GUIDE

Guide for Assessing the Impact of a Total Market Approach to Family Planning Programs

Dominique Meekers
Sarah C. Haynes

December 2019

USAID FROM THE AMERICAN PEOPLE
Evidence
PATH
MEASURE Evaluation

Guide for Assessing the Impact of a Total Market Approach to Family Planning Programs

Dominique Meekers, MEASURE Evaluation, Tulane University
Sarah C. Haynes, MEASURE Evaluation

MEASURE Evaluation
University of North Carolina at Chapel Hill
123 West Franklin Street, Suite 330
Chapel Hill, NC 27516 USA
Phone: +1 919-445-9350
measure@unc.edu
www.measureevaluation.org

This publication was produced with the support of the United States Agency for International Development (USAID) under the terms of MEASURE Evaluation cooperative agreement AID-OAA-L-14-00004. MEASURE Evaluation is implemented by the Carolina Population Center, University of North Carolina at Chapel Hill in partnership with ICF International; John Snow, Inc.; Management Sciences for Health; Palladium; and Tulane University. Views expressed are not necessarily those of USAID or the United States government. MS-19-180
ISBN: 978-1-64232-211-8

ACKNOWLEDGMENTS

We thank the United States Agency for International Development (USAID) for its support of this work.

The authors are grateful for input and feedback on the content of this guide provided by Francoise Armand (Abt Associates), Andrea Bare (William Davidson Institute, University of Michigan), Christopher Brady (Christopher Brady Consulting), Cindi Cisek (Palladium Group), Rebecca Husband (Population Services International), and Amani Selim (USAID), and to Annie Rhea Cisneros and Jacqueline Willits (Tulane University) for research assistance.

Finally, we acknowledge MEASURE Evaluation's knowledge management team at the University of North Carolina at Chapel Hill for editorial, design, and production services.

Suggested citation: Meekers, D., & Haynes, S.C. (2019). Guide for Assessing the Impact of a Total Market Approach to Family Planning Programs. Chapel Hill, NC, USA: MEASURE Evaluation.

Cover photo: A family sits together beneath their home on stilts in rural Cambodia. © 2009, Edson E. Whitney, courtesy of Photoshare.

CONTENTS

FIGURES

ABBREVIATIONS

CPR	contraceptive prevalence rate
CYP	couple-years of protection
DFID	Department for International Development (United Kingdom)
DHS	Demographic and Health Survey
FP	family planning
FPET	Family Planning Estimation Tool
FSW	female sex worker
IUD	intrauterine device
KFW	Kreditanstalt fur Wiederaufbau
LAPM	long-acting and permanent contraceptive method
LARC	long-acting reversible contraception
M&E	monitoring and evaluation
mCPR	modern method of contraception
MSI	Marie Stopes International
MSM	men who have sex with men
NGO	nongovernmental organizations
NHS	National Health Survey
PATH	Program for Appropriate Technology in Health
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PSI	Population Services International
RH	reproductive health
SQHN	Sun Quality Health Network
TMA	Total Market Approach
UNFPA	United Nations Population Fund
USAID	United States Agency for International Development

1. INTRODUCTION

A Total Market Approach (TMA) is an approach to coordinating family planning (FP) services among health planners and facilities, commodity suppliers, and funders from governmental, commercial, and private or nongovernmental sectors. With a TMA, these sectors work together to increase the market for and the availability of FP services and methods. Despite the growing popularity of TMA, the development of TMA strategies and their implementation are still in their infancy. As yet, there are no studies or agreed-upon approaches to assess the medium- and long-term impact of TMA on FP outcomes. This guide will describe a step-by-step approach for evaluating the impact of FP programs that are TMA-based. As such, it is complementary to earlier guides for designing and monitoring TMA programs (Brady, Wedeen, Hutchings, & Parks, 2016; Meekers, Haynes, & Kampa, 2016a). The United States Agency for International Development (USAID)-funded MEASURE Evaluation project produced this guide to help TMA implementers to properly plan for evaluations of the medium- to long-term impact of their TMA programs, which will also enable them to contribute to the evidence base on the impact of TMA programs.

Why Is This Guide Needed?

The 2012 USAID-funded primer on Total Market Initiatives for Reproductive Health notes that there is a need for research on the long-term effects of TMA on FP and reproductive health (RH) outcomes, and that such studies may need an evaluation that continues beyond the length of the TMA implementation (Barnes, Vail, & Crosby, 2012). The 2014 primer on Health Markets for Global Health also pointed out that traditional monitoring and evaluation (M&E) approaches are not able to capture the full impact of TMA interventions, and that there is a need for a methodological approach to measure whether these interventions are working (Lin & Wilson, 2014). Because of the high cost of such long-term studies, a study design that can be applied using rapid assessment techniques is preferred.

A cursory review of project documents shows that TMA-based projects can differ in terms of their specific objectives, and in terms of their strategy and approach to achieving those objectives. Consequently, there is a particular need for evidence that demonstrates which TMA variants work best, under which circumstances each variant works best, and who benefits from each type of approach. This requires a larger evidence base. Since an increasing number of TMA implementers are collecting data to design and monitor the TMA interventions, there is a unique opportunity to design impact evaluations. To facilitate this, it is essential that TMA implementers have guidelines to plan evaluations of the medium- to long-term impact of their TMA interventions.

An improved capacity to evaluate the effectiveness of TMA programs will also help implementers make evidence-based decisions to scale up and replicate effective approaches. Expanding effective TMA approaches can help reduce the unmet need for FP, which, through improved timing and spacing of pregnancies, supports USAID's goal of reducing preventable child and maternal deaths.

What Is Impact Evaluation?

Comprehensive program evaluations can include a wide range of activities, including program monitoring, process evaluations, cost-effectiveness analyses, and impact evaluations (Baker, 2000; Ezemenari, et al., 2000; Gertler, Martinez, Premand, Rawlings, & Vermeersch, 2016). Each of these components serves a specific purpose. Program monitoring assesses whether a program is being implemented as it was originally planned, with the aim of providing regular feedback that can be used to address specific problems or weaknesses that have been identified. Program monitoring typically includes continuous tracking of inputs and outputs. Programs can address many outcomes, but only target a few of them. The latter is useful to assess whether the program

objectives were met. Process evaluations assess how a program operates with the aim of increasing efficiency, while cost-effectiveness analyses aim to assess costs of producing program benefits, compared to possible alternative programs. Finally, impact evaluations aim to measure the *net* effect of a specific program on the intended outcomes. In other words, although program monitoring may assess whether there is improvement in key outcome measures, impact evaluations are a specific type of evaluation that is interested in cause-and-effect questions. Specifically, impact evaluations are concerned with measuring to what extent improvements in key outcome measures can be attributed to the program itself, as opposed to other factors. To determine whether TMA caused changes in the outcomes, an impact evaluation method is needed that rules out the possibility that the changes were caused by other factors.

The objectives of this guide are fourfold: (1) to compare definitions of TMA currently being used by implementing organizations; (2) to identify objectives and components of current or past TMA implementations for FP; (3) to outline methods for an impact evaluation of TMA programs; and (4) to demonstrate how a TMA might be evaluated for impact using an example from Cambodia. To achieve these objectives, we conducted a systematic review of current and past implementations of TMA programs. To supplement the findings of this review, we followed up with authors and conducted key informant interviews with implementers and supporters of TMA projects.

2. SYSTEMATIC REVIEW OF THE TMA LITERATURE AND KEY INFORMANT INTERVIEWS

To understand how impact of TMA could be evaluated, it is first necessary to determine which types of activities and objectives typically compose a TMA. We synthesized findings from a systematic literature review to present an overview of the current TMA landscape. We triangulate the results from the literature review with findings from interviews with key informants.

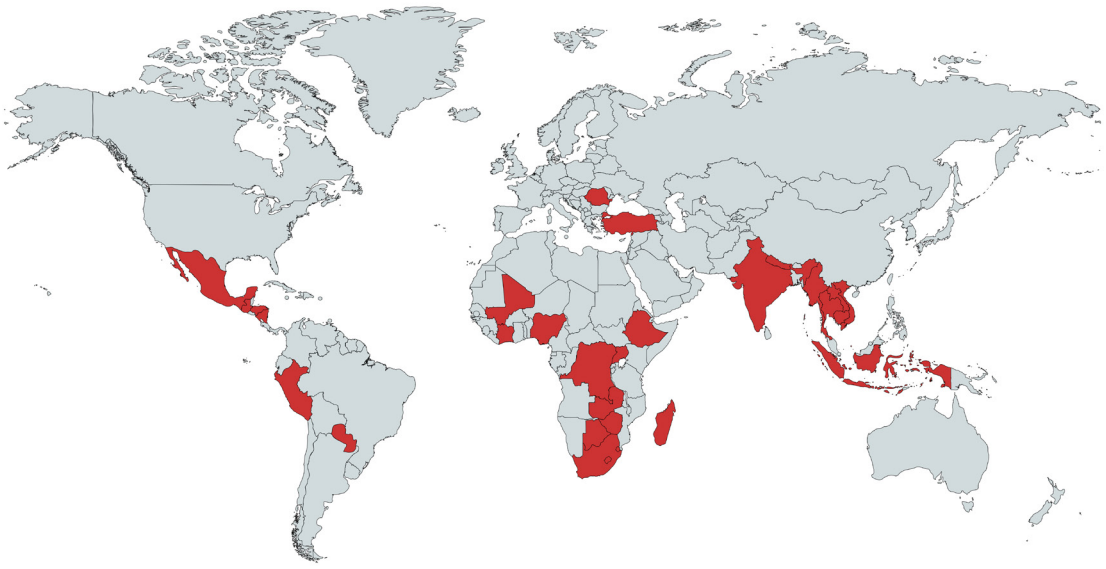
To obtain a comprehensive body of literature on TMA, we searched PubMed, Web of Science, JSTOR, Psychinfo, Google Scholar, and the Cochrane Library databases. Because such databases may not include all the gray literature (i.e., not distributed by commercial or academic publishers), we also searched the websites of all 20 organizations represented at the USAID Total Market Approach Working Group. For the literature review, we searched for all documents that explicitly referred to either "Total Market Approach" or "Total Market Initiative," in combination with "family planning." Further details about the method for the systematic review and the list of document sources are included in Appendices 1 and 2.

Of the 17 key informants who were invited to participate, eight agreed to a phone interview. Informants also answered follow-up questions by email. Key informants came mainly from nongovernmental organizations (NGOs) with experience implementing or coordinating TMA efforts. The interview guide and a list of key informants are included in Appendices 3 and 4.

Setting of TMA Initiatives

Twenty-nine countries were represented in the included studies (several studies reported on TMA implementation in more than one country). TMA was documented in 13 African countries, 9 Asian countries, five Latin American countries, and two Eastern European countries. Countries represented by more than one study included Vietnam (four studies), Nicaragua (three studies), Honduras (three studies), Paraguay (three studies), Zambia (two studies), and Indonesia (two studies). Interviews revealed that there are TMA efforts occurring in numerous other countries not found by a systematic search, confirming that much TMA documentation is not publicly available.

Figure 1. Countries represented in the systematic literature review



FP Methods

Thirteen of the 31 TMAs described in the review focused on male condoms only, usually with the purpose of promoting dual protection from both unplanned pregnancy and HIV (Barnes, Armand, Callahan, & Revuz, 2015; Evans, Taruberekera, Longfield, & Snider, 2011; Longfield, et al., 2014; Pallin, Meekers, Lupu, & Longfield, 2013a, 2013b, 2013c, 2013d; PATH, 2016). One program focused on intrauterine devices (IUDs) alone (White & Corker, 2016), and one focused only on and oral contraceptives and male condoms (Palladium Group, 2017). The remaining studies' TMAs focused on all FP methods or did not specify specific methods.

Definition of TMA

The definition of TMA varies between reports and organizations. Some definitions specify that the three sectors providing products or services (public, social marketing, and commercial) deliberately coordinate to segment the market (Barnes, et al., 2012; Brunner, Merida, Dawn Crosby, & Miles, 2014; Htat, Longfield, Mundy, Win, & Montagu, 2015; Pallin, et al., 2013a, 2013b, 2013c, 2013d). Others specify the important role of government stewardship over the process (Drake, 2011, 2013; Drake, Vail, & Stewart, 2014). Other definitions focus on market shaping but do not specify coordination or stewardship (Palladium Group, 2017; Pandit-Rajani, Cisek, Dunn, Chanda, & Zulu, 2017).

These differences were also seen in key informant interviews—with around two-thirds of participants saying that deliberate coordination was essential to TMA, while others defined TMA as a "reengineering" of the market in a way that promotes efficiency and equity, whether coordinated or not. Most publications and interviewees agree that sustainability and equity are the two main objectives of TMA, with segmentation and cross-sectoral coordination being the main tools to achieve these objectives.

Target Populations

Although many studies mentioned the use of market segmentation for improved targeting, most studies did not specify if implementation activities were targeted at specific groups. Nevertheless, there are examples of TMA programs that do target specific groups. Five of the programs described targeted high-risk groups for HIV prevention, including female sex workers (FSWs) and male clients, men who have sex with men (MSM), and persons living with HIV. Such programs were implemented in Nepal, Myanmar, and Vietnam (Barnes, et al., 2015; Longfield, et al., 2014; PATH, 2016). A TMA targeting adolescents was implemented in Nicaragua (USAID Deliver Project, 2014). Targeting of specific populations was not discussed in the informant interviews, although all key informants emphasized the importance of reaching low-income and vulnerable populations.

TMA Planning Activities

Most publications described some type of planning stage prior to the beginning of TMA implementation. Planning activities included workshops, market analyses, and the creation of working groups. Initial workshops are most often organized by social marketing organizations, donor organizations, or other NGOs and may include participants from public-sector agencies, commercial suppliers and distributors, and other local organizations involved in FP.

Planning workshops for TMAs take on two forms; in some cases, workshops occur prior to any other TMA planning activities as a first step for stakeholders to discuss a potential plan. More frequently, workshops occur as a second step following market analyses to present and discuss results among stakeholders, and to generate an action plan to address the needs highlighted in the market analyses. Although many studies mentioned the creation of working groups, none included details on specific activities, meeting schedules, or accomplishments of working groups over time.

Market analyses were done using a variety of data types, including stakeholder interviews (Barnes, et al., 2012; Brunner, et al., 2014; Drake, et al., 2011; Parks, 2016), market segmentation analyses (Barnes, et al., 2012; Drake, 2011, 2013; Drake, et al., 2014; Parks, 2016; PATH, 2016; Population Services International, 2011; USAID Deliver Project, 2014), health survey data (Barnes, et al., 2012; Drake, et al., 2014; Pallin, et al., 2013a, 2013b, 2013c, 2013d; Pandit-Rajani, et al., 2017; PATH, 2016; Population Services International, 2011; USAID Deliver Project, 2014), and distribution/retail data (Barnes, et al., 2015; Barnes, et al., 2012; Drake, 2011; Drake, et al., 2014; Pallin, et al., 2013a, 2013b, 2013c, 2013d; Parks, 2016; Population Services International, 2011; USAID Deliver Project, 2014).

Intersectoral Coordination and TMA Stewardship

NGOs, including Population Services International (PSI), PATH, Marie Stopes International (MSI), the Palladium Group, and Abt Associates, most often acted as the coordinators of the TMAs described in the included studies (Barnes, et al., 2015; Barnes, et al., 2012; Brunner, et al., 2014; Drake, et al., 2014; Htat, et al., 2015; Longfield, et al., 2014; Palladium Group, 2016; Pallin, et al., 2013a, 2013b, 2013c, 2013d; Parks, 2016; PATH, 2015a, 2015b; Population Services International, 2011; White & Corker, 2016; Winfrey, 2011). Donor agencies such as USAID and UNFPA were also commonly involved in establishing coordination of partnerships, and occasionally also in implementation (Barnes, et al., 2015; USAID Deliver Project, 2014). In several cases, the government led TMA efforts. This was the case in Indonesia and Thailand, where government boards or councils deliberately engaged NGO and private sector partners to strengthen the overall FP market (Drake, et al., 2014). For some, the concept of government stewardship of the TMA was seen as essential for the sustainability of the TMA. In Vietnam, leadership of the TMA effort transitioned over time from PATH to the government (Drake, 2011; PATH, 2016).

While it is frequently acknowledged in the studies that a TMA includes government, nonprofit, and commercial sectors, only a few demonstrated active engagement of all three sectors. In Honduras, USAID worked with the ministry of health as well as two commercial condom distributors and five drug companies to create a TMA (Barnes, et al., 2012). Similarly, commercial sector and government partners were included in USAID's TMA efforts in Nicaragua, Nigeria, and Nepal (Barnes, et al., 2015; USAID Deliver Project, 2014). In Myanmar, PSI developed relationships with several commercial companies as well as with government stakeholders to create a TMA for male condoms (Longfield, et al., 2014). While there are cases where the commercial sector has acted as an active and important participant in TMA (Barnes, et al., 2015; Longfield, et al., 2014), there are numerous instances where the commercial sector was not directly involved in the TMA, either because they refused to participate or had not been engaged. In such cases, use of the term "Total Market Approach" appears overstretched, although the original intent may have been to involve all three sectors.

TMA Implementation Activities

The studies included in the review show variation in the types of activities implemented in TMAs. Activities identified in these studies included price changes, introduction of new brands, commercial sector support, segmentation and targeting, controls on public-sector products, and policy changes.

