



**USAID**  
FROM THE AMERICAN PEOPLE

# CAN PRIVATE SECTOR PROVIDERS ACCURATELY ASSESS THE QUALITY OF SERVICES THEY PROVIDE? EVIDENCE FROM PRIVATE MIDWIVES IN UGANDA

September 2009

This publication was produced for review by the United States Agency for International Development. It was prepared by Mai Do for the Private Sector Partnerships-*One* project.



**PSP-*One***

PRIVATE SECTOR PARTNERSHIPS FOR BETTER HEALTH

## Country Report

**Country Report Series:** PSP-One Country Reports were developed to inform specific PSP-One country program operations, but they also contain results that may be of interest to a wider audience. All papers in the series were reviewed by PSP-One technical staff in the field and in Washington, DC, as well as by relevant PSP-One program management staff.

**Recommended Citation:** Do, Mai. September 2009. *Can Private Sector Providers Accurately Assess the Quality of Services they Provide? Evidence from Private Midwives in Uganda*. Bethesda, MD: Private Sector Partnerships-One project, Abt Associates Inc.

**Download:** Download copies of PSP-One publications at: [www.psp-one.com](http://www.psp-one.com)

**Contract/Project No.:** GPO-I-00-04-00007-00

**Submitted to:** Patricia Mengech, CTO  
Bureau of Global Health  
Global Health/Population and Reproductive Health/Service Delivery Improvement  
Center for Population, Health and Nutrition  
Bureau for Global Programs, Field Support and Research  
United States Agency for International Development



Abt Associates Inc. ■ 4550 Montgomery Ave, Suite 800 North ■  
Bethesda, Maryland 20814 ■ Tel: 301/913-0500 ■ Fax: 301/652-3916  
■ [www.PSP-One.com](http://www.PSP-One.com) ■ [www.abtassoc.com](http://www.abtassoc.com)

*In collaboration with:*

Banyan Global ■ Dillon Allman and Partners ■ Family Health  
International ■ Forum One Communications ■ IntraHealth International  
■ O'Hanlon Health Consulting ■ Population Services International  
■ Tulane University's School of Public Health and Tropical Medicine

# **CAN PRIVATE SECTOR PROVIDERS ACCURATELY ASSESS THE QUALITY OF SERVICES THEY PROVIDE? EVIDENCE FROM PRIVATE MIDWIVES IN UGANDA**

## **DISCLAIMER**

The authors' views expressed in this publication do not necessarily reflect the views of the United States Agency for International Development (USAID) or the United States Government.



# CONTENTS

<b>Acronyms.....</b>	<b>vii</b>
<b>Abstract .....</b>	<b>ix</b>
<b>1. Introduction.....</b>	<b>1</b>
<b>2. Methods.....</b>	<b>3</b>
<b>3. Findings .....</b>	<b>9</b>
<b>4. Discussion.....</b>	<b>13</b>
<b>References .....</b>	<b>15</b>



# ACRONYMS

<b>ANC</b>	Antenatal Care
<b>COPE</b>	Client-Oriented Provider-Efficient
<b>FP</b>	Family Planning
<b>HIV</b>	Human Immunodeficiency Virus
<b>PNC</b>	Postnatal Care
<b>PSP-One</b>	Private Sector Partnerships-One Project
<b>QI</b>	Quality Improvement
<b>STI</b>	Sexually Transmitted Infection
<b>UPMA</b>	Uganda Private Midwives Association
<b>VDRL</b>	Venereal Disease Research Laboratory





# ABSTRACT

Self-assessment tools have become increasingly important for quality improvement in settings where external supervision and evaluation of the quality of care may not always be feasible. The evidence has been mixed regarding the level of agreement between health care providers' self-assessment of quality and the actual quality of services observed. This paper examines this level of agreement for private midwives who provide reproductive health services in Uganda, including family planning (FP), antenatal care (ANC), and postnatal care (PNC). Self-assessment was part of a quality improvement package implemented by the Private Sector Partnerships-One (PSP-One) project, introduced to help midwives identify problems of service delivery and to work on solutions. The package comprises the following elements: a form to review service statistics; a provider self-assessment tool; a linked action plan to help solve issues identified by the self-assessment; and a tool to enable the supervisor to find solutions to problems identified by the provider. The self-assessment tool is meant to be used by private practitioners in independent practice for quality assurance that does not rely on external supervision. A total of 248 midwives were interviewed about the quality of their services before and after the intervention took place; some of them received training on the self-assessment tool while others did not. The actual quality of services was observed with clients of the same midwives before and after the intervention. The study found a moderate level of agreement between midwives' self-assessment of the quality and the observed quality of ANC services, while no statistically significant agreement was found with FP and PNC services. In addition, there was evidence that the training on the self-assessment tool improved the agreement between the midwives' self-assessment and the actual quality observed. The study underlines the potential of a provider's self-assessment tool for quality of care assessment, as well as quality improvement, in the private sector in developing country settings. Further research is needed to identify the types of services that may be best assessed by such a tool.



