USAID
FROM THE AMERICAN PEOPLE

Health Finance & Governance
Expanding Access. Improving Health.

EVALUATING THE COSTS AND EFFICIENCY OF INTEGRATING FAMILY PLANNING SERVICES INTO HIV AND AIDS TREATMENT SERVICES IN ZAMBIA

October 2015
This publication was produced for review by the United States Agency for International Development.
It was prepared for the Health Finance and Governance Project by Dr. Sophie Faye, Dr. Benjamin Johns, Dr. Elaine Baruwa, and Ms. Kelley Ambrose.

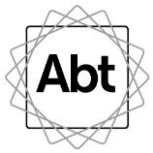
The Health Finance and Governance Project
USAID's Health Finance and Governance (HFG) project improves health in developing countries by expanding people's access to health care. Led by Abt Associates, the project team works with partner countries to increase their domestic resources for health, manage those precious resources more effectively, and make wise purchasing decisions. As a result, this five-year, $209 million global project increases the use of both primary and priority health services, including HIV/AIDS, tuberculosis, malaria, and reproductive health services. Designed to fundamentally strengthen health systems, HFG supports countries as they navigate the economic transitions needed to achieve universal health care.

OCTOBER 2015

Cooperative Agreement No: AID-OAA-A-12-00080

Submitted to: Scott Stewart, AOR
Office of Health Systems
Bureau for Global Health

Recommended Citation: Faye, Sophie, Benjamin Johns, Elaine Baruwa, and Kelley Ambrose**.** October 2015. *Evaluating the Costs and Efficiency of Integrating Family Planning Services into HIV and AIDS Treatment Services in Zambia.* Bethesda, MD: Health Finance and Governance Project, Abt Associates



Abt Associates| 4550 Montgomery Avenue, Suite 800 North | Bethesda, Maryland 20814
T: 301.347.5000 | F: 301.652.3916 | www.abtassociates.com

Avenir Health | Broad Branch Associates | Development Alternatives Inc. (DAI) | Johns Hopkins Bloomberg School of Public Health (JHSPH) | Results for Development Institute (R4D) | RTI International | Training Resources Group, Inc. (TRG)

EVALUATING THE COSTS AND EFFICIENCY OF INTEGRATING FAMILY PLANNING SERVICES INTO HIV AND AIDS TREATMENT SERVICES IN ZAMBIA

CONTENTS

List of Tables

ACRONYMS

ART	Antiretroviral therapy
CIDRZ	Center for Infectious Disease Research in Zambia
FP	Family planning
HFG	Health Finance and Governance project
HMIS	Health management information system
IR	Internal referral
MOH	Ministry of Health
OSS	One-stop shop model
PEPFAR	President's Emergency Plan for AIDS Relief
PMTCT	Prevention of mother to child transmission
USAID	United States Agency for International Development
VCT	Voluntary counseling and testing
ZMK	Zambian kwacha
ZPCT II	Zambia Prevention, Care and Treatment Partnership II project

ACKNOWLEDGMENTS

This work was funded by the USAID Office of HIV/AIDS. The study was executed by the Health Finance and Governance project, and the report was prepared by Dr. Sophie Faye, Dr. Benjamin Johns, Dr. Elaine Baruwa, and Ms. Kelley Ambrose. The authors are grateful to the people who made this study possible, including Dr. Peter Cowley, Dr. Laurel Hatt and Ms. Marianne El-Khoury from Abt Associates, who gave much-valued advice on the study objectives and design.

The contributions of Dr. Mwangelwa Mubiana and Mrs. Maureen Mzumara of the Center for Infectious Disease Research in Zambia, and Dr. Prisca Kasonde of the Zambia Prevention, Care and Treatment Partnership II project, were essential when collecting information from the sites.

The authors also thank the health team at USAID Zambia, especially Mr. Colin Zulu, for their important comments and insights on the study proposal.

We also gratefully acknowledge the contributions of Dr. D. M. Chikamata from the Permanent Secretary of the Ministry of Health, and Mr. Simmy Chapula, Director of Planning for the Ministry of Community Development, Mother and Child Health, as well as the District Medical Officers for Lusaka, Mongu, and Kabwe. The inputs and support of all staff at the Scaling Up Family Planning project in Lusaka, particularly Dr. Christopher Mazimba, who provided general support to this study, were also critical.

Finally, we wish to acknowledge the dedication of the data collection team lead by Mr Petan Hamazakaza, in country study coordinator, who travelled from site to site to collect the necessary data, and all the health facility staff that assisted them.

EXECUTIVE SUMMARY

Background

Integrating the delivery of health services is viewed as a priority in the fight for an AIDS-free generation, because this integration has the potential to improve access to HIV, family planning (FP), and other services and provide continuity of care for those living with HIV.

In particular, providing FP services to HIV-positive individuals as part of their health care can address some of the unique needs of this population, and can improve their overall well-being. In addition, such integration may make health service delivery more sustainable by increasing the efficiency with which resources are used, though evidence demonstrating the impact of integration on efficiency is limited. An initial review of the existing literature found qualitative evidence of increased efficiency from integration, at the programmatic level; but costing and cost-effectiveness evidence are lacking. Studies on integration often focused on measurement of outputs without necessarily considering the inputs required to obtain those outputs.

At the request of USAID's Office of HIV/AIDS and the USAID Zambia mission, the Health Finance and Governance (HFG) project conducted a study examining the costs and efficiencies involved in integrating family planning and antiretroviral therapy (ART) services. The specific objectives of the study were to:

1. Propose indicators that can measure the performance of the integrated programs with respect to efficiency, and assess their practical feasibility
2. Quantitatively assess the relative efficiency of different models of integration of HIV and FP services, using the defined indicators
3. Qualitatively identify potential barriers and facilitators to efficiency improvement

HFG worked with the FHI360-led Zambia Prevention Care and Treatment Partnership (ZPCT-II) project and the Centre for Infectious Disease Research in Zambia (CIDRZ). These implementing partners' programs supported the sites where FP services had been integrated into HIV clinics.

Approach and methods

A conceptual framework was developed to describe how integration might lead to different opportunities to increase the efficiency with which services are delivered, either by lowering costs/inputs required to deliver a given number of outputs or services delivered (lowering costs), or by increasing the number of services that can be delivered with a given set of inputs (increasing productivity). The framework also identified which inputs/outputs could be combined into indicators of efficiency. Three indicators described below were chosen to compare efficiency across models of integration of FP into HIV services.

The specific service-delivery points within a health facility that this study considers are the ART clinic and the FP clinic. The unit of analysis with respect to efficiency for this study is the ART clinic. An ART clinic could use one of two models of integration:

- The "internal referral" (IR) model, where FP *counseling* is offered in the ART clinic but the patient is referred to the FP clinic (in the same facility) for further FP services (particularly method provision)
- The "one-stop shop" (OSS) model, where FP counseling is offered in the ART clinic and the patient can also receive some FP methods from the same ART clinic

This analysis describes a cross-sectional, non-randomized comparison of the efficiency of these two models of integration, OSS and IR. Ten sites were purposively selected based on the implementing partners' recommendations. Data were collected for the 12-month period of October 2013 to September 2014. Service utilization data were collected from routine health management information systems (HMIS) and from a record review of 90 patient files per site, and supplemented by a patient exit interview of 15 randomly chosen patients per site. A time-motion study was also conducted to assess the length of a visit with and without provision of FP services, along with a patient flow assessment to determine how busy providers are at different times of the day. At each site, two key informants were interviewed on their current experience with integration and the potential barriers or facilitators to the process.

Cost data were collected for each health facility, as well as information on the cost of integration training provided to the staff. The cost data were mainly input quantities and prices, primarily related to labor, drugs, and medical supplies as well as outputs.

Findings

The indicators proposed for assessing the efficiency of the integration programs are mainly concerned with the direct inputs used in the production process and the outputs that result from the process. The rationale is to minimize inputs and/or maximize outputs to achieve better efficiency. The proposed indicators are each discussed in turn below in terms of the first two study objectives: feasibility and comparing program efficiency.

Percentage of missed opportunities at the ART clinic

A missed opportunity is defined as not offering FP counseling to an 18- to 49-year-old woman during an ART visit. Minimizing this percentage will maximize service output with a given set of inputs, and therefore increase efficiency.

Feasibility findings: In 6 of the 10 sites there was no mention of FP in any of the patient records reviewed, and record-keeping was not consistent at the other 4 sites. Thus, it was not possible to tell whether patients (a) were not receiving FP counseling or (b) were receiving FP counseling but this counseling was not being documented in their records. The patient exit interviews suggest that FP counseling does occur in the ART clinic but is not being systematically recorded when it happens. However, the indicator will be feasible if its components can be directly extracted from the routine data captured in patient files. These files should therefore be structured to contain data on the delivery of FP counseling where integrated services are being delivered.

Quantitative findings: The results from the patient exit interviews were used to determine the percentage of missed opportunities for all sites. We estimated the percentage of missed opportunities for different populations based on the exit interviews, and found a wide range in the number of missed opportunities across the sites: from 0 to 100 percent of patients were being counseled. Among non-FP users with an identified need for FP, there was an average of 47 percent counseled in the IR model and

14 percent in the OSS model. However, when looking across all women including current FP users, there was no statistically significant difference in the missed opportunity percentage between the two models.

Nurse/counselor time per ART patient counseled on FP

Under integration, FP services are tailored to individual patients, and are delivered during an ART visit, which has the potential to increase the time spent with each patient. The added time may have a large impact on the entire ART clinic's functioning, as it can mean extended working hours for the staff, or require a change in the staffing schedule to accommodate the extra work, both of which have costs.

Twenty-seven percent of interviewed patients across sites received FP counseling, but only 4 out of the 150 patients interviewed were given an FP method on the interview day. This small size of the group that received a method means that the data are insufficient to enable us to compare time of visit for ART+FP method provision across OSS sites.

Feasibility findings: This indicator cannot be easily used on a routine basis because it requires additional data collection effort: a time-motion study. Ideally, the time-motion study should be linked to an exit interview to determine exactly what services were offered, and a larger sample size of women should be interviewed to capture as much as possible the diversity in the FP methods provided across OSS sites.

Quantitative findings: There is no benchmark per se for an "efficient amount of time." Rather, we compared the relative efficiency across models and sites. We compared the average time spent on an ART visit with FP counseling across models, as well as comparing the average time spent on ART visits without such counseling. For the IR model the average time per ART visit without FP counseling was 9 minutes; with counseling it was 12 minutes. For the OSS model the average times were 10 and 13 minutes respectively. The addition of counseling adds very little time to each visit, approximately 3 minutes, and there was no statistically significant difference (p-value=0.65) between the two models, which is unsurprising, since they are doing the same thing.

Unit cost per ART patient counseled on FP

The unit cost indicator combines inputs and outputs into a single metric. Adding FP counseling to a regular ART visit should increase the cost of the visit. Health workers need to be trained to provide the additional services. Similarly to counseling, providing FP methods on-site (in the OSS model) will also add costs to the regular ART care. This extra cost contains training costs as well, but also depends on the type of FP method provided, because of FP commodities costs.

Feasibility findings: Again, the main challenge is the availability of data on the number of patients counseled on FP in the ART clinic. As noted, the patient exit interviews had to be used to estimate that number. However, if all the services provided per visit were being accurately recorded, then the information necessary for these indicators should be available from the HMIS routine data, from the facility management system (payroll data, procurement data, etc.), and from implementing partners (training and supervising costs). Compiling the data will require effort, but a template could be developed to automate analysis and estimation.

Quantitative findings: The calculated unit cost per patient for ART care is a direct measure that includes only labor, drugs, and supplies. For FP counseling the additional costs for training staff were added, as were further additional costs for method provision, including corresponding training costs and the cost of FP commodities. Table ES-1shows ART care costs with these additional costs.

TABLE ES-1: UNIT COST PER ART PATIENT PROVIDED WITH FP SERVICE

	Zambian kwacha (ZMK)*			US dollars		
	ART only	**ART + FP counseling**	**ART + FP counseling + method**	**ART only**	**ART + FP counseling**	**ART + FP counseling + method**
IR model	1,629	1,636	1,680	259	260	267
OSS model	1,619	1,623	1,640	257	258	260

*6.3 ZMK: 1USD.

When the percentage of missed opportunities is low for FP counseling, the ART clinic can potentially benefit from economies of scale and decrease its unit costs, because it is producing more outputs (ART care and FP counseling). As expected, the cost analysis results suggest that there are some efficiency gains from the OSS as opposed to obtaining FP at a referral clinic: ~$7 per patient. But this result appears small, because ART costs dwarf FP costs. In addition, we note that cumulatively, these "savings" could increase in size as missed opportunities decrease and more patients get FP counseling; but recall that the FP clinic will still have to function without ART patient referrals, because there is still an HIV-negative population to serve. A more appropriate way to view these results would be to say that providing a more comprehensive package of care to HIV-positive women, and increasing their access to FP services, costs relatively little regardless of which integration model is used; in the IR model it is an additional $8 and in the OSS model it is an additional $3 (no statistical difference between models). From this point of view, it is important to note that the societal benefit for the women of not having to make an additional clinic visit for FP services could be important, but was not assessed in this study.

Barriers and facilitators to the integration process

Potential staff shortages: Qualitative provider interviews repeatedly found the same main concern with integration: a shortage of staff. All of the health workers and managers interviewed noted that staff are overworked and that more staff are needed to successfully integrate FP into the ART clinic. However, the provider time assessment results suggest that all patients are attended to before closing time. Mornings are the busiest time of the day, and by the middle of the day the number of ART patients waiting for treatment is less than half of the number of patients waiting when the clinic opened, and in most clinics it is substantially less than half. Furthermore, the time-motion study found that only an additional 3 minutes of provider time, on average, appears to be used for visits where FP services are provided. Therefore, it may be more cost-effective overall to manage patient flow better using existing resources/inputs than to increase costs by adding staff that may not actually be necessary. However, this study did not address quality, and it could be posited that an additional 3 minutes of provider time for FP counseling is inadequate and actually reflects a staff shortage issue that has a negative impact on quality.