Price changes are one of the most common activities implemented in TMAs. In many countries, social marketing companies have increased the prices of subsidized condoms in an effort to avoid undercutting commercial products (Evans, et al., 2011; Pallin, et al., 2013a, 2013b, 2013c, 2013d; Population Services International, 2011). In Southern Africa and Cambodia, several social marketing brands have increased prices to achieve full cost recovery in an effort to increase sustainability (Pallin, et al., 2013a, 2013b, 2013c, 2013d; Population Services International, 2011). It is noted that promotion and marketing activities may still be subsidized. We found one example of a temporary price reduction for a commercial product (*Karol* condoms in Vietnam) to increase uptake of the brand (Longfield, et al., 2014).

In addition to price changes, introduction of new brands can be used to close price gaps and move consumers who are willing to pay away from free products. Several TMAs included the introduction of new mid-priced social marketing brands (Barnes, et al., 2015; Mahaffey, 2012; Palladium Group, 2017; Pallin, et al., 2013a, 2013b, 2013c, 2013d; Population Services International, 2011) or new commercial brands (Htat, et al., 2015; PATH, 2015a, 2015b, 2015c). We found one instance of a social marketing company partnering with a commercial company to create a mid-priced brand in Nigeria (Barnes, et al., 2015).

To engage the commercial sector in TMA efforts, some programs have provided direct support for commercial partners. In one case, a donor agency matched funds for three commercial partners' distribution of condoms to high-risk groups, increasing stock levels, and point-of-purchase promotion (Barnes, et al., 2015). In another, a social marketing company supported a commercial partner by offering product promotion and distribution (Longfield, et al., 2014). One key informant described a program in Nigeria in which the government provided free products to private FP service providers so that providers could keep costs of services low, thereby decreasing the price gap between private and public providers and providing more efficient delivery of services than would be possible with public-sector expansion alone.

Although several programs specified segmentation and targeting specific groups as a primary activity in the TMA, there is a paucity of detail on how this occurred in practice. Several TMA programs created profiles of population segments but it is unclear how these data were used (Barnes, et al., 2012; UNFPA & PATH, 2014; USAID Deliver Project, 2014). Programs in Indonesia and Turkey focused on ability to pay for FP products in their segmentation strategy (Drake, et al., 2014; Pollard, 2014). One study mentioned that better targeting of high-risk places had occurred as a result of TMA (Longfield, et al., 2014) while another mentioned better targeting of free condoms (Htat, et al., 2015). We also found one example of a donor-funded program's staff offering technical assistance to commercial condom manufacturers in exchange for increasing availability and visibility at new distribution points (Barnes, et al., 2015).

Commonly used TMA activities include:

- Price changes
- Introducing new FP brands
- Supporting the commercial sector
- Segmenting the FP market
- Targeting specific groups
- Introducing controls on public-sector FP products
- Policy advocacy

Because the availability of free products can decrease willingness to pay for products, changes in the way that public-sector products are distributed may be important for the success of TMA. In Indonesia, the government implemented a sliding price scale for public-sector FP commodities based on ability to pay in order to recover costs of products and allow commercial competition (Pollard, 2014). This same TMA effort made use of a distribution network of midwives to sell FP products at a profit. One study described how donors partnered to provide a volume guarantee to the manufacturer in exchange for more competitive pricing for public-sector purchases (McCarthy, Ramarao, & Taboada, 2015). Several programs have focused on increasing the availability of long-acting reversible contraception (LARC) methods by providing marketing, increasing stock levels, and increasing training for public and private sector providers (Population Services International, 2011; Rademacher, 2015; White & Corker, 2016). In some countries, resale of free products undermines the market for commercial products. Two TMA programs described strategies for preventing the sale of free condoms—one by adding a "not for sale" stamp on free products (Longfield, et al., 2014), and the other by improving targeting of free condoms (Htat, et al., 2015). Interviews also revealed that public-sector leakage is a main concern for many programs.

Policy change can be a powerful tool for shaping markets and increasing access to FP products. Several TMA programs have focused efforts on removing taxes on contraceptive products that negatively affect commercial sector suppliers (Barnes, et al., 2012; UNFPA & PATH, 2013, 2014). In addition, several countries' TMA efforts have focused on adding FP products to the national list of essential drugs to ensure coverage by national insurance programs (UNFPA & PATH, 2013, 2014). One program advocated a quality seal on condoms to

combat public skepticism about the quality of the condoms (PATH, 2016). Another worked to expand the definition of authorized providers of IUDs to include nurses and midwives (White & Corker, 2016). Such task shifting can be an important tool for improving contraceptive access (World Health Organization, 2013). Finally, several programs have focused on increasing government budgets for FP (UNFPA & PATH, 2014).

Monitoring and Evaluation of TMA Interventions

Many studies did not report on monitoring or evaluation for the TMA. For those that reported data, common indicators for monitoring progress included contraceptive prevalence by wealth quintiles and other characteristics, unmet need for FP, market share of each sector, total market value, number of brands on the market, total number of subsidies, and method mix (for an overview of recommended TMA indicators, see Meekers, et al., 2016a and Pallin & Meekers, 2014). These indicators were mainly measured using Demographic and Health Survey (DHS) data, although some studies also included data from retail surveys, surveillance data, and smaller survey programs. One program conducted a qualitative pre-post study to evaluate the impact of TMA from a stakeholder perspective (Brunner, et al., 2014). One program conducted a survey of men ages 18–49 to determine how brand satisfaction and loyalty differed between branded and free condoms (Evans, et al., 2011). One program used geographic information system mapping of retail data to measure coverage of FP products and services (Mahaffey, 2012). Only one study referred to a completed evaluation of the TMA, concluding that the policies implemented in the approach resulted in a change in method mix and growth of the private sector (Pollard, 2014).

Key informant interviews confirmed that TMA efforts typically do not include a formal evaluation plan or budget. Additionally, many TMA landscaping assessments have been funded without plans for further funding and without creation of an ongoing coordination mechanism.

Discussion

Many governments, social marketing organizations, and donor agencies have conducted market activities for FP that are not specifically labeled "TMA" despite being a potential TMA activity. Indeed, TMA draws upon theories and strategies long used in social and commercial marketing; market segmentation, market growth, and public-private partnerships are not novel strategies. However, TMA contributes several strengths to these existing approaches. First, TMA merges these market strategies with meaningful and deliberate coordination across sectors. Second, TMA brings a focus on equity as well as market sustainability. TMA initiatives should consider how these strengths can be used when planning and implementing activities.

One limitation to TMA research is that many TMA efforts have not been documented in published literature. Many project documents remain unpublished and are therefore not available to provide guidance to other potential implementers of TMA. For example, although key informant interviews revealed that TMA efforts have been implemented in Ethiopia, Ghana, Haiti, India, Kenya, Nepal, Nigeria, Tanzania, Uganda, and Zimbabwe, detailed written documentation of these efforts was not available. When possible, reports of these efforts should be made available to inform future TMA work.

For publications that do exist, many are planning documents that lack specific information on activities, monitoring, and evaluation. Among the documents that met the criteria for our systematic review, only one published document explained TMA objectives and activities in detail and attempted an evaluation (Longfield, et al., 2014). Key informants affirmed that TMA planning activities often do not include long-term plans for funding to implement and sustain planned coordination or activities.

Because TMA activities may vary, it is essential to define specific activities and keep track of when activities were in fact implemented. Coordination and commitment from stakeholders are important components but alone do not impact the market. Therefore, the start of TMA activities, rather than the start of coordination, should be considered the "starting point" for TMA evaluations. Creating a TMA timeline should be part of every TMA evaluation.

3. METHODOLOGICAL APPROACHES TO EVALUATING THE IMPACT OF FAMILY PLANNING PROGRAMS

How Do We Measure Impact?

Because the outcomes targeted by a program can be influenced by factors that are not caused by the program itself, it does not suffice to simply measure whether key program outcomes improved over the course of the program. The aim is to assess whether changes in the key outcome measures can be attributed to the program. To measure to what extent the program caused changes in the outcomes, an impact evaluation must compare these outcomes with what would have happened if the program had never been implemented.[1] In TMA projects, impact evaluations must therefore estimate what would have happened in case the TMA approach had never been adopted.

In technical terms, the estimate of what the outcome measures would have been in the absence of the intervention is called the *counterfactual*. Because the counterfactual cannot be observed, it always must be estimated. This involves identifying a comparison group to estimate what would have happened in the absence of the TMA approach. Such a comparison group must not have adopted TMA, but must otherwise have similar characteristics.

Program impact is then estimated by means of baseline and end line surveys conducted among both the intervention population and the comparison group. If the program was effective, we anticipate that the targeted outcome measures will improve more rapidly in the intervention location than in the comparison group. The magnitude of the program impact is calculated by measuring how much the outcome measures in the intervention population improved over the course of the intervention, compared to the improvement in the comparison group during the same period. This comparison of the trends in outcomes in the intervention and comparison groups is often referred to as a double-difference or difference-in-differences (DID) approach. If the intervention and comparison group are identical, this measure produces a valid estimate of the impact of the intervention. However, if the two groups are not completely identical, then it is possible that factors other than the intervention also affected observed changes in the outcome measures. Thus, the better the estimate of the counterfactual (i.e., the outcome measure in the comparison group), the more accurate the estimate of program impact.

> "The key challenge, then, is to identify a valid comparison group that has the same characteristics as the treatment group."
> (Gertler, 2016)

What Is an Appropriate Comparison Group?

Typically, the counterfactual is estimated by means of a comparison group (or "control" group). Irrespective of which type of program is being evaluated, the core of any program evaluation consists of identifying a valid comparison group that represents what would have happened in the program's absence. Ideally, the comparison group would be identical to the intervention group in every way (Baker, 2000; Gertler, et al., 2016; Khandker, Koolwal, & Samad, 2009). Since a perfect clone of the intervention group never exists, the challenge for all evaluations is to identify or create a convincing and reasonable comparison group. As stated by Gertler, et al. (2016), it is important that the intervention and comparison groups:

- have identical characteristics (on average), except for exposure to the program;
- are expected to respond to the program in the same way; and
- have similar exposure to other interventions during the course of the study period covered by the impact evaluation.

[1] Impact evaluations can be classified into those that measure the impact on everyone to which the intervention has been offered, regardless of whether they were exposed to or participated in the program. This so-called "intention-to-treat" (ITT) estimate measures the average impact of the program on the entire key population. Alternatively, evaluations can estimate the impact on the population that was actually exposed to program, referred to as the effect of "treatment-on-the-treated" (TOT). When everyone in the key population is exposed to the program, as is the case with TMA programs, both estimates are identical (Gertler, et al., 2016).

If the above conditions are true, the only difference between the two groups is the existence of the program. Therefore, any differences in the key outcome measures between the two groups can only be attributed to the program, thereby proving that the program caused the difference. However, when the above conditions are not fully met, differences in the outcome measures between the two groups will measure the combined effect of exposure to the program and any other differences between the intervention and comparison group.

When identifying a comparison group, evaluators must also take into account that the coverage area of the program may affect whether an appropriate comparison group is available, that comparison groups may become contaminated by movement of people to or from the intervention group, and that contemporaneous events in either the intervention or comparison group could affect their comparability (Ezemenari, et al., 2000).

Comparing Study Designs for TMA Impact Evaluation

Impact evaluations use different types of comparison groups, including—among others—experimental (randomized) control groups, quasi-experimental groups, and reflexive control groups (Baker, 2000; Gertler, et al., 2016; Khandker, et al., 2009). Some evaluations also use statistical controls. These types of controls vary in terms of their feasibility as well as methodological rigor. Because none of the methods are perfect, program evaluators must make decisions about their relative advantages and disadvantages. Program evaluators normally aim to use the most rigorous approach that is feasible, given the nature of the intervention, data availability, and other constraints (Oldsman & Hallerg, 2004). The key features of the different approaches and the feasibility of applying them to evaluate the impact of TMA is discussed below.

Experimental (Randomized) Designs

It is generally agreed that the most rigorous evaluation design consists of the experimental design, also known as a randomized controlled trial. In this design, a comparison group is created by using a random selection procedure (e.g., a coin toss) to determine who gets the intervention and who does not. Assuming the target population is large enough, random assignment of the intervention will ensure that on average the intervention group and the comparison group have identical characteristics. This reduces potential selection bias that occurs when the intervention groups has unmeasured characteristics that are related to the intended program outcomes. This is the procedure normally used in clinical trials to test the effectiveness of new drugs. In such studies, randomization is used to determine which patients receive the new drug, and which receive a placebo.

While randomized controlled trials are considered the gold standard for impact evaluation because they provide the strongest evidence of program impact, there are many instances where it is not possible to randomize an intervention at the individual level. One such example consists of programs that have nationwide coverage (Ezemenari, et al., 2000), as is the case for TMA programs. When programs have nationwide coverage, it is not possible to randomly assign individuals to intervention and comparison groups. Therefore, it is not feasible to evaluate TMA programs using individual-level randomized controls.

When programs have nationwide coverage, as is the case for TMA programs, it is not possible to randomly assign individuals to intervention and comparison groups. Therefore, it is not feasible to evaluate TMA programs using individual-level randomized controls.

While randomized controlled trials usually randomize at the level of the individual, it is also possible to randomize at a group level. In so-called cluster randomized controlled trials, groups of individuals are randomly assigned to either the intervention or comparison group. For example, for a school-based intervention it would be possible to obtain a list of all schools and to then randomly select schools that will receive the intervention and schools that will not. Using a cluster randomized controlled trial to evaluate

TMA impact would be very difficult. Since TMA programs are implemented at the country level, an evaluation using a cluster randomized controlled trial would require studying a large number of countries—and from those randomly selecting those countries that will use TMA and those that will not. Since implementing TMA requires extensive negotiation with actors from the public, social marketing, and commercial sectors, such random assignment is not feasible. Moreover, even if countries could be randomly assigned to either apply or not apply TMA, the number of countries involved would almost certainly be too small to produce intervention countries and comparison countries that would have comparable characteristics.[2] Advances in communication technology and transportation also imply that there would be a high risk of contamination of the comparison group.

Designs with a Matched Control Group

Often it is not possible to randomly assign who will receive the intervention. If a randomized controlled trial is not feasible, a common alternative is to evaluate the program using a quasi-experimental design with a nonequivalent control group (Baker, 2000). In this study design, a control group is purposively selected that has similar characteristics as the intervention group. That is, it is attempted to match the intervention and comparison group on a fixed number of characteristics.

Because the comparison group is never fully equivalent, statistical procedures are often used to correct for *known* differences between the intervention and control group. However, this implies that the relevant variables must be known and measured in advance. Because it is difficult to know all the different factors that can influence the outcome measures, this approach cannot rule out that there may be additional unobserved differences between the two groups that could affect the estimate of program impact. Therefore, the validity of this study design is affected by the extent to which the impact evaluation controls for the effect of relevant external factors that may influence the outcomes (Planning and Management Consultants, 1992).

Quasi-experimental designs with a matched comparison group are often used to assess the impact of interventions that target a relatively small region, such as a specific district. In such cases, it may be possible to identify another district that is similar (in terms of population, culture, economic status, health status, exposure to other interventions, etc.) that can serve as a control group, and to use statistical controls to correct for any known differences.

Evaluating the impact of a TMA program using a matched control group would require identifying a country that not only has nearly identical socioeconomic, cultural, and demographic characteristics, but that is also similar in terms of other relevant factors, such as its FP market (e.g., market share, etc.), fertility preferences, FP access and use, FP policies, etc. Provided the data are available, statistical methods can control for known individual level differences (such as socioeconomic status, fertility preferences, etc.). However, it can be both challenging and expensive to obtain the necessary data (including information on other interventions, etc.) for a comparison country where the TMA program is not being implemented. Moreover, statistical methods cannot control for country-level differences, such as differences in FP policies or characteristics of the FP market. Unless one can identify a comparison country that provides a very good match, and where the necessary data for statistical controls can be obtained, the impact evaluation would not be able to determine to what extent improvement in the outcome measures can be attributed to TMA, rather than by other differences. Except in very rare cases, it would therefore not be recommended to use another country as a matched control group to assess the impact of a TMA program.

[2] Theoretically, very large countries may be able to implement TMA at a subnational level (e.g., at the state level). However, as shown in the systematic review of the literature, there are no examples of countries where this has been the case.

Designs with Reflexive Control Groups

When experimental or quasi-experimental designs are not feasible, it may be possible to use nonexperimental controls, such as reflexive controls (Khandker, et al., 2009; Oldsman & Hallerg, 2004; Rossi, Lipsey, & Henry, 2018; Vedung, 2017). Reflexive controls are often used for evaluations of programs that have nationwide coverage, which rules out having a comparison group. Reflexive control designs usually do not use a separate comparison group, but rather involve a one-group design where the intervention group also serves as its own comparison group.

The most basic form of the reflexive control group design consists of using the outcome measure of the target population before the intervention as the control for the outcomes after the intervention. However, an important weakness of this pre-post design is that it cannot be ruled out that changes in the outcome measures were caused by external confounding factors, rather than by the intervention itself. Unless these external factors are controlled for, there is a risk that changes in the outcomes are incorrectly attributed observed changes to the intervention, while they were actually caused by other factors. Therefore, it is important to have a comprehensive baseline survey that includes a wide range of external factors that may influence the outcomes, including exposure to other programs (Khandker, et al., 2009). However, it is noted that unlike impact evaluations with a matched control group, basic pre-post reflexive designs do not control for secular trends in the outcomes. Therefore, simple pre-post designs that only use two data points (i.e., a baseline and end line survey) are not very useful for impact evaluation and should only be used as a last resort when more robust designs are not possible (Baker, 2000).

