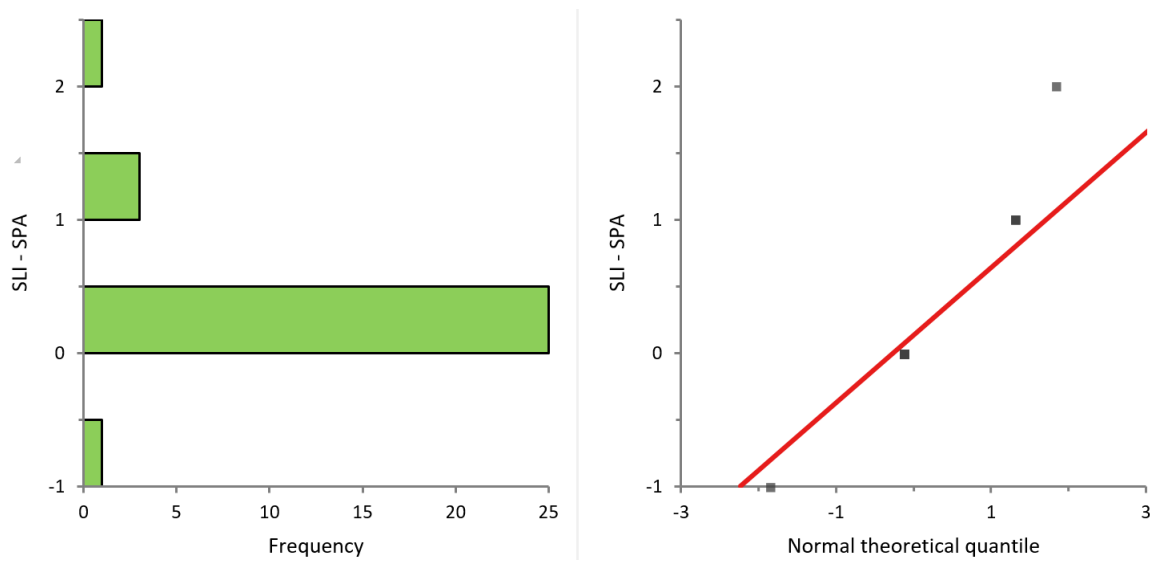


Fig-1:

Plot of Histogram and normal probability plot of the differences (Evaluation of assumptions). Using the Shapiro-Wilk test of normality, the W-test is 0.55 (p -value < 0.0001) which indicates distribution of the population is not normal.

