

One-Year Client Impacts of Quality of Care Improvements Achieved in Peru

**Federico R. León, Santiago Roca, Alex Ríos,
Adriana Zumarán, Ana Rosa Feijoo**

This study was funded by the **U.S. Agency for International Development (USAID)** under the terms of the Frontiers in Reproductive Health Program Cooperative Agreement Number HRN-A-00-98-00012-00 and Population Council Subagreement number AI02.09A with the Escuela de Administración de Negocios para Graduados (ESAN). The opinions expressed herein are those of the authors and do not necessarily reflect the view of USAID.

EXECUTIVE SUMMARY

Context. A study with 24 health directorates of the Peru MOH as units of experimentation and analysis that was conducted in 2000-2001 showed that two-day provider training plus one-day re-training on the job aids-assisted *Balanced Counseling Strategy* resulted in significant enhancement of the quality of care, an increase of four minutes in session length, and improved client knowledge of the method chosen when this was an IUD or hormonal method. This study presents results of a one-year follow-up.

Methodology. New-event family planning clients were recruited as they exited counseling sessions at treated and control clinics. The clients had chosen a contraceptive method and stated intentions to use contraception for at least one year. A questionnaire that included the calendar module of DHS was applied twice in home interviews to tap six and 12-month outcomes. Twelve-month data from 215 clients were analyzed in depth.

Results. Knowledge of the method chosen when this was the IUD or a hormonal method was greater in the treated cohort ($p < .05$, one-tailed, $df = 18$) at the end of the follow-up period. Mean monthly ($k = 12$) family planning use rates over health-directorates ($n = 12$) were obtained for each cohort. The latter months showed less use of needed contraception ($p < .01$, $df = 11,11$) and the treated cohort greater use (.81) than the control cohort (.78, $p < .01$, $df = 1,11$) when contemporaneous perceptions of need were considered.

Cumulative continuation rates from life tables did not differ significantly between cohorts according to the Wilcoxon (Gehan) statistic. Proportions of method shifting were similar in both cohorts. Attainment of contemporaneous reproductive goal was significantly greater in the treated cohort (.96) than the control cohort (.95, $p < .01$, $df = 1,11$).

Conclusions and Recommendations. The intervention caused consistent, yet small, improvements in client outcomes. Correlational findings suggested that better impacts can be achieved in provincial primary health facilities than in larger settings like clinics or hospitals in capital cities. The results are important because: 1) other studies have failed to show significant impacts of controlled quality-of-care improvements on the use of contraception; and 2) the *Balanced Counseling Strategy* was not implemented to its full extent. Greater impacts can be expected when the *Strategy's* potential is fully realized.

ACKNOWLEDGEMENTS

The authors are indebted to Luis Hernando Ochoa, Alberto Padilla, and Luis Ulloa for their assistance in questionnaire design, database management, and use of life tables. Jim Foreit, Ricardo Vernon, and Susan Adamchak are thanked for their valuable editorial comments.

CONTENTS

Executive Summary	i
Acknowledgements	ii
Contents	iii
I. Study Background	1
Effects of Quality of Care on Family Planning Use and Continuation.....	2
Methodological Requirements for Experimental Studies	3
Further Considerations Concerning the Measurement of Continuation.....	4
Contraceptive Continuation versus Contraceptive Use	5
Client's Attainment of Reproductive Goals.....	6
Long-term Effects on Client Knowledge	7
Study Objectives and Hypotheses.....	7
III. Design and Conduct of the Study	8
Research Design.....	8
Recruitment of Client Cohorts at Posttest.....	9
Decisions on Client Cohorts	9
Fieldwork	10
Data Collected.....	10
Data Analysis	12
Hypothesis Testing.....	13
IV. Findings	14
Use of Family Planning at Month 12 of Follow-Up	14
Use of Family Planning Over the Whole Follow-Up Period	16
Trend Analysis of Crude Family Planning Use Rates	18
Trend Analysis of Quality-Related Use Rates	22
Trend Analysis of Client's Index of Attainment of Reproductive Goals	23
All-Method Cumulative Continuation Rates from Life Tables	25
Estimation of Method Switching	27
Client Knowledge Concerning the Method Used	28
Methodological Discussion.....	29
V. Conclusions and Recommendations	34
Theoretical Issues.....	34
Methodological Issues.....	37
Practical Issues	39
References	41
Appendix 1. Calendar Module of DHS	45

I. STUDY BACKGROUND

Seeking to improve the quality of family planning care at clinics of the Peru Ministry of Health (MOH), the Frontiers in Reproductive Health Program cooperated with the MOH in the development of a job aids-assisted *Balanced Counseling Strategy* that puts client need diagnosis as central to method choice, avoids client information overload, and forces the provider to focus in-depth on the method chosen by the client. An intervention involving providers, clients and system was designed, and the project produced user-friendly method cards for providers, method pamphlets with exhaustive information for clients, and monitoring guidelines for supervisors. The intervention was implemented by means of two workshops for each of 12 experimental health directorates (DISAs), one offering two-day training and the other one-day retraining for providers, and entailed the dissemination of the new counseling model and tools to 214 clinics and 74 supervisors.

The posttest corresponding to the first phase of this project encompassed the providers of the control and treated groups found at their delivery points and their clients, regardless of whether the providers of the treated clinics had attended two, one, or zero workshops (León *et al.*, 2002*b*). A recent re-analysis of the data compared the control providers and their clients with the 155 providers of the treated clinics who received the full intervention and their clients (León, Ríos, and Zumarán. 2003).

The three-day intervention on providers enhanced the quality of care by more than two standard deviations, caused an increase of four minutes in session length, and significantly improved client knowledge of the method chosen when this was an IUD or hormonal method (León, Ríos, and Zumarán, 2003). The benefits for clients were less marked when the providers received less than the three-day training (León *et al.*, 2002*b*). The study took place in 12 experimental and 12 control DISAs that had been randomly assigned to the treatments after matching. DISA averages on the quality and outcome indicators were the units of analysis ($n = 24$).

The study reported here represents the second phase of the Peru MOH project and deals with the impacts of the quality-of-care intervention on such long-term client

outcomes as one-year contraceptive use, the continuation rate, attainment of reproductive goals, and knowledge of the method chosen in the consultation with a provider.

Effects of Quality of Care on Family Planning Use and Continuation

The quality of care is valued because of ethical considerations and because it is assumed to result in positive consequences for clients and the society at large - essentially, reduction of family planning discontinuation, achievement of individual fertility goals, and attainment of better health and satisfaction. Jain (1989) and Bruce (1990) hypothesized that the quality of care is causally related to contraceptive discontinuation. For example, women who have not been informed about the side effects of their method may discontinue using it because they are not prepared to tolerate adverse reactions; those who use the method incorrectly are at risk of an unintended pregnancy. Jain (1989) further speculated that, by enhancing continuation, the quality of care positively affects contraceptive prevalence and thus reduces fertility rates. More recently, Jain (1999), on the basis of an analysis of Peruvian data, concluded that some family planning programs may be more effective if they emphasize eliminating unintended pregnancies among women who are already practicing contraception than if they focus on persuading nonusers to become users.

Some empirical studies have linked discontinuation to specific components of the quality of care. Pariani *et al.* (1991) showed that women who had not received their originally requested method had lower continuation rates. Other studies have demonstrated a link between the amount of information received and continuation (Cotten *et al.*, 1992; Lei *et al.*, 1996; Patel *et al.*, 1999). Koenig *et al.* (1997) showed that perceptions of the quality received were significantly correlated with continuation.

Decision-makers, however, need evidence concerning the programmatic control of continuation and this implies demonstrating that program interventions that improve the quality of care also cause changes in long-term client outcomes. In an experimental study in the Philippines, Lacuesta *et al.* (2001) found a significant correlation between the quality of the services received, as perceived by the client, and the subsequent use of family planning reported by the same client, yet failed to show that the intervention improved the use of family planning. A non-experimental study in Senegal presented

similar results (Sanogo *et al.*, 2003). In this case, the study showed that better quality was provided at reference centers than at health centers, yet attendance at a reference center did not significantly increase the odds of subsequent contraceptive use. Only when the client perceived better quality did contraceptive use significantly increase. These studies addressed contraceptive use and did not study continuation. This distinction is discussed later in this section.

Thus, the empirical evidence on the causal relationship between improvements of the quality of care and contraceptive use and continuation is mostly indirect and non-experimental and remains weak. Still lacking is an objective demonstration that interventions that improve the quality of care also enhance subsequent contraceptive use and continuation. This study was designed to produce such evidence.

Methodological Requirements for Experimental Studies

Three methodological requirements must be met to rigorously assess the effects of quality-of-care interventions on contraceptive use and continuation. First, an adequate approach to the assessment of the quality of care is needed. Having both the quality and continuation variables measured on the basis of a single source (e.g., client self-reports) may easily create spurious relationships. In the first part of this study, the quality of care was measured by means of: 1) a *Service Test* implemented by trained simulated clients; and 2) direct observations of the client-provider interactions by a third party who sat with them in the counseling session (León *et al.*, 2002b).

Second, the strength of the intervention must go beyond mere statistical significance. Trivial quality of care improvements may attain statistical significance on the basis of a large number of cases but are likely to have trivial effects on contraceptive use and continuation. Moreover, the researcher must anticipate that the impact of an intervention may lose strength at each step of the causal link, from the quality of care achieved by the intervention to the immediate effects on clients to longer-term client effects. In the first part of this study, a three-unit increase of the quality of care seemed needed to cause a one-unit effect on client knowledge pertaining to the method chosen when this was the IUD or a hormonal method.