# I. INTRODUCTION

The goal of this paper is to document the performance of a provider's self-assessment tool in assessing the quality of health care, apart from its use as a quality improvement tool. Improving the quality of health care services is not an easy task in either the public or the private sectors. Recently, quality improvement approaches that rely on providers' taking ownership of their practice, such as the Client-Oriented Provider-Efficient (COPE) approach, which include self-assessment, action plans and supervision, have proven useful in the public sector (Bradley and Gras, 2005). Evaluations have shown positive impacts of these approaches. For example, a study in four African countries (Ghana, Kenya, Nigeria, and Uganda) showed that the COPE approach empowered staff, improved their morale and commitment to quality improvement, and enabled them to address the majority of problems without outside help (Lynam, Rabinovitz, and Shobowale, 1993). In Mali, self-assessment, when used regularly, was found to increase a provider's compliance with technical standards (Kelley et al., 2003). Self-assessment is believed to reliably assess a provider's learning needs, give health providers an opportunity to learn from experience, gain insight into their performance, promote reflection on personal performance, foster authority to change within providers, and help providers function more effectively (Beyeler et al., 2004; Evans et al., 2002; Hays et al., 2002; Marienau, 1990).

While a provider's self-assessment tool may be useful in improving the quality of health care, it is unclear how useful it is to assess the actual quality of services provided. A number of methods have been widely used to assess the quality of health care – each with its own strengths and weaknesses, relating to cost, the degree of intrusiveness, and the situations where it may work best (Miller, 1993). Direct observation of the process of health service delivery is often done by an outside observer who records on a checklist what is done during a consultation. This provides one of the most complete pictures of what takes place during the visit and is often considered a “gold standard” for evaluating other assessment methods (Franco et al., 1997; Huntington, Miller, and Mensch, 1996). If a self-assessment tool performs adequately, it could be a considerably less costly and logistically more feasible alternative to the direct observation method, and could be highly applicable in settings where external evaluation of the quality of care by peers and/or supervisors is not easily implemented.

The literature on health provider self-assessment, used as a tool to assess the quality of health care, is lacking and inconsistent in both the developed and developing worlds. Earlier studies examined the agreement between provider interview and direct observation (both by an outsider) of the quality of care, with varying results.

Franco et al. (1997) found varying levels of agreement between direct observation and provider interview for many aspects of sexually transmitted infection (STI) management in Malawi. For example, the level of agreement was low with regard to history taking, physical exam, and counseling, but fair to good with respect to treatment. In addition, providers' initial responses to interview questions often under-reported what was actually observed, while their follow-up responses over-reported when probed (Franco et al., 1997). A similar comparison of assessments of the quality of pediatric outpatient health care, also in Malawi, showed that providers did not reliably report what they did in most cases of common child illnesses; for rare events, however, the agreement between providers' report and observation was fair (Franco et al., 2002).

A recent review of 725 published papers in the U.S., the U.K., Canada, and Australia on physicians' self-assessment reported only 17 papers (describing 20 cases) that compared self-assessment with external observation of care (Davis et al., 2006). Of the 20 comparisons, only seven showed positive relationships between self-assessment and the observed quality of care; the majority (13) showed little or no – or even an inverse – relationship. The performance of self-assessment of the quality of care was worst among the least skilled and the most confident physicians (Davis et al., 2006). A similar phenomenon, in which weaker professionals tend to overrate themselves, has been documented in a number of studies elsewhere (Antonelli, 1997; Evans et al., 2002; Orsmond et al., 1997; Sullivan and Hall, 1997). Langer et al. (1998) also attempted to examine the correlation between providers' and users' perceptions of the quality of antenatal care in three countries (Cuba, Thailand, and Argentina). The published initial results of users' focus group discussions and in-depth interviews showed that service users assigned value to non-technical aspects of quality of care that were often neglected by providers; no results from providers' data were reported (Langer et al., 1998).

In developing countries, most of the evaluations were in the public sector. Because the scale of operation of most private, for-profit sector providers is usually smaller and they are often isolated from their peers, quality improvement instruments may not always be useful or feasible if they require substantial external supervision to assess a provider's performance, identify problems, and implement solutions. In these situations, a tool that relies on a provider's self-assessment to identify gaps may be effective in improving quality of care in the private, for-profit sector – especially if it is supported by a supervisor who helps find resources and solutions to those gaps (Agha, 2009). While the private sector is usually smaller in scale compared with the public sector in the developing world, it also tends to be more heterogeneous in both the types of providers and the quality of care. Such heterogeneity often makes it more difficult to assess the quality of care in the private sector (Brugha and Zwi, 1998). With increasing attention by policy makers to quality improvement in the private sector, it is important to have an understanding of how a private provider's self-assessment may perform as a quality assessment tool.