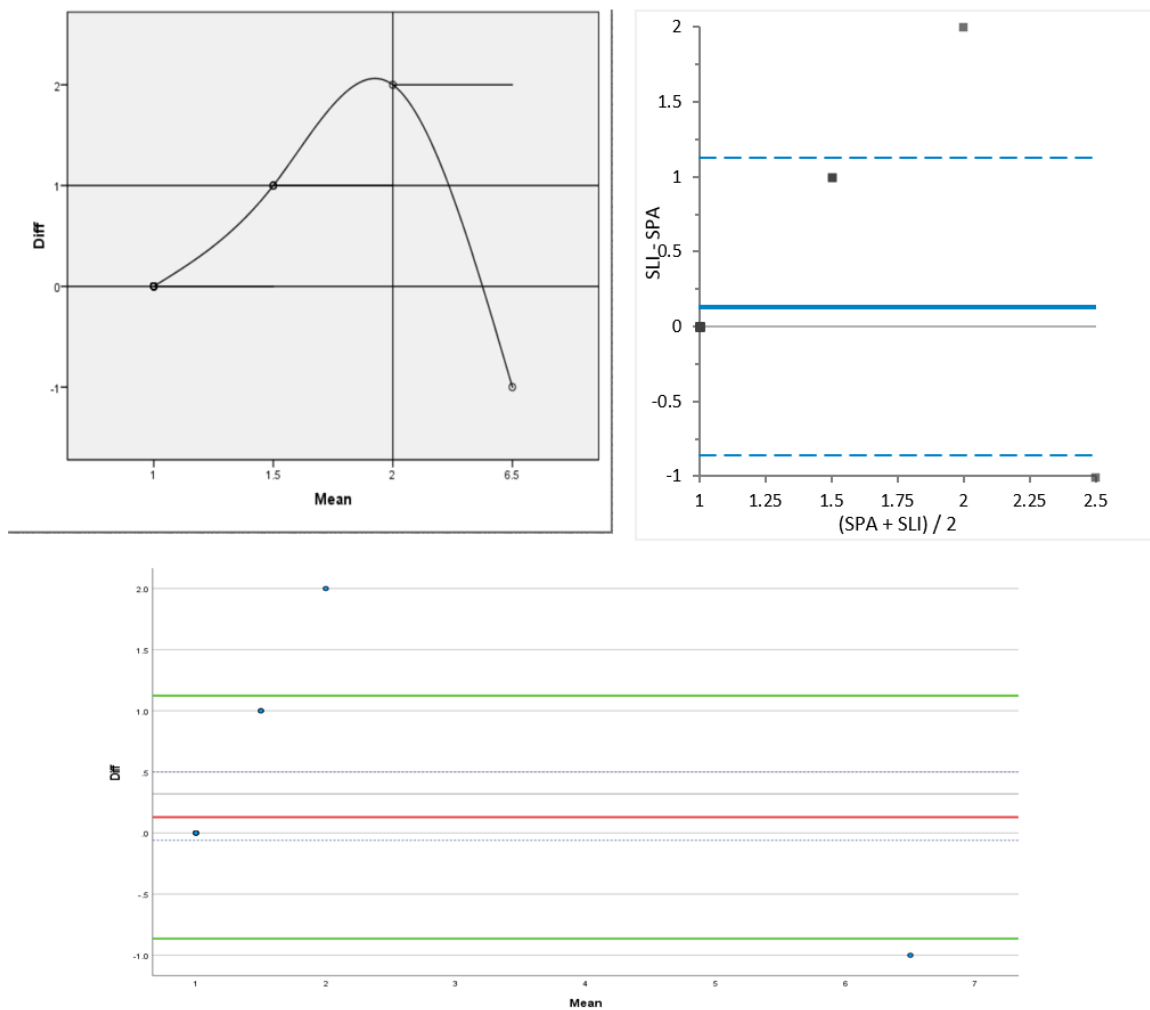


Fig-2: Bland Altman Plots

Legend: Solid red horizontal line shows the mean difference, dotted blue line shows the 95 % confidence interval of the mean difference, solid green lines show limits of agreement (± 1.96 standard deviation of the differences)