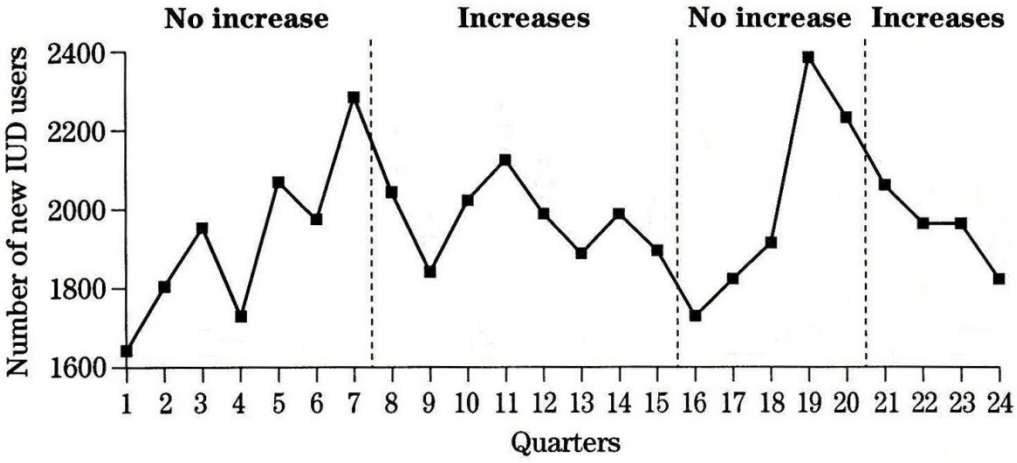
Reflexive control designs can be made much stronger by extending the basic pre-post design to create an interrupted time series design (Cook, 2000). In an interrupted time series design, data are collected at multiple points in time, both before and after the intervention. If the intervention was effective, we expect to observe a break in the trend in the outcome measures. If a break in the trend is observed before the start of the intervention, it cannot have been caused by the intervention and must have been caused by other factors. Likewise, if a break occurs a long time after the start of the intervention it is unlikely to have been the result of the intervention. If the timing of the break corresponds with other factors that typically affect the outcome, rather than with the start of the intervention, then there may be reason to believe that this other factor caused the change (Vedung, 2017).

> "The greatest limitation [of the time series design] is that many pre-intervention data points are required for rigorous analysis of the pre-intervention trend."
>
> (Ezemenari, Rudqvist, & Subbarao, 2000)

Measurement of the outcome measures at multiple time points before the start of the intervention also enables estimation of existing pre-intervention trends in the outcome measure. This in turn makes it possible to test if there were any deviations from that trend after the start of the intervention (Cook, 2000). An important limitation of time series designs is the need to obtain many pre-intervention data points to facilitate a rigorous analysis of the pre-intervention trends (Ezemenari, et al., 2000; Rossi, et al., 2018). Indeed, the design can only be used when extensive pre- and post-intervention observations of the key outcome measures are available (Rossi, et al., 2018).

Having a large number of data points and detailed information about the timing of specific intervention components can provide powerful evidence of program impact. For example, Leon and Cuesta used this approach to illustrate the effect of IUD price changes on the number of new IUD users in Guayaquil, Ecuador (Leon & Cuesta, 1993). As illustrated in Figure 2, the interrupted time series shows that periods of constant IUD prices experienced increases in the number of new users, while price increases resulted in a decrease in the number of new users.

Figure 2. Quarterly number of new IUD users at the pilot clinic in Guayaquil, Ecuador, during periods of price constancy and price increases, 1987–1992



Source: Adapted from Leon & Cuesta (1993).

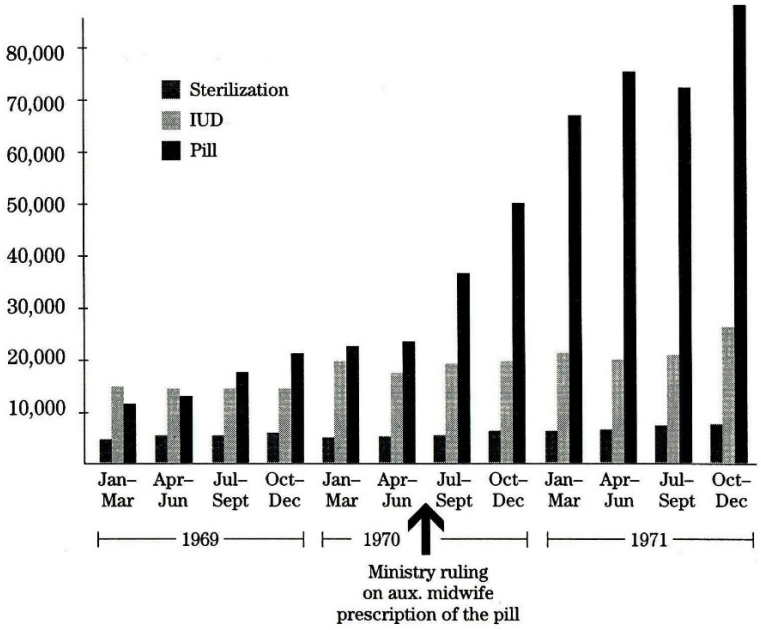
Since almost all TMA programs are full-coverage programs for which it is not possible to establish an appropriate control group, a well-constructed interrupted time series design is likely to be an acceptable and feasible solution (Wynn, Dutta, & Nelson, 2005). However, to be able to attribute changes in the outcome variables to the TMA program, rather than to existing natural fluctuations and trends, a long time series of data will be needed. Unfortunately, in most developing countries such data are not readily available (major health surveys are typically conducted only at five-year intervals). Consequently, conducting an impact evaluation of a TMA program will typically require planning for more frequent data collection, both before and after the switch to a TMA. This is feasible when the time horizon is sufficiently long and adequate funding is available.

Designs with Reflexive Control Groups Combined with a Comparison Group

It is generally recommended to use multiple impact evaluation techniques whenever possible, as the use of different techniques can strengthen the robustness and credibility of the results (Baker, 2000; Oldsman and Hallerg, 2004). Obviously, an interrupted time series design would be strengthened considerably by adding a comparison population (Cook, 2000).

For those TMA programs that target only one specific FP product, it is worth considering whether trends in other FP methods can serve as a generic control for the interrupted time series. For example, if the TMA objective is to facilitate use of injectables, it may make sense to compare the observed trends in use of injectables with the trends in use of contraceptive implants or IUDs. Assuming the program was effective, one would expect that after TMA implementation injectable use would increase more rapidly than use of either implants or IUDs. This approach has been successfully used to demonstrate the effect of policy changes on use of FP. For example, Rosenfield and Limcharoen show the effect of a ruling by the Thailand Ministry of Health that allowed auxiliary midwives to prescribe oral contraceptives by using an interrupted time series design that compares the trends in the number of oral contraceptive adopters with the trends in the number of sterilization and IUD adopters (Rosenfield & Limcharoen, 1972). As shown in Figure 3, after the policy change allowing midwives to prescribe the pill, the number of pill adopters increased rapidly, while the number of IUD and sterilization adapters remained constant. This provides very convincing evidence that the policy allowing midwives to prescribe oral contraceptives was largely responsible for the increase in pill acceptors.

Figure 3. Number of sterilization, IUD, and pill acceptors in Thailand, 1969–1971



Source: Rosenfield & Limcharoen (1972).

An alternative option for adding a comparison group to an interrupted time series design is to generate a counterfactual projection using multiple assumptions (Baker, 2000; Planning and Management Consultants, 1992; Rossi, et al., 2018). This design aims to compare improvement in the outcomes in the intervention group with what the expected outcomes would have been without the intervention. In practice, this implies using simulations or projections to estimate what would have happened in the absence of the intervention. This approach has been successfully used to estimate to what extent the Family Planning 2020 (FP2020) focus countries made progress toward increasing contraceptive use, compared to what would have been expected based on pre-FP2020 trends (Cahill, et al., 2018). Comparisons of the observed contraceptive prevalence with the counterfactual projections enables the authors to determine whether the progress had been faster or slower than expected.

Applying this approach to assess the impact of TMA would involve generating a counterfactual by projecting the pre-TMA trend into the future. The strength of this approach will therefore depend on the validity of the assumptions used to project the pre-intervention trend. As with some of the other study designs, this method is sensitive to potential external factors that may influence the outcomes. Therefore, it is important to measure and statistically control for known external factors that are related to the outcomes.

Assuming multiple pre-intervention data points are available, it is likely that it will be possible to use projections to estimate what the outcome measures would have been in the absence of a switch to a TMA approach. While very sophisticated counterfactual projections may not always be possible, simple projections can be conducted using Microsoft Excel spreadsheets. Excel includes a function for generating projections, using one of several options regarding the expected trend (e.g., whether the trend is expected to be linear or exponential). Hence, whenever data for an interrupted time series are available, it should be feasible to supplement this information with a counterfactual based on simple projections. The difference between the observed outcome measures and the projected counterfactual outcomes provides a *rough* estimate of how big an impact the program had. However, the accuracy of the estimated impact will depend on the validity of the assumptions used in the projections. Therefore, it is essential to clearly specify the assumptions that underlie the projections. Furthermore, it is recommended to conduct counterfactual projections using a range of reasonable assumptions, which will produce high and low estimates of the likely impact of the TMA program.

Using qualitative data in TMA impact evaluations

Comprehensive impact evaluations can also include qualitative methods (Baker, 2000; Cook, 2000; Ezemenari, et al., 2000). While measuring the counterfactual (i.e., what would have happened without the intervention) is central to quantitative impact evaluations, qualitative methods can strengthen the evaluation by providing information about the perspectives of programs' implementers and intended program beneficiaries, especially how they perceive the intervention and how they are affected by it. Qualitative data can enhance our understanding of the processes and causal relationships that may have affected the observed changes in the key outcome measures and may clarify how the context in which the intervention was implemented affected those processes. Such information can be invaluable to guide the interpretation of the quantitative findings.

Because triangulation of findings from different types of evaluation data enhances the validity of the results, there is a growing recognition that comprehensive impact evaluations benefit from combining quantitative and qualitative approaches. A growing body of literature describes the use of mixed-methods impact evaluations (Adato, 2012; Bamberger, 2012; Stern, et al., 2012).

4. PROCEDURE FOR EVALUATING TMA USING REFLEXIVE CONTROLS

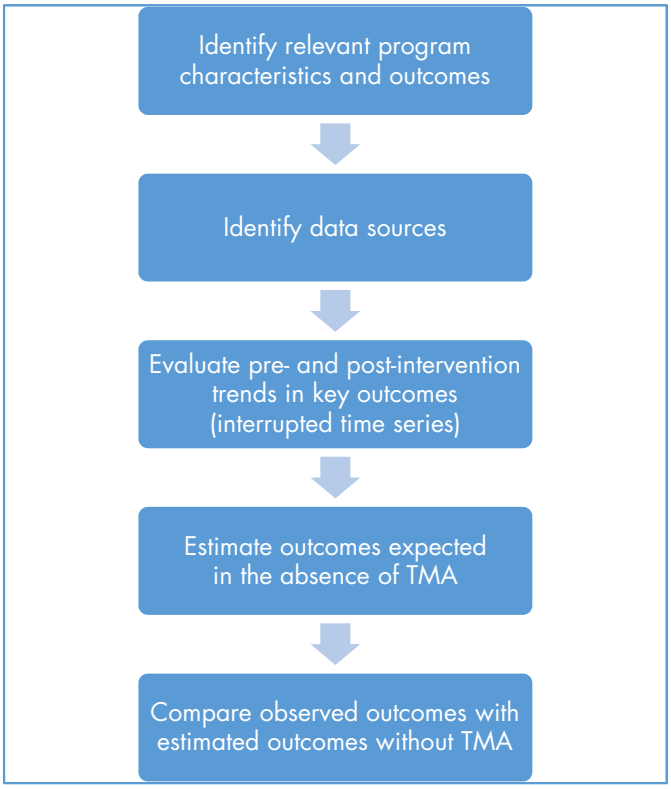
When you are considering the different evaluation designs it is important to recognize what is feasible (Oldsman & Hallerg, 2004). Even though impact evaluations should always aim to use the most robust study design, this must be done within existing constraints—such as the level of control over the intervention location, data availability, and budgets. Likely, the most feasible study design to evaluate the impact of adopting a TMA approach consists of an interrupted time series design, supplemented with a comparison group. It is anticipated that in virtually all cases, this comparison group will need to be artificially constructed based on the trends in the outcome measures that existed before the adoption of TMA (i.e., a counterfactual projection). The following sections outline the basic steps involved in applying this study design.

Step 1. Identify Relevant Program Characteristics and Outcomes

Most TMA programs have multiple objectives, including changes in ultimate higher-level outcomes as well as changes in intermediate outcomes. Hence, programs typically track progress in a wide range of indicators related to market size, accessibility, and sustainability, as well as equity. Detailed descriptions of the recommended indicators are available in the existing literature (Meekers, Haynes, & Kampa, 2016b; Pallin & Meekers, 2014). While an impact evaluation could assess the impact on each of these indicators, this is likely to become overwhelming and cloud the evaluation. Therefore, it is usually recommended that the impact evaluation focus on the main outcome indicators that the program is trying to improve (Ezemenari, et al., 2000).

The indicators selected for the impact evaluation must match the stated objectives of the TMA program. As indicated in the systematic review of the TMA literature, the most commonly cited objectives include (1) growing the overall FP market; (2) growing commercial market share; and (3) improving equitable access to FP products and services. Indicators that measure change in these objectives may include the modern contraceptive prevalence rate (by wealth quintile), the percentage of current FP users who obtain their method from a commercial source, and the market share of commercial products, among others.

Figure 4. Interrupted time series TMA impact evaluation



If time and resources permit, the evaluation can be expanded to include an analysis of the project impact on selected intermediate-level indicators (Meekers, et al., 2016b; Pallin & Meekers, 2014). In cases where the evaluation shows no evidence of impact on the ultimate project outcomes, analyzing the impact on intermediate indicators can be helpful to assess whether it is realistic to assume that impact is likely to occur after more time has passed. Analyzing intermediate indicators may also be necessary for relatively new TMA programs, as not enough time may have passed for the program to affect significant changes in the higher-level outcomes.

One of the strengths of the interrupted time series design is that comparing the timing of notable changes in the outcome indicators with the timing of the intervention can help identify changes that are caused by other external factors. Specifically, changes that either occurred prior to the TMA intervention or very long afterward are likely to have been caused by other factors. Hence, it is important to track when specific TMA activities were implemented.

The review of the TMA literature and key informant interviews have shown that the implementation of TMA is often a lengthy process that involves advocacy, market analyses, cross-sectoral coordination and planning, stewardship, and implementation of agreed-upon strategy changes. Therefore, it is often unclear when TMA activities began. Moreover, the TMA starting date indicated by project documents and informants may differ, depending on which aspect of TMA they are referring to. For the purposes of a TMA impact evaluation, the aspects of interest are those that are expected to have a direct effect on the outcomes. Activities such as TMA advocacy, consensus-building meetings with stakeholders, and developing a formal TMA plan will not by themselves affect the outcomes measures. Hence, for the impact evaluation the *de facto* start of TMA is not until at least one of the three FP distribution sectors formally changes the way they operate. If possible, it will be helpful to identify the dates when concrete changes in operation procedures took place, such as price changes or changes in the target population. Appendix 5 includes an example of a tool that can be used to track TMA activities.

It can be very helpful to also document the timing of important changes in external factors (unrelated to TMA) that could potentially influence the outcome measures. For instance, changes in insurance coverage, restrictions on FP providers, taxes and tariffs on FP products, one-time donations of FP products or services, and other regional or national marketing campaigns could all have a significant impact on outcomes and should be taken into account. We recommend that TMA monitoring include an annual checklist to track these types of policies and external factors. The tool in Appendix 3 includes an example of such a checklist.

Step 2. Identify the Data Sources

Interrupted time series designs require multiple data points before the intervention, as well as after the start of the intervention. Increasing the number of data points will substantially increase the robustness of the evaluation results. To obtain a reliable estimate of pre-intervention trends in the outcome indicators, data will be needed for a sufficiently long period. Moreover, to be able to assess fluctuations in the pre-intervention trend you will need at least three data points.

Ideally, annual data on the key outcome measures would be available for at least five years prior to the start of TMA activities. However, very few countries have annual data collection.[3] For many TMA impact evaluations, the main data source for the key outcome indicators (e.g., the modern contraceptive prevalence rate) will consist of large-scale survey programs, such as the DHS or the UNICEF Multiple Indicator Cluster Surveys. Such surveys are typically scheduled at five-year intervals. Therefore, to have three data points the pre-intervention period considered will need to cover more than ten years.[4] If at all possible, additional data points for this pre-intervention period should be obtained from other sources. For example, it may be possible to use data from nationally representative surveys conducted as part of a program evaluation, provided that they were implemented by a reputable data collection agency.

[3] There are exceptions, such as Senegal, which had a continuous annual DHS survey since 2012 (https://dhsprogram.com/pubs/pdf/DM34/DM34.pdf). Some countries also have intermediate DHS surveys. In addition, the PMA2020 program conducts annual surveys in a select number of countries (https://www.pma2020.org/pma2020-datasets-list).