This study is based on data collected as part of an evaluation of the impact of a Quality Improvement (QI) package on the quality of reproductive health services provided by private sector midwives in Uganda. The intervention was implemented by the Private Sector Partnerships-One (PSP-One) project, which aimed to improve the private sector's involvement in reproductive health (RH) service provision and service quality in developing countries. The package has been shown to be effective in improving the quality of services (Agha, 2009). The self-assessment tool includes two questions: midwives were asked to assess, first, the overall quality of services, and, second, the technical quality of services that they themselves provided. The actual quality of service delivery was also observed by the same independent observer for an average of three clients per midwife.

This paper has the following objectives:

- 1) To examine the agreement between providers' self-assessment of the quality of services and the actual quality of services observed, before and after a quality improvement intervention took place.
- 2) To assess whether the use of the self-assessment tool itself improved the agreement between providers' self-assessment of the quality of services and the observed quality of services, between baseline and follow up.

## 2. METHODS

### DATA

Data for this study came from a pre-test post-test quasi-experimental research study that was conducted with private midwives who were members of the Uganda Private Midwives Association (UPMA), designed to evaluate a quality improvement approach using a QI package. The package includes the following: a form to review service statistics, a provider self-assessment tool, a linked action plan for problem solving, and a tool to enable supervisors to assist in solving problems identified by midwives. The tool asked midwives to assess various dimensions of their service provision, ranging from equipment available at the facility, to the technical quality of care, to business planning and marketing. The tool is aimed to assist midwives to identify and solve problems in each of these quality dimensions using available resources.

Three midwife groups were included: one comparison group, comprised of midwives in Kampala, the capital city; and two intervention groups in the Central, Eastern and Western regions. More than 500 private midwives are members of the 12 UPMA branches in these regions. While the majority of clinics (84-86%) in Kampala and Eastern regions were in urban areas, about 60% of clinics in Central and Western regions were in rural areas. Midwives were randomly assigned to two intervention groups, at the branch level. In intervention group A, midwives received a one-day training in using the self-assessment tool and completing an action plan, but their supervisors were not given training in problem solving and mobilizing external resources to assist midwives in problem solving. In intervention group B, midwives received the same training in the self-assessment tool and action plan, and their supervisors were trained in problem solving and resource mobilization. Midwives in groups A and B used the self-assessment tool monthly to assess a variety of service aspects, and then met with their supervisors to work on solutions to problems identified.

It was intended that 100 midwives would be recruited for each of the three groups, and interviews and observations would be conducted with the same midwives at baseline in 2006 and follow-up in 2007 (the time interval between baseline and follow-up varied between three and six months). However (despite having an updated list of UPMA midwives), incomplete addresses and changes in the location of midwife clinics yielded 276 midwives recruited at baseline; 248 (or 90%) of them were followed up. Except for 82 midwives in Kampala, the rest were roughly equally divided between intervention groups A and B.

It was intended that on average three client interactions would be observed for each midwife on the day of clinic visit; thus, the clients observed at baseline and follow-up were not likely to be the same. A total of 772 observations of client-provider interactions were conducted at baseline and 776 at follow-up. All clients whose interactions with midwives were observed were also interviewed as they left the clinics. This study included only data obtained from provider interviews and observations of the 248 midwives who were followed up (i.e., before and after the intervention). Baseline and follow-up data were pooled to assess the correlations between subjective and objective assessments of quality and how these correlations might be improved by the interventions.

## DEPENDENT VARIABLES

The actual quality of RH services was measured by two main indicators: counseling score and technical quality score. Data collected through direct observation of the service delivery process were considered an objective assessment of the quality of care. Table 1 presents definitions of the objective quality indicators.

*Counseling score* assessed the client-provider interaction in terms of the degree to which providers treated clients with respect and understanding, the degree to which the interaction was participatory, and the extent that providers used the opportunity to discuss other health services available that may meet clients' needs. A composite score was constructed as an additive summary of yes or no responses to eight questions. Internal reliability was high for both the baseline and follow-up data: Cronbach alpha was .73 and .78, respectively.

**TABLE 1: DEFINITIONS OF INTERPERSONAL AND TECHNICAL QUALITY INDICATORS**

Interpersonal	Definition of Indicators
Counseling	Provider did the following during consultation: informed client of her right to privacy and confidentiality, asked client questions regarding how she felt and listened attentively, encouraged client to ask questions, provided client with relevant information to make health-related decisions, ensured that client understood the information provided by asking follow-up questions, asked client what she thought about the services provided, used opportunity to discuss additional health issues, discussed additional services provided at clinic (out of 8)
Technical	Definition of Indicators
Family planning	Provider explained how the reproductive systems works, discussed client's needs, counseled client based on her unique needs, explained which contraceptive methods are available, provided information on where to obtain desired method if unavailable at midwife clinic, explained benefits of selected method, explained risks of selected method, explained contraindications of selected method, explained side effects of selected method, discussed how selected method works, discussed how to use selected method, explained what to do in case of side effects, discussed resupply of selected method, explained when client should return for follow-up, discussed option of changing method if it does not work for client, recommended use of condom for dual protection, encouraged client to have partner participate in counseling and family planning (FP) decision making (out of 17, rescaled to be out of 30)