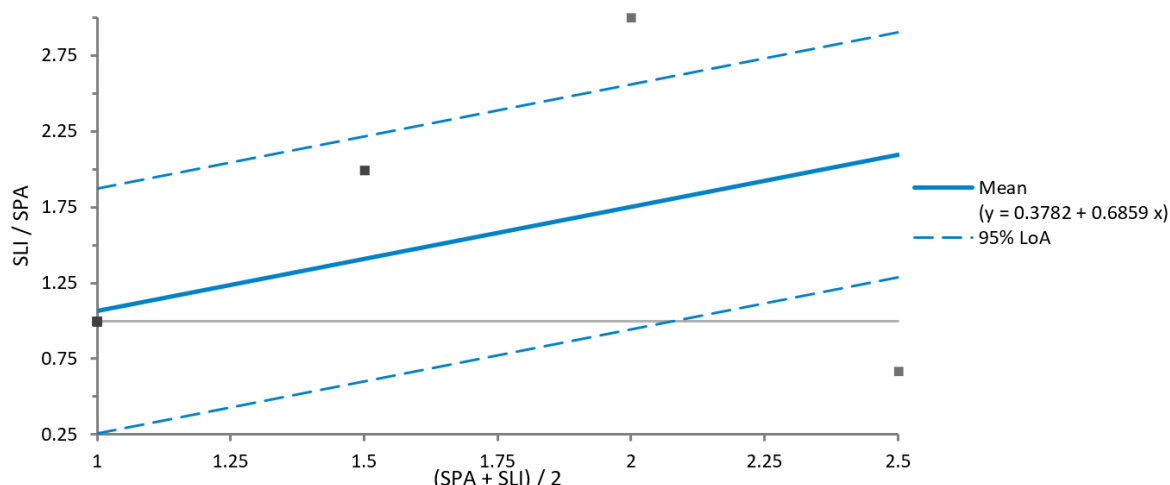


Fig-3: Inter-Tool Agreement Analysis

Table-7: Fit Differences

Parameter	Estimate	95% CI		SE
Mean difference	0.1	-0.06	to 0.32	0.09
95% Lower LoA	-0.9	-1.19	to -0.53	0.16
95% Upper LoA	1.1	0.8	to 1.46	0.16
SD	0.5			
Parameter	Estimate	95% CI		SE
Mean difference	0.1	-0.06	to 0.32	0.09
95% Lower LoA	-0.9	-1.19	to -0.53	0.16
95% Upper LoA	1.1	0.8	to 1.46	0.16
SD	0.5			

Fig-3 illustrates the Bland-Altman plot used for the analysis of interrater agreement (n = 30). This plot displays the limits of agreement along with 95% confidence intervals and a regression fit of the differences against the means. The 95% limits of agreement (95% LOA) serve to assess the extent to which the tools exhibit sufficient agreement for the purpose of evaluating ANC service provision. As detailed in Table-7, the 95% LOA between the two tools (-1.19 to 1.46) signifies that, for 95% of the observations, the scores attributed to ANC service provision using the SLI tool ranged from 1.19% less to 1.46% more than the scores obtained using the reference (SPA) tool.

Discussion

The primary purpose of this study was to assess the level of agreement between the Service Level Index (SLI) tool and the ANC module of the Service Provision

Assessment (SPA) tool (reference tool). The aim was to demonstrate the applicability of an alternative tool for assessing ANC service provision.

It was not possible to identify similar studies reporting results of levels of agreement between the MEASURE-DHS Service Provision Assessment tool and other tools. However, this study has demonstrated a close agreement between the SLI tool and the ANC module of the SPA tool. Ideally, Bland-Altman results should exhibit very small bias, narrow limits of agreement, and few outliers [16]. The current study found no evidence of proportional bias ($\beta = -0.205$, p-value = 0.276), restricted limits of agreement (-1.19 to 1.46), and a mean difference of 0.13, which is close to zero. For 83% of the assessed measures of ANC service provision that corresponded on both tools, there was no difference in health facility scores.

However, in three observations, the SLI tool consistently registered a higher score than the ANC module in the SPA tool by one point. These three observations were found under the ANC range of services domain and included urine protein testing, urine glucose testing, and diagnosis and treatment of STIs. Differences in scores for these observations are likely due to design differences between the tools. For example, the ANC module of the SPA tool requires assessor observation of urine testing for protein and provides three responses: action observed (score 1), action reported but not seen (score 2), and action not routinely done (score 3). On the other hand, the SLI tool relies solely on health provider responses for the assessment of the same service, providing three responses: yes, always provided (score 1), sometimes provided (score 2), and not provided at all (score 3). This means that if the action was observed using the SPA tool, the facility would score 1 point. However, if the same service was only sometimes provided despite being observed at the time of conducting the study, the SLI would assign a score of 2 for the same service. For urine glucose testing and STI diagnosis and treatment, the ANC module of the SPA tool assesses the availability of tests as a basis for scoring service content, while the SLI tool again relies on provider responses regarding the routine and consistent provision of the specific services.

Conversely, when assessing equipment availability under the service availability domain, the SLI tool combines the availability and functionality of medical equipment into a single score. In contrast, the SPA tool separates the availability and functionality scores for individual equipment, making the final scoring more cumbersome for the assessor. For instance, consider a stethoscope: the SLI provides three responses and scores—available and functional (score 1), available but not functional (score 2), and not available (score 3). Consequently, using the SLI tool, any equipment not physically verified by the assessor is scored as not available (3). On the other hand, the SPA tool offers five responses and scores for the same equipment: observed (score 1), reported but not seen (score 2), not available (score 3), if observed or reported and not seen; yes, functional (score 1) or no, not functional (score 2). With this example, the SPA tool would assign a score of 2 for a stethoscope that was reported as

available but not seen by the assessor, and then assign a score of 1 if the equipment was reported as being functional. However, the SLI tool would assign a score of 3 (not available) if the availability of this equipment could not be physically verified by the assessor during the study.