Weak referral tracking: An effective, formal referral tracking system was not part of the integration design and implementation, regardless of the model, and is a potential area for improvement. Health workers noted that it was difficult for them to track the patients they referred to the FP clinic, because there were no feedback mechanisms between the two clinics. Without a system that can track patients between the ART and FP clinics, it is difficult to accurately evaluate the true impact of integration in terms of FP method uptake for ART patients, and efficiency.

In terms of facilitators to integration, providers noted that integration of FP into ART care is a major change in the way the clinic has operated, and it needs to be discussed, understood, and owned by the staff for it to work. Having enough orientation and information up front about the integration was

identified as one of the necessary elements for success. Also, providers who received training about FP integration, FP counseling, and/or FP method provision greatly appreciated the new skills they acquired, and stressed the importance of this knowledge in caring for their ART patients.

Conclusion and recommendations

This study identified potential indicators that could be used to assess the relative efficiency of different models of integrating FP services into ART care, and tested these indicators' feasibility. Indicators used were the percentage of missed opportunities, the additional time it took to provide the FP services, and the unit cost of providing the services. We conclude that while collecting some of these indicators is feasible, all of them would require extra data collection efforts. Some of the indicators—such as the percentage of missed opportunities—are critical data needed for overall program monitoring, and are not needed uniquely for measuring efficiency.

The study found no significant difference in efficiency between the OSS and the IR models of integration for any of the proposed indicators. The drivers of efficiency appear to be at the facility level, not at the implementation model level. However, the analyses were based on a small, cross-sectional, and purposive sample, and confounding of the results due to selection bias or other factors is possible. In particular, the absence of complete FP service-delivery data at the patient level within the ART clinics and the absence of referral and counter-referral data requiring the use of patient exit interviews, mean that reaching definitive conclusions about the relative efficiency of either model was not possible in this study. Based on this finding, recommendations for future work aiming at improving integration of FP and HIV services in general are summarized below.

1. More effort is required to ensure that health workers systematically provide FP services in the ART clinic, as expected under integration. These health workers need to be educated in adequately recording the services they provide at the time of delivery and in terms of counter-referrals.
2. The HMIS should be adapted to be able to produce readily available statistics that can be used to monitor integrated services at the facility.
3. An effective, formal referral system should be part of the integration program design, to strengthen program monitoring and evaluation and patient record information.

To provide more-accurate information on the impact of integration on costs in general and on the use of services, a pre-post design would have been more suitable. However, at the time of the study, integration was already very widespread in Zambia. HFG is undertaking such a pre-post study on the efficiency of integration in Tanzania, and the results are expected to better inform us on the impact of integration.

1. INTRODUCTION

1.1 Rationale

Integrating the delivery of health services has the potential to improve health outcomes while also reducing the costs of delivering the services (WHO/USAID/FHI360, 2009). Increasing efficiency and maximizing impact through integration is a core principle of the Global Health Initiative and the President's Emergency Plan for AIDS Relief (PEPFAR; Global Health Initiative, 2012. Integration is also viewed as a priority in the fight for an AIDS-free generation, because it has the potential to increase access to HIV services and provide continuity of care for those living with HIV. In particular, FP services can address some of the unique needs of HIV-positive individuals. For instance, if patients in ART clinics become familiar with and confident in the staff there, it may create an opportunity for these staff to shift the focus from disease treatment only to a broader effort to improve patients' well-being. This broader effort would include promoting uptake of and adherence to FP methods, including use of condoms.

Despite a clear rationale and the recognized benefits of integration, evidence demonstrating the impact of integration on efficiency is limited. An initial review of the existing literature included the following key findings:

- **Qualitative evidence exists at the programmatic level.** A wide range of evidence demonstrates that the integration of HIV services has both clinical and service-delivery benefits (UNAIDS, UNFPA, FHI, 2004). The evidence also shows the benefits to patients in terms of continuity of care and increased access to HIV services. A common assumption is that integration can improve program efficiency (Church and Mayhew, 2009), yet the evidence supporting this claim remains vague, notwithstanding the numerous reviews that focus on HIV integration more broadly.
- Costing and cost-effectiveness **evidence is lacking.** A Cochrane systematic review on integration emphasizes the vital need for studies of the cost and cost-effectiveness of integrated service delivery (Lindegren et al., 2012). More specifically, studies on integration of FP and HIV services across several countries (Bollinger and DeCormier Plosky, 2013) show that while increased efficiency is a common argument for linking FP/Sexual Reproductive Health and HIV services, there is a dearth of data or analyses that support this assertion. A recent literature review focused on costs and efficiency of integrated HIV and other health services is a report prepared for Integra, a five-year, Gates-funded research project. Out of 46 studies reviewed (35 peer-reviewed and 11 from gray literature), only four considered potential cost savings through provision of FP services to HIV-positive individuals via integrating FP services within prevention of mother to child transmission (PMTCT) or HIV care and treatment programs (Sweeney et al., 2012). Moreover, all four of these studies modeled costs at the national level in the context of generalized epidemics; no studies were found that empirically evaluated the integration of FP with ART services at the program level, despite the widespread existence of such programs.
- **Studies focused on measurement of outputs without adjustment for inputs**. No studies reviewed by Integra researchers compared the unit costs (inputs) of integrated versus stand-alone FP or HIV care and treatment services relative to outputs, or examined the comparative costs of different models of integration. The main barrier cited was the scarcity of cost data from low- and middle-income countries.

1.2 Objectives of this study

Based on the above literature review and gaps in evidence related to the cost of integrated service provision, the HFG project conducted this study examining the costs and efficiency of FP and ART services integration at the request of USAID's Office of HIV and AIDS and the USAID Zambia mission. The main purpose of this activity is to identify potential indicators to assess and measure efficiency from FP/HIV integration across different models and over time (although the study did not address this final purpose). Programs could then use the information provided by these indicators to devise or evaluate options and strategies for improving efficiency.

The specific objectives of the proposed study are to:

1. Propose indicators that can measure the performance of the integrated programs with respect to efficiency and assess their practical feasibility
2. Quantitatively assess the relative efficiency of different models of integration of HIV and FP services using the defined indicators
3. Qualitatively identify potential barriers and facilitators to efficiency improvement

The indicators, methodology, and analysis generated by the study will support the following end-user objectives:

End user objective is to ….	**Study will support this objective by ….**
Support integration policy	Identifying a potential set of indicators to accompany integration activities and measure their efficiency
Support decision-making, programming, and budgeting around the choice of integration models	Identifying the differences in potential efficiency gains and efficiency pathways across integration models
Shape integration programs	Identifying barriers and enabling factors in implemented integration approaches

Furthermore, this study aims to support any programs benefiting from increased funding for the acceleration of FP/HIV integration activities in five selected African countries (Malawi, Nigeria, Tanzania, Uganda, and Zambia). The indicators this study finds most feasible could support the programming of the supplemental funding by identifying sources of efficiencies related to FP integration into HIV and AIDS services—e.g., reduced labor costs, or gains from shared overhead costs—and evaluating whether efficiency gains have been achieved.

HFG worked with the FHI360-led Zambia Prevention Care and Treatment Partnership (ZPCT-II) project and the CIDRZ. These implementing partners' programs supported the sites where FP services had been integrated into HIV clinics.

1.2.1 Background on FP/HIV integration in Zambia

In Zambia, the HIV epidemic is one of the country's major public health problems. According to 2014 estimates from the Joint United Nations Program on HIV and AIDS, adult HIV prevalence is 12.4 percent[1]. Ninety percent of new HIV infections in Zambia are driven by structural and biomedical

[1] http://www.unaids.org/en/regionscountries/countries/zambia/

factors such as multiple and concurrent sexual partnerships, mother to child transmission, low and inconsistent condom use, low levels of male circumcision, and mobility and labor migration.

The Zambian government's rapid scale-up of HIV prevention, care, and treatment services has been among the most successful in Africa, and now, 10 years into the response, major progress has been made. Access to lifesaving antiretroviral treatment has been expanded, prevention programs are making impact, and losses in life expectancy have begun to reverse (Republic of Zambia, 2014). These achievements are the result of strong political leadership and considerable financial support from international donors such as PEPFAR.

The Zambian government recognizes that FP is an important HIV prevention strategy, particularly with unmet need for FP estimated at 27 percent, and supports work that strengthens FP and integrates it into HIV clinical services. In Zambia, the HIV prevalence in women ages 15–49 is slightly higher than for men in the same age range; therefore, there is a significant population of HIV-positive women of reproductive age with FP needs. Two such projects were included in these analyses; they are each described next.

1.2.2 The Zambia Prevention, Care and Treatment Partnership II project

With a budget of $124 million funded by PEPFAR over five years (2009–2014), the Zambia Prevention, Care and Treatment Partnership II project (ZPCT II) supported the Ministry of Health (MOH) in strengthening and expanding HIV clinical and prevention services in six provinces—Central, Copperbelt, Luapula, Muchinga, Northern, and North Western. Working in more than 380 health facilities, ZPCT II provided technical and management support to improve and scale up PMTCT, counseling and testing, and clinical care services that include ART and male circumcision. The ZPCT II project represents a large-scale HIV program where FP/HIV integration happened in an intentional and substantial way.

Preventing unintended pregnancies among women living with HIV is a priority intervention in Zambia's national PMTCT guidelines, which are based on the World Health Organization's guidelines for preventing vertical transmission of HIV. A key objective in the guidelines is "To reduce the unmet need for FP by 50 percent from the current levels of 27 percent by 2015[2]." ZPCT II supported the MOH in rolling out the government's HIV guidelines by including FP as an element of the project.

ZPCT II used a referral-based model of FP/HIV integration. At most sites, FP counseling was integrated into voluntary counseling and testing (VCT), PMTCT, and ART services, and women who desired an FP method were referred to the FP provider on site. An OSS model was also piloted in some sites. The project supported the integration of FP and HIV services by incorporating content or messages on FP into its core project activities, mainly through training of providers. Modules on FP were incorporated into ZPCT II-supported trainings for VCT, PMTCT, and ART providers. FP providers were also trained on VCT (but this study does not address the integration of VCT into FP services).

1.2.3 The Center for Infectious Disease Research in Zambia

Since its inception, CIDRZ has been an active partner of the Zambian Ministry of Health, and today its health care service programs support more than 330 Government of Zambia clinics located in all 10 provinces. CIDRZ activities include HIV prevention and treatment; combatting tuberculosis; promoting

[2] https://www.k4health.org/sites/default/files/National%20PMTCT%20Protocol%20Guidelines.pdf

women's health and newborn and child health; community health promotion; and health systems strengthening.

With the generous help of many international donors, most notably PEPFAR through the U.S. Centers for Disease Control and Prevention in Zambia, CIDRZ has been able to support the financial and technical local ownership of service delivery in collaboration with the government of Zambia. Its focus is on increasing access to quality health care, and strengthening complementary and integrated health services across a range of priority areas.

The CIDRZ HIV prevention and treatment program service units support public health facilities in 3 of the 10 provinces of Zambia (Eastern, Western and Lusaka) by offering PMTCT services and HIV prevention, care, and treatment services; and by integrating HIV testing services in other disease screening and treatment programs.

CIDRZ support in these provinces focuses on HIV counseling and testing, HIV care and treatment, and condom distribution, as well as training in the use of condoms both for HIV prevention and for FP. At the health center level, the support includes: payments in the form of stipends to counselors; training of clinical staff and counselors at different levels and on different subjects; mentoring and technical support for health workers in the implementation of the National Health Guidelines; and support with drug logistics (contracting and procurement).

2. METHODS

2.1 Study design

The specific service-delivery points this study considers are the ART clinic and the FP clinic. The unit of analysis with respect to efficiency for this study is the ART clinic. However, those two clinics do not stand alone, and are part of a larger health facility[3] providing other medical services.

An ART clinic could be using one of two models of integration:

- The IR model, where FP counseling is offered in the ART clinic but the patient is referred to the FP clinic (in the same facility) for further FP services (particularly method provision).
- The OSS model, where the patients are counseled on FP in the ART clinic and can also receive some FP methods from that same clinic. The FP methods most commonly available include short-term methods such as pills, condoms, and injectables, but in rare cases implants and IUDs are available.

This study is a cross-sectional, non-randomized comparison of the efficiency of OSS and IR. The integration process and its possible outputs (X1 to X6) are shown in Figure 1.