[4] Typically, three successive DHS surveys cover a period of eleven calendar years. However, since the timing of the surveys may not exactly coincide with the start of TMA, the pre-intervention study period covered by the evaluation may be slightly longer.

Furthermore, exclusive reliance on large-scale surveys that are implemented only every five years implies that it will take a very long time to have sufficient post-intervention data points to enable an impact evaluation. Hence, it is important to plan to conduct regular rapid assessment surveys during the post-intervention period. This is feasible, provided it is carefully planned and there is adequate funding.

It is preferable that all data sources used are comparable (e.g., all are nationally representative, refer to the same age group, etc.). Because data sources that did use probability sampling may not be fully representative of the target population, it may not be possible to use them unless appropriate adjustments can be made (for example, by weighting the data).

Analyses based on survey data tend to suffer from having too few pre- and post-intervention data points. Hence, it can be helpful to triangulate the findings with other data sources whenever feasible. Such data sources can include qualitative studies or service statistics. While detailed services statistics from the public and commercial sectors are not routinely available, DKT International publishes data on sales of contraceptive products by social marketing programs (DKT International, 2019). The DKT database covers all the major social marketing organizations and includes most countries. For each country, annual sales numbers for each type of contraceptive method is provided, from 1995 through present, along with the corresponding estimate of the number of couple-years of protection (CYP).[5] These data can be used to examine trends in CYP provided by the social marketing sector. In addition, they can be used to examine trends in the method mix for contraceptives provided by the social marketing sector.

Step 3. Evaluate Pre- and Post-Intervention Trends in the Outcomes

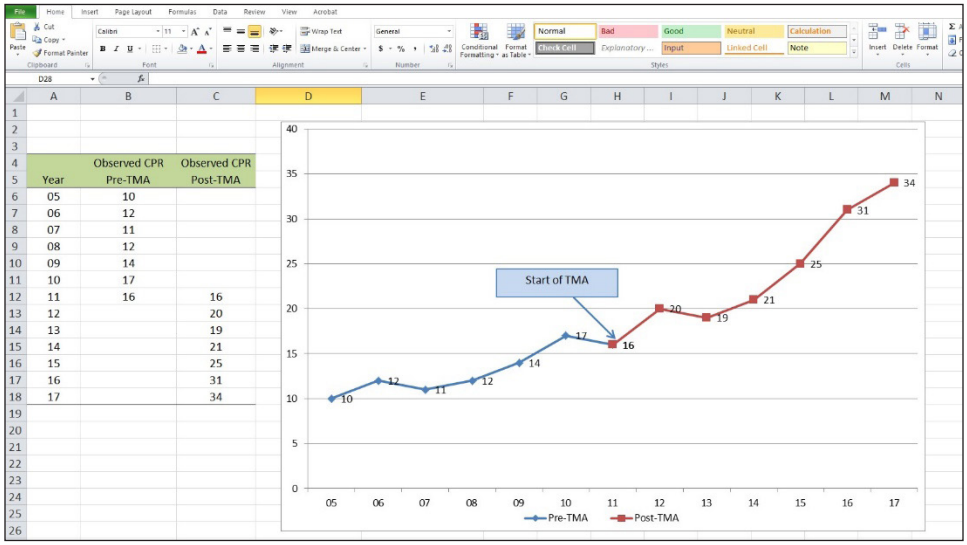
The initial step in the impact evaluation is the interrupted time series. The most basic version of this analysis simply involves graphing the observed data for each of the key outcome measures addressed by the TMA intervention over the course of the pre- and post-intervention periods. Adding the start of TMA activities (such as the formal adoption of the TMA plan, prices changes, etc.) to the graph can help clarify whether the intervention is likely to have caused the change in outcomes. Simple graphs of the interrupted time series for TMA evaluations can easily be produced using a spreadsheet program such as Excel.

Let's take the example of a hypothetical country *X* for which we have annual contraceptive prevalence data from 2005 through 2017. In country *X*, TMA coordination began in 2010, and the implementation of various TMA activities started in 2011—a mid-priced commercial brand was introduced in 2011, the leading social marketing organization increased prices for their oral contraceptives and injectables in 2012, and a new strategy to address the public sector began in early 2013.

Figure 5 shows the source data and a line graph with the resulting interrupted time series. The start of the TMA activities has been marked manually with a textbox. Note that the data for the pre-TMA and post-TMA periods were deliberately entered in separate columns, as this makes it easy to use different data point markers and line colors for the two periods.

[5] The measure CYP estimates the level of protection provided by contraceptive products or services during a one-year period, based on the volume of contraceptive products and services sold or distributed. The calculation involves a conversion factor that reflects how many units of a contraceptive method are needed to protect a couple from pregnancy for a full year.

Figure 5. Using Excel to graph an interrupted time series for the contraceptive prevalence rate (hypothetical example)



Since the interrupted time series design can be affected by changes in external factors that can influence FP use, it may be necessary to control for such changes. When the outcome measures are obtained from comprehensive health surveys, it is typically possible to control for variations in factors such as socioeconomic status, level of education, and even fertility preferences. Doing this will require merging data from the different survey years, and then using regression methods to estimate, for example, the probability of using modern contraceptives, controlling for these external factors.[6] The predicted probability of using a modern method can then be graphed using a spreadsheet as previously described.

Step 4. Estimate Outcomes Expected without TMA, Based on Pre-Intervention Trends

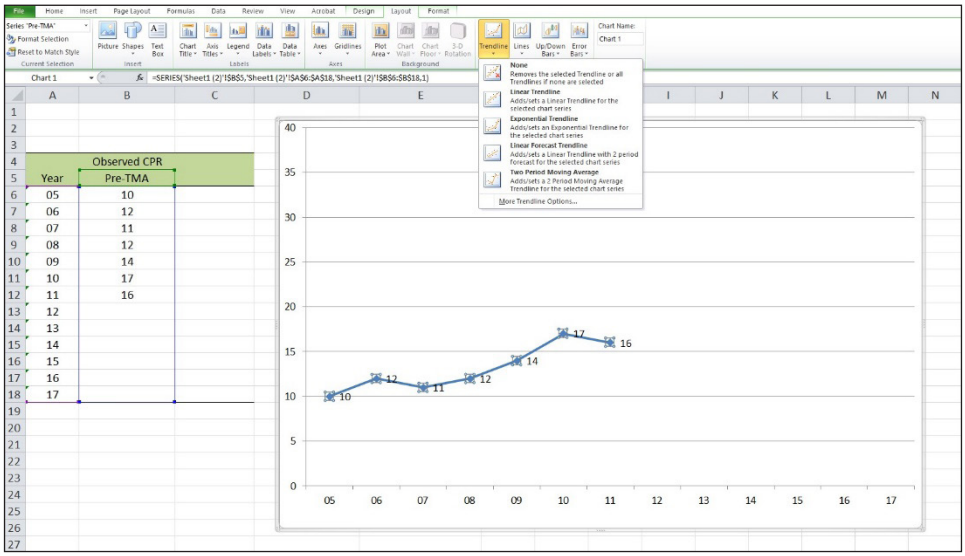
Microsoft Excel includes functions to automatically project future values of an indicator based on a series of existing data (Microsoft, 2019). For example, let's assume we created an Excel chart that shows the trend in the modern contraceptive prevalence rate (CPR) for the pre-TMA period. Excel can add a trendline that projects the modern contraceptive prevalence for subsequent years, based on the pre-TMA data. In other words, it will project how the CPR was expected to change without TMA. Hence, it serves as a counterfactual estimate of the outcome measure for a hypothetical comparison group in which the pre-intervention trend continued (i.e., a counterfactual projection). Obviously, the validity of this counterfactual estimate will depend on the accuracy of the assumptions used in the projection. Excel offers several different options for estimating the existing trend, including linear, exponential, and logarithmic trends.

To project a trend in Excel, apply the following steps (see Figure 6).

- Create a line chart that shows the trend in the CPR for the pre-TMA data points.
- Click on the trendline.
- Scroll to the tab "Layout" at the top of the screen, select "Trendline," and select the preferred type of trendline from the dropdown menu.

[6] Note that other external factors may be harder to control for. For example, Uganda has had an uneven supply chain system for medicines, including FP products, which has resulted in methods expiring before they reach the consumer (Advance Family Planning, 2017; Ministry of Health, 2012). Trend data on the number of products that expire in the pipeline are typically not available.

Figure 6. Using Excel to create a counterfactual projection of an outcome, based on pre-TMA data



The linear trend option will generate a straight trendline. In other words, it assumes that the CPR increases by the same amount each year. Linear trends in contraceptive use tend to be observed in countries where opposition to use of FP is strong. The exponential option will produce a trendline that assumes that increases in the CPR will accelerate over time, and hence that the annual increase will get larger over time. This pattern is fairly typical for countries with an unmet need for FP. The logarithmic option will create a trendline that initially increases fairly rapidly, but then tapers off over time [Alkema, 2013]. This pattern is expected in countries with high levels of contraceptive use, and where the supply is starting to meet the demand for FP.

It is recommended to produce trendlines for each of the three options mentioned above (linear, exponential, and logarithmic). For each option selected, Excel will show the regression equation used to plot the trendline, as well as a statistic (R^2 or "R squared") that measures how well the trendline fits the data.[7] The type of trendline with the highest R^2 value provides the best fit for the observed pre-TMA data points. However, there is no guarantee that this trendline will most accurately predict the CPR for future years. For example, a country may be experiencing exponential growth in the CPR, but once the demand for FP is met the CPR is likely to taper off. In this case, the logarithmic option may more accurately predict the CPR for future years.

The resulting trendlines with the estimated CPR for each of the three forecast options for our hypothetical example are shown below, in Figures 7–9. In each case, the projected CPR values have been superimposed over the observed post-TMA CPR values to provide a visual presentation of the difference between the two situations (with and without TMA). However, note that the trendlines shown are based only on the pre-TMA values (as noted earlier, this is similar to the counterfactual projection used to assess the FP2020 countries; see Cahill, et al. [2018]).

[7] The goodness-of-fit statistic shown is the R^2 value ("R squared"). R^2 can range from 0 to 1. The larger the value of R^2, the better the trendline fits the observed data. A value of 1 indicates that the trendline explains all of the variability in the outcome measure (i.e., it fits the data perfectly).

Figure 7. CPR projection based on pre-TMA data, linear counterfactual projection

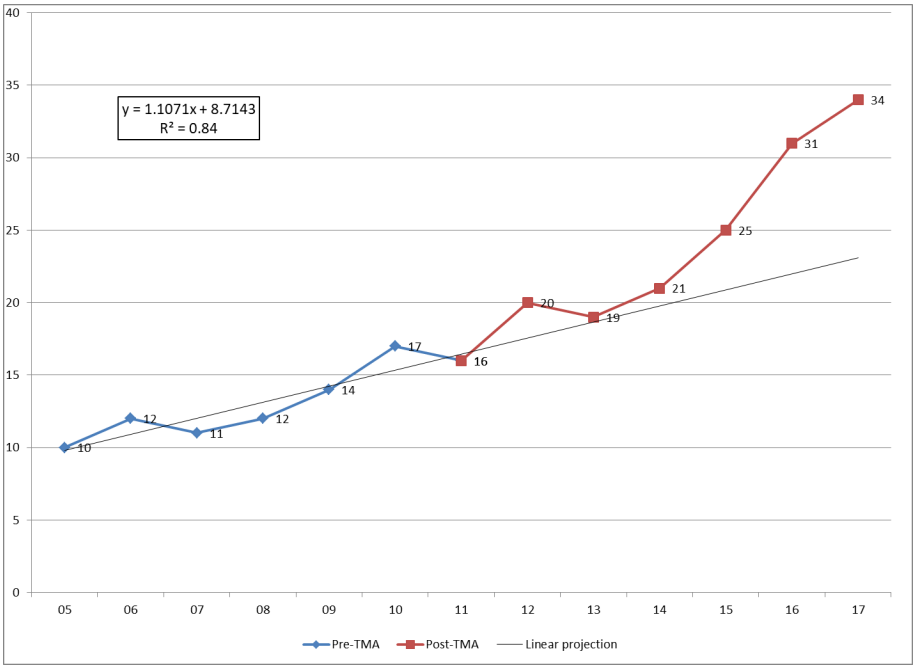


Figure 8. CPR projection based on pre-TMA data, exponential counterfactual projection

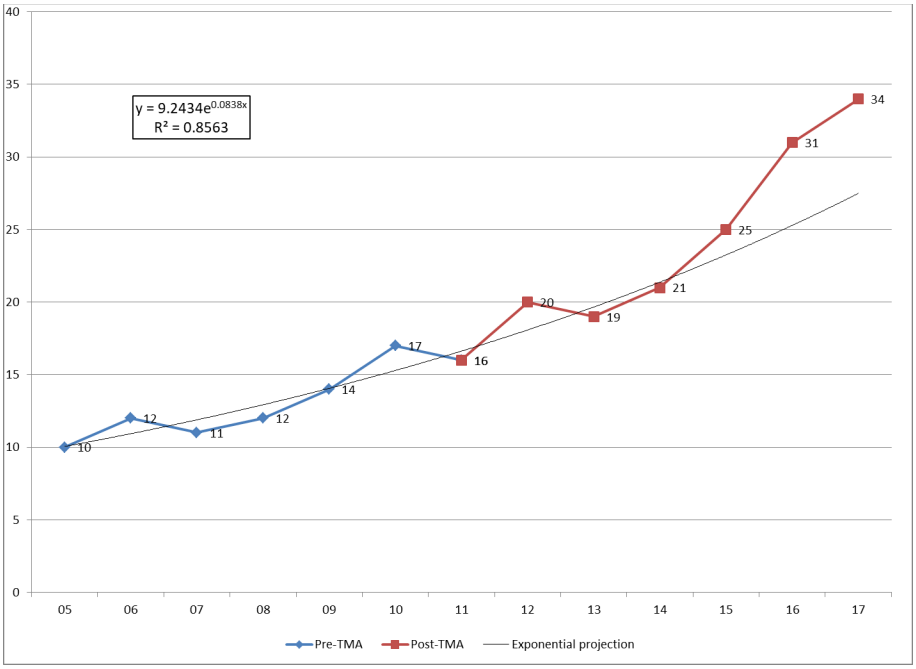
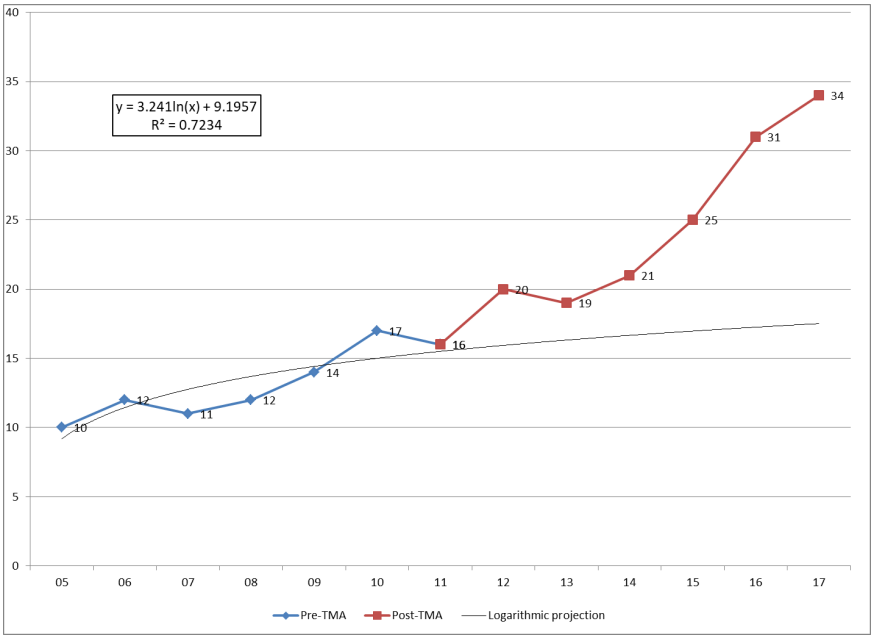


Figure 9. CPR projection based on pre-TMA data, logarithmic counterfactual projection



The R^2 values indicate that the linear and exponential projections provide the best fit for the observed CPR values during the pre-TMA period. In our hypothetical example, both of these models suggest that the increase in CPR from about 2014–2015 onward was notably higher than what would have been expected based on the pre-TMA trends. The logarithmic projection, which assumes that observed increases in CPR will taper off, results in a much larger difference between the observed and predicted CPR. However, the R^2 value (0.72) indicates that this model is a poorer fit for the pre-TMA data.

The results from the three different trendline options can subsequently be used to calculate the range within which future CPR is likely to be (i.e., the lowest and highest predicted CPR for a given calendar year).

Step 5. Compare Observed Outcomes with Outcomes Expected without TMA

The trendlines are very useful to have a visual presentation of what the CPR was expected to be in the absence of TMA. Unfortunately, the trendlines do not show the exact value of the predicted CPR for a given calendar year. To obtain the predicted value of the CPR, they must be calculated using the regression equations used to calculate the trendline. The regression equations are shown on the trendline charts.

To calculate the predicted value of the CPR, take the following steps (see Figure 10).

- In the Excel datasheet, add a column that represents the observation number. In our example, the first calendar year (2005) is observation 1, 2006 is observation 2, etc. The observation numbers are shown in column A on the spreadsheet in Figure 10.
- Add a column for the predicted CPR based on the linear trendline (see column E)
- In the cell that corresponds to the first year of TMA (cell E13), enter the formula that corresponds to the regression equation shown on the linear trendline chart. Still using the hypothetical example from Figure 7, the equation is as follows:

  $y = 1.1071x + 8.7143$

  where y stands for the outcome (CPR) and x stands for observation number.