There was a single observation assessed under the timing of services domain: the number of days per week that the facility provided ANC services to clients. This observation resulted in an outlier value, likely explained by the design effect. Interestingly, the SPA tool does not assign a score to this observation, whereas the SLI tool assigns a specific score to the number of service provision days. This difference in handling these observations can lead to errors in analysis and interpretation, contributing to the disagreement between the tools for these observations.

An important finding from the study is that the Service Level Indicator (SLI) tool offers several advantages over the ANC module of the SPA tool. The SLI tool is solely focused on ANC service provision assessment, thereby providing a more comprehensive coverage of this aspect of service provision.

Conclusion

The potential to establish agreement between existing and novel health facility assessment tools holds the promise of reducing the expenses associated with health facility evaluations and augmenting the utilization of assessment data by micro-level health managers to enhance service provision. This study contributes significantly to this field through the introduction of a tool that seamlessly complements an already established counterpart. This augmentation allows for the collection of more comprehensive information, specifically focused on ANC service provision. Moreover, this study lays the groundwork for the development of analogous tools concentrated on other dimensions of service provision, all meticulously crafted and validated to align with existing tools such as the MEASURE-DHS SPA. These developments carry implications for service providers, researchers, and policy makers, as they expand the arsenal of available tools that can facilitate decision-making and fortify healthcare systems.

Further research is imperative to refine the design of new facility assessment tools and to evaluate their levels of agreement with established and validated tools. Future research endeavors should strive to demonstrate the necessary rigor and standards required to instill confidence in new, simplified assessment tools.

In conclusion, the present study evaluated the degree of agreement between two tools for measuring ANC service provision, revealing that these tools can be used interchangeably. Furthermore, the results showcased the potential of the Service Level Index tool as an alternative to the ANC module of the Service Provision Assessment tool for evaluating ANC services at the micro-level.

Declarations

Study Limitations:

While some scholars recommend a minimum of 30 observations (sample size), it's noteworthy that Bland Altman analysis is sensitive to outliers in the data, a limitation that can be mitigated by a larger number of observations [17]. The primary limitation of this study is the utilization of a relatively small number of observations to assess the agreement level of the scores obtained from the ANC service provision assessment tools. This limitation stems from the necessity to align with the number of corresponding measures on the tools being compared.

Acknowledgements:

The author expresses gratitude to the dissertation committee at Walden University for their invaluable assistance in reviewing and refining the manuscript. Additionally, the author extends thanks to the district health team of Iganga district in Uganda for their aid in facilitating data access and the data collection processes.

Funding Source:

No external funding was received for this study.

Competing Interests:

The author declares the absence of any competing interests.

Ethical Approval:

Ethical approval for this study was obtained from both the TASO-Uganda Institutional Review Board and the Walden University Institutional Review Board. All protocols for obtaining consent were carried out in accordance with the regulations set forth by the TASO-Uganda Institutional Review Board, the Walden University Institutional Review Board, and the principles outlined in the Declaration of Helsinki. Administrative clearance to conduct the study in Iganga District was granted by the District Health Officer. The selected study sites were informed about the study aims by the District Health Officers. Prior to commencing data collection, each data collector presented copies of the approval letter from the District Health Office and the Institutional Review Board to the medical officer/health facility in-charge at each site.

Informed Consent:

Written informed consent was obtained from the medical officer/health facility in-charge and documented as an integral part of the study. Before data collection or interviews, data collectors thoroughly explained the study outline to respondents to ensure a clear understanding of the requests being made. Respondents were informed that they had the right to decline answering or terminate the questionnaire at any point without facing any adverse consequences.