<p>Antenatal care</p>	<p>Provider discussed need for four spaced antenatal visits, provider informed client about due date, explained importance of personal hygiene and nutrition during pregnancy, discussed how to prevent malaria, discussed how to avoid exposure to STI/Human Immunodeficiency Virus (HIV) by being faithful and asking partner to use condom, discussed how client can involve partner in preventing STI/HIV, reviewed danger signs of pregnancy, encouraged pregnant woman and partner to come for HIV counseling and testing, provided information about health problems and appropriate treatment, discussed need to develop a birth plan (including arrangements for emergency transportation), discussed need and options for family planning, discussed what client should bring to clinic for delivery, discussed unsafe traditional practices, discussed signs and symptoms of labor and what to expect during labor</p> <p>Provider did the following during the first antenatal visit: recorded height, weight, and blood pressure, determined expected date of delivery, performed or referred clients for syphilis blood test [Venereal Disease Research Laboratory (VDRL)] and hemoglobin cross-matching, listened to fetal heart tone and recorded results, inspected and palpated breasts, prescribed/dispensed iron and folic acid tablets and other preventive medication, determined tetanus toxoid status and vaccinated for tetanus toxoid or referred for vaccination</p> <p>Provider did the following during repeat antenatal visit: recorded weight and noted changes, recorded fundal height and noted changes, recorded blood pressure and noted changes, listened for and recorded presence of fetal heart beat, checked for following danger signs: vaginal bleeding, severe headache, visual changes or epigastric pain, swelling of face or hands, leaking amniotic fluid, severe nausea or vomiting, high temperature, severe abdominal pain, lack of fetal movement (out of 30)</p>
<p>Postnatal care</p>	<p>Provider discussed personal hygiene, nutrition and infant feeding, care of perineum and breasts, family support, FP and avoiding unwanted pregnancy, benefits of exclusive breastfeeding</p> <p>Provider assessed mother's knowledge of and ability to breastfeed</p> <p>Provider asked client about postpartum danger signs, including excessive vaginal bleeding, vaginal discharge with odor, severe abdominal pain, worsening perineal pain, high temperature, continuous nausea and vomiting, redness or pain in breasts, pain in urination or difficulty in voiding</p> <p>Provider asked client if she had noticed danger signs in infant, including infant not sleeping well, sleeping all the time, vomiting or spitting, watery, dark green stool, breathing too fast, stiffness or convulsions, yellow skin and eyes, redness around or foul discharge from umbilicus or from eyes (out of 30)</p>
<p>Family planning, antenatal care, and postnatal care</p>	<p>Summary score across FP, antenatal (ANC), and postnatal clients (out of 30)</p>

*Technical quality score* was constructed in a similar way. The score was first constructed separately for three different types of services that midwives provided: ANC, FP, and postnatal care (PNC). For ANC and PNC services, technical quality score measured the degree to which providers discussed specific information with clients: the need for health care in the antenatal or postnatal period, danger signs, and essential clinical exams and laboratory tests. For FP services, the measure focused on discussion of methods that could meet clients' RH needs, information related to the use of a particular methods, condom use for STI prevention, and partner's involvement in FP decision-making. Each of the three scores was re-scaled to a scale between 0 and 30. The three scores were then summed up to yield a total score of technical quality for all services provided. Internal reliability coefficients for these technical quality scores were also reasonably high, ranging from .78 to .90 at baseline and from .67 to .76 at follow-up.

## **INDEPENDENT VARIABLES**

The main independent variable of interest in this study is the midwife's self-assessment of quality of services provided. This measure comes from responses to two questions in the self-assessment tool. The first question was "In your opinion, on a scale from 0 to 2, where 0 = poor, 1 = fair, and 2 = good, how would you rate the overall quality of services that you provided?" Responses were dichotomized into two quality groups, good and poor/fair. The second question dealt directly with the technical quality: "In your opinion, on a scale from 0 to 2, where 0 = no improvement needed, 1 = some improvement needed and 2 = a lot of improvement needed, how much improvement is needed in [technical competence/skills]?" Midwives were grouped into two categories depending on whether or not they responded that their technical skills needed a lot of improvement. The midwife's self-assessment was considered a subjective assessment of the quality of care.

Other covariates were included in the analysis to control for factors that may be hypothetically related to both the subjective and objective assessments of quality. A midwife's involvement in work outside of the clinic might hamper the actual quality of services as well as her ability to accurately assess her service quality, due to potentially divided attention. The number of staff at clinics is hypothesized to be positively related to the actual quality, though it may be negatively related to midwife's assessment of the quality. The provider's number of years of experience working as a midwife, on the other hand, could improve both the actual quality and midwife's ability to assess the quality of services provided. Similarly, a midwife's asking clients for their opinions about services provided should be positively related to both perceived and actual quality of services.