To calculate the predicted value for 2012 (the start of TMA), enter the following in cell E13:

= 1.1071 * A6 + 8.7143

- Calculate the estimated CPR values for the remaining years by copying cell E13 onto cells E14 through E18.

Figure 10. Calculating the predicted values of the CPR, based on the trendlines

E13 f_x = 1.1071*A13 + 8.7143

	A	B	C	D	E	F	G
1							
2							
3							
4	Observation		Observed CPR	Observed CPR	Predicted CPR	Predicted CPR	Predicted CPR
5	Number	Year	Pre-TMA	Post-TMA	(linear)	(Exponential)	(logarithmic)
6	1	05	10				
7	2	06	12				
8	3	07	11				
9	4	08	12				
10	5	09	14				
11	6	10	17				
12	7	11	16				
13	8	12		20	17.6	18.1	15.9
14	9	13		19	18.7	19.7	16.3
15	10	14		21	19.8	21.4	16.7
16	11	15		25	20.9	23.2	17.0
17	12	16		31	22.0	25.3	17.2
18	13	17		34	23.1	27.5	17.5
19							
20							

A similar procedure is used to calculate the predicted CPR values based on the exponential and logarithmic trendlines. Using the example from Figure 8, the exponential trendline is represented by the equation $y = 9.2434\ e^{0.0838x}$. The corresponding predicted CPR values can be calculated by entering the formula = 9.2434 * EXP(0.0838 * A13) in cell F13.

Similarly, the formula for the logarithmic trendline is y = 3.241 ln(x) +9.1957 (see Figure 9). To calculate the corresponding predicted values in Excel, enter the formula " = 3.241 * LN(A13) + 9.1957" in cell G13.

The three different counterfactual projections of the CPR can now be used to calculate the likely range of the impact of the TMA program. To illustrate, in our hypothetical TMA program the observed CPR for 2017 is 34.0 percent (cell D18). The predicted CPR in the absence of TMA ranges for a low of 17.5 percent (logarithmic model, cell G18) to a high of 27.5 percent (exponential model, cell F18). Therefore, the estimated impact of the TMA program on the CPR ranges from 6.5 percentage points (=34.0% - 27.5%) to 16.5 percentage points (= 34.0% - 17.5%).

As noted previously, even though the interrupted time series design takes into account pre-existing trends in the outcomes, the results can still be affected by external factors that influence the outcome during the intervention period. This can make it difficult to separate the effects of the program from those other factors. Hence, the results should be used with caution and a good understanding of nonintervention factors that can influence the outcomes is needed to help interpret the results (Ezemenari, et al., 2000; Gertler, et al., 2016; Vedung, 2017). Conceptual frameworks that identify the range of nonintervention factors that can influence the outcomes can provide guidance for confounding factors that should be tracked to facilitate the interpretation of the results of the impact evaluation.

5. PLANNING A FAMILY PLANNING TMA IMPACT EVALUATION: AN ILLUSTRATIVE EXAMPLE

Considering that TMA is a relatively new approach, most countries are in the early stages. As shown in the systematic literature review and in the results of the in-depth interviews, while many countries have conducted market analyses and are developing plans to adopt TMA, as yet very few countries have started the implementation of concrete TMA activities.

For most countries that have started TMA-induced changes in FP provision, the implementation period is still too short to assess the full impact of these changes. Although a small number of countries have already been actively implementing TMA activities for years, they do not yet have sufficient post-intervention data points to accurately assess whether these activities changed existing pre-intervention trends in key outcome measures.

This section illustrates the application of an interrupted time series design to conduct an initial preliminary assessment of the impact of TMA on FP outcomes in Cambodia and makes recommendations for additional steps that will be required for conducting a full impact evaluation.

Program Overview

In Cambodia, modern contraceptive use has increased steadily over the last two decades. The percentage of married women ages 15–49 who currently use a modern contraceptive method rapidly increased from 18.8 percent in 2000 to 27.2 percent in 2005, and to 34.9 percent in 2010. Between 2010 and 2014, the pace of this increase in modern contraceptive prevalence diminished somewhat, reaching 38.8 percent in 2014 (DHS Program, 2019). The method mix is tilted heavily toward short-term FP methods, with oral contraceptives accounting for almost half of all modern method use. In 2014, 17.6 percent of married women were using oral contraceptives (the daily pill). By contrast, only 9.1 percent were using injectables, 4.4 percent the IUD, 2.2 percent contraceptive implants, and 2.1 percent were using condoms (National Institute of Statistics & ICF International, 2015).

The FP market has been dominated by the public-sector and the NGO sector. It has been estimated that the commercial sector accounts for less than one percent of the market for oral contraceptives, and even less for longer-acting methods, such as IUDs and implants (Population Services International, 2011). According to the 2014 DHS data, 47.2 percent of women reported that they obtained their most recent contraceptive method from a public-sector source (National Institute of Statistics & ICF International, 2015). The public sector distributes mostly short-acting FP methods, such as oral contraceptives and injectables (Population Services International, 2011).

Most NGO distribution of FP products is done through social marketing. Cambodia has a long history of social marketing for FP. Population Services International (PSI) has social marketed family planning products in Cambodia since 1993 through its local subsidiary, Population Services Khmer (Population Services International, 2019). In 2002, PSI started a social franchise network of private-sector providers, the Sun Quality Health Network (SQHN). The franchise network offers high quality health services and also distributes PSI-brand contraceptives. PSI currently markets *OK Pill* and *Eva Marvelon* brands of oral contraceptives, *OK* injections, *Jadelle* implants, *OK* IUDs, and *Next 72* emergency contraceptive pills. They also market *OK*, *OK Plus*, and *Number 1* brands of condoms. The 2014 DHS data show that among women who were using oral contraceptives 94.8 percent were using a social marketing brand. Similarly, 87.7 percent of women reported using a social marketing condom brand.

Since the 1990s, funding for FP supplies for the public sector and for PSI has been provided by the British Department for International Development (DFID) and the German Kreditanstalt fur Wiederaufbau (KfW) (Population Services International, 2011). However, in 2007 KfW announced that it would cease funding short-acting contraceptive methods (such as oral contraceptives) effective at the end of 2012.

To avoid the risk that this change in the funding environment would reduce access to contraceptive supplies, PSI/Khmer, the Cambodia National Reproductive Health Program, and the Ministry of Health's Contraceptive Security Working Group collaborated on the development of a plan to ensure the continuous supply of essential FP products and to develop the FP market. This marked a shift toward a TMA strategy.

Key TMA Program Objectives and Activities

One of the first steps to plan any program impact evaluation consists of identifying the key program objectives, and the specific program activities that are expected to help accomplish those objectives. In the case of TMA for FP, the specific objectives of the TMA approach itself can usually be gleaned from project documents, but it may be necessary to supplement this with information from key informants.

A review of the available literature about FP in Cambodia provides insights into the main TMA objectives and activities (Population Services International, 2011; Sievers, 2017). The primary objectives of the FP TMA in Cambodia are as follows:

- Sustain use of modern FP methods among women of reproductive age
- Increase the share of long-acting and permanent contraceptive methods (LAPMs) in the overall contraceptive method mix
- Improve focus on public-sector resources by supporting delivery of LAPMs, especially IUDs and sterilization
- Improve market segmentation to increase the role of the NGO (social marketing) and commercial sectors

Although the government committed to the Millennium Development Goal of achieving a modern contraceptive prevalence rate of 60 percent by 2015 (Population Services International, 2011), reaching this goal was not the primary objective of the switch to a TMA. The primary impetus for adopting TMA was to avoid a reduction in contraceptive supplies due to the impending reduction in donor funding for FP, which in turn may lower contraceptive use. In other words, despite the fact that the milestone set forth in the Millennial Development Goals had not yet been reached, the short- to medium-term objective of Cambodia's TMA was not to increase the modern contraceptive prevalence rate, but rather to prevent it from declining from its existing levels.

Specific activities that have been implemented to achieve the various TMA objectives include:

- Development and use of the "Reproductive Health Total Market Approach Calculator," a tool to forecast the total need for FP commodities, and to assess the potential contribution of each FP sector (2010–2011)
- Agreement to include increasing the share of LAPMs in the contraceptive method mix as a formal goal in the five-year National Health Strategic Plan (2011)
- Price increases for socially marketed oral contraceptives and injectables to facilitate commercial-sector growth for these products (2011–2012)

Data

Multiple nationally representative health surveys have been conducted in Cambodia, including the 1998 National Health Survey and the 2000, 2005, 2010, and 2014 Demographic and Health Surveys (National Institute of Public Health and ORC Macro, 2006; National Institute of Statistics and ICF International, 2015; National Institute of Statistics and ICF Macro, 2011; National Institute of Statistics & ORC Macro, 2001). Raw data are available for all these surveys (https://dhsprogram.com/).

Given that several TMA activities took place in 2011–2012, these existing data sources provide four pre-intervention data points, but only one post-intervention data point. Hence, planning for a full evaluation of the impact of TMA on FP in Cambodia will require obtaining additional data sources. Subsequent DHS surveys have not yet been announced, but, assuming that the regular five-year interval is maintained, data collection would be expected for 2019 and 2024, with the data not becoming available until about 18 months later. Hence, to be able to estimate the full impact of TMA there would be a need to identify additional data sources and/or to plan for additional data collection.

The existing National Health Survey (NHS) and DHS collected data that enable the calculation of several indicators that are directly related to the main TMA objectives:

- Percentage of married women ages 15–49 who report using a modern contraceptive method
- Percentage of married women ages 15–49 who report using a LAPM method
- Percentage of public-sector FP clients who report using a LAPM method (married women ages 15–49 only)

Although DHS data on the source of supply allow identification of public-sector sources, it is not possible to distinguish between sources for socially marketed and commercial-sector products. The 2005, 2010, and 2014 DHS also include data on the brand of condom and oral contraceptives used. Because this information allows identification of social marketing brands, it can be used to estimate the marketing share of the social marketing programs. Specially, it can be used to calculate the percentage of condom and oral contraceptive users who are using a socially marketed brand. However, with only two pre-TMA data points, it is not possible to obtain a reliable estimate of the pre-TMA trend in use of socially marketed brands.

The DKT International database on sales of contraceptive products by social marketing programs (DKT International, 2019) enables us to triangulate findings from the survey data with information on trends in the number of CYP provided by each of the contraceptive methods distributed by the social marketing program.

Initial Effects of TMA on Trends in Use of Modern Contraceptives

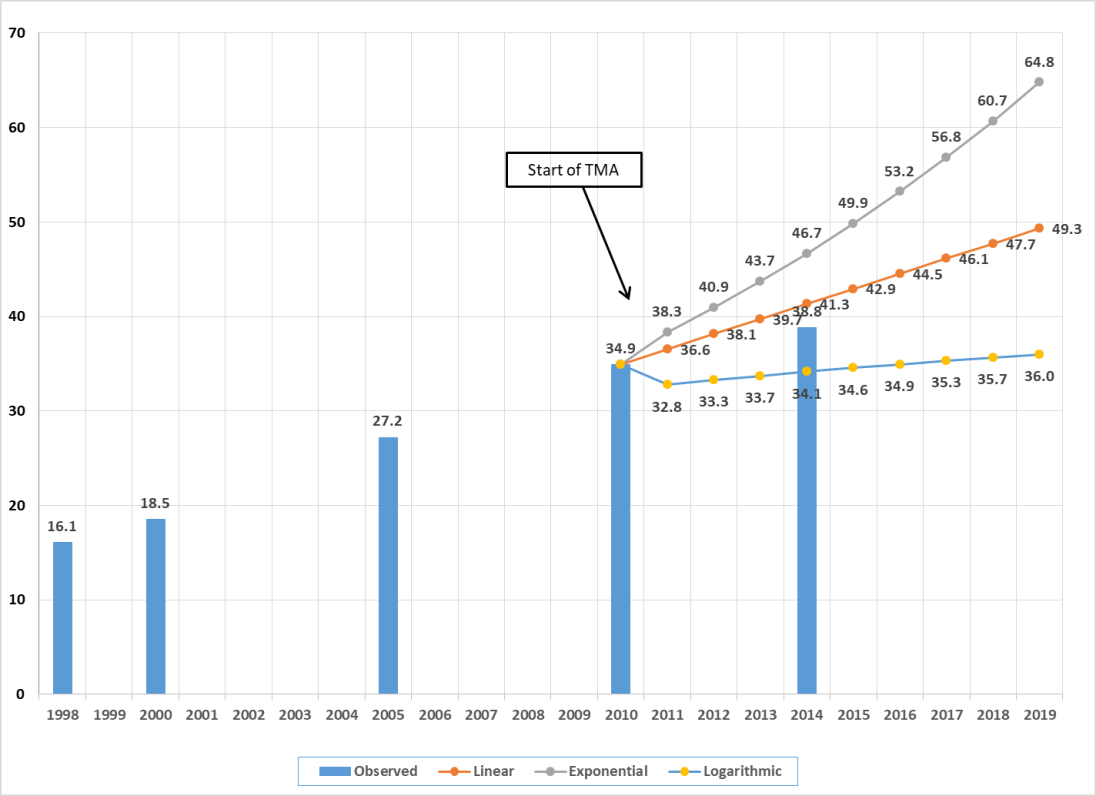
Figure 11 shows the percentage of married women who currently use a modern method of contraception (mCPR). During the pre-intervention period (1998–2010), the percentage of married women who were using a modern method of FP steadily increased from 16.1 percent in 1998 to 27.2 percent in 2005, and to 34.9 percent in 2010. TMA was subsequently introduced to avert a decline in the contraceptive prevalence rate that could potentially be triggered by reduction in KfW funding that would take place at the end of 2012. The data show that after the start of TMA activities around 2011, the percentage of married women who were using a modern method further increased from 34.9 percent to 38.8 percent between 2010 and 2014. This latter finding shows that despite the reduction in donor funding in 2012, the contraceptive prevalence rate did not decline.

However, counterfactual projections of the mCPR trend based on the pre-TMA period also suggest that the rate of increase in the mCPR slowed after 2010. The linear and exponential models provide the best fit for the pre-TMA data (R^2 = .998 and .990, respectively). According to these models, had the existing trends continued, the

mCPR would have been between 41.3 percent and 46.7 percent by 2014, which is higher than the observed 38.8 percent. Only the logarithmic model, which assumes that the mCPR will gradually taper off, yields an estimate below the observed mCPR. However, because the logarithmic model does not fit the pre-TMA as well as the other two models (R^2 = .887), this scenario seems less likely.

Thus, the available data indicate that the adoption of a TMA was successful in preventing the significant drop in the use of modern contraceptives that had been expected to occur due to a funding-related shortage of short-term methods of contraception. However, as yet there is no evidence that TMA has been able to grow the contraceptive market. Additional data will be needed to assess the longer-term effects on the mCPR.

Figure 11. Observed and projected counterfactual trends in the percentage of married women who use any modern method of contraception, 1998–2019

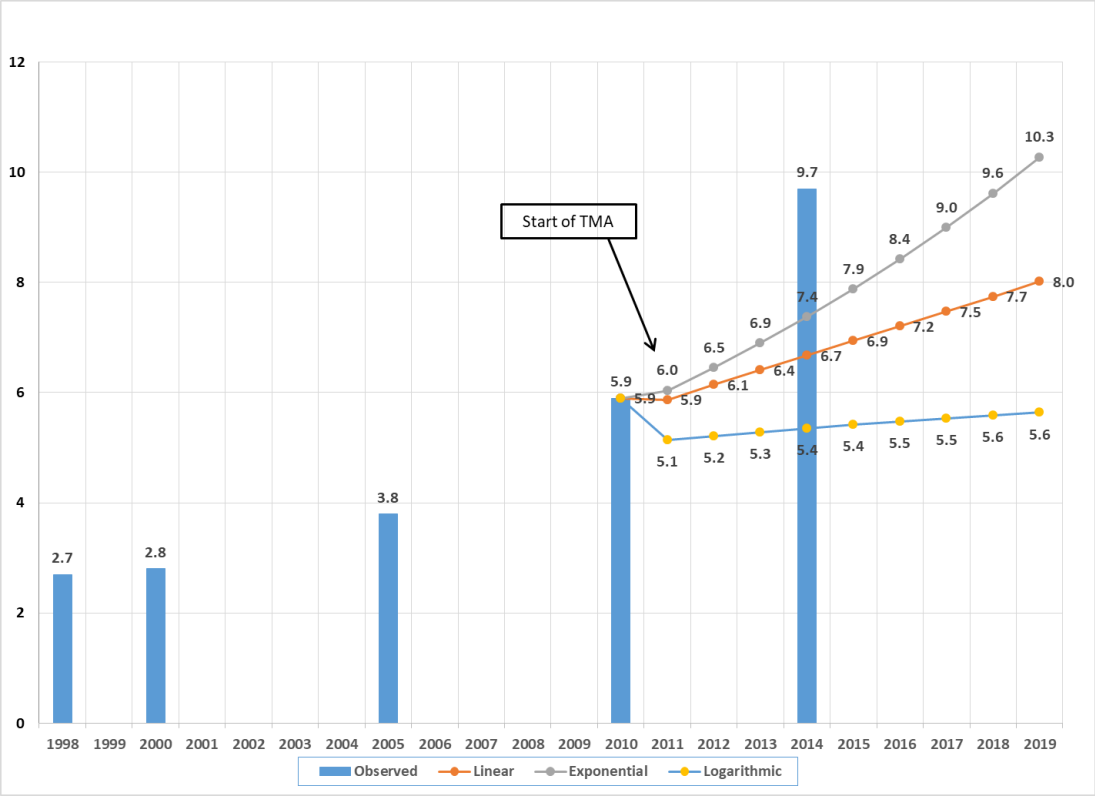


Initial Effects of TMA on Trends in Contraceptive Method Mix

To assess whether the TMA approach had an effect on the contraceptive method mix, we examine the percentage of married women who are using LAPMs (defined as IUD, implants, and female or male sterilization). The results are shown in Figure 12. Before the TMA intervention, the percentage of married women using LAPMs had been increasing substantially. LAPM use was very low in 1998 and 2000 (2.7% and 2.8%, respectively), but increased to 3.8 percent by 2005 and 5.9 percent by 2010. Examination of this pre-TMA trend in LAPMs shows that the exponential model provides the best fit (R^2 = .972), followed by the linear model (R^2 = .937). By contrast, the logarithmic model does not fit the data as well (R^2 = .719).