Third, it is fundamental to assess the effects of the intervention on the basis of univariate data. When multiple regression analysis or multivariate tests are used, a significant coefficient represents a virtual reality of the type “Quality Would Affect Continuation If All Other Factors Were Held Constant”. Program administrators may be misled if they are not clearly told that such results represent mathematical constructions based on assumptions of variable levels of reality. We measured the strength of the intervention effects calculating univariate effect sizes, i.e., posttest differences between experimental and control groups expressed in terms of standard deviation units, and the next phase of the study maintains the focus on univariate data.

Further Considerations Concerning the Measurement of Continuation

The measurement of family planning continuation must be consistent with an explicit theory concerning the relationships between quality of care and continuation. For example, method shifting must be differentiated from the abandonment of family planning. Method shifting is a legitimate option, especially for women who are experimenting with contraceptives until finding the one that is satisfactory for their particular case. Hence, the all-method continuation rate is the appropriate indicator when the effects of quality-of-care interventions on contraceptive use and continuation are studied. In the calculation of such rates method shifting does not count as a discontinuation; discontinuation of all methods of family planning does.

Blanc *et al.* (1999) considered that the client’s circumstances must be taken into account. They distinguished between two all-method continuation rates according to the reasons for stopping use of family planning. The first category includes reasons that imply a reduced need for contraception including: wanting to become pregnant, having infrequent sex or husband/partner away, being menopausal or subfecund, and marital dissolution or separation. These are reasons that according to Blanc *et al.* are not related to characteristics of the method or the service environment. Since the discontinuation due to these reasons is not expected to be influenced by the quality of care, the reduced-need discontinuation rate would seem to be irrelevant to the assessment of the effects of quality-of-care interventions on contraceptive discontinuation.

The second category consists of all other reasons for discontinuation including:

contraceptive failure, husband's disapproval, side effects, health concerns, lack of access, cost, inconvenience of using the method, being fatalistic, and other (unspecified) reasons. Blanc *et al.* referred to these reasons as "quality-related reasons" since the level of discontinuation for these reasons is expected to be directly related to the quality of care.

In this study we measured both a general contraceptive continuation and a specific quality-related discontinuation. The first indicator does not take into account reasons for stopping use of family planning. Hence, it implicitly includes reduced need factors as well as quality-related factors as reasons for discontinuation. Contrary to the view implicit in the Blanc *et al.* formulation, we assume that quality-of-care interventions may affect reduced-need discontinuation. For example, the method pamphlets of the job aids-assisted *Balanced Counseling Strategy* tell the client that she may abandon the use of a reversible method if she wishes to become pregnant; i.e., they make clear to the client that it is her right to change her reproductive intentions at any time. This may have effects on the evolution of her intentions, thus affecting reduced-need discontinuation. Or the provider may explain to the IUD-using wife of a recently enrolled migrant laborer that it is all right to discontinue use of the IUD and instead use pills or condoms when the husband is back. The wife may tolerate the side effects of the IUD when she has the husband around all the time, yet may see no point in continuing with the IUD given the temporal separation. In this case, good quality of care will cause several segments of reduced-need discontinuation. The quality-related discontinuation rate is oblivious to the impacts via reduced need; hence, it must be complemented by a measure of general discontinuation that is indifferent to the reasons for discontinuation.

Contraceptive Continuation versus Contraceptive Use

Contraceptive continuation is a specific technical concept that refers to continuous segments of contraceptive use, typically analyzed using a life table. A segment of use ends when the client stops using family planning. If she reinitiates use, a new segment is entered into the life table. For example, the seven-month cumulative continuation rate is calculated as the percentage of segments involving the continuous use of family planning from month 0 through month 7, the denominator being the total number of segments. We have two problems with this coefficient. First, by definition, it fails to take into account

segments of zero length. These are relevant to the concerns of the present study, for a percentage of clients who choose a family planning method in the consultation with a provider may fail to implement the decision to initiate the use of that method or any other method. An evaluation of the impact of quality-of-care enhancements on contraceptive continuation that only uses the continuation rate will ignore this contingency. Since the focus will be limited to cases that have initiated use of a method, the failures to initiate method use will not be counted as a negative contribution to the prevalence of contraception.

Second, the conventional calculation of continuation rates requires processing segments of use by means of life tables. Our study had the health directorate (DISA) as the unit of experimentation and analysis. Not only the small total number of these units ($N = 24$) makes them unsuited for life-table analysis. Each DISA encompasses a number of clients. While the dichotomous concept of Use-No Use can be applied to the single client, it is meaningless in reference to a DISA central tendency.

Consequently, the main analyses of this study use a different technical concept, the use rate, that takes into account failures to initiate use of a method and is amenable to statistical operations having the DISA as the unit of analysis. This concept is similar to the prevalence rate of demographic studies. Nonetheless, to satisfy the expectations of an audience used to dealing with continuation rates, we also analyze the data by means of life tables and report standard continuation rates based on use segments.

Client's Attainment of Reproductive Goals

Neither contraceptive use nor contraceptive continuation are analogous to attainment of reproductive goals. Meeting fertility goals is a two-sided and dynamic outcome that depends on the evolution of the client's reproductive intentions. For a woman who does not wish to become pregnant in the short term, success is defined by the avoidance of pregnancy and failure by being pregnant. For a user who has changed her outlook concerning reproduction and desires an immediate pregnancy, success is defined by pregnancy and failure by its opposite.

The quality of care is expected to affect the ability of the client to attain this type of reproductive goal. For example, a provider may fail to inform the woman that the use of DMPA may cause temporary infertility that could last from six to 12 months after discontinuing its use. Consider a woman who chose DMPA and used it for a number of months, yet now wants a pregnancy. She will undergo from six to 12 months of failure to attain her new reproductive goal. Had she known, she might have chosen the pill and be able to meet her reproductive intentions sooner.

Long-Term Effects on Client Knowledge

Two mechanisms whereby the quality of care affects contraceptive use can be distinguished. One is the adequacy or suitability of the method chosen by the client. The *Balanced Counseling Strategy* is expected to positively influence the choice process by reducing client information overload and allowing the client to choose the method that best suits her needs. The second component pertains to the extent, accuracy, and relevance of the knowledge concerning the method chosen that the client acquires as she interacts with a provider. The Peru MOH's quality-of-care intervention improved client knowledge of the method chosen when this was an IUD or hormonal method, which account for about 70 percent of the prevalence at MOH facilities (León, Ríos, and Zumarán, 2003). This knowledge is expected to help the client face the contingencies of method use with a problem-solving approach and facilitate behaviors consistent with her reproductive goals. Additionally, the client can be expected to further improve her knowledge concerning the method chosen through frequent review of the method pamphlet that she takes home after the consultation.

Study Objectives and Hypotheses

The main objective of this study was to test whether an intervention that caused substantial quality-of-care improvements and significant improvements in client knowledge of the method chosen also enhances subsequent use of family planning. The hypotheses are that the job aids-assisted *Balanced Counseling Strategy* will:

- Improve one-year contraceptive use and contraceptive continuation.

- Improve one-year quality-related contraceptive use and contraceptive continuation.
- Improve attainment of one-year reproductive goals.
- Increase one-year knowledge concerning the method chosen.

II. DESIGN AND CONDUCT OF THE STUDY

Research Design

This study represents the second part of a research project whose first part was implemented in six phases:

1. Twenty-four health directorates (DISAs) of the Peru MOH were randomly assigned to experimental ($n = 12$) and control ($n = 12$) conditions (June 2000).
2. A pretest involving 334 facilities was conducted in experimental and control health centers (June-September 2000). Group equivalence at the baseline was confirmed.
3. A two-day workshop per DISA was offered to the 12 experimental DISAs (June-October 2000). The job aids-assisted *Balanced Counseling Strategy* was taught to 279 providers and observers who received method cards and method checklists.
4. Simulated clients implementing the *Service Test* offered feedback to the research team that led to adjustments in the intervention (October-December 2000).
5. A one-day retraining workshop per DISA was conducted for providers at the 12 experimental DISAs (December 2000 – March 2001). Method pamphlets replaced the method checklists. 155 providers received both the two-day training and one-day retraining. More than 100 providers received only one of them.
6. A posttest was conducted in experimental and control facilities (July-September 2001). It was found that the three-day intervention on providers enhanced the quality of care by more than two standard deviations, caused an increase of four minutes in session length, and significantly improved client knowledge of the

method chosen when this was an IUD or hormonal method (León, Ríos, and Zumarán, 2003). The benefits for clients were less marked when the providers received only one or two days or no training. Moreover, the providers who used the *Strategy's* job aids in the interactions with clients (37 percent) showed the best results. The effects at treated clinics were negligible when the providers did not use the job aids (León *et al.*, 2002b).

Recruitment of Client Cohorts at Posttest

Clients of experimental ($N = 159$) and control MOH clinics ($N = 179$) who had chosen a family planning method were recruited for this study (month 0 of cohort age) from 9 July through 23 September 2001 as part of the prior study posttest, i.e., in exit interviews immediately after participating in counseling sessions that had been observed by a third party. All of them were new-event users at the time of recruitment, i.e., clients who used family planning for the first time, switched from one method to another, or reinitiated use of a method after six months. The users stated that they did not want to be pregnant for at least one year, gave their consent to be interviewed twice in the following 13 months, and provided precise home addresses.