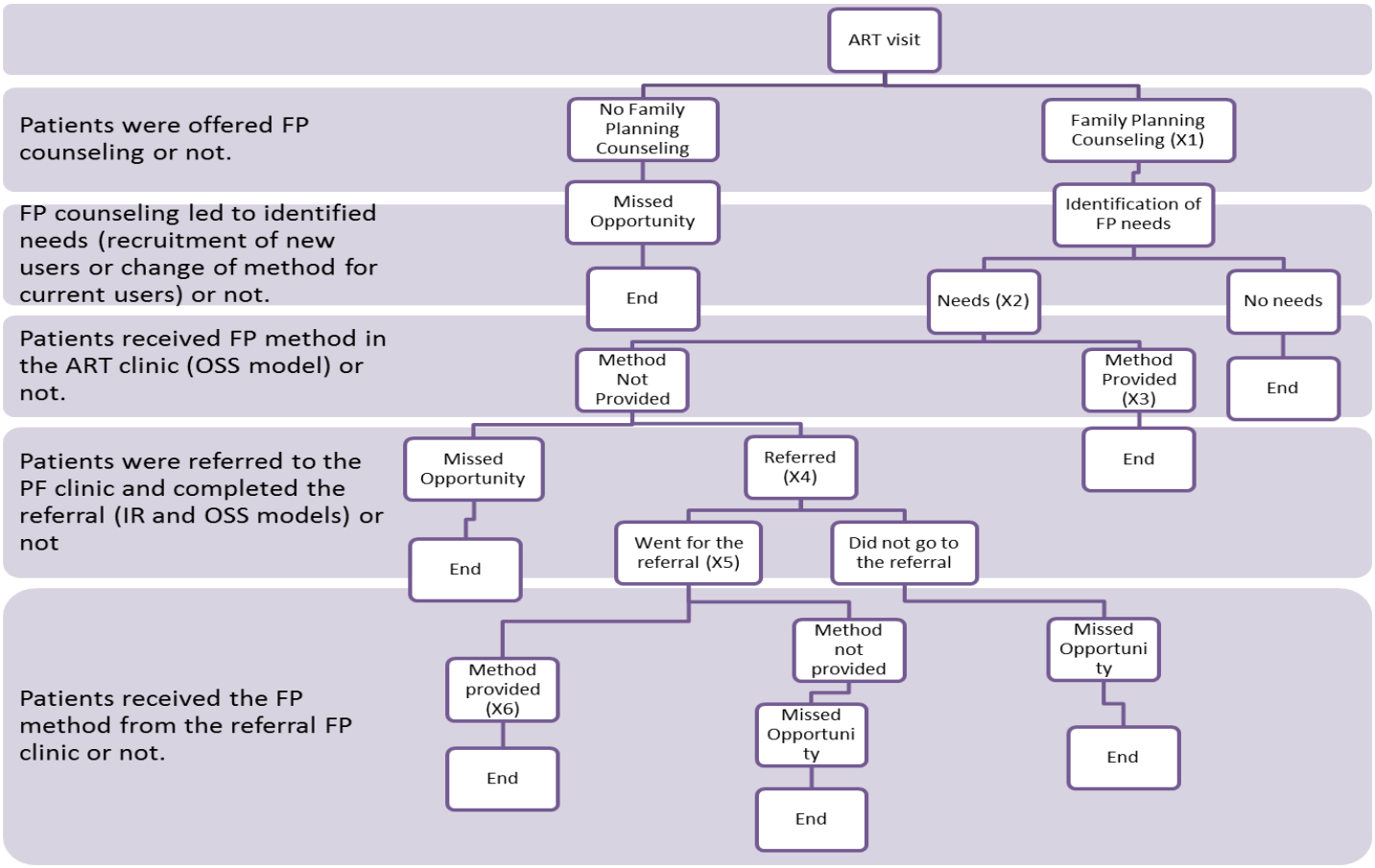
With integration of FP services, a visit to the ART clinic can have several possible different results:

- FP counseling is provided, and X1 represents the number of those services, regardless of the model of integration.
- FP counseling leads to identified needs—the patient wants to start using an FP method, or wants to change her current FP method; X2 represents the number of patients with identified needs for both models of integration.
- If we have the OSS model, the number of patients who received the FP method they expressed a need for will be measured by X3. If that method is not available, the patient should be referred to an FP clinic, and the number of referrals made is represented by X4.
- If we have the IR model, the patients are directly referred to the FP clinic and X4 also represents the number of referrals made.
- For both models the number of completed referrals—patients who goes to the FP clinic for the services they were referred for—is represented by X5.
- The number of patients who received the desired FP method (or another method) from the FP clinic is represented by X6.

At each level of output, when the corresponding service is not offered as it should be, there is a missed opportunity.

Figure 1: Framework of integration of ART and FP services

[3] All of the health facilities were health centers that offer ART and FP services but also other health services (child care, immunizations, maternity care, general outpatient care, etc.). Although the focus of this study is ART and FP, some data about the facility in general were collected (general information, staff, activities, etc.).



2.2 Site selection

Three districts, in three different provinces, were purposively selected in consultation with the implementing partners ZPCT II and CIDRZ. Taking into account time and resource constraints, the analysis is based on a sample of 10 facilities. The three districts visited were Lusaka, Mongu and Kabwe. The implementing partners' presence and type of support differed across districts, so in each district the facilities visited were chosen based on the implementing partner recommendations.

TABLE 1: STUDY SITES BY IMPLEMENTING PARTNER

Implementing partner	Province	District	Facility name	Integration model
ZPCT-II	Central	Kabwe	Natuseko	IR
	Central	Kabwe	Katondo	OSS
	Central	Kabwe	Mahatma Gandhi Memorial	IR
	Central	Kabwe	Makululu	OSS
	Central	Kabwe	Kasanda	OSS
	Central	Kabwe	Ngungu	IR
CIDRZ	Lusaka	Lusaka	George Clinic	IR
			Matero	IR
	Western	Mongu	Liloyelo	IR
			Limulunga	IR

This purposive selection was necessary given the relatively few sites where partners have introduced the formal integration of FP into HIV services. While this approach limits the generalizability of the findings, the results will still provide valuable information on the differences in efficiency of the different models of integrated services in an early stage of development.

2.3 Data collection process

Cost and utilization data were collected for a period of 12 months from October 2013 to September 2014. Retrospective data collection took place in January–February 2015. One research coordinator and 10 data collectors attended a two-day training session before data collection. After the training the tools were piloted in Chipata Health Center in Lusaka. Data collectors were divided into three teams, one for each of the three districts. The research coordinator supervised each team during data collection in their assigned district. Data from surveys were double-entered in Microsoft Excel; other data were collected in the field by the research coordinator and directly recorded, including qualitative interview results.

The study used six main data collection methods: medical record review, patient exit interviews, length-of-visit assessment (time-motion study), provider time use assessment, semi-structured interviews with health care providers, and a service cost assessment. The different components of the data collection process are intended to provide information on the inputs and outputs of the integration process, and are each described below.

2.3.1 Medical record review

A sample of patient records from the ART clinic were reviewed to determine the number of ART patients counseled on FP or provided with an FP method in the ART clinic, as well as the number referred to the FP clinic for further services. This record review enabled an assessment of missed opportunities—patients who did not receive any FP counseling services at the ART clinic—and to establish whether the referral made to the FP clinic had been completed.

Assessing the number of completed referrals and counter-referrals necessitated gathering information from the FP clinic about HIV patients who were referred there. (A counter-referral is defined as a process by which the service provider at the receiving service sends the patient back to the referring service with adequate information about the services provided at the receiving service.) Unique identification numbers were not available in the visited sites to match patients from the ART and FP clinics, so, where necessary, information contained in sampled patient records (name, address, date of visit) was used to do the match. The resulting data made it possible to determine how many of the referred patients went to the FP clinic following their referral and which services they received.

For the medical record review, all ART patients that received services during the study period were considered. For each ART clinic, a random sample of 90 patient records was selected.[4]

To be part of the sample a patient had to meet the following inclusion criteria:

- Women 18–49 years of age

[4] The sample size was calculated based on a chosen confidence level of 95%, and a desired margin of error of 10% for the sample proportion of ART patients who received FP counseling. In the absence of a Zambia-specific prior study on the proportion of ART patients counseled on FP, we used the finding from a study in Kenya (Adamchak et al., 2010) where this proportion was 38%. With α= 5%, p=0.38 and E=10%, the formula is $n >= (Z_{\alpha/2})^2 * ((p\ (1-p)/E^2)$.

- Actively on ART or pre-ART
- Made at least 1 visit to the ART clinic since the integration of services started

Because records were paper-based, the desired 90 cases were divided by the number of shelves or file cabinets (however the records were stored), and the corresponding number of records were pulled out and reviewed against the above sampling criteria. This process was repeated until the desired sample size was achieved, upon which each record was used to fill in the extraction template. This process took place in the ART clinic.

For the patients referred to the FP clinic in our sample, there were two scenarios:

1. The ART clinic kept records of the counter-referral in individual patient files, and we could recover the information about what services had been received at the FP clinic.
2. The ART clinic did not keep a record of counter-referral, and we needed to find out at the FP clinic whether the referral had been completed and what services had been provided.

For the second scenario, the data collection team went to the FP clinic to recover the matching information and complete the extraction template (note that the FP referral should be internal, occurring within the same health facility).

2.3.2 Patient exit interview

For the ART providers that were offering some level of FP services (counseling or method provision), we assess the extent to which those FP services were being systematically provided. The patient exit interview was designed to determine whether/when FP counseling had been offered, as well as to provide estimates about the percentage of missed opportunities. Trained interviewers asked patients exiting from an ART visit specific questions related to the services offered during the visit on that day or on previous visits to the ART clinic.

The study team aimed to interview 15 patients per site; the number was chosen based on an interview time of 25 minutes per patient and a one-day data collection period. Patients were randomly selected on the day of the site visit and approached to complete the patient exit interview. Systematic sampling was used, which involves the selection of every "nth" case from the target population. On the data collection day, the team randomly selected every second or fourth[5] patient arriving for care, so that the data collection team had enough time to finish one interview before the next patient was ready to be interviewed. The survey was administered to women 18 to 49 years old[6] who agreed to the exercise after having been read the consent form and been given a chance to ask questions; males and younger/older women were excluded.

Once a woman had met the inclusion criteria and given her informed consent, she was given a card with a number between 1 and 15. This study-unique identification number was used to link the patient exit interview and the length-of-visit assessment (discussed next). After the ART visit, the paper-based exit questionnaire was administered to that patient.

[5] The exact chosen number depended on the patient flow at that particular site.
[6] Based on judgment of interviewer and then verified during informed consent.

2.3.3 Length-of-visit (time-motion) assessment

The time that providers spend for ART visits that include FP counseling and/or service provision is important in order to assess staff time needs that may arise from integration. A time-motion study instrument was used to determine how long an average visit lasted (with and without FP counseling and/or service provision) under the integration model in each site. Once a patient selected for the exit interview had entered the consultation room for her visit, the time (rounded to the nearest minute) was recorded on the time-motion sheet; the time she exited the room was recorded as well.

2.3.4 Provider use-of-time assessment

Determining how busy providers are at different times of the day may help determine how best to organize work under integration. For example, information on how busy an ART clinic is can suggest whether the provision of certain FP methods is possible in the clinic, or whether it will delay patients' access to regular ART care, so that it might be better to rely on referrals unless additional staff resources can be made available. These data could also provide insight into possible changes in operating hours that could improve efficiency. A provider time-use sheet was developed to capture information on the flow of patients. The sheet recorded the number of ART patients in the waiting area at different times of the day (early morning, noon, midafternoon) to determine when the provider was likely to be very busy, reasonably busy, or not busy.

2.3.5 Semi-structured provider interviews

At each site, key informants were asked to participate in the study after giving their informed consent. For each site, the key informants included the responsible officer-in-charge of the ART clinic, and a health worker directly interacting with patients in the ART clinic. In total 20 staff were interviewed. To minimize work disruption at the clinic, the interviews were held after operating hours or at the convenience of the interviewee. A standardized guide was used to lead points of discussion. The interviewers took notes and digital audio recording was also used.

Information was collected on the following: current integration intervention (how it works, what training was offered, etc.), the successes of integration, potential barriers to effective integration, and perceived gaps and challenges to completing effective referrals. The results from these interviews were used to document the integration process, its successes, and its challenges from the provider perspective.

2.3.6 Costing data

One potential way of assessing efficiency is to compare the unit cost of each integration model. In order to calculate the unit cost, information is needed on both the inputs and the outputs of integration. Input information (quantities of inputs used, and prices) was mainly for labor, drugs, and medical supplies. Input types were fairly consistent across sites and were collected either from the health facility or from the District Medical Office (both for the ART clinic and FP clinic).

However, depending on the model of integration, we can have different outputs from the ART clinics:

- Number of patients counseled on FP (both models)
- Number of patients provided with an FP method (both models)
- Number of patients referred to an FP clinic (IR)
- Number of patients who completed referrals (IR)

A costing template was developed that captured information on the level of inputs and outputs in each health facility (health center), with a particular focus on the ART and FP clinics. The template also collected general information on the health facility (ownership, partners, etc.) and on integration training provided to the staff. HMIS data were the main source of information on the overall use of the health facility and the ART and FP clinics. The team also reached out to implementing partners (ZPTC II and CIDRZ) to obtain the costs related to training and supervision for the integration intervention.

Estimates of unit and total cost of providing each type of services for ART patients were calculated for each site. The cost of services received by ART patients (ART, FP counseling, FP method provision) is assessed in three steps. First, the volume of services provided is assessed. Second, total cost is determined using a mixed approach:

- A top-down approach is used for labor[7] cost, where the volume of patients seen in the ART clinic (compared to the facility as a whole) is used as the allocation criterion.
- A top-down approach is used for the training costs. For example, when staff from different health facilities attended the same FP training session, the corresponding cost to each facility depended on the number of staff from that facility who participated in the training. Training costs were depreciated over 5 years.
- A bottom-up approach is used for drugs and medical supplies. In most facilities the drugs and supplies procurements are centralized, and we needed to consider only the part of the costs that could be allocated to ART and/or FP care service-delivery points.