## **STATISTICAL ANALYSIS**

Bivariate analyses were carried out to provide a preliminary assessment of the correlations between subjective and objective assessments of quality of services and how they may have changed after the interventions took place. In this analysis, overall and service-specific technical quality scores were compared between groups of midwives with differing subjective assessments of quality (i.e., good versus poor/fair, a lot of improvement needed versus some or no improvement needed) at both baseline and follow-up.

In the second part of the analysis, potential confounders that may be related to both subjective and objective assessments of quality were included in multivariate linear regression models. These models took a general form as follows:



$$y_{ij} = \beta_0 + \beta_1 x_{1ij} + \beta_2 x_{2ij} + \beta_3 x_{1ij} x_{2ij} + \beta_4 x_{3i} + \beta_5 z_{ij} + \varepsilon_{ij}$$

where:

$y_{ij}$  = actual quality of care observed of midwife  $i$  in group  $j$

$x_{1ij}$  = midwife's assessment of quality of care

$x_{2ij}$  = time period (baseline vs. follow-up)

$x_{3i}$  = intervention group

$z_{ij}$  = a set of confounding variables, including whether the midwife worked outside of the clinic (dichotomous), whether the midwife asked clients for opinions about services (dichotomous), the number of years of experience (continuous), and the number of staff at the clinic (continuous).

$\varepsilon_{ij}$  = random error

The same models were computed for counseling and overall and service-specific technical competence scores. All analyses were carried out at the client level, taking into account the clustering of clients at the midwife level. Stata/SE version 10 was used for all statistical analyses (StataCorp, 2008).

In these models, the  $\beta_1$  coefficient indicates the correlation between midwives' assessment of quality and the actual quality observed at baseline. A positive, significant  $\beta_1$  indicates a high ability to accurately assess the quality of services they provided, whereas a negative, significant  $\beta_1$  indicates a poor ability to assess the quality of services. A non-significant  $\beta_1$  suggests no correlation between midwives' subjective assessment of the quality and the actual quality observed.

The  $\beta_3$  coefficient indicates the effects of the interventions on the correlation between subjective and objective assessments of quality. A positive, significant  $\beta_3$  indicates an improvement in midwives' ability to assess quality of services due to the interventions, whereas a negative, significant  $\beta_3$  indicates a negative impact of the interventions on midwives' ability to accurately assess their service quality. A non-significant  $\beta_3$  indicates no change in the correlation between the subjective and objective quality assessments that could be attributed to the interventions. While  $\beta_1$  indicates the correlation between midwives' assessment of the quality and the actual service quality observed at baseline, the sum of  $\beta_1$  and  $\beta_3$  indicates the correlation between subjective and objective quality assessments at follow-up.



### 3. FINDINGS

Table 2 shows changes in midwives' subjective assessment of the quality of services at baseline and at follow-up, for all services and for each type of service they provided. Overall, the majority of midwives considered their services to be of good quality and/or felt that there was no or very little improvement needed (between 60% and 70%, depending on the type of services). There was little change in their assessment over time: for ANC services, a higher percentage of midwives reported that their technical competence needed "a lot" of improvement at follow-up compared to baseline (40% versus 32%,  $p < .05$ ). No differences were observed for other services. When data were stratified by intervention group, results were similar (not shown). Significant changes were observed for ANC technical competence: more midwives reported that improvements were needed at follow-up compared to baseline, especially among the comparison group (50% versus 26%,  $p < .01$ ).

**TABLE 2: MIDWIVES' ASSESSMENT OF QUALITY OF CARE BY SERVICE PROVIDED, UGANDA, 2006 - 2007**

	ANC		FP		PNC		ALL SERVICES	
	Baseline %	Follow-up %	Baseline %	Follow-up %	Baseline %	Follow-up %	Baseline %	Follow-up %
Overall quality								
Poor/Fair	31.1	35.3	31.8	31.4	30.0	31.8	31.0	33.1
Good	68.9	64.7	68.2	68.6	70.0	68.2	69.0	66.9
Technical competence needed improvement								
A lot of improvement needed	31.7*	39.6*	36.1	31.4	30.9	27.1	32.9	33.6
No or some improvement needed	68.3*	60.4*	63.9	68.6	69.1	72.9	67.1	66.4
N (clients)	284	323	255	239	233	214	772	776

\*  $p < .05$ ; \*\*  $p < .01$ ; \*\*\*  $p < .001$

Table 3 compares actual quality scores (technical competence and counseling) with midwives' subjective assessment at baseline and follow-up. With regard to subjective assessment of the overall quality of services, there was a general agreement at both baseline and follow-up between observed quality scores and the subjective assessment: i.e., observed quality scores were higher among providers who said their services were good. With the exception of FP technical score and counseling score, all of the differences in technical scores between the good and poor/fair overall subjective quality groups were statistically significant.