Decisions on Client Cohorts

The recruitment process targeted the clients exiting counseling sessions at the time of the visit by the data collectors, regardless of the individual status of the provider from whom the services were received. The idea was to assess the effects of the intervention on clinics that had been given the opportunity to send providers to the workshops and had taken advantage of this opportunity. But the individual providers who attended the workshops could have been absent at the moment of the posttest and the clients recruited at the treated clinic might have received the counseling from a provider who attended only one or none of the intervention workshops. Under this research design, the intervention did not achieve statistically significant effects on client knowledge of the method chosen (León *et al.*, 2002b).

However, statistical significance was achieved concerning effects on client knowledge of the method chosen when this was the IUD or a hormonal method if the

analysis excluded the treated-group providers who had not received the full intervention. In this instance, León, Ríos, and Zumarán (2003) compared the clients of control-group providers with clients of the providers that were present at their clinics during the posttest and belonged in the group of 155 providers that received both the two-day training and one-day retraining.

Assessing the impacts on contraceptive use and continuation of an intervention that failed to affect first-level client outcomes made little sense. Hence, we decided to perform the one-year follow-up study considering the clients of control group providers and the clients of treated-group providers who had received the full intervention.

Fieldwork

Excluding the 65 clients of the treated clinics who had received services from providers who did not attend both workshops, the sample sought in the second interview had an $N = 273$. Of these, 233 interviews were completed at month 10 on average (85 percent follow-up). The cases lost were due to addresses not found ($n = 11$), client unknown at address ($n = 7$), client moved away ($n = 11$), rejections ($n = 2$), and others ($n = 9$). Nine interviews were discarded as the clients' responses had been incoherent.

The sample sought in the second interview at month 15.5 on average also had an $N = 273$. Five incoherent interviews were excluded from analysis. 215 interviews were completed and used (79 percent follow-up). The cases lost were due to addresses not found ($n = 12$), client unknown at address ($n = 14$), client moved away ($n = 12$), and rejections ($n = 15$). The percentage of cases lost was 20 percent in the control cohort and 22 percent in the treatment cohort. In the final samples, the proportions choosing hormonal methods, IUD, and barrier/natural methods were .64, .05, and .31 for the treatment cohort and .63, .05, and .31 for the control cohort.

Data Collected

Use of Family Planning. We employed the Calendar Module of the Demographic and Health Survey (DHS) in the follow-up study. The calendar helps respondents place events in time. The client is asked to recall births, pregnancies, abortions, and contraceptive use and discontinuation of specific methods month by month

starting with the present. The reasons for discontinuation are recorded in a separate column. We made the calendar more precise by dividing months in half. For example, consider the first half and the second half of July 2001. If the date of client recruitment registered in the posttest fell on the first half, the month was defined as month 1. If it fell on the second half, the next month was defined as month 1. A discontinuation or lack of use was registered for the month if it occurred at either of its halves. A column concerning the reproductive goals of the client in each month period was added.

Client Knowledge. Items from the client exit interview were utilized to measure client knowledge pertaining to the method in use. If the client was not using a method, the questions asked pertained to the method discontinued. If the client had not started using a method, the questions asked pertained to the method chosen at the consultation more than one year earlier. The areas tapped were: general information; use instructions, contraindications, side effects, and alarm signs of the method chosen; and follow-up. The open-ended questions were of the type, “What are the possible side effects of the pill?” The questionnaire contained the correct response (“headaches, nausea, breast tenderness, weight gain”) and the interviewer registered a positive mark if the correct response was contained in full in the client’s spontaneous recall. Three dichotomous items were defined and summed for each area: *a*) One or more correct responses (Yes = 1, No = 0); *b*) Two or more correct responses (Yes = 1, No = 0); *c*) Three or more correct responses (Yes = 1, No = 0). The reliability of the summed scores in the posttest had been satisfactory (Cronbach’s $\alpha = .85$). But a maximum score of 18 could be obtained on the knowledge scale only if the client chose the IUD or a hormonal method. Barrier and natural methods practically have no contraindications, and certainly no side effects or alarm signs. Hence, clients who chose these methods could only obtain a maximum score of 9. To avoid distortions, the decision was made to deal separately with these cases despite the moderate reliability of the barrier/natural methods scale ($\alpha = .57$). Throughout the report, clients who chose IUD/hormonal methods are called “users of IUD/hormonal methods” despite that some could change method along the 12-month follow-up period. The case is similar for clients who chose barrier/natural methods.

Data Analysis

Reasons for Discontinuation. In data analyses we distinguished reasons for discontinuation between reduced need (i.e. wanting to become pregnant, having infrequent sex or husband/partner away, being menopausal or subfecund, and marital dissolution or separation) and quality-related (i.e. contraceptive failure, husband's disapproval, side effects, health concerns, lack of access, cost, inconvenience of using the method, being fatalistic, and other). In this process we detected that 32 clients who had been recruited as new-event family planning users reported monthly events as if they had been continuers at the moment of recruitment. Since we lacked criteria to ascertain the source of the confusion and exclusion of these cases would have led to the loss of some DISAs, we decided to retain them as part of the client cohorts at the cost of increasing the measurement error.

Use and Continuation Rates. We computed for each DISA a use rate per month as the number of cases using family planning in a given month, divided by the total number of cases in that DISA. This was performed independently for each month. Hence, one client that discontinued the use of family planning at a given month could be counted as a user at another month if she restarted using a method. The use rate did not reach 1.0 at month one because not all of the clients who chose a method at the posttest and were recruited for the follow-up implemented their decision to use a method. Family planning continuation rates were obtained from life tables in which the unit of analysis was the segment of use. The life-table analyses only included cases that had effectively initiated use of family planning according to the calendar module of the DHS. With clients of providers who received less than three days of training excluded from the analysis, and clients who did not initiate use also excluded, each of 190 clients contributed between one and four use segments, making a total of 238 events.

Attainment of Reproductive Goals. For a woman who does not wish to become pregnant in the short term, success (1 point) was defined by the avoidance of pregnancy and failure (0 points) by being pregnant. For a user who had changed outlook concerning reproduction and desired an immediate pregnancy, success was defined by being pregnant (1 point assigned) and failure by its opposite (0 points assigned).

Knowledge Scores. The summated scores for the client knowledge measures were normally distributed according to Kolmogorov-Smirnov tests.

Hypothesis Testing

We considered with reservations the binary decision procedure of null hypothesis testing whereby one “rejects” or “fails to reject” the null on the basis of a p level, typically .05. As was the case of the final report of Phase I of this project, concerned with the first-level outcomes of the intervention (León *et al.*, 2002b), we evaluated the long-term impacts on the basis of effect sizes. The effect size expresses the difference between the treatment (M_t) and control means (M_c), or d , in pooled standard deviation units (or σ). In the calculation of d , the $M_t - M_c$ difference is divided by $([s_t^2 + s_c^2] / 2)^{1/2}$, where s is the sample standard deviation (Lipsey, 1990). This indicator protected the conclusions of the study against the error of taking a statistically significant p value to imply a scientifically important finding. Only if the effect size is large and significant can one conclude that the findings are important. In the context of an analysis of variance, the effect size is obtained as the square root of a ratio of the variance of the means to the variance within-groups, i.e., $ES = (s^2_{\text{Means}}/s^2_{\text{Within}})^{1/2}$ (Rosenthal and Rubin, 1994).

Second, given the very small number of cases in the study ($N = 24$), there was a risk in concluding that failure to reject the null implied that there was no relationship between the independent and dependent variables. Consequently, we considered the counternull value of the obtained effect size (Rosenthal and Rubin 1994). This is “that nonnull magnitude of effect size that is supported by exactly the same amount of evidence as is the null value of the effect size. In other words, if the counternull value were taken as the null hypothesis, the resulting p value would be the same as the obtained p value for the actual null hypothesis” (Rosenthal and Rubin, 1994, p. 329). The case could arise that the study results fail to reject the null hypothesis of zero effects and also fail to reject the nonnull hypothesis that the effect size has a sizable value greater than zero.

An additional effort to reduce the probability of committing a Type II error was made. It consisted of the use of trend analyses that allowed us to detect specific time segments of the follow-up period in which the study hypotheses were, or failed to be, supported. This allowed us to state conclusions pertaining to significant differences in family planning use between the control and treated cohorts in specific semesters or months of the follow-up period. With a more static approach to hypotheses testing, we would have focused only on the entire follow-up period, ignored intervention x time interactions, and committed a Type II error if such an interaction, in fact, existed.

III. FINDINGS

Use of Family Planning at Month 12 of Follow-Up

Lacuesta *et al.* (2001) and Sanogo *et al.* (2003) presented their results in terms of impacts on family planning use at a discrete point in time: the end of the follow-up period. Table 1 reports findings obtained under the same scheme, though having the DISA rather than the individual client as the unit of analysis.

Indicators	DISAs in the two groups (N)	Mean Family Planning Use Rate at Month 12		<i>t</i> ^a	Effect Size (ES)	Counternull ES ^b
		Control Cohort	Treated Cohort			
Crude Use Rate: All Methods	24	.71	.75	.50	.20	.40
Quality-Related Use: All Methods	24	.71	.73	.47	.10	.20
Index of Goal Attainment: Crude	24	.92	.95	.76	.31	.62
Index of Goal Attainment: Adjusted	24	.95	.95	.43	.01	.02

^a This is the one-tailed *t*-test for differences between the treatment and control means.