Based on the exponential and linear models, which provide the best fit, continuation of the pre-intervention trend would predict LARC use to be between 6.7 percent and 7.4 percent by 2014. The finding that the observed rate of LARC use was 9.7 percent indicates that the TMA strategies were successful in shifting the method mix toward LAPMs. Even with the most optimistic counterfactual projection, this level of LARC use would not have been expected until about 2018.

Figure 12. Observed and projected counterfactual trends in the percentage of married women who use a long-acting or permanent contraceptive method, 1998–2019



Initial Effects of TMA on Trends in Public-Sector Resource Allocation

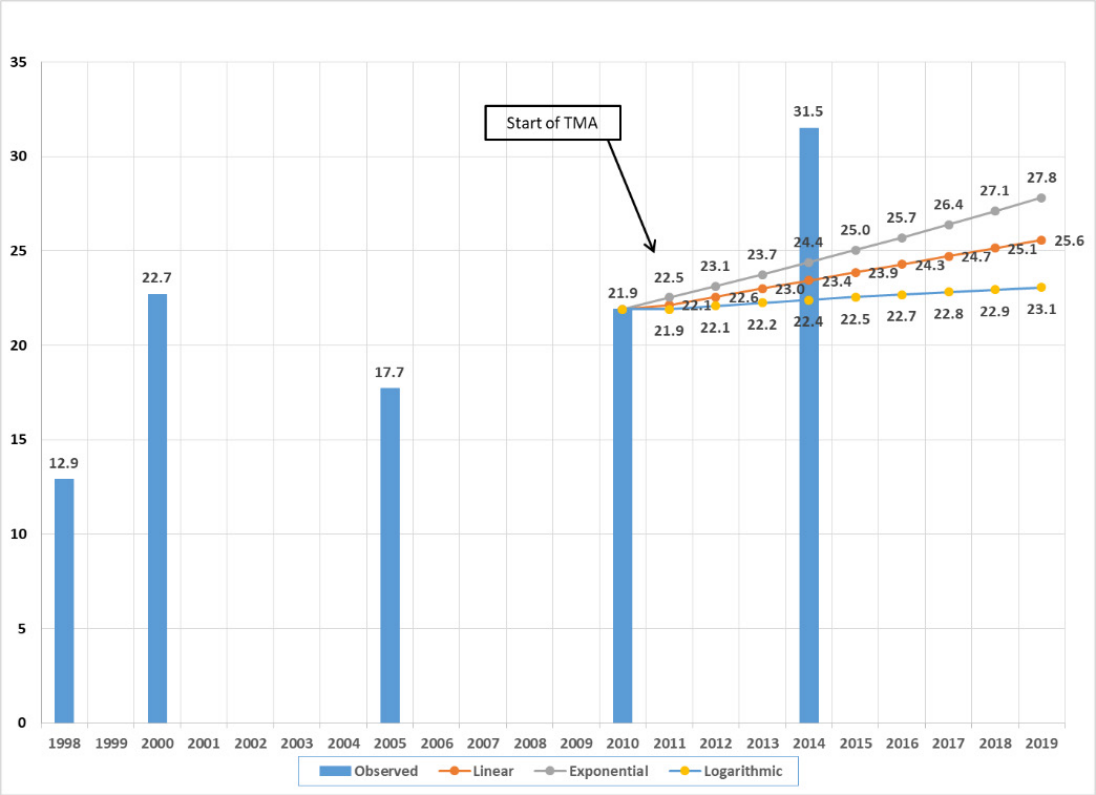
The TMA strategy also aimed to address public-sector resources toward LAPMs, which in turn is expected to improve market segmentation by increasing opportunities to develop the commercial market for short-term methods. If the strategy were successful, we would expect to see a change in the contraceptive method mix toward LAPM methods among public-sector users. To assess whether efforts to shift public-sector resources toward LAPM were successful, Figure 13 examines the observed and projected counterfactual trend in LAPM use among those married women who obtain their contraceptive method from a public-sector source.

The data reveal a fairly inconsistent pattern of LAPM use among public-sector clients during the pre-TMA period. LAPM use among public-sector clients increased from 12.9 percent in 1998 to 22.7 percent in 2000, but then dropped to 17.7 percent by 2005. Between 2005 and 2010 LAPM use increased again, reaching to 21.9 percent in 2010. Despite these fluctuations, it is evident that LAPM use among public-sector clients was already increasing before the adoption of the TMA strategy.

The erratic pre-TMA pattern highlights an important limitation of the interrupted time series methods. When the number of pre-intervention data points is small, the method is very sensitive to outliers in the data. Indeed, none of the three models provide a good fit for the data (the best fit is for the logarithmic model, $R^2 = 0.421$). Hence, the projected values are not reliable. Without the outlier for the year 2000, there would have been a very steady increase in LAPM use among public-sector clients between 1998 and 2010, which would in turn result in much higher projected values.

However, even without the outlier for the year 2000, the projected values would still not reach the observed 31.5 percent. Given the previous evidence that overall LAPM use increased, it therefore seems likely that the TMA strategy helped accelerate the shift toward LAPM in the public sector. However, the data in Figure 13 are insufficient to clarify the magnitude of that impact.

Figure 13. Observed and projected counterfactual trends in the percentage of married women who use a long-acting or permanent method of contraception, among users of a public-sector source, 1998–2019



Initial Effects of TMA-Driven Price Increases on the Contraceptive Method Mix for Socially Marketed Products

In an effort to further stimulate growth of the commercial contraceptive market, the TMA activities include price increases for short-term contraceptives. These price increases were intended to ensure that social marketing companies stopped undercutting the commercial sector by offering products at highly subsidized prices that commercial companies could not compete with. As part of the TMA strategy, PSI increased the price of their socially marketed oral and injectable contraceptives in 2011–2012. If successful, this should lead to increased commercial sector investments in short-term contraceptives. To the extent that the price increases caused customers to either shift to commercial brands or to other types of contraceptives, such as LARCs, it should also result in a shift in the contraceptive method mix for socially marketed products.

Figure 14 shows trends in the number of CYP provided by the major types of socially marketed contraceptives. Examination of the trends in CYP for the period before the implementation of TMA activities shows that condoms dominated the method mix for socially marketed contraceptives. The number of CYP provided by condoms increased rapidly up to about 2008, after which it started declining. Because this decline started several years before TMA was adopted, it was caused by factors other than TMA, possibly the rapid increase in use of socially marketed contraceptive pills.

The annual number of CYP provided by socially marketed contraceptive pills increased rapidly and steadily, from less than 50,000 prior to 2001 to about 300,000 in 2011. After the price increases of 2011/2012, the number of CYP briefly continued to increase, reaching about 350,000 by 2013. This was expected, as it would take some time to clear oral contraceptive stocks that were already in the supply chain pipeline. The number of CYP did decline noticeably in 2014, presumably as a result of the price increases. However, the data show this effect was only temporary. By 2016 the CYP levels for pills had returned to their 2013 levels, and they continued to increase after that.

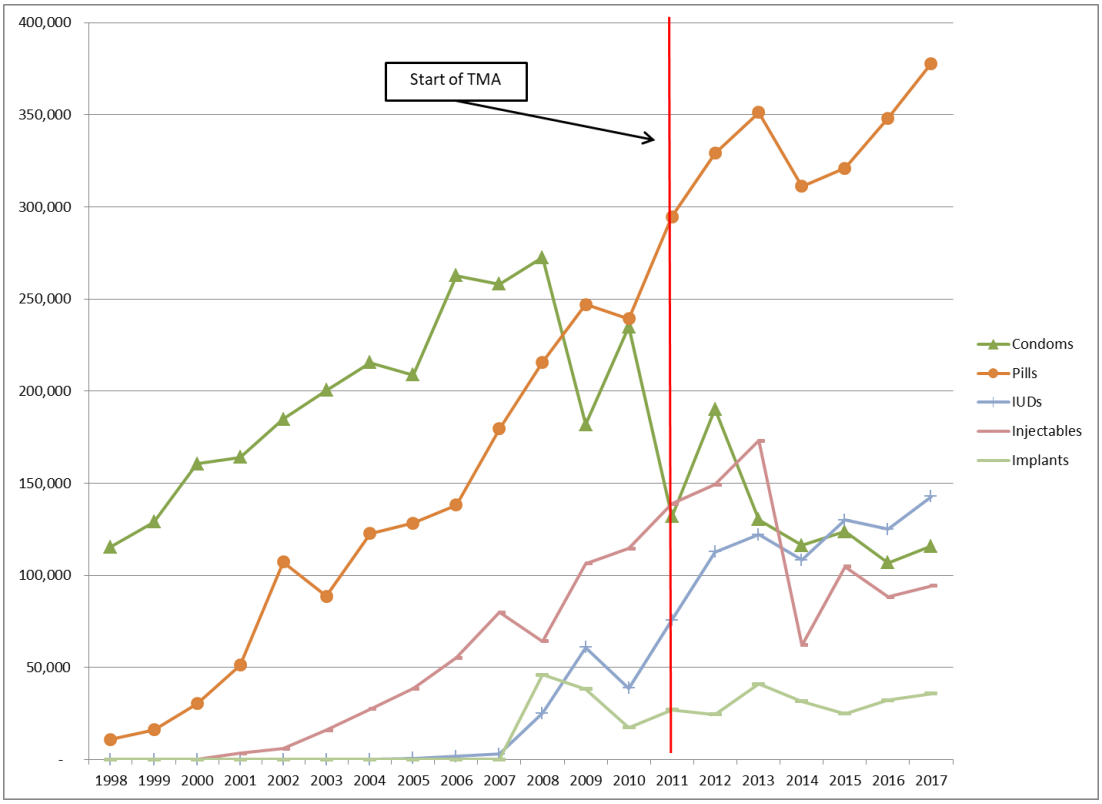
The annual CYP for injectables also increased steadily during the pre-TMA period, ranging from less than 50,000 up to 2005 to almost 150,000 for 2011. As was the case for contraceptive pills, the CYP for injectables continued to increase until 2013, reaching about 170,000. After 2013, the CYP for injectables dropped substantially to about 65,000, which is a level roughly the same as in 2006. After 2014, the CYP for injectables increased a little, but never regained its earlier levels.

The results for socially marketed implants show that the number of CYP has been roughly constant since the introduction of these implants in 2008. In other words, there is no indication that the price increases for contraceptive pills and injectables led to a shift toward socially marketed implants.

The number of CYP provided by IUDs increased rapidly from 2007–2012. After 2013 the number of CYP for IUDs continued to increase, but at a slower rate. Although the slowdown in CYP is fairly small, it is consistent with the TMA objective of shifting the public-sector focus toward IUDs.

The total CYP for all socially marketed methods combined increased steadily, from just over 126,000 in 1998 to over 1.1 million in 2013 (not shown). Although the total CYP for socially marketed products temporarily dropped to 951,000 in 2014 (reflecting the reduced use of injectables and oral contraceptives), it increased steadily afterward, reaching nearly 1.3 million in 2017.

Figure 14. Trends in CYP provided by social marketing, selected methods (1998–2017)



Assessing Which Subgroups Benefited Most from the TMA Strategy

Although the Cambodia TMA strategy was not designed to address any specific subgroups, increased coordination by the three FP supply sectors was assumed to help ensure that the FP needs of underserved populations, such as the poor and rural populations, would be adequately addressed. Hence, it can be helpful to assess whether these groups gained from the switch to TMA. Provided the sample size is sufficiently large, this can be done by conducting separate interrupted time series analyses for these groups.

Analysis of the pre-TMA data for rural areas shows that the percentage of married women who currently used a modern method of contraception steadily increased from 15.5 percent in 1998 to 35.8 percent in 2010 (not shown). Following the switch to TMA, the mCPR for rural married women increased to 39.9 percent by 2014. However, based on the exponential and linear counterfactual projections, which best fit the data, the mCPR for rural married women was expected to be between 42.3 percent and 48.4 percent by 2014. These findings show that while the mCPR for rural women did not decline after the change in donor funding, the switch to TMA was not sufficient to help sustain the growth of the rural FP market.

Among urban women, the mCPR increased only modestly during the pre-intervention period. Although the mCPR increased from 22.4 percent in 1998 to 30.6 percent in 2005, it subsequently stagnated by reaching only 30.7 percent by 2010. By 2014, roughly three years after the start of TMA activities, the urban mCPR reached only 32.8 percent. Based on this very slow increase, the mCPR was projected to reach between 32.0 percent and 35.1 percent. Although TMA may have prevented the change in donor funding from reducing access to FP, it is striking that by 2014 the rural mCPR exceeded the urban mCPR (39.9% vs. 32.8%, respectively). This suggests that the TMA approach may have been more beneficial for rural areas.

Lessons Learned

Although TMAs for FP in many countries aim to increase the overall size of the market, increase the commercial market share, or improve equity in FP access, the specific objectives of TMA can vary considerably from country to country. In Cambodia, TMA was triggered by an impending cut in donor funding for short-term contraceptive methods. Hence, the primary objectives of the TMA were to avoid a drop in the CPR, and to increase sustainability by improved market segmentation. Specifically, the aim was to have the public-sector focus more on LAPMs, and to increase the role of social marketing programs and the commercial sector in supplying short-term methods. The example of Cambodia illustrates the need to clearly identify the specific objectives of the switch to a TMA approach, as this determines which outcome measures should be evaluated.

The interrupted time series design requires several data points both before and after the adoption of TMA to calculate robust estimates of how much TMA affected the key outcome measures, which is not always possible. For our analysis of the Cambodia TMA, we only had access to three pre-intervention and one post-intervention data points. Although the Cambodia example benefited from the fact that a DHS survey was conducted just before the implementation of the TMA strategy, it shows that a lot of insights about TMA program impact can be obtained even when data availability is fairly limited.

Triangulating survey results with other data sources, such as service statistics, can help strengthen the conclusions of the evaluation. Since the Cambodia TMA strategy aimed to alter the contraceptive method mix provided by the three supply sectors, the DKT contraceptive social marketing statistics database proved very valuable. Even though the database only contains data about one of the sectors—the social marketing sector—it showed that the TMA-induced price increases for socially marketed injectables decreased sales as intended, thereby potentially providing more opportunities to grow the commercial market for injectables. However, while price increases for oral contraceptives also lowered sales volumes, they did so only temporarily. The sales data further show that the rate of increase of sales of socially marketed IUDs decreased after the implementation of TMA. This is consistent with survey findings showing the increased role of the public sector in providing LAPMs.

Even when the TMA strategy does not identify any specific groups, as was the case in Cambodia, it can be important to examine whether all groups benefited equally from the new strategy. The large sample size of the DHS surveys makes it possible to examine differences across subgroups. By simply analyzing the Cambodia survey data separately for rural and urban respondents, we were able to show that during the post-intervention period the urban mCPR was virtually stagnant, while the rural mCPR continued to increase, albeit at a slower rate than previously. These findings suggest that rural populations may have benefited more from TMA than urban populations.

Nevertheless, our analysis has important limitations. The lack of multiple post-intervention data points, for example from annual rapid assessment surveys, prevents us from determining whether the post-intervention trends are truly different from the pre-TMA situation. This is important, considering that the social marketing sales data suggest that the price increases for some socially marketed short-term methods only had a temporary effect on sales volumes. Furthermore, while the literature provided information about major changes that affected family planning supplies, it must be recognized that it is likely that many other changes may have been taking place at the same time. Key informant interviews could be a useful tool for obtaining a more comprehensive overview of what happened on the ground, and to facilitate the interpretation of findings.

6. RECOMMENDATIONS FOR PLANNING TMA IMPACT EVALUATIONS

We recommend that countries that are planning to conduct an impact evaluation of their TMA approach for FP service delivery consider the following.