The third step is the calculation of unit cost by dividing the total cost of each type of service by the volume of patients. The cost per patient provided with an FP method from the FP clinic is calculated similarly, but only for short-term methods.[8]

It is important to note that FP services provided in the ART clinic were not always accurately recorded. Where the volume of those services was not available from the HMIS to calculate unit cost as described above, the percentage of patients who received those services was estimated from the exit interview and HMIS data on the number of patients under ART treatment. In estimating these numbers it was assumed that 54 percent of those ART patients were the target population for FP services, because in Zambia that is the estimated proportion of adult women (as opposed to men) among those receiving ART (Republic of Zambia, 2014).

[7] Supporting and administrative staff.

[8] So that we can compare the IR sites to the OSS sites, which offer short-term methods only.

Table 2 summarizes the data collection process described above, linking the array of variables needed and the different metrics used to assess efficiency.

TABLE 2: STUDY INDICATORS, VARIABLES, AND DATA SOURCES

Indicators	Variables	Data sources
Unit cost per ART patient	• Input quantities and prices (labor, drugs) • Volume of ART patients	• Service costing data collection • Record review data • Length-of-visit assessment
Unit cost per ART patient counseled on FP	• Input quantities and prices (labor, drugs) • Integration training and supervision costs • Number of ART patients counseled (whether by same provider in OSS or another provider after IR)	• Service costing data collection • Record review data • Length-of-visit assessment •
Unit cost per ART patient provided with an FP method	• Input quantities and prices (labor, drugs) • Integration training and supervision costs • Number of ART patients provided with FP (by type)	• Service costing data collection • Record review data • Length-of-visit assessment
Percentage of missed opportunities	• Missed opportunities determined through ART patient age, marital status, sexual activity, current contraceptive use, fertility intentions	• Client exit interview • Record review
Input/output ratios	• Input quantities and prices (labor, drugs) • Number of counseling sessions • Number of methods provided (by type) • Number of referrals made/completed*	• Service costing data collection • Record review data • Length-of-visit assessment •
Barriers or enabling factors to FP-ART integration	• Reports of commodity, drug, and test kit stock-outs • Availability of necessary materials and equipment • Provider time use • Training provided • Frequency of supervision	• Provider time use assessment • Manager interviews • Health worker interviews

* For IR model only.

2.4 Data analysis

Cost data were calculated in Microsoft Excel. Survey and patient record data were analyzed using Stata 12.0 Means and proportions were calculated for the sample population, and a robust standard error adjusting for clustering at the facility level was used to calculate the 95 percent confidence interval when aggregating averages across facilities. For the statistical analysis of the collected data, in addition to descriptive statistics, a non-parametric Mann-Whitney U test is used to compare quantitative variables across implementation models. Tests will be considered statistically significant if the two-sided p-value is below 0.05.

For semi-structured interviews, the more frequently identified obstacles and facilitators to the integration process are presented and discussed below.

2.5 Ethics

This study received approval from the Abt Associates Institutional Review Board and from the ERES CONVERGE Institutional Review Board in Lusaka, Zambia. Oral informed consent was obtained from patients as well as from facility managers and workers on the day of the study team's visit and before any interviews took place. All respondents gave oral informed consent before being interviewed and/or observed. The study was judged to have minimal risk to participants, and most data collected were not sensitive in nature. Written interview papers did not include the names of the patient or the individual staff member, and paper interviews were identified with a code for the district and the health facility where applicable.

3. FINDINGS

In the first section of this chapter we discuss the process of obtaining each proposed indicator and the feasibility of collecting the data needed to assess those indicators; no results other than the feasibility of collecting the data needed for each indicator are discussed. In the second section, results are discussed and compared across integration models for the indicators that were feasible given the data situation in the visited clinics. The final section presents some barriers and facilitators to integration that were identified through conversations with providers.

3.1 Objective 1: Propose efficiency indicators and assess them for feasibility

In production terms, health service-delivery programs combine a fairly standard set of resources in a production function to produce an output. Specifically, the inputs include labor, infrastructure, overhead costs, drugs and supplies, etc. Measurements or indicators based on the production process should provide information on how different models achieve different results, by highlighting differences in their production process.[9] Furthermore, analysis of these indicators should indicate bottlenecks in the production process that prevent programs from maximizing their results, and which could lead to efficiency gains if addressed.

A conceptual framework was developed to describe how integration might lead to different opportunities to increase the efficiency with which services are delivered, either by lower costs/inputs required to deliver a given number of outputs or services delivered (lowering costs), or by increasing the number of services that can be delivered with a given set if inputs (increasing productivity). The framework also identified which inputs/outputs could be combined into indicators of efficiency. Given the integration program being examined in Zambia, three indicators were chosen to compare efficiency across models of integration of FP into HIV services that mainly concern the direct inputs used in the FP/ART service production process and the outputs that result from the process. Recall, however, that depending on the model of integration, we can have different outputs from the ART clinics:

- Number of patients counseled on FP (both models)
- Number of patients provided with an FP method (OSS)
- Number of patients referred to an FP clinic (IR)
- Number of patients who completed referrals (IR)

The idea is to minimize inputs and/or maximize output to increase efficiency. The ease of influencing those inputs and outputs in the short term should also be considered when proposing the indicators, i.e., we need to prioritize the measurement of inputs and processes that we can actually affect programmatically. Therefore, given the programs being studied, the proposed indicators considered in this analysis are:

[9] In this study, the analysis will be cross-sectional, but the indicators should also be appropriate for monitoring efficiency over time.

1. *Percentage of missed opportunities at the ART clinic.* This indicator concerns the maximization of outputs in order to increase efficiency. Given the level of fixed costs that can arise in the production of care, the level of potential output is reduced each time an opportunity is missed in terms of not providing FP counseling, not providing an FP method, or not referring a patient. With a given level of inputs, not maximizing output leads to lower efficiency.
2. *Nurse/counselor time per ART patient counseled on FP or provided with an FP method.* This indicator measures a direct input that can be changed in the short term: staff time. Assuming that quality is constant (a strong assumption, which must be questioned programmatically but is not addressed in this study), the additional time spent for FP services under integration will have a cost. This cost can be measured in monetary terms if additional staff are hired or overtime is paid to existing staff. The additional time spent could also increase the waiting time of patients. Spending more time with patients means that the level of resources to serve the same number of patients is increased (again, assuming quality has not changed), and, hence, efficiency will be reduced.
3. *Unit cost per ART patient counseled on FP or provided with an FP method.* This cost measure is an indicator that combines inputs and outputs, standardizing the measurement of efficiency. For example, the cost of providing FP counseling to an ART patient can be compared across models/sites. Assuming a constant quality, the lower that cost is for a site/model the more efficient it is.

3.1.1 Process and feasibility

The feasibility of these indicators depends on (i) the data needed to calculate the indicator being available on a regular basis and (ii) the ease of compiling the data. We will analyze each of the above indicators for feasibility based on those two criteria and using our experience collecting data in the ART clinics visited.

3.1.1.1 Percentage of missed opportunities at the ART clinic

Process description

All of the chosen sites have integrated FP into ART care, and the patients are offered at least FP counseling, and in OSS sites they are offered FP methods (mostly short-term methods) in the ART clinic. A missed opportunity is defined here as not providing FP counseling[10] to an 18- to 49-year-old woman during an ART visit, or not providing an FP method to that women if a need is identified (only for the OSS sites). The percentage of missed opportunities is then calculated by dividing the number of missed opportunities found in a sample by the total sample of women considered. This metric was compiled from two sources: the patient record review and the patient exit interview.

The main difficulty in compiling the indicator from the patient record review was the lack of data on FP counseling in the ART patient records**.** All sites, regardless of the integration model, use a standard official patient record template, which has a space for mentioning whether FP information has been given to the patient. However, in 85 percent of records reviewed, the information on FP counseling was not present, and so it was not possible to determine whether the lack of information was due to a lack of reporting or whether the service was not offered to patients.

[10] That is, counselling for recruitment or to discuss satisfaction with current FP method.

The patient exit interview results helped determine whether the lack of FP counseling data in the patient record review was because the service was not offered or because the staff often failed to record it. Compared to the results from the record review, a more positive percentage of patients counseled on FP was found in each site, which suggests that rather than not systematically happening, FP counseling in the ART clinic is not systematically recorded when it happens. This, in turn, suggests a problem with the information system in the context of integration.

Feasibility of the indicator

From these two separate data collection exercises, it is clear that with the observed level of information available in the visited clinics, it will be difficult to monitor the percentage of missed opportunities. However, the indicator is not impossible to compute, as its components can be directly extracted from the routine data that should be captured in patient files; if this were done, the indicator would not necessitate a separate recording effort. The patient exit interviews suggest that the problem is a failure of facility staff to completely record all services, including FP, provided to patients in the ART clinic. Interventions are necessary to sensitize and motivate health workers to better record the needed information. In the meantime, surveys directly involving patients, such as a patient exit interview, can be used occasionally to monitor progress on data reporting in the context of integration.

3.1.1.2 Nurse/counselor time per ART patient counseled on FP or provided with an FP method

Process description

Under integration, FP services are tailored to individual patients and are given during an ART visit. This has the potential to increase the time spent with each patient, and the amount of added time spent depends on what is discussed during the counseling and/or what method is provided. The added time could impact the entire ART clinic's functioning; for example, it could potentially mean extended working hours for the staff, or a change in schedule to accommodate the extra work, longer waiting times for patients, etc. Furthermore, if staff are pressed for time, the quality of care during FP counseling and method provision may suffer, and additional work is needed to determine how much time would be necessary to efficiently provide a quality service and on measuring the quality of FP counseling and method provision.

For this study, we had no benchmark for an "efficient amount of time," but we wanted to compare the situation across different models and different sites, which is to say we measure relative efficiency. The lengths of the visits were compared for patients coming for an ART visit where they received FP counseling versus those where the patient did not receive FP counseling. The assumption was that FP counseling should be happening in the same way across models, following the guidelines provided to the clinic staff. The guidelines specify the following steps in an FP counseling session: greet the client in a respectful manner; ask about her needs, concerns, and reproductive goals; tell the patient about FP methods (different available choices, how they work, effectiveness, potential side effects, etc.); answer any questions (e.g., to clarify something); and help the patient reach an informed decision. We assume these basic steps are followed across sites (although we can check this assumption with data collected in the patient interviews), and, thus, any significant difference in the time of visit could be partially attributed to efficiency factors related to the individual facility.

Feasibility of the indicator

For this indicator, the length-of-visit assessment performed alongside each client exit interview was the main source of information. For FP counseling it was possible to use the collected data to estimate the difference in time. However, although 27 percent of interviewed patients across sites received FP counseling, only 4 out of the 150 patients interviewed were given an FP method. Because of that lack of

data, comparing the amount of time per visit for ART+FP method provision was not possible, as there was not enough information across different models or sites for different types of FP method.

In terms of feasibility, this indicator cannot be easily used on a routine basis because it requires an additional data collection effort. Moreover, to be able to compare the time it takes for method provision with the time it takes for counseling only, a larger sample size of women would need to be interviewed to capture as much diversity as possible in the FP methods provided across OSS sites. However, as with the patient exit interview activity, it could be performed occasionally to monitor integration and (with some additional observation or mystery client adaptations) could be used to monitor quality. Ideally, it should be coupled with the patient exit interview so that information is obtained on the exact services provided during that visit.

3.1.1.3 Unit cost per ART patient counseled on FP or provided with an FP method

Process description

The previous indicators looked at inputs and outputs in nominal terms that may not reveal structural differences: for example, 10 minutes of nurse-time for an FP counseling session can have a different cost in different sites. Thus, an indicator combining input and output in monetary terms can enhance the nominal comparisons. In this study we first calculated the unit cost per ART patient across sites and across models. This unit cost took into account only labor and drugs, and supplies used in the production process; infrastructure and utilities costs were not considered, since they were reasonably assumed to be the same within sites.

Adding FP counseling to a regular ART visit should increase the cost of the ART visit; it could possibly lengthen the time needed for each visit, and health workers need to be trained to provide the additional services. Moreover, depending on the way the FP counseling is provided (with or without visuals aids, flyers, or other supports), the costs can be further increased. Similarly, providing FP methods in an OSS model will add costs to the regular ART care. This extra cost depends on the type of FP method provided. For method provision the added costs will consist mainly of training, but also will include procurement costs, because with integration the ART clinic will need to have FP commodities readily available. To the extent that FP provision in the ART clinic is additional to what would have happened in the absence of integration, the additional costs would also be associated with additional benefits.

In this study, we calculate the total cost of counseling ART patients on FP and the cost per patient. As discussed previously, most of the visited sites had very poor recording of the number of ART patients counseled on FP. Thus, we had to estimate the number of patients counseled on FP using results from the patient exit interview (as explained in the methods section).

Feasibility

Assuming that all the services provided are accurately recorded, all information necessary for this indicator should be available from routine HMIS data, facility management systems (payroll data, procurement data, etc.), and from implementing partners (training and supervising costs). Therefore, this indicator is feasible, as it does not require a primary data collection effort but rather a compiling of existing data from different sources. Therefore, compared to simply using data available routinely from the HMIS, calculating the costs may require a higher level of effort.

However, unlike with the previous proposed indicators, collecting and analyzing cost data does require specific technical skills. For example, the calculations involved developing an allocation processes, because some of the costs are shared among clinics inside the facility. Some staff working in the ART clinic were also working in other clinics, and this allocation process was necessary to estimate the cost

of the time they spend in the ART clinic only. Because of the added complexity of cost indicators, it might be more suitable for an implementing partner or the MOH to collect and monitor them than for every site to calculate its own value. In addition, the process could be simplified by using templates describing which data are needed and a basic Excel or paper-based tool to conduct the calculations.

3.2 Objective 2: Quantitatively assess relative efficiency using the proposed indicators

For the interpretation of results under this objective, "patient" refers to the population of adult female ART patients that are the main focus of FP services provision and data collection for this study.