^b See text concerning definition of the counternull value of the effect size.

* *p* < .05

The cases corresponding to the minority of providers who used the *Balanced Counseling Strategy*'s job aids in the consultation with clients were not singled out for special study. The number of DISAs lost to follow-up would have made the results extremely unreliable. In several DISAs, none of the clients had been seen by a provider

who used the job aids. Similar was the case of clients who chose barrier/natural methods. Hence, the table only refers to the results in the undivided treated and control cohorts. Data were available at the rate of two through 18 clients per DISA ($N = 24$ DISAs). The results pertain to the second follow-up interview. The first follow-up interview was designed to tap only the first six months after recruitment.

The crude use rate for a given DISA is the number of cases using family planning at the 12th month, divided by the total number of cases in that DISA. The crude use rate was non-significantly larger in the treated cohort than in the control cohort ($t = .50, p < .32$, one-tailed). The correlation between use rate and DISA size was negative and non-significant ($r = -.29, df = 22$). However, the fact that negative correlations were observed either within the control ($r = -.21, df = 10$) or treated ($r = -.37, df = 10$) cohorts suggests that the lack of significance was due to the small sample sizes and not to a lack of relationship between the two variables.

The quality-related use rate is a crude use rate adjusted considering current use intentions, i.e., excluding all the discontinuation segments associated with reduced need as the stated reason for the discontinuation. Since the lack of initiation of use of the method chosen is not a discontinuation properly, the calendar module of the questionnaire lacked a question on the reasons for not starting use. For the sake of rigor, we assumed that all these cases were due to quality-related factors. With the reduced-need discontinuations treated as continuations in the month ratio, we recalculated the use rates and thus obtained quality-related use rates. The difference between cohorts was non-significant ($t = .47, p < .42$, one-tailed)

In calculating the crude index of attainment of reproductive goals it is assumed that the respondent has maintained the reproductive goal stated at recruitment time. Consequently, the absence of a pregnancy or an abortion at month 12 of follow-up counts as goal attainment (a 1 score). The adjusted index took into account the reproductive goals stated month by month in the calendar module of the questionnaire. Hence, having had an abortion or being pregnant could receive a 1 or a 0 depending on the reproductive goal stated for month 12. It received a 1 if the abortion or pregnancy coincided with the expressed desire for being pregnant but a 0 if the client had not changed her reproductive

intentions. Similarly, not having had an abortion or been pregnant could receive a 1 (if the client maintained her original reproductive goal) or a 0 (if the reproductive goal had changed). The differences between cohorts in terms of either index failed to reach significance.

All the effect sizes in Table 1 were smaller than the significant effect size obtained in Phase I of the study for knowledge of the method chosen when this was an IUD or hormonal method ($ES = .83, p < .05$; see León, Ríos, & Zumarán, 2003) and considerably smaller than the effects of the intervention on the quality of care ($ES = 2.65, p < .001$). As for the counternull value of the observed effect size, this is obtained as: $ES_{\text{counternull}} = 2ES_{\text{obtained}} - ES_{\text{null}}$. Because the effect size expected under the null is zero, the value of the counternull is simply twice the obtained effect size. This value merited special consideration in two cases: the crude index of reproductive goal attainment ($ES = .31$) and the crude use rate ($ES = .20$). The evidence of this study strongly indicates that these effect sizes were not significantly different from zero as they indicate that the effect sizes were not significantly different from .62 or .40, respectively.

Use of Family Planning Over the Whole Follow-Up Period

One problem with the Lacuesta *et al.* (2001) and Sanogo *et al.* (2003) analytic approach followed in the construction of Table 1 is that, by focusing on the n^{th} month of follow-up, it ignores what went on in the past $n - 1$ months. If the effects of our intervention lasted for less than one year, the measurements at month 12 would utterly fail to register the impact in the prior months and a Type II error would be committed by concluding that the intervention had no effects. Another problem entails the reliability of measurement. By averaging subjects within DISAs the reliability attained is greater than by using the single client as the unit of analysis. Yet, an average based on just two clients (as was the case in one DISA) or a slightly higher number does not bring substantial gain.

To parsimoniously and more reliably capture the effects of the intervention over the entire 12-month follow-up period, we appealed to a solution that was used in one of the preliminary data analyses (see León *et al.*, 2003d). We integrated each set of 12 DISA averages into a single mean cohort score for each given month. In other words, with respect to each indicator, we obtained for each cohort an average score across DISAs per

month, i.e., 12 monthly experimental means and 12 monthly control means. Therefore, the DISA disappeared as the unit of analysis and was replaced by a higher-order unit that is expected to be more reliable. The analysis would have encompassed 12 months x 24 DISAs = 288 DISA means, too, if we had averaged the 12 months for each DISA instead of the 12 DISAs for each month. Yet, the month averages are based on a greater *minimal* number of informants (12 times the average DISA size) than the DISA averages, one of which would be based on 12 times the minimal DISA size and others on only slightly greater numbers of informants.

The data analyzed stem from the second follow-up and refer to the one-year history of the clients. For each of the four indicators of Table 1, we submitted the 24 monthly data points to a randomized-block analysis of variance (Kirk, 1968) in which the months were handled as treatments ($k = 12$) and the levels of the intervention as blocks ($n = 2$). Under the randomized block approach, the underlying structure is a mixed model comprising fixed effects for the repeated measurements factor and random effects for the blocks. The fixed-effects model is appropriate for situations in which all treatment levels about which inferences are to be drawn are included in the experiment (1-year family planning use following method choice). If the experiment were replicated, the same treatment levels would be included in the replication. Under these conditions, conclusions drawn from the experiment apply only to the k treatment levels included in the experiment. On the other hand, the blocks are conceived to represent a random sample from a population of interventions. In the present case, one block (the experimental cohort) represents an intervention that caused 37 percent of the providers to fully use the *Balanced Counseling Strategy* in their interactions with clients (see León *et al.*, 2002b) and the other (the control cohort) one in which none of the providers implemented the *Strategy* (the providers neither received training on the *Strategy* nor the *Strategy*'s job aids). Other blocks can be imagined, e.g., an intervention that causes more than 70 percent of the providers to implement the *Strategy*, as we achieved in Guatemala (León *et al.*, 2003a). The population of interventions is assumed to have a normal distribution around a mean of 50 percent of providers implementing the *Strategy*. Table 2 presents the main results of the randomized block analysis.

The effects of the month factor were significant in three of the four analyses. Generally, the latter months of the follow-up showed less family planning use than the initial months. Reflecting the greater reliability of the monthly data and the increased experimental power of the randomized block design, three of the comparisons between treated and control cohorts reached statistical significance despite that the differences between means were no larger than those observed in Table 1. In fact, the largest mean difference, pertaining to the quality-related use rate, only amounted to 3.1 prevalence points. The difference in crude use rate amounted to 2.6 prevalence points and the difference in adjusted goal attainment only reached 1.7 points.

Table 2. Effects of the Intervention over the Whole Follow-Up Period (Monthly Averages^a)

Indicators	F _{Months} ^b	Mean Monthly Family Planning Use Rate Average		F _{Intervention} ^c	Effect Size (ES)	Counternull ES
		Control Cohort	Treated Cohort			
Crude Use Rate: All Methods	13.92**	.78	.80	5.37*	.47	.94
Quality-Related Use: All Methods	43.23**	.78	.81	31.69**	1.69	3.38
Index of Goal Attainment: Crude	5.74**	.96	.97	3.02	.35	.71
Index of Goal Attainment: Adjusted	2.73	.95	.96	12.12**	.71	1.42

^a Whereas the unit of analysis for the data in Table 1 was the DISA, in this case the unit of analysis is the monthly average across DISAs. We had 12 such values for the treated cohort and 12 for the control cohort.

^b This is the F-ratio of the analysis of variance for the first factor. Degrees of freedom for the month effects are 11, 11.

^c This is the F-ratio for the second factor. The degrees of freedom for the intervention effects are 1,11.

* $p < .05$, *

* $p < .01$.

The effect size for the crude use rate was significantly greater than 0σ and significantly smaller than $.94\sigma$. The effect size for the quality-related use rate was significantly greater than 0σ and significantly smaller than 3.38σ . The effect size for adjusted reproductive goal attainment was significantly greater than 0σ and significantly smaller than 1.42σ .

Trend Analysis of Crude Family Planning Use Rates

The randomized block analysis of variance used in the previous section offered the advantage of dealing with such a reliable indicator as the monthly average, but lacked

an interaction term. Hence, we could not know whether the lack of significant differences between means over the 12 months of follow-up that was observed in two comparisons was due to a pattern of results that prevailed along the whole period or reflected a cancellation of positive effects in one segment of time by negative effects in another segment. To obtain a more dynamic account of the results over time we submitted the monthly averages to a trend analysis.

The *raw* crude family planning use rate for the treated cohort was smaller than that for the control cohort in months 1, 2, and 3, equal in month 5, and greater in months 4, 6, and thereafter. The probability for the *raw* use rate to be greater in the treated than in the control cohort in each month of the second semester ($n = 6$) is only .016 on the basis of chance alone, according to the sign test/binomial distribution (Siegel, 1956). In contrast, there was no significant difference during the first semester ($p < .99$, two-tailed, $n = 6$). That is, we detected a significant interaction between semesters and intervention. These results are more easily seen in the smoothed curves of Figure 1, that show a cross-over at month 3.5. This interaction was a reliable finding. It can be observed again in Figure 2, that presents results from the first follow-up interview. Stemming from interviews conducted one semester earlier, such findings confirm a lagged effect of the intervention that could not be explained on the basis of the possible distorting effects of distant recall. Given the redundancy between the 12-month and six-month follow-up data, only the data from the second follow-up are considered in subsequent analyses.