**TABLE 3: ACTUAL QUALITY OF CARE OBSERVED BY MIDWIVES' SUBJECTIVE ASSESSMENT OF QUALITY OF CARE, UGANDA, 2006-2007**

	Subjective assessment of overall quality				Subjective assessment of technical competence			
	Baseline (n=772)		Follow-up (n=776)		Baseline (n=772)		Follow-up (n=776)	
	Poor/Fair %	Good %	Poor/Fair %	Good %	Lots of improvement needed %	No/some improvement needed %	Lots of improvement needed %	No/some improvement needed %
Technical competence score	19.2**	20.8**	21.5***	24.7***	20.1	20.4	23.1	23.9
ANC technical score	13.3***	15.5***	21.1***	24.3***	13.9	15.2	22.7	23.5
FP technical score	24.2	25.0	24.2**	26.1**	25.2	24.5	24.9	25.8
PNC technical score	21.0*	22.8*	19.1***	23.6***	21.3	22.6	21.8	22.3
Counseling score	5.9	6.1	6.1***	6.8***	5.8*	6.2*	6.4*	6.7*

\* p<.05; \*\* p<.01; \*\*\* p<.001

We found weaker agreement between subjective assessment and objective assessment regarding the need for improvement in technical quality. Only the counseling score showed a significant and consistent agreement with midwives' assessment of technical quality: the counseling score was higher among midwives who said that their technical skills needed "no" or "little" improvement than among those who said they needed "a lot" of improvement, at both baseline and follow-up (p<.05). For the other observed technical quality scores, none of the differences were statistically significant.

In this section, we discuss the correlations between subjective and objective assessments of service quality and how these correlations may have changed after the interventions took place, taking into account factors that may affect both midwives' subjective assessment and the objective assessment of the quality of services. Table 4 presents results from the multivariate models. Here, the main independent variable of interest is the midwives' assessment of overall quality of services, given in response to the first question. (The second question dealt specifically with the technical quality.) Only for ANC service was the midwives' assessment of the overall quality positively associated with observed technical quality: the ANC technical score was marginally higher for midwives who stated

that their services were of good quality than for those who responded otherwise ( $p < .10$ ). In most cases, the interventions did not influence the associations between the subjective and objective quality assessments. The only area where the interventions seemed to improve the correlation between the subjective and objective assessments of quality was in counseling ( $p < .05$ ).

**TABLE 4: ASSOCIATIONS BETWEEN MIDWIVES' ASSESSMENT OF OVERALL QUALITY AND ACTUAL QUALITY OF CARE, CONTROLLING FOR OTHER FACTORS, UGANDA, 2006-2007**

	Technical competence <sup>1</sup>	ANC technical score	FP technical score	PNC technical score	Counseling score <sup>1</sup>
	Coefficient (s.d.)	Coefficient (s.d.)	Coefficient (s.d.)	Coefficient (s.d.)	Coefficient (s.d.)
Midwives' assessment of overall quality					
Poor/Fair	—	—	—	—	—
Good	1.16 (.80)	1.58 (.85) <sup>†</sup>	.43 (.80)	1.25 (1.37)	-.02 (.28)
Time					
Baseline	—	—	—	—	—
Follow-up	2.68 (.65) <sup>***</sup>	7.94 (.78) <sup>***</sup>	.13 (.84)	-1.88 (1.25)	.18 (.23)
Time * midwives' assessment	1.28 (.81)	.76 (.98)	1.18 (1.03)	2.62 (1.59)	.54 (.27) <sup>*</sup>
R-squared	.29	.43	.14	.14	.13
N	1,547	606	494	447	1,547

All models control for intervention group, whether midwife also worked outside of clinic, number of years of experience, number of staff at clinic, and whether midwife asked clients for opinions about services.

<sup>1</sup> Controlling for services provided.

<sup>†</sup>  $p < .10$ ; \*  $p < .05$ ; \*\*  $p < .01$ ; \*\*\*  $p < .001$

Table 5 shows midwives' assessment of the *technical* quality of services. Similarly to results from the bivariate analysis, the associations between subjective assessment of the technical quality and the observed technical quality scores were weak. There were no statistically significant associations between midwives' assessment of the *technical* quality and the actual quality observed. The interventions seemed to improve the correlation between midwives' assessment of quality and the actual FP service technical score: the association between subjective technical quality assessment and the objective FP technical score was marginally higher at follow-up than at baseline ( $p < .10$ ).

When data were stratified by intervention group, similar results were found. However, the only statistically significant results were observed among intervention group A, where midwives relied only on the self-assessment tool and the action plan to improve the quality of their services (with no additional training for their supervisors). Also within group A, there was evidence of improved correlations between subjective and objective assessments of service quality for all services: ANC, PNC, FP, and counseling (results not shown).