3.2.1 Percentage of missed opportunities in the ART clinic

We first discuss the measurement of missed opportunities from the patient record review, and then discuss results from the patient exit interview. As previously discussed, data on the FP services provided to ART patients were missing in many patient files. This was the case for Mahatma Ghandi, Ngungu, Natuseko, Kasanda, Katondo, and Makululu Clinics. A patient was considered to have received FP counseling if the counseling was noted on the person's patient record during the last year.

TABLE 3: LEVELS OF FP SERVICES ASSESSED FROM THE PATIENT RECORD REVIEW

Integration model	Facility	Patients counseled on FP from reviewed records[11]	Percentage of missed opportunities for FP counseling[12]	95% confidence interval	
				Lower bound	Upper bound
IR	Liloyelo	38 (42%)	58%	47%	67%
	Limulunga	32 (36%)	64%	54%	74%
	George	21 (23%)	77%	67%	85%
	Matero	35 (39%)	61%	51%	71%
	Mahatma	3 (3%)	NA	NA	NA
	Ngungu	0 (0%)	NA	NA	NA
	Natuseko	0 (0%)	NA	NA	NA
OSS	Katondo	0 (0%)	NA	NA	NA
	Kasanda	0 (0%)	NA	NA	NA
	Makululu	1(1%)	NA	NA	NA

Table 3 shows zero or very few (less than 5 percent) recorded FP counseling sessions for patients in six ART sites that are integrating FP services into ART clinics. Three of those sites are OSS and three are IR sites; furthermore, with one exception, all of those sites are located in Kabwe district. The remaining sites, which are all IR sites and supported by CIDRZ, appear to have recorded the information on FP counseling: the Liloyelo and Limulunga Clinics in Mongo; the George and Matero Clinics in Lusaka. None of the OSS sites reported FP method provision in the reviewed records. Also, very few referrals appear to have been made: only Liloyelo, Limulunga, George and Matero Clinics reported referrals[13] made to

[11] Recall that 90 records were reviewed in each site.
[12] Defined only for the four facilities with good enough data.
[13] Out of those counselled on FP.

the FP clinic. The record review on 90 patient files found 1 referral made in each of the 3 first clinics and 2 referrals made in Matero Clinic.

Overall, the lack of recorded data on FP services at the patient level does not seem to be related to the model of integration, but rather to other factors that will need to be better explored beyond this study.

Defining the missed opportunities from the patient record review was possible only for the four sites that had good enough data. Table 3 shows the results for the percentage of missed opportunities calculated as the ratio of patients who did not receive counseling to the sample of patient records. In the sites where data were available, the percentage of missed opportunities for FP counseling was more than 50 percent in every site: however, even though data are available, it is still possible that FP counseling was under-reported in patient records.

The lack of data about FP services from the patients' records appears more acute when it comes to referral or method provision than for counseling. For the sites with data on FP counseling in the records (Liloyelo, Limulunga, Georges, and Matero Clinics), the records contained information for very few FP referrals: between 0 and 3 percent of the patients counseled on FP. If indeed no referrals were made for those patients that were counseled, this could represent a second level of missed opportunities, since those women might not get the FP services for which they might have expressed a need during counseling.

There are at least two potential scenarios explaining the missing data in patient records, each of which might require a different solution:

- FP counseling is offered (systematically or not) in the ART clinic, but the health workers fail to record the counseling on the patient files. If this is the case, in order to collect data on program implementation, more effort is needed to sensitize staff on the importance of completely filling out patient records and the importance of the data that can be collected through those records. These data are not only important for program monitoring and evaluation, they are also critical for providers delivering care, who need to understand what patients have or have not done over the course of their last visits to the ART clinic.
- FP counseling is not being offered (systematically or not) to ART patients as intended by the integration intervention. This case represents a failure of the integration model, and the process of FP/ART integration should be revised; also, there should be a deeper analysis looking at the reasons why integration is not happening as planned.

The patient exit interview helps determine which one of the above scenarios is more likely the case. The information that we attempted to collect in the patient record review relating to FP services provision was also collected during the patient exit interview. The results are shown in Table 4. Recall that in each site we interviewed 15 patients; also note that not all patients needed counseling on FP.

TABLE 4: LEVELS OF FP SERVICES ASSESSED FROM THE PATIENT EXIT INTERVIEW

Integration model	Facility	Patients counseled on FP	95% confidence interval	
			Lower bound	Upper bound
IR	Liloyelo	7 (47%)	22%	72%
	Limulunga	3 (20%)	0%	40%
	George	11(73%)	51%	95%
	Matero	13(87%)	70%	100%
	Mahatma	8(53%)	28%	78%
	Ngungu	10(67%)	43%	91%
	Natuseko	7(47%)	22%	72%
Average		*59(56%)*	*46%*	*65%*
OSS	Katondo	8(53%)	28%	78%
	Kasanda	9(60%)	35%	85%
	Makululu	13(87%)	70%	100%
Average		*30(67%)*	*53%*	*81%*

Based on the results from the exit interview results, FP counseling is happening at all sites, with variation across sites. However, referrals were reported only in Mahatma (4 out of the 8 counseled patients), Ngungu (1 out of the 10 counseled patients) and Kasanda (3 out of the 9 counseled patients), with no patients at the other sites reporting having been referred for FP. For method provision, 1 patient in Katondo, 2 patients in Kasanda, and 1 patient in Makululu stated having received a method.

We also use the patient exit interview to estimate the percentage of missed opportunities. The missed opportunities analysis considers two different types of patients: those who are not using any FP methods and those currently using an FP method (not including condoms[14]). The percentage of missed opportunities is thus calculated in two different ways: the usual ratio of the number of patients not counseled on FP over the sample size for the study period, and a more refined metric calculated as the ratio of patients who did not receive FP counseling out of the patients who were identified as needing FP services based on their characteristics.[15] The refined estimation method concerns only interviewed patients who were not using any FP method at the time of the study. Table 5 and 6 below show the results of these calculations for each clinic.

[14] Condoms were excluded because they are more likely to be available in the ART clinic even without FP services integration.
[15] A woman married/with a partner or sexually active in the last 3 months, who did not want a child in the next 2 years, and was not currently using an FP method was considered as needing FP.

TABLE 5: SAMPLE STATISTICS FOR THE FP AND NON-FP USERS FROM THE EXIT INTERVIEWS[16]

Integration model	Facility name	Non-FP users in sample	Non-FP users with identified FP counseling need	FP users in the sample
IR	Liloyelo	7(47%)	2(29%)	8 (53%)
	Limulunga	8(53%)	3(38%)	7(47%)
	George	13(87%)	3(23%)	2(13%)
	Matero	13(87%)	4(31%)	2(13%)
	Mahatma	14(93%)	5(36%)	1(7%)
	Ngungu	9(60%)	3(33%)	6(40%)
	Natuseko	9(60%)	1(11%)	6(40%)
OSS	Katondo	8(53%)	2(25%)	7(47%)
	Kasanda	11(73%)	7(64%)	4(27%)
	Makululu	9(60%)	3(33%)	6(40%)

TABLE 6: PERCENTAGE OF MISSED OPPORTUNITIES FOR FP COUNSELING FROM THE PATIENT EXIT INTERVIEW

Integration model	Facility	% of missed opportunities among non-FP users*	% of missed opportunities among non-FP users with identified needs	% of missed opportunities among FP users
IR	Liloyelo	57%	50%	50%
	Limulunga	88%	67%	71%
	George	23%	0%	50%
	Matero	8%	0%	50%
	Mahatma	50%	40%	0%
	Ngungu	22%	33%	50%
	Natuseko	67%	100%	33%
OSS	Katondo	38%	0%	57%
	Kasanda	36%	43%	50%
	Makululu	11%	0%	17%

* Confidence intervals were not reported because of the small sample size for each of these three measures.

The results in Table 6 show a wide range in the percentage of missed opportunities across the sites, with four facilities counseling all respondents with identified FP needs, and one facility not counseling any of the respondents with identified FP needs. However, based on the small sample size, it is difficult to draw overall generalizations about the extent of missed opportunities, other than to say that some women are being missed at some sites. It is also problematic to compare patient interview data directly with the results from the patient records, although it can be noted that some clinics with data in patient records showed marked differences in the results between the two methods (most notably, George and Matero).

[16] Confidence intervals are not reported because of the small sample size for each of these three measures.

To compare efficiency across the two models, as measured by the percentage of missed opportunities, a Mann-Whitney U test was performed for each of the three measures of missed opportunities above.[17] At a 5 percent significance level, the null hypothesis of no difference in the population average was accepted, meaning there was no significant difference between the average percentage of missed opportunities between the two models, hence no efficiency difference.

Other factors seem to be influencing integration aside from the model used or the partner supporting the process. Results of the analysis under Objective 3, which looks at the barriers and facilitators for the integration process using providers' interviews, provide more information on the context of integration at the clinic level.

3.2.2 Nurse/counselor time per ART patient counseled on FP

In this study we compare the average time spent on an ART visit with FP counseling across models, as well as comparing the average time of an ART visit with and without counseling on FP within each model. For ART visit + FP method provision, the data collected across the OSS sites was not enough to make a comparison among those sites for each type of FP method. As reported above, only 4 patients (1 each in Katondo and in Makululu, and 2 in Kasanda) of the 45 interviewed for OSS sites received a method on that day, and they received different methods.[18]

The length-of-visit assessment/time-motion study was coupled with the patient exit interview, so that we could get time information on FP counseling provided during the visits on that day. Table 7 presents the results for the length-of-visit assessment.

TABLE 7: AVERAGE VISIT LENGTH: WITH AND WITHOUT FP COUNSELING

Model of integration	Facility name	Average visit time with FP counseling (minutes)	95% confidence interval (minutes)	Average visit time without FP counseling (minutes)	95% confidence interval (minutes)
IR	Liloyelo	12	8–6	6	5–7
	Limulunga	11	8–4	7	6–8
	George	11	10–13	10	7–13
	Matero	9	7–11	8	5–11
	Mahatma	16	13–18	12	9–15
	Ngungu	13	10–6	12	8–17
	Natuseko	11	6–15	11	6–15
OSS	Katondo	12	7–16	12	8–15
	Kasanda	14	10–17	10	7–12
	Makululu	13	10–15	11	10–12

[17] For non-FP users, the p-value was 0.43; for non-FP users with needs it was 0.3; and for FP users it was 0.9. In all cases the p-value is higher than 0.05, so we fail to reject the null hypothesis.

[18] Implants, pills, and injectables.

On average, an ART visit with FP counseling lasted the longest in Mahatma Clinic, with 16 minutes per visit, and the shortest in Matero Clinic, with 9 minutes per visit. When we look at the average time of ART with FP counseling across models, there is not a significant difference.[19] However, visits with FP counseling last longer on average than visits without counseling. A Mann-Whitney U test[20] of the average time across sites for visits with and without counseling showed a significant time difference between those two types of visits, as would be expected. The average difference in minutes was 2 minutes, with a 95 percent confidence interval of 1 to 4 minutes.

To put these findings in context, some scenario analysis was conducted. A scenario was designed to estimate the additional full-time equivalent workers that this added time could represent. If we take the upper bound of 4 minutes and consider that a patient is counseled at least 2 times a year, we will have an additional time of 8 minutes per patient per year. If we consider a hypothetical number of active ART patients of 10,000 (all the sites visited had a lower number of active ART patients; we use 10,000 to explore a maximal impact in a hypothetical clinic), and suppose that 54 percent (Republic of Zambia, 2014) of those patients are adult women and need FP counseling, we calculate that 10,000 patients x 54 percent needing FP counseling, x 8 additional minutes per year for FP counseling equals a total of 43,200 minutes per year (720 hours). Given that a full-time equivalent worker works 240 days per year and 8 hours per day, which corresponds to 1,920 hours per year, the added visit time represents a 0.37 full-time equivalent staff in a very busy clinic. For the clinics visited in our sample, the average number of patients was about 3,900. In these clinics, taking the average of 2 minutes and 18 seconds as the mean additional time per visit for FP counseling, the results suggest that about 0.08 full-time equivalents—i.e., less than 10 percent of one person's time—is required for FP counseling. Due to lack of data, we cannot calculate similar numbers for FP provision. But this scenario analysis suggests that concerns regarding staffing availability when integrating services maybe allayed with some data, or there may be a (negative) quality dimension to the estimated additional time requirements that has not been considered in this study.

3.2.3 Unit costs for stand-alone ART and FP care

The unit costs per patient for stand-alone ART and FP care presented here are direct measures that include only include labor, drugs, and supplies. ART care is provided from the ART clinic, and FP services are provided from the FP clinic for the IR model.

The cost per patient receiving an FP method considers only short-term methods: female condoms, injectables, pills. Ideally the costs should have been estimated for each type of method, because they are different in terms of costs: a dose of injectable costs more than a package of contraceptive pills. However, the activity data in the ART and FP clinics were not disaggregated in terms of number of patients provided with each type of method, but rather grouped into short-term and long-term methods. This resulting cost should then be interpreted as the average cost of method provision. Tables 8 and 9 show the total costs of separately providing ART and FP care for all sites.

[19] From a Mann-Whitney U test the p-value was 0.65. Therefore, the difference is not significant at 5%.

[20] The p-value of the Mann-Whitney U test is 0.03, so at 5% we reject the null hypothesis of no difference in average.