To better understand these results, we obtained trends for sub-samples of clients at the cost of dealing with a reduced number of DISAs: clients that had chosen an IUD or hormonal method at recruitment, available at 22 DISAs, and clients that had chosen a barrier or natural method, available at 20 DISAs. (If one DISA had zero cases with these characteristics, its paired DISA was discarded too). Figures 3 and 4 present the one-year monthly family planning use rate *trends* for each group. Intervention x time interactions are observed in both cases. Also noteworthy is that users of the IUD or hormonal methods generated higher use rates that started at 1.0 or close to 1.0 at month one and did not fall below .70 at month 12 whereas users of barrier or natural methods generated use rates that started below .80 at month one and fell below .70 at month 12.

Figure 1. Trends for Contraceptive Use Rate in Experimental (EC) and Control Cohorts (CC):
Second Follow-Up, All Methods

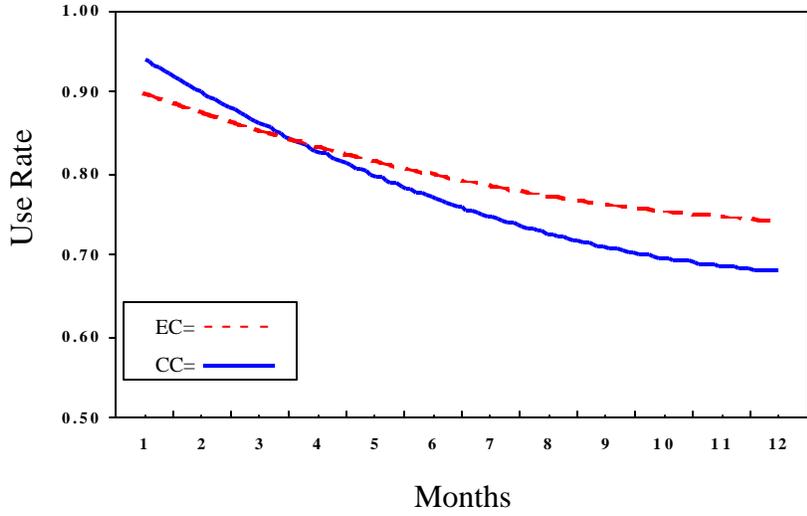


Figure 2. Trends for Contraceptive Use Rate in Experimental (EC) and Control Cohorts (CC):
First Follow-Up, All Methods

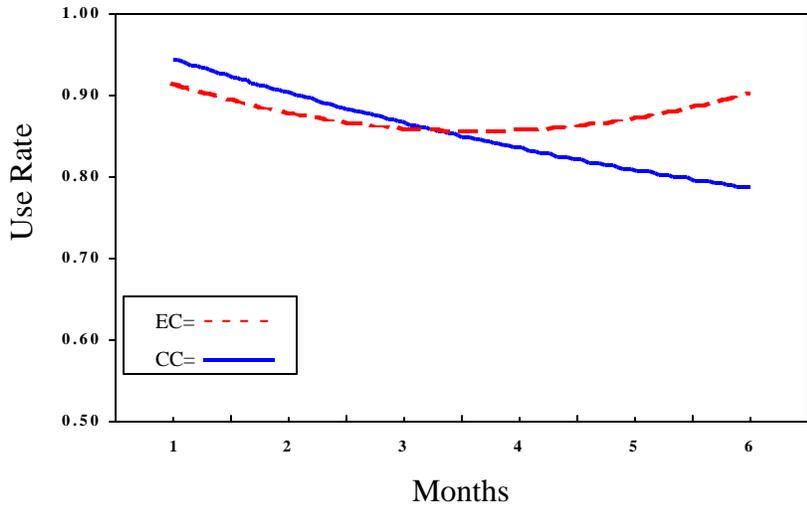


Figure 3. Trends for Contraceptive Use Rate in Experimental (EC) and Control Cohorts (CC): Second Follow-Up, IUD/Hormonal Methods

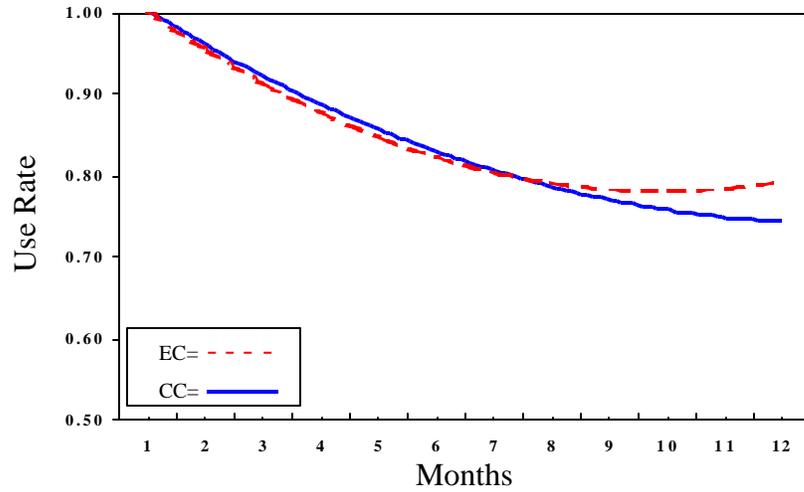
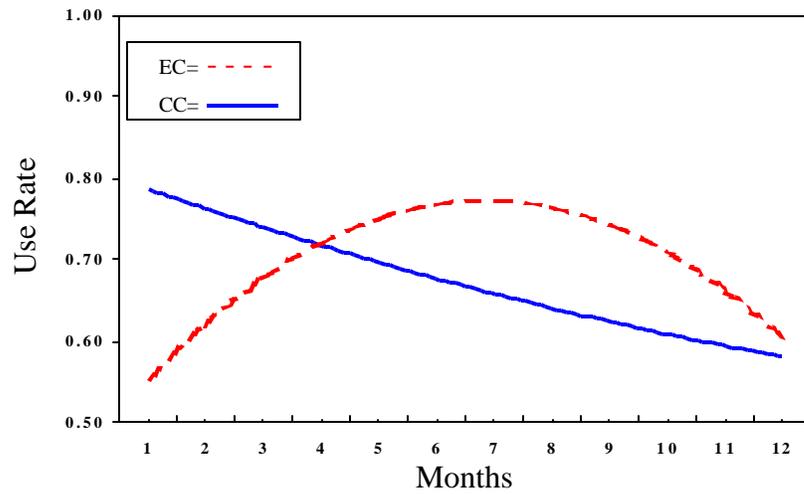


Figure 4. Trends for Contraceptive Use Rate in Experimental (EC) and Control Cohorts (CC): Second Follow-Up, Barrier/Natural Methods



The failures to initiate use of a method by month one shown by the curves in Figure 1 are accounted for by the behavior of the clients who chose to initiate barrier or natural methods after the period. The clients who chose the IUD or a hormonal method immediately implemented their decision. Whereas the IUD, DMPA, and pill can be considered in use once inserted, injected, or taken regardless of the user's sexual activity, the use of barrier/natural methods is contingent on sexual activity and partner cooperation. Consequently, the findings suggest that a number of clients miscalculated their likelihood of sexual activity and/or the required partner cooperation. Alternatively, some clients might have received a condom as a temporary method and neither used it, nor returned for IUD insertion or a hormonal method when they were ready for it.

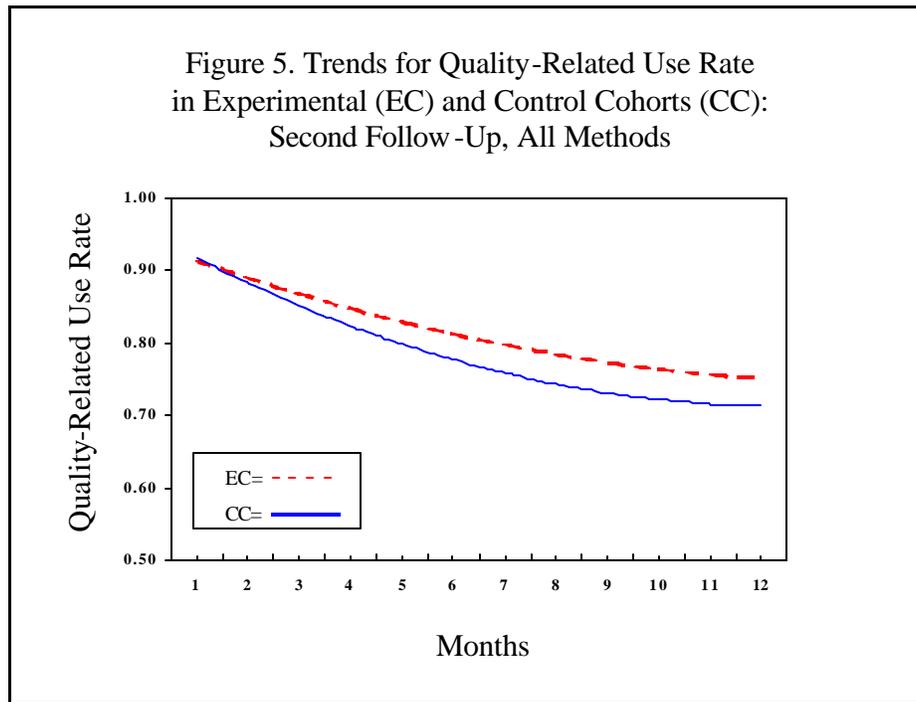
The trends per method also explain the differences between cohorts in Figure 1. The greater use of contraception by the control clients before the cross-over of the curves in that figure is explained by their greater use of barrier/natural methods during the first quarter of follow-up. At the other extreme, the greater use of contraception by the treated clients during the last quarter of follow-up is explained by their greater use of IUD/hormonal methods. The greater use of contraception by the treated cohort in the midst of the follow-up period is explained by the greater use of barrier/natural methods.