- Assessing the impact of adopting a TMA strategy to supply FP products and services requires that the specific objectives of that strategy are clearly identified, as this will determine which indicators should be prioritized. To facilitate this, it is recommended that the agreed-upon TMA objectives are clearly documented, for example in the form of a publicly available TMA plan or by incorporating them in the national reproductive health policy.
- Except in very rare cases, an interrupted time series design is likely to be the most rigorous study design that is feasible to assess the impact of adopting a TMA strategy. The basic study design can be supplemented with an artificially constructed comparison group that uses pre-intervention trends to project what likely would have happened in the absence of TMA (i.e., a counterfactual projection).
- Sufficient pre-intervention data points are needed to enable reliable estimates of pre-intervention trends in the key outcome measures. In the event that fewer than three pre-intervention data points are available, we recommend forgoing attempts to evaluate the net impact (added value) of the switch to TMA. In such cases, it may be more productive to instead focus on whether the TMA approach succeeded in achieving its key objectives.
- To assess how much was accomplished without TMA, the final pre-intervention data should be collected shortly before the start of TMA implementation. When pre-intervention data are based on large national surveys, the gap between the last pre-intervention data point and the start of TMA implementation activities can be several years. This makes it more difficult to accurately assess differences between the pre- and post-intervention trends. In this case, it is recommended to conduct a rapid assessment survey on the key outcome measures immediately before the start of TMA implementation activities.
- Since a full impact evaluation also requires multiple post-intervention data points, the data source used affects when an impact evaluation can be started. Relying exclusively on national survey programs that collect data at five-year intervals is likely to imply that a full evaluation cannot be conducted until five to ten years after the start of TMA. To get more timely information about the impact of TMA, it is recommended to consider conducting annual or biannual rapid assessment surveys during the post-intervention period. This will enable a timelier start of the impact evaluation, while the increased number of data points will also make the impact estimates more robust. Because rapid assessment surveys may not be fully comparable to existing surveys conducted during the pre-intervention period, such surveys should collect information on socioeconomic and demographic variables that can facilitate weighting data from the different surveys.
- Interpreting the results from an interrupted time series design requires information about the timing of when TMA interventions were implemented. Because TMA is a process, detailed information is needed about the timing of different TMA activities. Because this information can be difficult to obtain retroactively, it is essential that TMA implementers regularly document when specific activities took place (e.g., price increases), for example in quarterly reports or other routine project documents. Likewise, it is important to document external confounding factors that took place and that may also have affected the key outcome measures.

REFERENCES

Adato, M. (2012). Combining Quantitative and Qualitative Method for Program Monitoring and Evaluation: Why Are Mixed Method Designs Best? In Lopez-Acevado, G., Krause, P., & Mackay, K. (Eds.), *Building Better Policies: The Nuts and Bolts of Monitoring and Evaluation Systems*. Washington, DC, USA: The World Bank.

Advance Family Planning. (2017). *Taking Family Planning Advocacy Efforts to Scale: Results from Nusa Dua*. Baltimore, MD: Advance Family Planning. Retrieved from http://gatesinstitute.org/sites/default/files/AFP_Diffusion_Results%20from%20Nusa%20Dua_Case_Study%5B1%5D.pdf.

Alkema, L., Kantorova, V., Menozzi, C., & Biddlecom, A. (2013). National, regional, and global rates and trends in contraceptive prevalence and unmet need for family planning between 1990 and 2015: a systematic and comprehensive analysis. *The Lancet; 381:*1642–1652.

Baker, J.L. (2000). *Evaluating the Impact of Development Projects on Poverty: A Handbook for Practitioners*. Washington, D.C.: The World Bank.

Bamberger, M. (2012). *Introduction to Mixed Methods in Impact Evaluation.* Retrieved from https://www.interaction.org/wp-content/uploads/2019/03/Mixed-Methods-in-Impact-Evaluation-English.pdf.

Barnes, J., Armand, F., Callahan, S., & Revuz, C. (2015). *Using total market approaches in condom programs*. Bethesda, MD: Strengthening Health Outcomes through the Private Sector Project, Abt Associates. Retrieved from https://www.rhsupplies.org/uploads/tx_rhscpublications/Using%20Total%20Market%20Approaches%20in%20Condom%20Programs%20v2-1.pdf.

Barnes, J., Vail, J., & Crosby, D. (2012). *Total market initiatives for reproductive health*. Bethesda, MD: Strengthening Health Outcomes through the Private Sector Project, Abt Associates. Retrieved from https://www.rhsupplies.org/uploads/tx_rhscpublications/Total%20Market%20Initiatives%20for%20Reproductive%20Health.pdf.

Brady, C., Wedeen, L., Hutchings, J., & Parks, J. (2016). *Planning Guide for a Total Market Approach to Increase Access to Family Planning. Module 1: Landscape Assessment.* Seattle, WA: PATH and Washington, DC, USA: The Population Council, Evidence Project. Retrieved from https://www.path.org/resources/planning-guide-for-a-total-market-approach-to-increase-access-to-family-planning/.

Brunner, B., Merida, M., Crosby, D., & Miles, L. (2014). Revisiting a Total Market Approach to Contraceptive Security in Honduras. *Cases in Public Health Communication and Marketing; 8:*S64–S85.

Cahill, N., Sonneveldt, E., Stover, J., Weinberger, M., Williamson, J., Wei, C., . . . Alkema, L. (2018). Modern contraceptive use, unmet need, and demand satisfied among women of reproductive age who are married or in a union in the focus countries of the Family Planning 2020 initiative: A systematic analysis using the Family Planning Estimation Tool. *The Lancet; 391*(10123):870–882.

Cook, T. (2000). Comment on "Impact Evaluation: Concepts and Methods". In Feinstein, O. & Piciotto, R. (Eds.), *Evaluation and Poverty Reduction: Proceedings from a World Bank Conference*. Washington, DC, USA: The World Bank.

DHS Program. (2019). STATcompiler. Retrieved from https://statcompiler.com/en/.

DKT International. (2019). Contraceptive Social Marketing Statistics. Retrieved from www.dktinternational.org/contraceptive-social-marketing-statistics.

Drake, J.K. (2011). *Developing a Total Market Plan for Family Planning in Vietnam.* Seattle, WA, USA: PATH. Retrieved from https://path.org/resources/developing-a-total-market-plan-for-family-planning-in-vietnam-an-innovative-public-and-private-collaboration-to-enhance-equity-and-sustainability/.

Drake, J.K. (2013). *Developing total market strategies for family planning.* Seattle, WA, USA: PATH.

Drake, J.K., Espinoza, H., Suraratdecha, C., Lacayo, Y., Keith, B.M., & Vail, J.G. (2011). Stakeholder perceptions of a total market approach to family planning in Nicaragua. *Revista panamericana de salud publica (Pan American journal of public health); 29*(5):329–336.

Drake, J.K., Vail, J.G., & Stewart, J.R. (2014). Best practices for fostering family planning total markets in Indonesia, Mexico, Romania, Thailand, and Turkey: A retrospective analysis. *Cases in Public Health Communication & Marketing;* 8(suppl1):S19–S41.

Evans, W.D., Taruberekera, N., Longfield, K., & Snider, J. (2011). Brand equity and willingness to pay for condoms in Zimbabwe. *Reproductive health; 8*(1):29.

Ezemenari, K., Rudqvist, A., & Subbarao, K. (2000). Impact evaluation: A note on concepts and methods. In Feinstein, O. & Piciotto, R. (Eds.), *Evaluation and Poverty Reduction: Proceedings from a World Bank Conference* (pp. 65–75). Washington, DC, USA: The World Bank.

Gertler, P.J., Martinez, S., Premand, P., Rawlings, L.B., & Vermeersch, C.M.J. (2016). *Impact Evaluation in Practice*. Washington, DC, USA: The World Bank.

Htat, H.W., Longfield, K., Mundy, G., Win, Z., & Montagu, D. (2015). A total market approach for condoms in Myanmar: the need for the private, public and socially marketed sectors to work together for a sustainable condom market for HIV prevention. *Health policy and planning;* 30(Suppl 1):i14–22. doi:10.1093/heapol/czu056

Khandker, S.P., Koolwal, G.B., & Samad, H.A. (2009). *Handbook on Impact Evaluation: Quantitative Methods and Practices*. Washington, DC, USA: The World Bank.

Leon, F.R., & Cuesta, A. (1993). The need for quasi-experimental methodology to evaluate pricing effects. *Studies in Family Planning;* 24(6):375–381.

Lin, A., & Wilson, J. (2014). *Healthy markets for global health: a market shaping primer*. Washington, D.C.: USAID.

Longfield, K., Ayers, J., Htat, H.W., Neukom, J., Lupu, O., & Walker, D. (2014). The role of social marketing organizations in strengthening the commercial sector: case studies for male condoms in Myanmar and VietNam. *Cases in Public Health Communication & Marketing; 8*(Suppl 1):S42–S63.

Magnani, R.J., Ross, J., Williamson, J., & Weinberger, M. (2018). Can family planning service statistics be used to track population-level outcomes? *Global Health: Science and Practice; 6*(1):93–102. Retrieved from https://doi.org/10.9745/GHSP-D-17-00341.

Mahaffey, E.E. (2012). *Privatizing Public Health: Social Marketing for HIV Prevention in Tanzania, East Africa*. Retrieved from Berkeley, CA, USA: UC Berkeley. Retrieved from https://escholarship.org/uc/item/0xt7x708.

McCarthy, K., Ramarao, S., & Taboada, H. (2015). New Dialogue for the Way Forward in Maternal Health: Addressing Market Inefficiencies. *Maternal and Child Health Journal; 19*(6):1173–1178.

Meekers, D., Haynes, S.C., & Kampa, K. (2016a). *Handbook for research on the family planning market. Volume 1: Using data to inform a total market approach to family planning*. Chapel Hill, NC, USA: MEASURE Evaluation, University of North Carolina. Retrieved from https://www.measureevaluation.org/resources/publications/tr-16-131a/.

Meekers, D., Haynes, S. C., & Kampa, K. (2016b). *Planning Guide for a Total Market Approach to Increase Access to Family Planning. Module 2: In-Depth Analysis of the Family Planning Market.* Retrieved from Chapel Hill, NC, USA: MEASURE Evaluation, University of North Carolina. Retrieved from https://www.measureevaluation.org/resources/publications/tr-16-131c-en.

Microsoft. (2019). Calculate trends by adding a trendline to a chart. Retrieved from https://support.office.com/en-us/article/project-values-in-a-series-5311f5cf-149e-4d06-81dd-5aaad87e5400.

Ministry of Health, Republic of Uganda. (2012). *Uganda National Redistribution Strategy for Prevention of Expiry and Handling of Expired Medicines and Health Supplies*. Kampala, Uganda: Ministry of Health. Retrieved from http://library.health.go.ug/publications/medicines/uganda-national-redistribution-strategy-prevention-expiry-and-handling.

National Institute of Public Health, National Institute of Statistics (Cambodia), & ORC Macro. (2006). *Cambodia Demographic and Health Survey 2005*. Phnom Penh, Cambodia and Calverton, MD, USA: National Institute of Public Health, National Institute of Statistics (Cambodia), and ORC Macro.

National Institute of Statistics, Directorate General for Health (Cambodia), & ICF International. (2015). *Cambodia Demographic and Health Survey 2014*. Phnom Penh, Cambodia and Rockville, MD, USA: National Institute of Statistics, Directorate General for Health, and ICF International.

National Institute of Statistics, Directorate General for Health (Cambodia), & ICF Macro. (2011). *Cambodia Demographic and Health Survey 2010*. Phnom Penh, Cambodia and Calverton, MD, USA: National Institute of Statistics, Directorate for Health, and ICF Macro.

National Institute of Statistics, Directorate General for Health (Cambodia), & ORC Macro. (2001). *Cambodia Demographic and Health Survey 2000*. Phnom Penh, Cambodia, and Calverton, MD, USA: National Institute of Statistics, Directorate General for Health, and ORC Macro.

Oldsman, E., & Hallerg, K. (2004). Evaluating business assistance programs. In *Evaluating Local Economic Development Assistance* (pp. 229–250). Paris, France: Organization for Economic Cooperation and Development.

Palladium. (2016). *Palladium's total market approach: a pathway for delivering health impact*. Washington, DC, USA: Palladium.

Palladium. (2017). *Zambia: Building momentum for a total market approach*. Washington, DC, USA: Palladium..

Pallin, S.C., & Meekers, D. (2014). Towards the Standardization of Total Market Approach Indicators for Male Condoms. *Cases in Public Health Communication & Marketing; 8*(1):S86–S103.

Pallin, S.C., Meekers, D., Lupu, O., & Longfield, K. (2013a). *Botswana: A total market approach for male condoms*. Population Services International (PSI)/United Nations Population Fund (UNFPA) Joint Studies on the Total Market for Male Condoms in Six African Countries. Washington, DC, USA: PSI/UNFPA.

Pallin, S.C., Meekers, D., Lupu, O., & Longfield, K. (2013b). *Lesotho: A total market approach for male condoms*. Population Services International (PSI)/United Nations Population Fund (UNFPA) Joint Studies on the Total Market for Male Condoms in Six African Countries. Washington, DC, USA: PSI/UNFPA.

Pallin, S.C., Meekers, D., Lupu, O., & Longfield, K. (2013c). *South Africa: A total market approach for male condoms*. Population Services International (PSI)/United Nations Population Fund (UNFPA) Joint Studies on the Total Market for Male Condoms in Six African Countries. Washington, DC, USA: PSI/UNFPA.

Pallin, S.C., Meekers, D., Lupu, O., & Longfield, K. (2013d). *Swaziland: A total market approach for male condoms*. Population Services International (PSI)/United Nations Population Fund (UNFPA) Joint Studies on the Total Market for Male Condoms in Six African Countries. Washington, DC, USA: PSI/UNFPA.

Pandit-Rajani, T., Cisek, C., Dunn, C., Chanda, M., & Zulu, H. (2017). *Zambia TMA Landscape Assessment*. Boston, MA: United States Agency for International Development/DISCOVER-Health Project, JSI Research &Training Institute, Inc.

Parks, J. (2016). *PATH Total Market Approach Activities in the Democratic Republic of Congo*. Seattle, WA: PATH.

PATH. (2015a). *Developing Sustainable Markets for the Woman's Condom*. Seattle, WA, USA: PATH.

PATH. (2015b). *Introducing O'Lavie Woman's Condom to China*. Seattle, WA, USA: PATH.

PATH. (2015c). *Introducing the V Condom to South Africa: Expanding the Female Condom Market*. Seattle, WA, USA: PATH.

PATH. (2016). *Making Markets Work for HIV Prevention: The Total Market Approach for Condom Security in Vietnam*. Seattle, WA, USA: PATH.

Planning and Management Consultants, Ltd. (1992). *Evaluation of Urban Water Conservation Programs: A Procedures Manual*. Carbondale, IL, USA: Planning and Management Consultants, Ltd.

Pollard, R. (2014). Total market approaches to the social marketing of health services and commodities in low-income countries. *Cases in Public Health Communication & Marketing; 8*(suppl 1):S4–S18.

Population Services International. (2011). *Total Market Approach to Family Planning in Cambodia: A Case Study*. Washington, DC, USA: Population Services International.

Population Services International. (2019). PSI Cambodia (website). Retrieved from www.psi.org/country/cambodia/#about.

Rademacher, K. (2015). *Expanding access to the LNG-IUS in Kenya: Development of a product introduction strategy*. Paper presented at the General Membership Meeting of the Reproductive Health Supplies Coalition, Oslo, Norway.

Rosenfield, A.G., & Limcharoen, C. (1972). Auxilliary midwife prescription of oral contraceptives: An experimental project in Thailand. *American Journal of Obstetrics and Gynecology; 114*(7):942–949.

Rossi, P.H., Lipsey, M.W., & Henry, G.T. (2018). *Evaluation: A Systematic Approach* (8th ed.). Thousand Oaks, CA, USA: Sage Publications.

Sievers, D. (2017). PSI Case Studies Complement Learnings at the World Social Marketing Conference. Retrieved from www.psi.org/2017/05/psi-case-studies-complement-learnings-at-the-world-social-marketing-conference.

Stern, E., Stame, N., Mayne, J., Forss, K., Davies, R., & Befani, B. (2012). *Broadening the Range of Designs and Methods for Impact Evaluation*. Department for International Development Working Paper 38. London, UK: Department for International Development.

Track20. (2015). Technical Brief: Family Planning Estimation Tool (FPET). Retrieved from http://www.track20.org/download/pdf/Track20%20Technical%20Briefs/english/Technical%20Brief_FPET%20(2015%2004%2013).pdf.

Track20. (2016). *How to use Track20's Family Planning Estimation Tool (FPET). Step-by-step instructions on how to use the online tool to produce annual estimates of key family planning indicators*. Glastonbury, CT, USA: The Track20 Project, AvenirHealth. Retrieved from http://track20.org/download/pdf/Track20%20FPET%20&%20Data%20Prep%20Instructions%20February%202016.pdf.

United Nations Population Fund (UNFPA) & PATH. (2013). *Road-mapping a total market approach: Eastern Europe and Central Asia workshops*. Seattle, WA, USA: UNFPA, & PATH.

United Nations Population Fund (UNFPA) & PATH. (2014). *Policy advocacy for total market approaches for family planning: Eastern Europe and Central Asia Regional Workshop May 2014*. Seattle, WA, USA: UNFPA, & PATH.

United States Agency for International Development (USAID) Deliver Project. (2014). *A Participatory Approach: Using Evidence to Support a Total Market Approach to Family Planning*. Arlington, VA, USA: USAID Deliver Project.

Vedung, E. (2017). *Public Policy and Program Evaluation*. New York, NY, USA: Routledge.

White, J.N., & Corker, J. (2016). Applying a total market lens: Increased IUD service delivery through complementary public-and private-sector interventions in 4 countries. *Global Health: Science and Practice; 4*(Suppl 2):S21–S32.

Winfrey, W. (2011). *Ability and Willingness to Pay for Family Planning in Vietnam*. Retrieved from Seattle, WA, USA: PATH.