TABLE 8: COSTS IN ZMK FOR STAND-ALONE ART SERVICES

Model of integration	Facility name	ART drugs costs (ZMK)	ART staff costs (ZMK)	ART care total costs (ZMK)	Number of ART patients	Cost per ART patient (ZMK)
IR	Liloyelo	3,912,993	219,373	4,132,365	2,536	1,629
	Limulunga	3,786,468	162,079	3,948,548	2,454	1,609
	George	5,511,518	380,537	5,892,055	3,572	1,650
	Matero	6,690,353	345,020	7,035,373	4,336	1,623
	Mahatma	9,842,658	298,695	10,141,353	6,379	1,590
	Ngungu	3,305,059	257,656	3,562,716	2,142	1,663
	Natuseko	3,447,013	227,980	3,674,993	2,234	1,645
OSS	Katondo	5,895,720	459,110	6,354,829	3,821	1,663
	Kasanda	8,312,023	409,795	8,721,818	5,387	1,619
	Makululu	9,294,900	203,161	9,498,062	6,024	1,577

TABLE 9: COSTS IN ZMK FOR STAND-ALONE FP SERVICES

Model of integration	Facility name	FP drugs costs (ZMK)	FP staff costs (ZMK)	FP care total costs (ZMK)	Number of patients given SACM[21]	Cost per patient given SACM (ZMK)
IR	Liloyelo	9,344	31,972	41,315	924	45
	Limulunga	15,259	65,946	81,205	1,509	54
	George	91,373	320,818	412,191	9,036	46
	Matero	43,866	189,008	232,874	4,338	54
	Mahatma	10,405	32,431	42,836	1,029	42
	Ngungu	17,150	46,805	63,955	1,696	38
	Natuseko	17,696	43,020	60,716	1,750	35
OSS	Katondo	6,068	38,289	55,611	1,713	32
	Kasanda	15,523	90,615	115,652	2,476	47
	Makululu	17,881	35,704	64,544	2,852	23

For both models and all sites, the level of drug costs depended largely on the number of patients being treated. Staff costs, however, do not always increase with the facility activity level (number of patients).

The cost per ART patient per year ranged from 1,577 ZKW (250 USD)[22] in Makululu to 1,663 ZKW (264 USD) Ngungu and Katondo. Cost per patient provided with a short-term FP method ranged from a minimum of 23ZKW (3.65 USD) in Makululu to a maximum of 54 ZKW (8.57 USD) in Matero and Limulunga. There was no significant difference for the average cost per patient across models for both ART and FP care (a Mann-Whitney U test showed p-values of 0.64 and 0.21, respectively).

[21] Short acting contraceptives methods (short term methods)

[22] Using the average exchange rate for 2014 obtained from the bank of Zambia website: http://www.boz.zm/

Additional unit cost per ART patient counseled on FP

With the introduction of FP services, the cost per ART patient will likely increase. However, the lowest increase per patient treated, keeping quality constant, determines the relative efficiency of a model/site compared with another model/site. That relative efficiency can depend in part on the number of patients counseled, where increasing numbers lead to potential economies of scale. Costs are reported for a period of one year.

For FP counseling in the ART clinic, the training cost for integration is the main additional cost. As seen earlier, the cost of the added staff time with ART counseling is minimal (likely less than 10 percent of one staff person's time), and is not included in these calculations. The training costs are added to the cost for regular ART care to obtain the current total cost per ART patient counseled on FP. Assuming that this training is renewed every five years, the depreciated yearly cost was considered as the estimated financial cost of the integration. However, the integration might have other nonmonetary costs that can results from the now-longer length of visits (with FP counseling): for example, extended wait time for patients, hence potential productivity loss. These costs were not evaluated in this study. Table 10 shows the additional unit costs of providing FP counseling in the ART clinic for all sites.

TABLE 10: COSTS IN ZMK FOR FP COUNSELING IN THE ART CLINIC

Model of integration	Facility name	Training costs (ZMK)	Number of patients counseled on FP*	Additional cost per patient counseled on FP (ZMK)	Current cost per patient counseled on FP (ZMK)	Cost per ART patient under 90% counseling on FP (ZMK)	Percentage decrease in unit costs
IR	Liloyelo	2,635	639	4	1,633	1,631	0.12%
	Limulunga	2,635	265	10	1,619	1,611	0.48%
	George	7,906	1,415	6	1,656	1,655	0.06%
	Matero	6,149	2,029	3	1,626	1,626	0.01%
	Mahatma	3,514	1,837	2	1,592	1,591	0.05%
	Ngungu	3,514	771	5	1,668	1,666	0.07%
	Natuseko	8,784	563	16	1,661	1,653	0.45%
OSS	Katondo	15,811	1,100	7	1,670	1,667	0.18%
	Kasanda	5,270	1,745	2	1,621	1,620	0.03%
	Makululu	15,811	2,819	3	1,580	1,580	0.01%

*Estimated using the proportion of female adults from the total number of ART patients (54%) and the patient exit interview data.

Adding counseling on FP to ART care resulted in a maximum increase of 16 ZMK (2.30 USD) per ART patient counseled on FP per year for Natuseko Clinic, but was as low as 2 ZMK (0.29 USD) per ART patient counseled on FP per year in Mahatma Clinic.

Table 10 also shows that cost depends on how many patients receive counseling. We looked at the effect of potential economies of scale if missed opportunities were to be reduced in the ART clinics. Unit costs were calculated for a scenario where 90 percent of the ART patients were counseled on FP. In this scenario, the depreciated cost of training is spread out to more patients, so the resulting additional cost associated with FP counseling is reduced per person per year, producing the envisaged potential efficiency gains. However, the efficiency gains were under 0.5 percent of the unit cost for all clinics. This is because the additional costs of FP counseling are already small compared to the overall costs of ART with current counseling rates. (FP counseling represents an increase of less than 1 percent in the cost per person per year across all clinics.) Thus, increasing the counseling level—while worth

doing—is more beneficial to the patients in terms of continuity of care than it is to programmatic cost reduction in terms of reduced unit costs or efficiency gains. (And it should be noted that if the current level of trained staff in some clinics is not enough to reach 90 percent counseling on FP, more staff would need to be trained, so training costs would increase.)

A Man Whitney U test performed on the cost per ART patient counseled in Table 10 showed that the hypothesis of no difference among the average unit costs per model could not be rejected.[23] Therefore, we conclude that the two models of integration do not appear to have different levels of efficiency.

Additional unit cost per ART patient provided with a short-term method

The additional cost of providing FP methods in the ART clinic (OSS model) includes training costs for the staff and costs of FP commodities. Note that the training for providing FP methods is different from the training for only counseling on FP; it requires more time and costs more. The depreciated unit cost of training plus the average cost per FP commodity were added to the unit cost of an ART patient counseled on FP to obtain the unit cost of an ART patient given an FP method. Table 11 presents the cost results for FP provision.

TABLE 11: COSTS IN ZMK PER ART PATIENT RECEIVING FP METHOD (OSS MODEL)

Model of integration	Facility name	Training costs (ZMK)	FP commodities costs (ZMK)	Number of patients given a method	Additional cost per patient given a method (ZMK)	Current cost per patient given a method (ZMK)	Cost per ART patient under 90% counseling on FP (ZMK)	Percentage decrease in unit costs
OSS	Katondo	7,906	6,068	600	23	1,693	1,685	0.49%
	Kasanda	2,635	15,523	1,535	12	1,633	1,631	0.07%
	Makululu	7,906	17,881	1,768	15	1,595	1,594	0.02%

Makululu Clinic had the highest number of patients given an FP method (1,768) among the three clinics, resulting in an additional cost per patient given an FP method of 15 ZMK (2.15 USD). However, Kasanda had fewer patients given an FP method (1,535) but also had a lower additional cost per patient receiving an FP method than Makululu, because of lower overall training costs. Following the analysis presented in Table 10, the unit cost per ART patient provided with a method decreases under a scenario where 90 percent of ART patients would be counseled on FP and the current levels of method provision would be applied to them (63 percent for Makululu, 55 percent for Katondo, and 88 percent for Kasanda). The potential efficiency gains from increasing FP method delivery in the ART clinics for the OSS model are similar to those for FP counseling alone, with unit costs decreasing less than 0.5 percent across the three facilities at higher patient volume compared to current patient volume.

In the IR model, ART patients get FP counseling in the ART clinic, but if a need is identified they are referred to the FP clinic in the same facility for method provision. Thus, to estimate the unit cost of FP

[23] P-value equal to 0.73 for cost per ART patient counselled on FP.

method provision in the IR model we aggregate the unit cost of FP method provision from the FP clinic with the unit cost of counseling ART patients in the ART clinic. The results are presented in Table 12.

TABLE 12: UNIT COST PER ART PATIENT PROVIDED WITH FP METHOD: IR VERSUS OSS MODEL

Model of integration	Facility name	Cost per ART patient given a method for the IR model (ZMK)	Cost per ART patient given a method for the OSS model (ZMK)
IR	Liloyelo	1,678	NA
	Limulunga	1,673	NA
	George	1,702	NA
	Matero	1,680	NA
	Mahatma	1,634	NA
	Ngungu	1,706	NA
	Natuseko	1,696	NA
OSS	Katondo	NA	1,693
	Kasanda	NA	1,633
	Makululu	NA	1,595

The average cost per person per year of providing a method for the OSS model was 1,640 ZMK (260 USD) inclusive of ART and FP costs, and for the IR model it was 1,680 ZMK (266 USD). This difference was not significant at 5 percent: a Man Whitney U test found a p-value of 0.13. Neither of the models was found to be more efficient than the other in terms of unit cost per patient provided with a method.

However, the slightly lower average cost of the OSS model suggests that this way of providing FP services to ART patients might not cost more than the traditional referral model. Moreover, if we factor the nonmonetary benefit that this model has for patients, including continuity of care, confidentiality of care, decreased loss of patients due to FP referral not being completed, etc., then such a model may contribute to improving care and treatment for the people living with HIV, at minimal cost.

3.3 Objective 3: Identify potential challenges and facilitators to efficiency improvement

The study also intended to identify potential barriers and facilitators to integration efficiency that can be tested in future research rather than testing *a priori* hypotheses about barriers and enabling factors.

3.3.1 Barriers to integration

From the ART providers' interviews, the main challenge in the integration process was the shortage of staff. All of the health workers and managers interviewed noted that the staff is overworked and that more staff was needed to successfully integrate FP into the ART clinic. The most basic FP service that can be made available in the ART clinic is counseling, and, from the results of this study, that service is not always being offered to ART patients. From the perspective of providers, one explanation of this situation is that there is not enough clinical staff on any given day.

The provider time assessment results reported in Table 13 aimed at providing some insights about the workload for the ART staff, but the results were not very conclusive in assessing the level of provider workload. The clinic daily staffing level does not always align with the percentage of FP counseling or the reported flow of patients.

TABLE 13: FLOW OF PATIENT ON A TYPICAL ART DAY

Model of integration	Facility name	Percentage of patients counseled on FP[24]	Number of clinical staff working on a typical day	Number of patients at opening[25]	Number of patients 4 hours later	Number of patients at closing
IR	Liloyelo	47%	2	57	22	2
	Limulunga	20%	2	36	17	0
	George	73%	4	135	7	0
	Matero	87%	4	53	9	0
	Mahatma	53%	5	47	21	0
	Ngungu	67%	2	38	4	0
	Natuseko	47%	2	40	5	0
OSS	Katondo	53%	2	35	5	0
	Kasanda	60%	2	36	11	0
	Makululu	87%	4	39	7	0

In most observed clinic sites, all patients are attended to before closing time, and mornings are the busiest time of the day. By the middle of the day, the number of ART patients waiting for treatment is less than half the number of patients that were waiting when the clinic opened; but in Liloyelo, Limulunga, and Mahatma, a queue of more than 15 patients were still waiting for service at midday. None of the clinics had completely cleared the queue by midday, suggesting that all clinics are seeing patients throughout the morning.

[24] Estimated from the patient exit interview.

[25] This value represents all patients, although we would expect variation from this average on any given day.

There is little evidence to suggest that long queues cause providers to "hurry" or otherwise shorten patient visits, although our data are far from conclusive. Data collected during exit interviews show only a weak correlation between having more patients waiting at opening time and spending less time per patient (correlation coefficient -0.36, p-value = 0.30). However, the data do suggest that longer patient visits will result in somewhat longer wait times for patients. At the same time, the fact that almost all clinics had cleared their patients by the scheduled closing time could suggests that providers are capable of incorporating FP counseling into ART visits, without working overtime . Note, however, that this is in a situation where the number of female patients counseled on FP is small in some clinics (high percentage of missed opportunities).

At the time of the study, the OSS sites offered primarily short-term FP methods. Long-term methods require more training and more time per visit; if those methods were to be introduced, providers would suffer more-acute time constraints. In such a situation, referral to the FP clinic for FP methods might be more efficient, but a strong referral tracking system would be needed ensure that the services were actually provided.

Staff/human resources constraints were not the only challenge in integrating services that facility staff mentioned. There also seems to be a health worker education aspect: the importance of providing FP counseling in addition to ART care was not always well understood by the clinical staff. As an interviewed ART clinic officer-in-charge from Lusaka said:

"*The integration intervention is not quite a success, because sometimes the clinical staff fail to give the FP information to the clients. It is as if they put more importance on the ART care. If you look at the patient files you can see that.*"

In contrast, another interviewed health worker from Mongu clearly articulated the benefit of providing FP counseling to ART female patients:

"*When we give them [female ART patients] FP information, they don't conceive every other year, and they are healthier, so as health professionals we deal with more-stable patients."*

Ensuring that all staff fully understands the benefits of FP for ART patients may be one route to decreasing the percentage of missed opportunities.

Infrastructure was another barrier that the providers mentioned, mainly the ones using an OSS model—for example, inadequate levels of privacy in the ART clinic. Some providers in the ART clinic were trained to provide long-term methods, but the integration of those services had not yet started, because of infrastructure constraints, such as the lack of private space in the ART clinic. One OSS clinic mentioned stock-outs of FP methods as a barrier to effective integration.