Trend Analysis of Quality-Related Use Rates

The trends so far presented not only supply evidence on the effects of the intervention on family planning use after method choice. They also speak to the rate of behavioral consistency with the reproductive goals that were stated by the clients one year earlier, at which time all the clients wanted to avoid pregnancy for at least 12 months. To obtain trends that reflected the failure to behave consistently with *contemporary* use intentions, we considered the monthly quality-related use rates, from which the discontinuations due to reduced need had been excluded.

Figure 5 presents the results. The *trends* for quality-related use rate differed from the pattern of the unadjusted *trends* in that the intervention x time interaction had an earlier point of intersection of the curves (month 1 rather than month 3.5). The *raw* quality-related use rate average was equal in the treated and control cohorts at months 1 and 4 but greater in the former at all the other months. The probability that this outcome

arises from chance alone is equal to .019 according to the sign test (one-tailed). This is consistent with the results of the statistical analyses of Table 2. Further breakdowns of the samples to address specific methods were not pursued here; they would have led to an excessive loss of DISAs.



Trend Analysis of Client’s Index of Attainment of Reproductive Goals

This variable yielded crude and adjusted indicators. Figures 6 and 7 present the respective trends. The crude index of attainment of reproductive goals exhibited a pattern that was similar to the pattern of contraceptive use. The descending scores over months were associated with a significant F ratio in Table 2. The superior levels of use shown by the treated cohort in the second semester were cancelled by the results of the first semester and this seems to have determined the lack of significance of the F ratio for the intervention in Table 2.

The raw monthly averages for adjusted goal attainment differed from those entailing contraceptive use in that they did not show a substantial decline over months (hence the lack of significance of F_{Months} in Table 2). On the other hand, at months 5, 6, and 8-through-12 the *raw* treated cohort average was above the *raw* control cohort

Figure 6. Trends for Crude Goal Satisfaction in Experimental (EC) and Control Cohorts (CC): Second Follow-Up

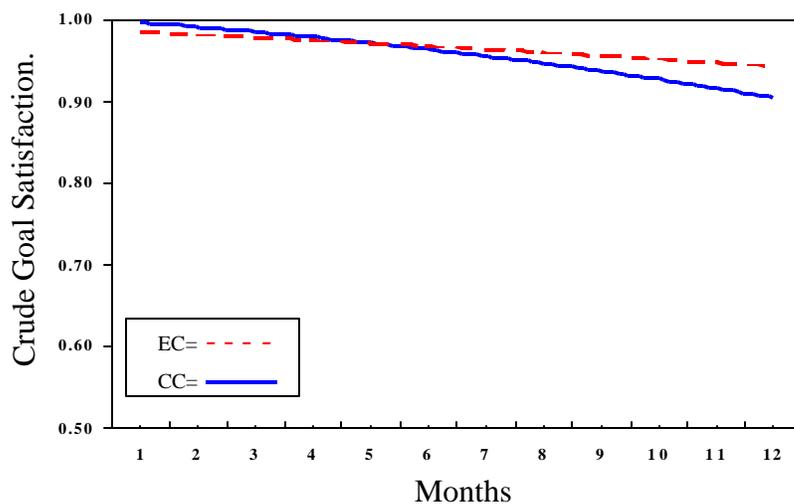
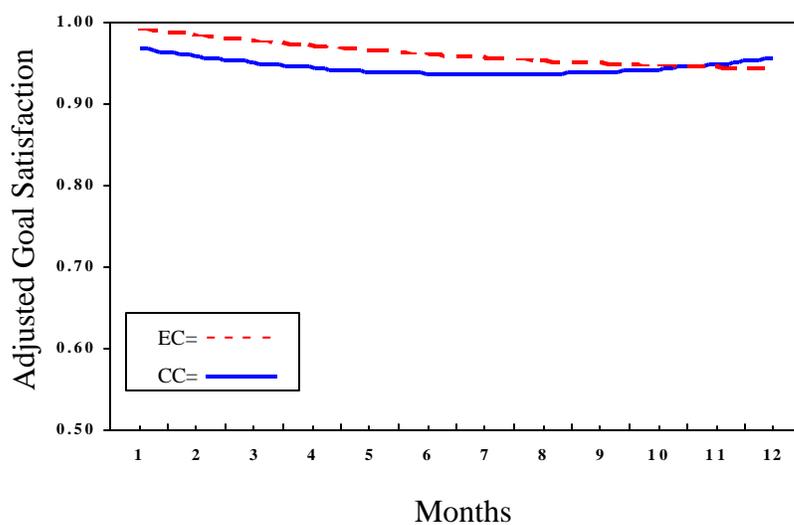


Figure 7. Trends for Adjusted Goal Satisfaction in Experimental (EC) and Control Cohorts (CC): Second Follow-Up



average (hence the significance of $F_{\text{Intervention}}$ in Table 2). That is, the intervention influenced the attainment of reproductive goals when these were contemporaneously defined but not when they were defined one year in advance.

All-Method Cumulative Continuation Rates from Life Tables

All the previous analyses had the DISA or a higher-order average as the unit of analysis. This section is concerned with the analysis of continuous segments of family planning use according to standard practice, i.e., having the segments as units of analysis. In addition to the fact that the segment of use, not the DISA, is the unit of analysis, three differences must be taken into account in the comparison of the continuation rate addressed in this section with the use rate addressed earlier.

- First, the cases that failed to initiate use of a method were excluded by default in the life-table analysis.
- Second, whereas the use rates referred to actual months elapsed since the time of client recruitment, the continuation rates pertain to segments of use with virtual initiation times. For example, a user who discontinued at month six, reinitiated use in month 10 and discontinued again at month 11, contributes two segments that are treated as if they came from different persons who initiate use of a method and thus enter into the cumulative process at month 0.
- Third, the denominator of the continuation rate is the total number of segments of continuous use, whereas the denominator of the use rate is the total number of clients. Consequently, the continuation rate will be almost always smaller than the use rate. It cannot be greater than the use rate. The two rates can be equal only when there are zero cases of reinitiation of contraceptive use after a discontinuation.

Table 3 presents, for each cohort, the life-table cumulative survival rate per month. The table also presents the cumulative proportion of discontinuations due to reduced need and due to quality-related reasons. The life-table from which these data were obtained actually had 2 lines per month, corresponding to the half-months of the calendar module used in the survey. The second half-month was selected for presentation

Table 3. Cumulative Family Planning Continuation and Discontinuation Rates

Ordinal Month	Surviving at End of Month		Discontinued, Reduced Need Reason		Discontinued, Quality-Related Reason	
	Control Cohort	Treated Cohort	Control Cohort	Treated Cohort	Control Cohort	Treated Cohort
1	.91	.89	.03	.04	.05	.07
2	.90	.85	.03	.05	.07	.09
3	.85	.77	.05	.07	.11	.16
4	.80	.71	.06	.08	.13	.21
5	.77	.70	.07	.08	.16	.22
6	.72	.68	.09	.08	.19	.24
7	.70	.66	.09	.08	.21	.26
8	.69	.61	.09	.09	.22	.30
9	.64	.58	.10	.10	.25	.31
10	.63	.55	.10	.10	.26	.34
11	.62	.55	.10	.10	.27	.34
12	.44	.48	.17	.10	.39	.42

here. The number of segments in the control cohort ranged from 155 entering month 0 through 75 entering month 12. The number of segments in the treated cohort ranged from 83 entering month 0 through 39 entering month 12. In these analyses, the segment represents a continuous use of contraception; method shifting does not count as discontinuation.

The cumulative continuation rate at the end of the follow-up period was .44 in the control cohort and .48 in the treated cohort. To test for differences between cohorts over the whole follow-up period, we compared the survival functions from half-month 0 through half-month 24 using the Wilcoxon (Gehan) test. This test is concerned with the agreement between two cumulative distributions. If the two sample cumulative distributions are “too far apart” at any point according to a statistical criterion, this suggests that the samples come from different populations.

Table 4 presents the results of the Wilcoxon (Gehan) test. The median survival time has been converted from half-months into months. The average score is calculated by comparing each case to all others and incrementing the score for a case by one if a case has a longer survival time than another case and decrementing it by one if the case has a shorter survival time. The first part of the table presents the results for all the use segments. The second part refers to the results of an analysis in which the reduced-need segments were treated as missing values and only quality-related discontinuations were

Table 4. Wilcoxon (Gehan) Test: Half-Monthly Distributions of Family Planning Use

Cohort	Total N	Median Survival Time	Average	Statistic	Two-tailed Probability
All Segments					
Control	155	12.0 months	4.68	.72	.40
Experimental	83	12.0 months	-8.73		
Reduced-Need Segments Treated as Missing Values					
Control	136	12.0 months	3.68	.56	.45
Experimental	74	12.0 months	-6.77		

considered. The signs of the averages suggest greater continuation in the control group, yet the differences between distributions were non-significant in both cases.