World Health Organization. (2013). *Task Shifting to Improve Access to Contraceptive Methods*. Geneva, Switzerland: World Health Organization.

Wynn, B.O., Dutta, A., & Nelson, M.I. (2005). *Challenges in Program Evaluation of Health Interventions in Developing Countries*. Retrieved from Santa Monica, CA, USA: RAND Corporation.

APPENDIX 1. METHODS USED FOR THE SYSTEMATIC REVIEW OF THE LITERATURE

Criteria for Considering Studies for Systematic Review

Because many TMA findings are not published in journal articles, we included reports, briefs, conference presentations, and project documents as well as peer-reviewed articles that discussed a TMA for family planning. Documents were considered if they were written in English, French, Dutch, or Spanish and published from 2008 onward. Studies were excluded if (1) there were no specific activities described with regard to the implementation of a TMA, (2) the approach described was not focused on family planning, or (3) there was not enough detail on the planned approach. Studies were included if planning activities had been completed, even if market strategies had not yet been implemented.

Search Methods for Identification of Studies

To obtain a comprehensive body of literature on TMA, we searched PubMed, Web of Science, JSTOR, Psychinfo, Google Scholar, and the Cochrane Library databases. Because such databases may not include all the gray literature (i.e., literature not distributed by commercial or academic publishers), we also searched the websites of all 20 organizations represented at the USAID TMA Working Group (see Appendix 2 for the full list).

All databases and websites were searched using the terms "Total Market Approach" *or* "Total Market Initiative" *and* "Family Planning." For websites that did not permit Boolean operators, the terms "Total Market Approach" and "Total Market Initiative" were searched separately and any documents that did not pertain to family planning were manually excluded. Our search was restricted to documents published from 2008 onward. This initial search identified 279 documents related to TMA, including 97 duplicates. The remaining 182 documents were imported for further review using Covidence software for systematic reviews (www.covidence.org).

Selection of Studies

Two authors independently reviewed studies identified from the search in two stages using Covidence software. In the first stage, authors assessed the potential relevance of titles and abstracts identified from the search. In the second stage, authors reviewed the full texts of remaining articles and assessed against the inclusion and exclusion criteria. Authors had in-depth discussions to resolve disagreements.

Data Extraction and Management

Two authors independently extracted data from the included studies using a template. The forms were checked against each other, and authors relied on discussion and review of the original papers to resolve discrepancies. In order to describe the characteristics of each TMA implementation described in the studies, we extracted the following elements:

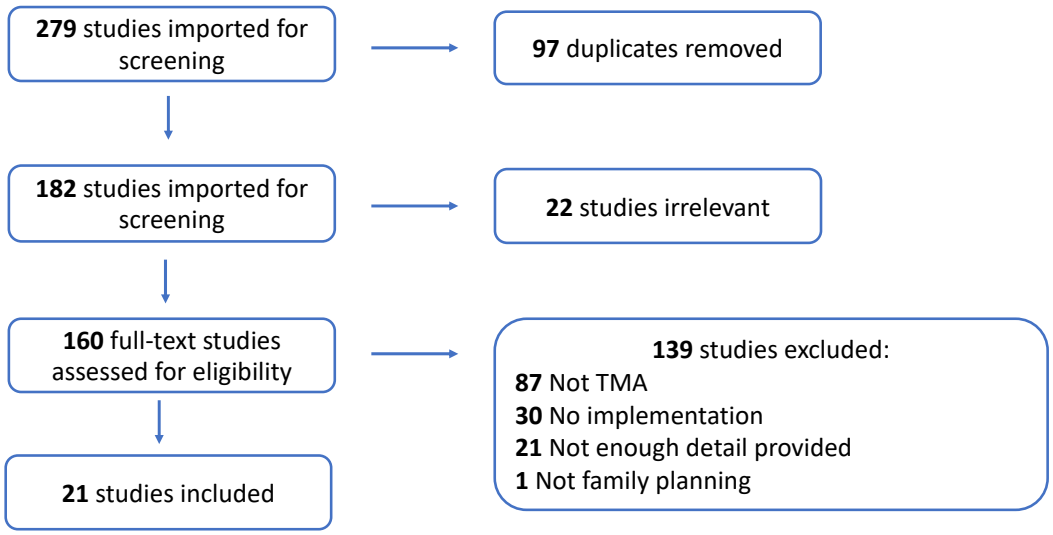
- Definition of TMA provided
- Focus countries
- Family planning methods targeted
- Years of TMA efforts
- Target groups
- Planning activities

- Coordination
- Data sharing activities
- Implementation activities
- Other strategies
- Monitoring and evaluation strategies
- Key indicators reported

Results of the Search

Our initial search identified 279 studies (see Figure 15). After 97 duplicates were removed, 182 studies remained. Twenty-two of these studies were found to be irrelevant during title and abstract screening. Of the remaining 160 documents that underwent full-text review, 127 were excluded for not describing TMA (87), including only theoretical discussion (but no reference to implementation) of TMA (30), lacking detail (21), and not family planning (1). The 21 studies remaining are included in this review. Figure 15 shows the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram for our review.

Figure 15. PRISMA flow diagram



Included Studies

We included six peer-reviewed journal articles, 13 published full reports, one conference presentation, and two brief reports. Seven studies describe more than one TMA. In total, these 21 studies describe 31 unique TMA programs. Six of these 31 programs are described in more than one publication.

APPENDIX 2. DOCUMENT SOURCES FOR THE SYSTEMATIC REVIEW OF THE LITERATURE

Databases and Document Repositories

- Development Experience Clearinghouse: https://dec.usaid.gov/dec/home/Default.aspx
- Google Scholar: https://scholar.google.com/
- JSTOR: https://www.jstor.org/
- Popline: https://www.popline.org/
- PubMed: https://www.ncbi.nlm.nih.gov/pubmed/
- Web of Science: www.webofknowledge.com/
- PsycInfo: http://www.apa.org/pubs/databases/psycinfo/index.aspx?tab=2

Organizations Represented on the USAID TMA Working Group

- Abt Associates: https://www.abtassociates.com/
- Avenir Health: http://avenirhealth.org/ ; http://avenirhealth.org/publications.php
- Clinton Health Access Initiative: https://clintonhealthaccess.org/programs/family-planning/
- DKT International: https://www.dktinternational.org/
- George Washington University: https://www.gwu.edu/
- International Planned Parenthood Federation (IPPF): https://www.ippf.org/
- Jhpiego: https://www.jhpiego.org/
- JHU Bloomberg School of Public Health: https://www.jhu.edu/ ; https://www.jhsph.edu/
- John Snow, Inc. (JSI): https://www.jsi.com/JSIInternet/
- Marie Stopes International: https://mariestopes.org/
- Meridian Group: http://www.meridian-group.com/
- Palladium Group: http://thepalladiumgroup.com/
- PATH: https://www.path.org/
- Pfizer: https://www.pfizer.com/
- Population Council: http://www.popcouncil.org/
- Population Reference Bureau: https://www.prb.org/
- Population Services International (PSI): https://www.psi.org/
- Reproductive Health Supplies Coalition (RHSC): https://www.rhsupplies.org/
- UNFPA: https://www.unfpa.org/
- USAID: https://www.usaid.gov/

Other Sources

- Personal repository of TMA documents from previous projects

APPENDIX 3. DISCUSSION GUIDE FOR IN-DEPTH INTERVIEWS

Project Implementation

- Does your organization implement family planning programs that use a TMA, or has it done so in the past?
- Can you describe what this TMA approach entails?
- How does a TMA program for family planning differ from other non-TMA family planning programs?
- What precisely are the main objectives that your organization is trying to achieve by using TMA? Do these objectives vary depending on the country or type of program (e.g., condoms only, IUDs only, etc.)?
- Do the FP programs that use TMA all use the exact same approach, or would you say there are different variants of the TMA approach (i.e. the approach varies from program to program)?
- Can you describe the main TMA variants and tell us how exactly they differ?
- When a country program uses a TMA approach for FP, does this imply that TMA is used for all types of FP products (e.g., pills, condoms, injectables,…), or does it occur that one FP product is marketed using a TMA approach, while other FP products are marketed using traditional social marketing approaches?
- [If it varies by product type] Can you explain why the approach varies by FP product?
- Can you direct us to some country directors or managers who are currently implementing (or previously implemented) TMAs?

Monitoring and Evaluation

- Does your organization usually collect or track specific indicators to measure progress toward those TMA objectives?
- What are those indicators? Do the indicators vary by country or objective?
- Are these data collected by your organization or do you rely on outside sources for data?
- Have there been evaluations for any of your organization's family planning TMA programs?
    - Can you describe the methodology that was used to conduct these evaluations, or refer me to 1) documents about these evaluations, 2) persons who are familiar with the evaluations
        - Obtain full references for impact evaluation documents
        - Obtain copies if possible
- Have there been *impact* evaluations for any of your organization's family planning TMA programs?
    - Can you describe the methodology that was used to conduct these impact evaluations, or refer me to 1) documents about these impact evaluations, 2) persons who are familiar with the impact evaluations
        - Obtain full references for impact evaluation documents
        - Obtain copies if possible
- Are there plans for additional impact evaluations for TMA FP programs?
- Are there any other people, either within your own organization or elsewhere, that you think I should speak with?

APPENDIX 4. LIST OF KEY INFORMANTS

- Patrick Aylward, Population Services International (PSI)
- Andrea Bare, Market Development Approaches Working Group of the Reproductive Health Supplies Coalition (RHSC)
- Jeff Barnes, Abt Associates
- Chris Brady, Palladium
- Cindi Cisek, Palladium
- Jennifer Drake, PATH
- Jennifer Gassner, Marie Stopes International (MSI)
- Michelle Weinberger, Avenir Health

APPENDIX 5. TOOL FOR MONITORING TMA ACTIVITIES AND EXTERNAL FACTORS

1. Have there been policy changes in the last year related to family planning products (taxes, tariffs, rules concerning import of products, reimbursement for family planning products, etc.)?
   Describe:
   Month and year change went into effect:
   Likely effect on family planning market:
2. Have there been policy changes in the last year related to the provision of family planning services (types of providers allowed to provide services, reimbursement for family planning services, etc.)?
   Describe:
   Month and year change went into effect:
   Likely effect on family planning market:
3. Have there been communication campaigns in the last year related to family planning or safe sex?
   Describe:
   Month and year change went into effect:
   Likely effect on family planning market:
4. Have there been donations of free family planning products in the last year (not including public sector provision):
   Describe:
   Month and year change went into effect:
   Likely effect on family planning market:
5. Have there been changes in the price of social marketing or commercial condoms?
   Describe:
   Month and year change went into effect:
   Likely effect on family planning market:
6. Have there been introduction of new family planning products on the market in the last year?
   Describe:
   Month and year change went into effect:
   Likely effect on family planning market:
7. Have there been changes in the distribution of public sector products (targeting, efforts to stop leakage, etc.)?
   Describe:
   Month and year change went into effect:
   Likely effect on family planning market:
8. Have there been other changes put in place in the last year which may affect the growth of the commercial sector?
   Describe:
   Month and year change went into effect:
   Likely effect on family planning market:
9. Have there been new organizations entering the family planning market in the last year?
   Describe:
   Month and year change went into effect:
   Likely effect on family planning:

APPENDIX 6. ALTERNATIVE APPROACHES FOR GENERATING COUNTERFACTUAL PROJECTIONS OF FAMILY PLANNING OUTCOMES IN THE ABSENCE OF TMA

This guide has described an approach for estimating the impact of TMA implementation on key TMA outcomes, such as the percentage of sexually active women who currently use a specific family planning method. The proposed approach can readily be extended to other TMA indicators, provided that an adequate number of pre- and post-TMA data points are available (for a list of recommended indicators, see Meekers, et al. [2016b]). One of the main advantages of the proposed method is that it can be applied using Excel and does not require advanced statistical software or expertise. Hence, if the data are available, most programs should be able to apply it.

As with all impact evaluation methods, the accuracy of the impact estimates depends largely on the accuracy of the estimates of what would have happened in the absence of the intervention, which in this case is the switch to a TMA. Since this is done using a counterfactual projection based on the pre-TMA trends, the larger the number of pre-TMA data points, the more accurate the estimate of the existing pre-TMA trend will be. However, the method also assumes that this pre-TMA trend would have continued had there not been a switch to TMA.

In many cases, perhaps most, this will be a reasonable assumption for the years that immediately follow the implementation of TMA activities. However, the more time that passes since the start of TMA activities, the less likely it becomes that the pre-TMA trend is a good estimate of what would have happened in the absence of TMA. For example, during the pre-TMA period the contraceptive prevalence rate may be increasing rapidly, and the trend may fit an exponential increase. While it is reasonable to assume that the CPR will continue to increase for some time, it is obvious that this rate of increase cannot continue indefinitely.

If there is reason to believe that the pre-intervention trend in the TMA outcome is unlikely to have continued during the post-TMA period, then a more realistic estimation of the counterfactual will be needed. In some cases, there may be justification for combining different projection methods to produce the counterfactual. For example, the pre-TMA data may show that the CPR is increasing exponentially, but data from neighboring countries may suggest that this exponential increase usually starts to taper off after the CPR reaches 30 percent. In such cases, the best estimate of the counterfactual may be an exponential increase up to a 30 percent CPR, and a logarithmic increase (i.e., tapering off) afterward.

In other cases, there may be too few pre-TMA surveys to produce reliable estimates of the trends that are likely to have occurred in the counterfactual. In this instance, it may be helpful to estimate the counterfactual trends using the Family Planning Estimation Tool (FPET) that was developed by Avenir Health's Track20 project (Track20, 2015).

Avenir Health's FPET

FPET is an online Excel-based tool that can produce projections of core family planning indicators, such as the overall contraceptive prevalence rate, the modern contraceptive prevalence rate, and the percentage of women with an unmet need for family planning (Alkema, Kantorova, Menozzi, & Biddlecom, 2013; Cahill; et al.; Track20, 2015). FPET fits a logistic growth curve to the available data to produce the most likely trends in the outcome measure. To model the mCPR, FPET uses a logistic growth curve that assumes that initially there will be a very gradual increase in the mCPR, which will be followed by an accelerated increase until a high prevalence level is reached, after which the increase will taper off (Cahill, et al., 2018; Magnani, Ross, Williamson, & Weinberger, 2018).

One of the important advantages of FPET is that it can use both survey data and service statistics to calculate the projected estimates. Although service statistics are not suitable for producing accurate stand-alone estimates of the mCPR level, trends in service statistics do tend to mirror trends in the mCPR (Magnani, et al., 2018). FPET only uses service statistics to help estimate the slope of the mCPR trend, not the level. In doing so, it can improve the accuracy of the pre-TMA mCPR trend, which in turn helps improve the accuracy of the counterfactual projections. For example, FPET can use service delivery statistics to estimate the trend in the mCPR during the three-to-five-year period since the last DHS survey.

Another important feature of FPET is that the tool produces a 95 percent confidence interval for the estimated trends. Data with a lower estimated error—such as DHS surveys—are given a greater weight in the calculations, which increases the reliability of the estimates (Track20, 2015).

How to Use FPET for TMA Counterfactual Projections

Detailed step-by-step instructions on how to use the online FPET tool are already available from the Track20 project (Track20, 2016). Only minor adjustments are needed to produce TMA counterfactual projections, as follows:

- Collect pre-TMA data on the key outcome measures (CPR, mCPR, or the percentage of women with an unmet need for family planning). For most countries, data from major survey programs such as the DHS, the UNICEF Multiple Indicator Cluster Surveys, and the PMP2020 surveys have already been preloaded. Hence, it is only necessary to obtain data from additional surveys and/or service statistics.
- Use the FPET Data Preparation Tool to prepare any additional data that should be added to the preloaded data (the tool is available from http://www.track20.org/pages/resources/track20_tools.php). Make sure to add only data that refer to the pre-TMA period, including the year that the first TMA activities started.
- Upload the resulting data file to the online FPET tool.
- In the FPET tool, select "Prepare data" and load the appropriate database and country. Because we need a counterfactual projection that is based on the pre-TMA situation, we need to create a reduced data set that includes only data from the period prior to the start of TMA activities. To do this, review the column labeled "Include?" and uncheck any data points that refer to the period after TMA activities have started. Assign a new name to this reduced dataset and click "save." FPET will confirm that the reduced data set has been saved.
- At the top of the page, select "Start run." In the left-hand panel select the reduced survey database that you had saved. If applicable, select any service statistics database that you created. Enter a run name for the projection, as well as the post-TMA years for which you want to generate a counterfactual projection (starting with the first year following the start of TMA activities). Enter a name for the run and select "start run."
- Upon completion of the run, navigate to the top of the page and select "View run." On this page, navigate to "Results" to view the results for the counterfactual projection. Select "Download all results" to download the counterfactual projection data as an Excel file.

To assess the impact of the TMA implementation, the values observed in post-TMA surveys should be compared with the median values of the counterfactual projection and the associated confidence interval. Observed values that are outside the lower and upper boundaries of the confidence intervals can be considered significantly different from what was expected to occur in the absence of TMA.

MEASURE Evaluation
University of North Carolina at Chapel Hill
123 West Franklin Street, Suite 330
Chapel Hill, NC 27516 USA
Phone: +1 919-445-9350
measure@unc.edu
www.measureevaluation.org

USAID FROM THE AMERICAN PEOPLE
Evidence
PATH
MEASURE Evaluation