The referral system also was seen as a challenge in the integration process. Health workers noted that they had difficulty in trying to track the patients they had referred to the FP clinic, owing to the absence of counter-referral systems between the two clinics. In all facilities, the official government generic referral form is supposed to be used to refer a patient between clinics. However, as could be seen from the data collection, that form is not always filled out when referring the ART patients to the FP clinic; or it is filled out but not included in the patient files. A formal referral tracking system was not part of the integration design and implementation, regardless of the model, and that is a potential area for improvement. Without a system that can track patients between the ART and FP clinics, it is not possible to evaluate the true impact of integration on FP method uptake for ART patients.

A pilot study from the ZPCT II project (Kasonde, et al. 2014) effectively illustrates the need for a formal tracking system in the context of integration. Between September and November 2013, ZPCT II piloted a new system of tracking referrals[26] from HIV to FP services in 15 clinics in three provinces, focusing on quantifying the number of clients referred by type of service (VCT, PMTCT, ART), and tracking the uptake of FP from such referrals. Of 8,746 clients seen in HIV services during this referral tracking period, 1,702 (19.4 percent) were referred for FP. Of those clients referred, 1,453 (85.4 percent) were reached and were seen for FP services. Over 91 percent of those that reached FP services received a method (1,327 out of 1,453). The pilot concluded that with an enhanced referral and tracking system, it is feasible to demonstrate very good uptake of FP services by clients accessing HIV services where fully integrated HIV-FP services (OSS model) are not available.

In summary, all of these challenges that providers raised are also recognized by implementing partners, as this quote from a ZPCT II project brief shows:

"The project does face challenges in ensuring that family planning remains a priority in the context of human resource constraints, budget limitations, and pressure to meet aggressive HIV-specific targets."

3.3.2 Facilitators to integration

Upfront engagement on the purpose and process of integration

A key facilitator for the integration of FP into ART care was found to be the provision of enough upfront information about the process, and adequate staff training. Integration of FP into ART care is a change in the way the clinic operates, and the process needs to be discussed, understood, and owned by the staff for it to work. Possessing enough orientation and information up front about the integration was identified as one of the necessary elements for success, and was mentioned by providers in 7 out of the 10 sites.

Desire for new skills/ability to provide better service

Providers who received training about FP integration, FP counseling, and/or FP method provision generally appreciated the new skills they acquired, and stressed the importance of this knowledge in caring for their ART patients. Depending on implementing partners and other institutions that are working with the visited sites, the types of training offered to the staff in the ART clinic differed. For example, ZPCT II-supported trainings for ART providers do not cover all technical aspects of counseling clients living with HIV and providing contraception to them. From the project perspective, time and budget limitations allowed only a basic overview of informed-choice contraceptive counseling, method options, dual method promotion, and the importance of FP for reducing unintended pregnancies. In some cases this received training was not judged sufficient by health workers, who would have liked more-formal and comprehensive training on FP services provision. Another facilitator that the providers identified was the number of trained staff inside a typical ART clinic: the higher the numbers of trained staff, the more people are able to provide the integrated service.

[26] A referral tracking form was developed and referral procedures revised; providers in HIV and FP services were oriented on these revisions.

Supervision

The supervision that the providers received from the implementing partners at the visited clinics was also acknowledged as highly instrumental in the integration process. Because adding FP services to the ART care constituted an important change in the functioning of the clinic in general, the providers appreciated the fact that the partners visited regularly to help with some aspect of implementation, especially monitoring and evaluation. A provider interviewed in Lusaka said:

"*The team from CIDRZ comes every quarter. They look at issues with supplies and staffing, and they also help us with how to enter the data." [interviewed provider in Lusaka].*

4. CONCLUSION

The goal of this study was to provide evidence-based decision making tools to the partners and government agencies that are monitoring current integration programs or thinking of integrating FP into ART. This study looks at potential indicators that could be used to assess the relative efficiency of different models of implementing integration of FP services into ART care. Indicators measuring the percentage of missed opportunities, the staff time associated with FP services, and the unit cost of providing the services were tested for feasibility in 10 health centers across three districts in Zambia. It is hoped that the indicators proposed here will help integrated programs improve their efficiency (for example, reduce missed opportunities, and hence increase output and minimize unit costs), and that the barriers and facilitators to integration identified will help in the design of future programs.

The three chosen districts were from three provinces (Eastern, Western, and Lusaka) and the chosen sites (10) were all in urban settings, so generalization of cost estimates to other parts of the country or to other countries is not possible. Furthermore, the data in general had important limitations: data were missing or unavailable at some facilities visited (such as missing stock cards in pharmacies, and incomplete patient records). Moreover, especially in the case of utilization data, the quality of the data can be questioned, which may affect the accuracy calculations and results. We have extensively documented the problems related to a lack of good and complete data in this report. In the current situation, the lack of readily available, high-quality data is the main challenge in compiling any indicator that would be used to monitor efficiency.

Efficiency was compared between an IR model and an OSS model. The proposed indicators were: percentage of missed opportunities, average time of ART visits with FP counseling, and unit cost per ART patient provided with FP counseling and with an FP method from the ART clinic. From the data collection experience and the state of information in the visited clinics, compiling any of those indicators was challenging, because the data were often missing. In theory, most of the needed data are supposed to be routinely collected or available on request, and, aside from collecting the average time for an ART with FP services visit, no additional data collection effort should be necessary to monitor integration using the proposed indicators. However, the reality is that the current HMIS processes for recording service delivery and counter-referrals at the patient level are not adequate to calculate most of the proposed indicators, and, if the efficiency of integrated services are to be monitored over time, these processes need to be improved or additional data collection efforts will be needed each time. Some of the indicators we proposed—such as the percentage of missed opportunities—are critical data needed for overall program monitoring, and are not needed uniquely for measuring efficiency. A further example: one of the major reservations staff reported having about integration of FP into ART services was the time it took; collecting data on the extent to which integration affects staffing time use would help to provide concrete evidence from which to base recommendations on how to alleviate this problem.

The study found no significant difference of efficiency between the OSS and IR method when we considered percentage of missed opportunities, visit time per ART+FP counseling, unit cost per patient counseled on FP, and unit cost per patient provided with an FP method. However, the OSS model does not seem to be more costly than the IR model, and given its added benefits for the patient, mainly continuity of care, it constitutes a desirable model for improving the care provided to people living with HIV.

From this study, the drivers of efficiency seemed to be at the facility level, not at the implementation model level. Interestingly, among clinics supported by the same implementing partners and receiving the same types of support for integration, we saw different values for our efficiency indicators. For example, Limulunga and Liloyelo clinics are in the same district and both receive support from CIDRZ for the integration process. However, the percentage of missing opportunities for non- FP users at the clinic in Limulunga was 88% while at Liloyelo it was 57%. Health facility staff reported that the common types of support received from implementing partners (training, supervision, and resources provision) were very instrumental in the success of the integration process, regardless of the model. The reported barriers to the integration intervention were a shortage of clinical staff, not having enough upfront information, and inadequate orientation on the importance of FP for ART patients.

Several findings and problems were encountered while collecting and analyzing the data for this project. Based on these problems and findings, recommendations for future work aiming at improving integration of FP and HIV services in general are summarized below.

- More effort is required to ensure that health workers systematically provide FP services in the ART clinic, as expected under integration; accurately record the services they provide at the time of delivery; and make the needed counter-referrals.
- The HMIS system should be adapted to be able to produce readily available statistics that can be used to monitor integrated services at the facility.
- An effective, formal referral system should be part of the integration program design, to strengthen program monitoring and evaluation and patient record information.

To provide more-accurate information on the impact of integration on costs in general and on the use of services, a pre-post design would have been more suitable. However, this was not possible in Zambia at the time of the study, because integration was already very widespread. HFG is undertaking such a pre-post study on efficiency of integration in Tanzania, and the results are expected to better inform us on the impact of integration.

One main limitation of this study is the small size of the sample hence the possibility that we may not have had an adequate sample size to accurately capture differences between the two models. Additionally, the analyses were based on a cross-sectional purposive sample, and confounding of the results due to selection bias or other factors is possible. The size of the sample in our study was small for the following reasons:

- In the study design, we planned to work with implementing partners already present in Zambia. From the three initial partners we targeted (CIDRZ, ZPCT II, and the Society for Family Health) only two were ready to participate in the study (CIDRZ and ZPCT II).
- Among the sites supported by these two partners, there were only a few OSS sites (ideally we would have the same number of OSS and IR sites), and we included all the OSS identified by the partners in the study.
- Even if we had a higher number of facilities, there was still a relatively small number of the population of interest amongst the ART patients in our sample (women in reproductive age who are not pregnant, and hence need FP).
- Finally we note that amongst this already small number of eligible patients, it was even harder to estimate efficiency because very few patients were reported to have actually gotten FP (further reducing the output of integration).This situation was a consequence of the weak referral/tracking systems, which we were not aware of before starting the study.

Future research in this area should take these issues into account when planning studies and calculating sample sizes needed.

5. REFERENCE LIST

Adamchak, S., G. Janowitz, J. Liku, E. Munyambanza, T. Grey, and E. Keyes. 2010. *Study of Family Planning and HIV Integrated Services in Five Countries: Final Report.* Research Triangle Park, North Carolina, Family Health International

Bollinger, L. and W. DeCormier Plosky. May 2013. *Costing of HIV Interventions: What Are the Critical Gaps?* Futures Institute. http://hiv-erg.org/publications/cte/criticalgaps.pdf. Accessed Sept 9, 2014.

Church, K., and S. H. Mayhew. 2009. Integration of STI and HIV Prevention, Care, and Treatment into Family Planning Services: A Review of the Literature. *Stud Fam Plann* 40:171-186.

Global Health Initiative. 2012. GHI Principle Paper on Integration in the Health Sector. U.S. Government. http://www.ghi.gov/principles/docs/principlePaperIntegration.pdf. Accessed Sept 9, 2014.

Lindegren, M. L., C. E. Kennedy, D. Bain-Brickley, H. Azman, A. A. Creanga, L. M. Butler, et al. 2012. *Integration of HIV/AIDS Services with Maternal, Neonatal and Child Health, Nutrition, and Family Planning Services*. Cochrane Database Syst 12:9:CD010119. doi: 10.1002/14651858.CD010119.

Republic of Zambia. 2014. *Zambia Country Report: Monitoring the Declaration of Commitment on HIV and AIDS and the Universal Access*. http://www.unaids.org/sites/default/files/country/documents/ZMB_narrative_report_2014.pdf Accessed July 9, 2015.

Sweeney, S., C. D. Obure, C. B. Maier, R. Greener, K. Dehne, and A. Vassal. 2012. Costs and Efficiency of Integrating HIV/AIDS Services with Other Health Services: A Systematic Review of Evidence and Experience. *Sex Transm Infect* 88:85-99.

UNAIDS, UNFPA, FHI. 2004. New York Call to Commitment: Linking HIV/AIDS and Sexual and Reproductive Health. http://www.unfpa.org/upload/lib_pub_file/321_filename_New%20York%20Call%20to%20Commitment.pdf Accessed Sept 9, 2014.

Waddington, C. and Egger, D. 2008. Integrated Health Services: What and Why? WHO: Making Health Systems Work. Technical brief no 1. http://www.who.int/healthsystems/technical_brief_final.pdf Accessed July 9, 2015.

WHO, USAID, FHI360. 2009. *Strategic Considerations for Strengthening the Linkages Between Family Planning and HIV/AIDS Policies, Programs, and Services*. http://www.who.int/reproductivehealth/publications/linkages/fp_hiv/en/. Accessed Sept 9, 2014.

P. Kasonde, P. Koni, T. Malebe, L. Dulli, N. Chishinga, L. Nyirenda, M. Welsh. 2014.Enhanced Systems for Tracking Referrals to FP from HIV Services—Does It Help Increase Uptake of FP Services? Poster presented at the AIDS2014 conference, Melbourne, July 20-25 2014.

ANNEX A: DEFINITION OF KEY TERMS

For the purpose of this study, the following WHO definition of program **integration** is used: (Waddington and Egger, 2008). "The management and delivery of health services so that clients receive a continuum of preventive and curative services, according to their needs over time and across different levels of the health system."

A **referral** is the process by which a health professional sends a patient to another health professional or health service for further diagnosis and/or treatment that the primary health professional is unable or unwilling to provide. For the purposes of this study we are interested in referrals between the ART services and the FP services. The service that first makes the referral is called the referring service (or point of initiation of the referral). The service to which the patient is referred is called the receiving service (or organization that completed the referral).

An **internal referral** is when the referring and the receiving services are located in the same health facility. An **external referral** is when the referring and the receiving services are located at different sites, i.e., different health facilities.

A **complete referral** is defined as one in which the patient who is referred to a specified service arrives at that receiving service according to the referring service health provider's documented instructions. A referral that does not meet the above definition for complete referral is defined as an incomplete referral, i.e., a referral not resulting in a visit to the receiving service.

A **counter-referral** is defined as a process by which the service provider at the receiving service sends the patient back to the referring service with adequate information about the services provided at the receiving service.