Estimation of Method Switching

To estimate the amount of method switching, we repeated the life-table analysis redefining the segment of use as one of continuous use of any given method. In these analyses, method switching counted as a discontinuation. Tables 5 and 6 present the results of the analysis.

Table 5. Cumulative Method Continuation and Discontinuation Rates

Ordinal Month	Surviving at End of Month		Discontinued, Reduced Need Reason		Discontinued, Quality-Related Reason	
	Control Cohort	Treated Cohort	Control Cohort	Treated Cohort	Control Cohort	Treated Cohort
1	.87	.78	.03	.05	.11	.17
2	.79	.74	.03	.06	.18	.20
3	.63	.60	.06	.07	.31	.33
4	.55	.53	.06	.08	.38	.40
5	.54	.52	.06	.08	.40	.41
6	.45	.50	.08	.08	.48	.42
7	.39	.48	.08	.08	.52	.44
8	.39	.43	.08	.09	.53	.48
9	.33	.35	.09	.11	.58	.54
10	.32	.33	.09	.11	.58	.57
11	.32	.31	.09	.11	.58	.58
12	.15	.20	.13	.11	.72	.69

Cohort	Total N	Median Survival Time	Average	Statistic	Two-tailed Probability
All Segments					
Control	241	6.15 months	2.46	.12	.73
Experimental	125	6.44 months	-4.74		
Reduced-Need Segments Treated as Missing Values					
Control	219	6.29 months	-.10	.00	.99
Experimental	114	7.88 months	.19		

The results were more favorable to the research hypothesis under the quality-related mode of analysis but the differences between the distributions were statistically non-significant. The algebraic difference between the *method* continuation rate (Table 6) and the *all-method* continuation rate (Table 4) yields an index of method switching. This was .29 in the control cohort and .26 in the experimental cohort at the end of the follow-up period, i.e., trivial differences in method switching were observed.

Client Knowledge Concerning the Method Used

At the end of the follow-up period, complete client knowledge scores concerning the method chosen when this was the IUD or a hormonal method were available in the experimental group at 10 of the 12 DISAs. (The *N* fell to 20 when the matched control DISAs were dropped from analysis.) The average score for the treated cohort was significantly greater than that for the control cohort (see Table 7).

Indicators	Number of DISAs ^a	Control Cohort		Treated Cohort		Effect Size ^b
		Subjects Per DISA	Mean Score	Subjects per DISA	Mean Score	
Knowledge IUD/Hormonal Methods	20	5.9	10.4	3.3	12.2	.88*
Knowledge Barrier/Natural Methods	6	1.0	5.0	2.0	5.8	.90

^aDegrees of freedom = *N* – 2.

^bThe asterisk refers to the level of significance of the respective t-test.

* *p* < .05.

Knowledge scores pertaining to the method chosen when this was a barrier or natural method were available only at a few treated DISAs and the difference between treated and control cohorts did not reach significance.

Methodological Discussion

The study design and implementation pose a number of problems in the interpretation of results and evaluation of the findings.

Reliability of Measurement

León *et al.* (2003e) reported results of a preliminary analysis concerned with months 1-6 of the follow-up period. Important inconsistencies in reported family planning use for this period were found at the individual level between the clients' responses obtained in the first interview, that encompassed months 1-6, and the second interview, that encompassed months 1-12. The study hypothesis concerning family planning use in months 1-6 was more strongly supported when we excluded from analysis the cases whose use history from the two interviews did not coincide exactly. In that analysis, the definition of month one only considered the month of recruitment. Regardless of exact date, if the client was recruited in July, July was defined as month one.

The results became more reliable when, following consultant advice, we changed the definition to the one described in the methodological section of this report. According to the new definition, July is defined as month one if the client was recruited during the first 15 days of July. If the client was recruited in the second half of that month, August is defined as month one. Client memory problems are expected to be more effectively controlled under this definition. Indeed, whereas the new definition did not eliminate the unreliability of the data at the individual level, it made the average results more consistent. For example, cross-over of the trends were nearly identical in Figures 1 and 2.

Statistical Validity: Client Knowledge

Testing the research hypothesis with respect to client knowledge was a simple operation. Knowledge of the method used when this was the IUD or a hormonal method

showed significantly greater scores in the experimental than the control cohorts by the end of the follow-up period. Knowledge of the method used when this was a barrier or natural method presented a similar tendency but had a presence only in three pairs of DISAs and did not attain significance.

Statistical Validity: Use Rate and Goal Attainment Index

When we analyzed family planning use and attainment of reproductive goals at a single point in time one year after client recruitment (end of month 12), the results were very similar to those of Sanogo *et al.* (2003): in the expected direction but failing to reach statistical significance. This, however, should not lead us to conclude that the intervention had no effects at the 12th month of follow-up. What the evidence obtained told us was that, for quality-related contraceptive use and adjusted index of reproductive goal attainment, the effects were neither significantly different from zero nor significantly different from about half a standard deviation. Whereas this distinction lacks practical value, it has considerable theoretical value.

When the reliability of the data was increased by averaging rates across DISAs, and the experimental power enhanced by submitting the one-year monthly averages to a randomized-block analysis of variance, statistically significant effects were detected. Carry-over effects were minimized by allowing discontinuers to reenter the data set if they restarted using contraception at later months. Yet, a requirement of the statistical model was violated when we defined each block as a treatment condition that receives all the levels of the month variable. The nature of the month variable precluded randomization of order. The fact that a sizable percentage of users (those of DMPA) were necessarily protected for three continuous months caused contiguous months to be correlated. Given that each treatment condition was used as its own control, it is unlikely that all the covariances were equal and it is possible that the error term of the randomized block was artificially reduced. One way of dealing with this problem is setting a more demanding level of statistical significance. With a .01 significance level, the results establish that crude family planning use and crude index of goal attainment failed to reach significance whereas the effects for quality-related contraceptive use and adjusted index of reproductive goal attainment were significant.

Statistical Validity: Continuation Rate

The methodological priority of this study was to ensure the equivalence of the cohorts at the baseline in order to maximize the legitimacy of attributing cause in the absence of a pretest concerning contraceptive use. Having decided on the DISA as the unit of experimentation to avoid the risk of contamination by information and/or materials crossed between providers or between their supervisors, the most parsimonious way of dealing with the data was having the DISA as the unit of analysis. Hence, the differences in client knowledge, use rate, and reproductive goal attainment expected in the follow-up study could be attributed to the intervention. This is why the research proposal for this study stated that the adequate unit of analysis was the DISA (León, 2002), and the DISA averages were given equal weight in the analysis of these indicators.

Since the DISA samples were of markedly different sizes, random assignment of DISAs to treatment conditions was, by definition, less efficient in attaining group equivalence insofar as the segments of continuous family planning were taken as the units of analysis. Whereas in the experimental design of the study the DISA averages had equal weight, in the life-table analyses the DISAs did not have equal representation. A DISA with 18 clients contributed a much larger number of family planning use segments than one with only two clients. Hence, in terms of the internal validity of the experiment, the smaller DISAs were under-represented in the analyses of use segments and the larger ones were over-represented. This probably had consequences given the negative correlation between family planning use and DISA size that was found within the control cohort as well as within the treated cohort. The fact that the absolute magnitude of the former was greater than that of the latter suggests that the study effects were stronger in the smaller DISAs. The under-representation of the smaller DISAs in the life-table analyses implies that the probability of confirming the research hypothesis was reduced. That is, the continuation rates, being over-determined by the larger DISAs, were biased against the research hypothesis of the study.

External Validity

By giving equal weight to each DISA average in the calculation of family planning use rates and indexes of goal attainment, we maximized the internal validity of

the findings but introduced a distortion in the external-validity dimension. Equal weighting implied over-representing in the study the less populated DISAs of Peru and under-representing the most populated ones. If the negative correlation between DISA size and client family planning use that was found in the two cohorts held at the population level, the absolute sizes of the use rates reported in this study would represent overestimates of the population values. That is, at the Peru MOH national level, the actual rate of family planning use 12 months after method choice probably is smaller than the rate = .71 found in the control group of this study. Likewise, the average use rate over the 12 months of follow-up probably is smaller than the .78 found in the control group. The results from the life-table analyses are free of this problem because the raw data were more representative of the general population. At most the error was augmented due to irregularities in the sampling of cases, but a bias was not introduced. However, the continuation rates must be appropriately interpreted. They are proportions of segments of continuous use, not proportions of clients.

The weighting issue forces us to take two stands with respect to generalizing *relationships* between the independent (intervention versus control group) and dependent variables of this study (the client behavior and outcomes observed). It is legitimate to generalize to all of Peru the findings of this study concerning impacts on use rate and attainment of reproductive goals insofar as we maintain the DISA as the unit of analysis. However, if we wished to talk about the general population of Peruvian women, special caution would be required. If the population correlation between family planning use and DISA size, in fact, is negative, the relationship between the intervention and client behavior and outcomes probably is weaker at the level of the general population of women than what we found here.