References

- [1] Hozumi D, Fronczak N, Minichiello SN, Buckner B, Fapohunda B, Kombe G, Searing H, Ramarao S, Ricca J. Profiles of Health Facility Assessment Methods. United States: Report of the International Health Facility Assessment Network (IHFAN), Measure Evaluation, USAID; 2008. Available from: <https://www.measureevaluation.org/resources/publications/tr-06-36/>
- [2] Donabedian A. Aspects of Medical Care Administration. Boston: Harvard University Press; 1973.
- [3] Dutton DB. Explaining the low use of health services by the poor: costs, attitudes, or delivery systems? *Am Sociol Rev*. 1978 Jun;43(3):348-68. [PMID: 686535]

- [4] Paxton A, Bailey P, Lobis S, Fry D. Global patterns in availability of emergency obstetric care. *International Journal of Gynecology & Obstetrics*. 2006 Jun 1;93(3):300-307.
- [5] World Health Organization. Standards for improving quality of care for maternal and new born care in health facilities. Geneva, Switzerland: World Health Organization; 2016. Available from: <https://apps.who.int/iris/bitstream/handle/10665/249155/9789241511216-eng.pdf>
- [6] Sheffel A, Karp C, Creanga AA. Use of Service Provision Assessments and Service Availability and Readiness Assessments for monitoring quality of maternal and newborn health services in low-income and middle-income countries. *BMJ Glob Health*. 2018 Nov 26;3(6):e001011. [PMID: 30555726]
- [7] Primary Health Care Performance Initiative. Methodology Note. PHCPI;2015. Available from: https://www.improvingphc.org/sites/default/files/PHCPI%20Methodology%20Note_o.pdf
- [8] Fort A. Service Provision Assessment. United States: UNFPA; 2014. Available from: <https://www.unfpa.org/resources/service-provision-assessment>
- [9] Nickerson JW, Adams O, Attaran A, Hatcher-Roberts J, Tugwell P. Monitoring the ability to deliver care in low- and middle-income countries: a systematic review of health facility assessment tools. *Health Policy Plan*. 2015 Jun;30(5):675-86. [PMID: 24895350]
- [10] Demographic & Health Surveys. Service Provision Assessment Questionnaires. United States: DHS Program. Available from: <https://dhsprogram.com/methodology/Survey-Types/SPA-Questionnaires.cfm>
- [11] World Health Organization. Service Availability and Readiness Assessment (SARA): an annual monitoring system for service delivery. Geneva, Switzerland: World Health Organization; 2015. Available from: <https://apps.who.int/iris/handle/10665/149025>
- [12] Dwivedi V, Drake M, Rawlins B, Strachan M, Tanvi M, Unfried K. A Review of the Maternal and Newborn Health Content of National Health Management Information Systems in 13 Countries in Sub-Saharan Africa and South Asia. United States: USAID Maternal and Child Health Integrated Program (MCHIP) and Maternal and Child Survival Program (MCSP); 2014. Available from: <https://www.mchip.net/sites/default/files/13%20country%20review%20of%20ANC%20and%20LandD.pdf>
- [13] Demographic & Health Surveys. SPA Methodology. United States: DHS Program. Available from: <https://dhsprogram.com/methodology/Survey-Types/SPA-Methodology.cfm>
- [14] Altman DG, Bland JM. Measurement in Medicine: The Analysis of Method Comparison Studies. *The Statistician*. 1983;32:307-17.
- [15] Bland JM, Altman DG. Measuring agreement in method comparison studies. *Stat Methods Med Res*. 1999 Jun;8(2):135-60. [PMID: 10501650]
- [16] Scott LE, Galpin JS, Glencross DK. Multiple method comparison: statistical model using percentage similarity. *Cytometry B Clin Cytom*. 2003 Jul;54(1):46-53. [PMID: 12827667]
- [17] Gerke O, Vilstrup MH, Halekoh U, Hildebrandt MG, Høilund-Carlsen PF. Group-sequential analysis may allow for early trial termination: illustration by an intra-observer repeatability study. *EJNMMI Res*. 2017 Sep 26;7(1):79. [PMID: 28952076]