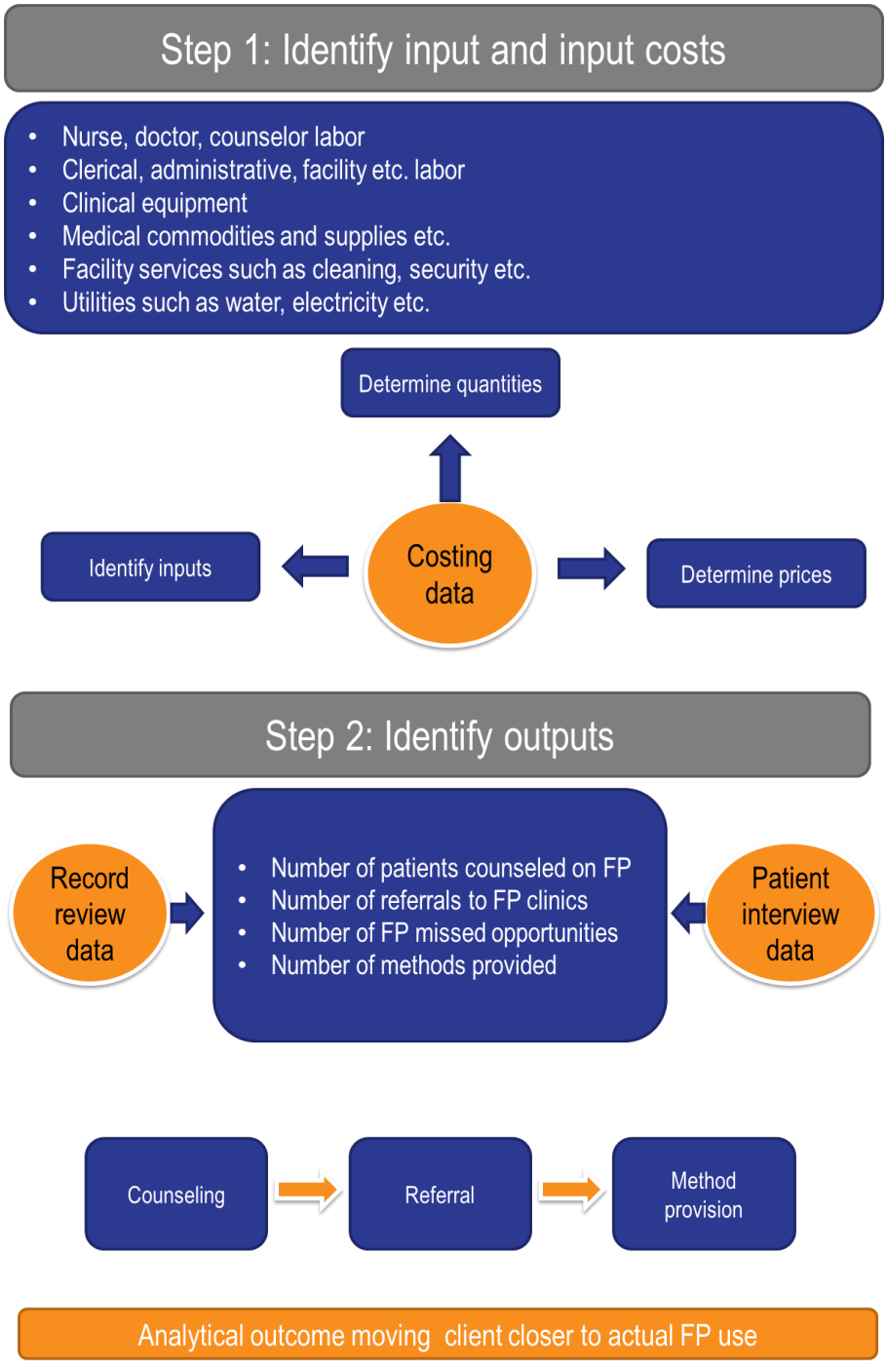
A **one-stop shop** is the process by which different kind of services—diagnosis, treatment, etc., for different conditions—traditionally provided from different point of service are combined and given at the same point of service. An example is FP counseling offered during an ART visit in the ART clinic.

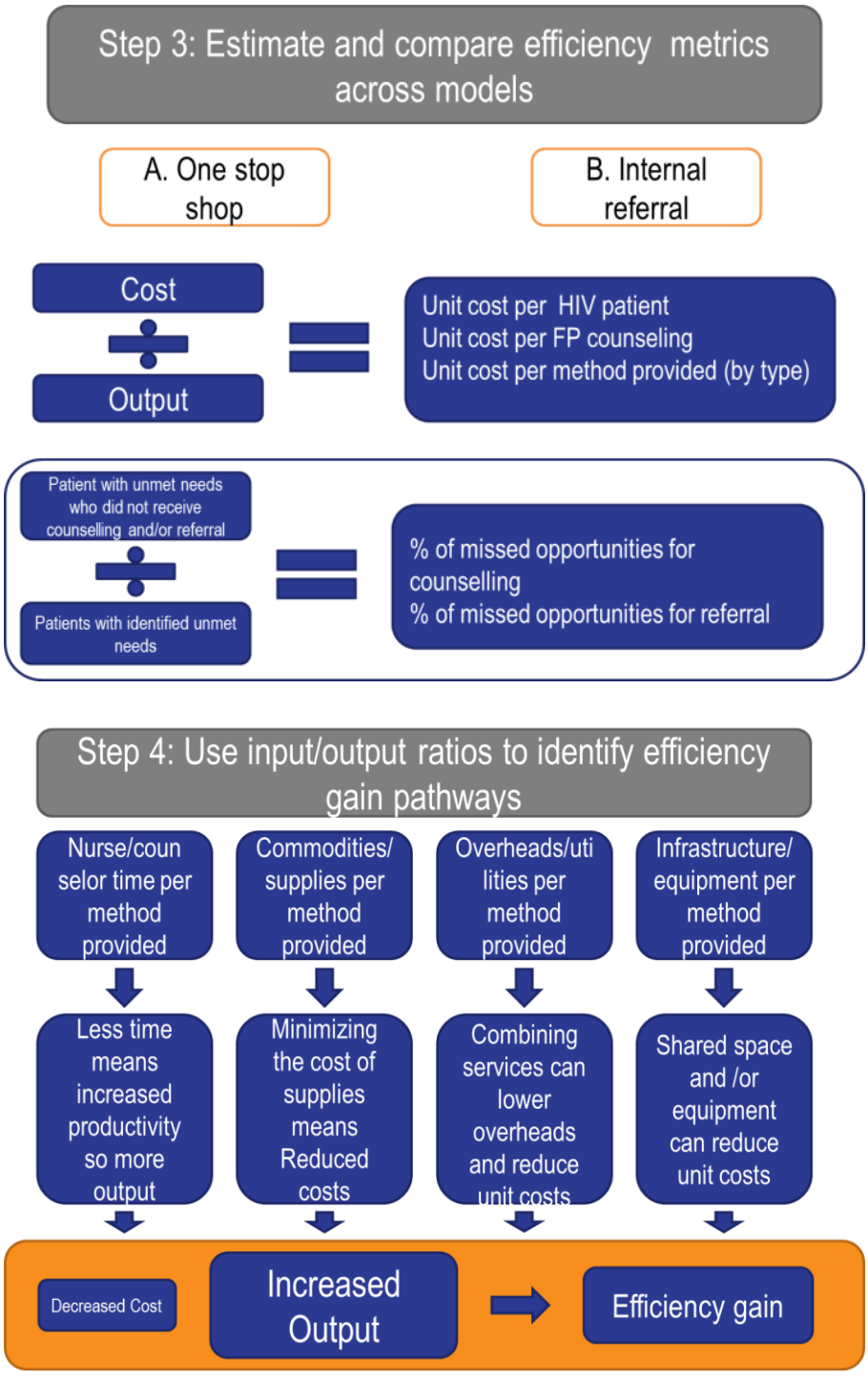
A **missed opportunity** in this study is defined as an instance in which a female HIV patient's FP needs[27] are not addressed by a provider. For a typical patient, it could be a need for information about FP, a need for first-time use of a modern FP method, or a need to change the woman's current FP method. The opportunity can be missed by not providing the FP counseling/FP method, or by failing to refer the patient to adequate FP services.

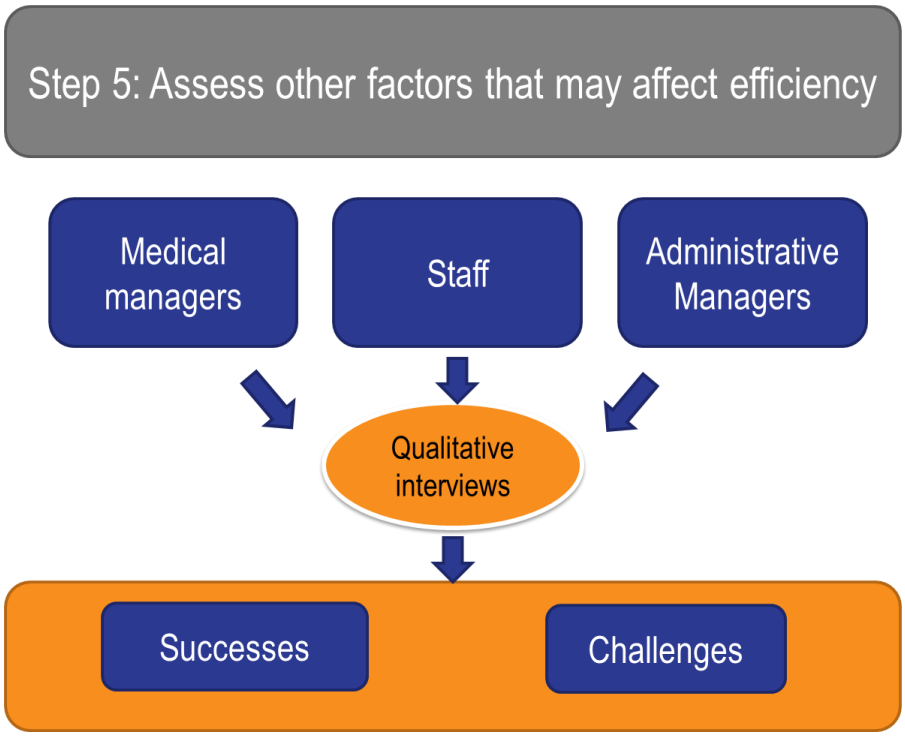
Before **efficiency** can be measured, it must first be clearly defined. One important distinction is between efficiency and changes in outputs. The literature review shows that integrating ART services with FP services can have "benefits" such as continuity in health services delivery, decreased missed opportunities, and change in health-seeking behavior. However, even with these positive results, such an intervention might not necessarily be efficient. To achieve efficiency, the *resources* used to obtain these outputs need to be taken into account as well. Ideally, ART/FP integration will have a positive impact on efficiency by either a) reducing the level of resources required per unit of output obtained; or b) producing more output per unit of resource used.

[27] Sexually active women who used no FP method and did not intend to become pregnant in the next two years.

ANNEX B: EFFICIENCY CONCEPTUAL MAP




Step 3: Estimate and compare efficiency metrics across models
A. One stop shop
B. Internal referral
Cost
Output
Unit cost per HIV patient
Unit cost per FP counseling
Unit cost per method provided (by type)
Patient with unmet needs who did not receive counselling and/or referral
Patients with identified unmet needs
% of missed opportunities for counselling
% of missed opportunities for referral
Step 4: Use input/output ratios to identify efficiency gain pathways
Nurse/counselor time per method provided
Commodities/ supplies per method provided
Overheads/utilities per method provided
Infrastructure/ equipment per method provided
Less time means increased productivity so more output
Minimizing the cost of supplies means Reduced costs
Combining services can lower overheads and reduce unit costs
Shared space and /or equipment can reduce unit costs
Decreased Cost
Increased Output
Efficiency gain


Step 5: Assess other factors that may affect efficiency
Medical managers
Staff
Administrative Managers
Qualitative interviews
Successes
Challenges

ANNEX C: ADDITIONAL PATIENT EXIT INTERVIEW DATA

Sample Descriptive Statistics

Province/district	Number	%
Central/Kabwe	90	60%
Lusaka/Lusaka	30	20%
Western/Mongu	30	20%
Total	150	100%
Age	**Number**	**%**
18–25	30	20%
26–35	64	43%
36–40	20	13%
41–50	36	24%
Total	150	100%
Sexually active in past 3 months	**Number**	**%**
Yes	109	73%
No	40	27%
Unsure	1	1%
Total	150	100%
Age of youngest child	**Number**	**%**
0–6 months	4	3%
6–24 months	31	22%
2–5 years	42	30%
More than 5 years	62	45%
Total	139	100%
Currently using an FP method	**Number**	**%**
Yes	75	50%
No	75	50%

Marital status	Number	%
Married	81	54%
Has a partner	28	19%
Does not have a partner	41	27%
Total	150	100%
Level of education	**Number**	**%**
None	5	3%
Primary	62	41%
Secondary	79	53%
Higher education	4	3%
Total	150	100%
Desire a child in the next 2 years	**Number**	**%**
Yes	40	27%
No	97	65%
Unsure	13	9%
Total	150	100%
Living children	**Number**	**%**
0	11	7%
1–3	89	59%
4–6	46	31%
7–10	4	3%
Total	150	100%
Currently using an FP method other than condoms	**Number**	**%**
Yes	49	33%
No	101	67%

Descriptive statistics for the key variables from the interview data

Variables	Sample proportion	95% confidence interval*	
		Lower bound	Upper bound
Aged between 18 and 35 years	63%	55%	71%
Married or had a partner	73%	66%	80%
Currently has a child less than 2 years old	25%	18%	32%
Sexually active in the last 3 months	73%	66%	80%
Do not desire a child in the next 2 years	65%	57%	72%
Not using any FP methods (including condoms)	50%	42%	58%
Not using any FP methods (excluding condoms)	67%	59%	74%

*95% confidence intervals based on robust standard errors.

BOLD THINKERS DRIVING
REAL-WORLD IMPACT