**TABLE 5: ASSOCIATIONS BETWEEN MIDWIVES' ASSESSMENT OF TECHNICAL QUALITY AND ACTUAL QUALITY OF CARE, CONTROLLING FOR OTHER FACTORS, UGANDA, 2006-2007**

	Technical competence <sup>1</sup>	ANC technical score	FP technical score	PNC technical score	Counseling score <sup>1</sup>
	Coefficient (s.d.)	Coefficient (s.d.)	Coefficient (s.d.)	Coefficient (s.d.)	Coefficient (s.d.)
Midwives' assessment of technical quality					
Poor/Fair	—	—	—	—	—
Good	-.01 (.75)	.46 (.80)	-1.01 (.79)	.77 (1.30)	.18 (.28)
Time					
Baseline	—	—	—	—	—
Follow-up	3.42 (.68)***	8.22 (.86)***	-.40 (.88)	.50 (1.20)	.50 (.24)
Time * midwives' assessment	.16 (.97)	.36 (1.17)	2.05 (1.15) †	-.88 (1.63)	.06 (.33)
R-squared	.27	.42	.13	.10	.13
N	1,548	607	494	447	1,548

All models control for intervention group, whether midwife also worked outside of clinic, number of years of experience, number of staff at clinic, and whether midwife asked clients for opinions about services.

<sup>1</sup> Controlling for services provided.

† p<.10; \* p<.05; \*\* p<.01; \*\*\* p<.001

## 4. DISCUSSION

This paper documents the performance of a self-assessment tool in assessing the quality of RH services provided by private midwives in Uganda. Self-assessment tools have become increasingly important for quality improvement in settings where external supervision and evaluation of the quality of care may not always be feasible.

Generally consistent with previous research, the study found a moderate agreement between midwives' self-assessment of the quality and the actual quality of ANC services observed; the agreement was not significant, however, when data were stratified by intervention group (results not shown). No statistically significant agreement was found with FP and PNC services. Results within each intervention group (A and B) were similar although not statistically significant (results not shown). In addition, there was some evidence that use of the self-assessment tool, and training on its use, improved the agreement between self-assessment of quality and the actual quality observed. Only intervention group A showed strong evidence of the effectiveness of the tool on the level of agreement.

Previous studies have shown that the degree of agreement could vary between different dimensions and domains of the service quality (Franco et al., 1997; Franco et al., 2002). Therefore, it was not entirely surprising to find agreement with respect to ANC services but not FP and PNC services. One possible explanation is that midwives, because of the nature of their job, pay more attention to the quality of ANC services than to the other two services. Midwives performed better in providing ANC services, and they may therefore be more accurate in assessing the quality of these services. This hypothesis is consistent with earlier studies, where it was found that providers who were least skilled also performed worst in self-assessment (Davis et al., 2006; Hodges, Regehr, and Martin, 2001; Kriger and Dunning, 1999).

The finding that the training on the self-assessment tool improved the correlation between the subjective and objective quality assessments also warrants some discussion. An earlier evaluation of the tool showed positive impacts of the tool on the actual technical quality (Agha, 2009). One possible interpretation is that the interventions improved the actual technical quality without producing similar improvements in midwives' ability to accurately assess the quality of services they provided. This would imply that midwives overrated the quality of their services at baseline and as the actual quality was increased, agreement improved. However, if this were the case, level of agreement should have improved for both ANC and FP services, as the earlier evaluation showed improved technical quality in these services (Agha, 2009). Yet, for FP services, there was no evidence of improved agreement between the subjective and objective quality assessments.

On the other hand, it is also possible that the training in the use of the tool in fact increased the midwives' ability to assess their service quality, either because they got better at using the tool or paid more attention to the quality of care. This is consistent with the finding that at follow-up, more midwives stated that their technical quality needed improvement than at baseline. The self-assessment tool may have enabled midwives to assess their services from different, multiple angles. This might also explain the strong evidence observed for intervention group A regarding the impact of the self-assessment tool on the level of agreement between subjective and objective quality assessments. Midwives in this group relied solely on the self-assessment tool to identify dimensions of quality that

needed improvement, working on possible solutions using the action plan without any outside help (from their supervisors, etc.). It is consistent, then, for the tool to show stronger impacts for this group than for the other two groups.

Because the midwives in this study received only one initial training in the use of the tool and were followed up a short time afterwards, it was especially impressive that they not only showed an improved quality of services but also a more accurate self-assessment of their service quality. Self-assessment therefore appears to be a potentially feasible and valuable tool, which could be used with training to both improve the quality of care and to measure it. Depending on the purpose of the quality assessment, a self-assessment tool could be used by itself to identify common quality issues and to develop strategies that do not require external resources to address them. Self-assessment could also be used as a screening tool to flag quality issues or to identify midwives who need further quality assessment and improvement interventions. The tool could also be used in combination with other methods of quality assessment for supervision purposes, or to identify and address problems that may require external resources.

Further research on self-assessment is needed to shed light on behavioral psychological processes when providers use such tools, and especially providers' capacity to change, in order to design tools and applications that may be most appropriate. A limitation of this study is that we were not able to examine in detail the process of self-assessment or to capture the insights that midwives might have into their services from using the tool. Another limitation of a self-assessment tool is that it cannot assess some aspects of the quality of care, such as interpersonal communication skills and clinical reasoning. The former would require direct observation by someone other than the provider; the latter would require more detailed scenario analysis. In addition, the relatively small sample sizes of midwives and clients may make some of the results insignificant when stratified by intervention group.

The study nevertheless underlines the potential of a provider's self-assessment as a cost- and time-saving tool for assessment and improvement of quality of care in the private sector, in a developing country setting. The ability to accurately assess a provider's quality of services is an important step in improving services. Training in the use of the tool is essential for providers to be able to accurately assess the quality of services. Further research is needed, however, to identify the types of services that are most appropriately assessed by such a tool.



# REFERENCES

**Agha, S. (2009)**

The impact of a quality improvement package on the quality of reproductive health services delivered by commercial health providers in Uganda. Bethesda, MD: Private Sector Partnerships-One, Abt Associates Inc.

**Antonelli, M.A. (1997)**

Accuracy of second-year medical students' self-assessment of clinical skills. *Academic Medicine* 72:563-565.

**Beyeler, C., Westkämper, R, Villager, P.M., and Aeschlimann, A. 2004**

Self assessment in continuous professional development: a valuable tool for individual physicians and scientific societies. *Annals of the Rheumatic Diseases* 63:1684-1686.

**Bradley, J. and Gras, S. (2005)**

Improving the quality of child health services: participatory action by providers. *International Journal of Quality in Health Care* 17:391-399.

**Brugha, R. and Zwi, A. (1998)**

Improving the delivery of public health services by private practitioners in low and middle-income countries. *Health Policy and Planning* 13:107-120.

**Davis, D.A., Manzmanian, P.E., Fordis, M. et al. (2006)**

Accuracy of physician self-assessment compared with observed measures of competence: a systematic review. *Journal of the American Medical Association* 296(9):1094-1102.

**Evans, A.W., Aghabeigi, B., Leeson, R., O'Sullivan, C., and Eliahoo, J. (2002)**

Are we really as good as we think we are? *Annals of the Royal College of Surgeons of England* 84:54-56.

**Evans, A.W., McKenna, C., and Oliver, M. (2002)**

Self-assessment in medical practice. *Journal of the Royal Society of Medicine* 95:511-513.

**Franco, L.M., Daly, C.C., Chilongozi, D., and Dallabetta, G. (1997)**

Quality of case management of sexually transmitted diseases: comparison of the methods for assessing the performance of providers. *Bulletin of the World Health Organization* 75(6): 523-532.

**Franco, L.M., Franco, C., Kumwenda, N., et al. (2002)**

Methods for assessing quality of provider performance in developing countries. *International Journal for Quality in Health Care* 14 (Supp. 1): 17-24.

**Hays, R.B., Jolly, B.C., Caldon, L.J.M., et al. (2002)**

Is insight important? Measuring capacity to change performance. *Medical Education* 36(10):965-971.

**Hodges, B., Regehr, G., and Martin, D. (2001)**

Difficulties in recognizing one's own incompetence: novice physicians who are unskilled and unaware of it. *Academic Medicine* 76:S87-S89.

**Huntington, D., Miller, K., and Mensch, B. (1996)**

The reliability of the situation analysis observation guide. *Studies in Family Planning* 27:277-282.

**Kelley, E., Kelley, A.G., Simpara, C.H.T., et al. (2003)**

The impact of self-assessment on provider performance in Mali. *International Journal of Planning and Management* 18:41-48.

**Kruger, J. and Dunning, D. (1999)**

Unskilled and unaware of it: how difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology* 77:1121-1134.

**Langer, A., Nigenda, G., Romero, M., et al. (1998)**

Conceptual bases and methodology for the evaluation of women's and providers' perception of the quality of antenatal care in the WHO Antenatal Care Randomised Controlled Trial. *Pediatric and Perinatal Epidemiology* 12 (Supp. 2):98-115.

**Lynam, P., Rabinovitz, L.M., Shobowale, M. (1993)**

Using self-assessment to improve the quality of family planning clinic services. *Studies in Family Planning* 24(4):252-260.

**Marienu, C. (1990)**

Self-assessment at work: outcomes of adult learners' reflections on practice. *Adult Education Quarterly* 49:135-146.

**Miller, G.E. (1993)**

Conference Summary. AAMC consensus conference on the use of standardized patients in the teaching and evaluation of clinical skills. *Academic Medicine* 68:471-476.

**Orsmond, P., Merry, S. and Reiling, K. (1997)**

A study in self-assessment: tutor and students' perceptions of performance criteria. *Assessment and Evaluation in Higher Education* 24:357-369.

**StataCorp. (2008)**

Stata Statistical Software: Release 10. College Station, TX: StataCorp LP.

**Sullivan, K., and Hall, C. (1997)**

Introducing students to self-assessment. *Assessment and Evaluation in Higher Education* 22:289-303.