Internal Validity

The research design, consisting in the random assignment of 24 paired health directorates to the experimental conditions of the study, promised strong internal validity to the extent that group equivalence at the baseline could be justified by the other study procedures and outcomes. If group equivalence was justified, we could interpret the

treatment-control differences observed in the follow-up as changes reflecting the effects of the intervention (Bauman, Viadro, and Tsui, 1994; Cook and Campbell, 1979).

Due to employee turnover, change of location, annual leave, ill health, and other factors, varying numbers of providers, depending on the indicator, were lost to follow-up in Phase 1 of the study. The sample attrition apparently was unbiased: analyses of baseline data limited to the providers who did participate in the posttest yielded non-significant treatment-control differences for each of the indicators covered in the pretest (León *et al.*, 2002b).

Another possible source of bias that requires evaluation is provider self-selected reception of the intervention. The follow-up study only considered clients of the treated-group providers who had received the full intervention, in addition to all the clients available in the control group. If the providers with poorer job performance self-selected themselves out of the intervention, the quality of care offered by the treated group, and consequently client use of family planning following method choice, would automatically be inflated. In this case it would be difficult to disentangle the effects of the intervention from the effects of provider self-selection. That is, we could conclude that the quality of care influenced the client outcomes but not that the intervention influenced the quality of care and, through it, the client outcomes.

Against this alternative hypothesis, however, is the fact that our results were not limited to simulated client data that in the posttest demonstrated significant differences in quality of care between control providers and providers who had received the full intervention. The direct observation of the client-provider interactions, too, demonstrated significant differences in quality of care between control providers and providers from the treated clinics regardless of whether they had participated in two, one, or none of the intervention workshops (León *et al.*, 2002b). That is, provider self-selection was controlled in this case and the results continued to be consistent with the hypothesis that the intervention would cause positive effects on the quality of care.

V. Conclusions and Recommendations

In a generalized misapplication of the quality-of-care concept, family planning providers often talk in excess about the methods available at the program and too little about the method chosen by the client. Consequently, the client can be expected to suffer information overload and impaired learning. Phase 1 of this study showed that three-day provider training on the job aids-assisted *Balanced Counseling Strategy* was enough to modify this situation. Provider behavior improved in the three general areas of individual counseling that were evaluated as well as with respect to the post-choice phases. Furthermore, a significant improvement was observed in client's spontaneous recall of attributes of the method chosen when this was an IUD or hormonal method (León, Ríos, and Zumarán, 2003).

The question addressed in Phase II of the study was whether these effects had any relevance for the long-term use of contraception and the satisfaction of the client's reproductive goals. Since the results have different implications in the theoretical, methodological, and practical realms, this section will present separate conclusions and recommendations in these areas.

Theoretical Issues

Controlled quality-of-care improvements modestly increase 1-year use of needed contraception following method choice

Clients who one year before had stated a desire to use contraception for at least one year and had chosen a method were interviewed concerning their contraceptive behavior and subjective status over this period. The results of the study showed that those who had received services from providers trained during three days on the *Balanced Counseling Strategy*, i.e., had received services of greater quality, used needed contraception to a modest but significantly greater extent than control clients ($p < .01$, $df = 1, 11$). Impacts were not observed at the discrete 12th month of follow-up but only when the whole 12-month period was considered. This reveals that the impacts were consistent but of small magnitude. However, the results are of scientific value considering the

longitudinal effect size observed, greater than 1.5 standard deviations. Trend analyses suggested that the impacts continued beyond the 12th month.

The effects of the intervention were not statistically significant at the .01 probability level when the impact was assessed considering the use intentions stated 12 months before, i.e., when all the observed discontinuations of use were counted regardless of the reason for discontinuation stated by the client at the follow-up interview. The results were significant at the .01 level only when the contemporaneous client perceptions of need for contraception were taken into account. That is, when discontinuations of use due to new situations of reduced need (e.g., partner absence, onset of menopause) were not counted as discontinuations and only those related to the quality of care (e.g., side effects, method failure) were counted.

Controlled quality-of-care improvements increase client's attainment of contemporaneous reproductive goal

This study made a distinction between use of contraception and reproductive goal attainment. The former refers to a behavior given a one-sided reproductive goal. The latter pertains to an outcome given two-sided goals: avoiding pregnancy if one does not desire it and achieving pregnancy if one desires it. The results of the study did not lend support to the hypothesis that the intervention would improve the attainment of reproductive goals when the outcome criterion was the reproductive goal stated one year earlier. The difference between cohorts, however, was positive and achieved statistical significance when the outcome was defined in terms of contemporary reproductive goals ($p < .01$, $df = 1, 11$).

That use of needed contraception following choice of a method and reproductive goal attainment in terms of contemporaneous goals require a differentiated conceptual treatment is demonstrated by the fact that they evolved differently over time. Whereas use of needed contraception systematically decreased from the first through the 12th month of follow-up ($p < .01$, $df = 1, 11$), reproductive goal attainment in terms of contemporaneous goals was indifferent to the passage of time.

The type of method chosen strongly influences the dynamics of family planning use

The use of IUD/hormonal methods chosen in the interaction with providers started in nearly 100 percent of the cases immediately after method choice. On the other hand, less than 80 percent of the clients implemented contraception immediately when their choice was a barrier or natural method.

This can be explained by the fact that the use of barrier or natural methods depends on the opportunities for coital events and requires partner cooperation. It would seem as if important percentages of clients miscalculated the opportunities for sex or the cooperation from the partner in the use of family planning. An alternative explanation is that some clients received condoms as a temporary method and never used them nor returned for an IUD insertion or when they were ready to receive a hormonal method. Attempting to explain the differences observed between treated and control cohorts in this area would be speculative.

In this study, the causal role of method choice could not be disentangled from that of using the Balance Counseling Strategy's method pamphlets

The positive effects of the intervention on long-term use of contraception and reproductive goal attainment can be attributed to the client's better choice of method afforded one year earlier in the interaction with a provider exposed to the *Balanced Counseling Strategy*. The simplification of the counseling task for the client by means of sequential decision-making made her less exposed to information overload pertaining to irrelevant methods and probably enhanced the choice of the method best suited for her.

It is also possible that the method pamphlet of the *Strategy*, being a memory aid for the client, played a positive role. One year after the family planning consultation the clients of the treated cohort still showed significantly superior knowledge concerning the method used when this was the IUD or a hormonal method ($p < .05$, one-tailed, $df = 18$). This is unlikely to have occurred unless the method pamphlet was repeatedly reviewed by the clients. Yet, there is a problem with this interpretation. The intervention did not cause a significant enhancement of client knowledge concerning the method used when this was a natural or barrier method, whereas the intervention's success entailing contraceptive use encompassed clients who had chosen a barrier or natural method.

Perhaps these clients effectively used of the pamphlet during the follow-up period but their increased knowledge concerning the method used was not demonstrated because of limitations of the measurement tools.

Methodological Issues

The study findings are not inconsistent with the relevant empirical literature

The results of this study do not contradict those of Lacuesta *et al.* (2001) or Sanogo *et al.* (2003). The Philippines findings can be parsimoniously explained referring to the notion that the intervention was not strong enough to cause substantial changes in quality of care. The Senegal findings were more similar to ours. Had the impact of the better quality of care provided by reference centers been evaluated over the whole follow-up period and quality-related use targeted as the outcome instead of general family planning use, statistically significant effects might have been observed in that study. Neither are the findings reported here inconsistent with those of an Egyptian study conducted in parallel with ours that reported non-significantly greater use rates in the control group than in the experimental group (Makhlouf *et al.*, 2003).

To reduce to zero the validity of alternative interpretations, the measurement of needed contraception and reproductive goals would have to be improved

It seems only logical to obtain results that are consistent with the hypothesis of a quality-of-care study only when discontinuations of family planning use due to quality-related reasons are counted and those pertaining to reduced need are ignored (quality-related use rate). When reduced need discontinuations were counted (crude use rate), the differences between cohorts failed to attain the .01 significance level. Similarly, the concurrent reproductive goal is more intuitively appealing than the goal stated one year earlier as a criterion to establish goal attainment.

Yet, reporting on reasons for stopping use of a method belongs in a subjective realm governed by complex attribution mechanisms (Jones *et al.*, 1972; Nisbett and Wilson, 1977; Sperber *et al.*, 1995). For example, owing to a need to reduce cognitive dissonance or just to save face before herself and/or others, a woman who became pregnant because she forgot to take the pill (quality-related discontinuation) may report

that she changed her reproductive intentions and actually wished to have children (reduced-need discontinuation). This implies that, in our study, the positive effects of the intervention on contemporaneously needed use of contraception could have been biased by attribution mechanisms. Similar is the case of the attainment of updated reproductive goals.

However, this is only a theoretical possibility. There are no reasons to expect greater attribution distortions from one cohort. Nonetheless, future studies should find ways to exert control on such plausible attribution mechanisms, distinguishing them from the real changes that may take place in use intentions and reproductive goals. A possible solution is to shorten the period encompassed by the prospective use intention or reproductive goal. Instead of asking about the intentions or goals in a one-year perspective, the questions may pertain to the following three months. This would require more frequently repeated interviews with the clients.

The continuation rate should be abandoned as priority outcome

Client's attainment of reproductive goals is an intrinsically valid outcome. If the reproductive-health program is at the service of clients' needs, it should strive to help them meet their reproductive goals insofar as this does not compromise other aspects of their general health. The quality of care is assumed to help in this process and this study lent support to this contention.

Why is family planning continuation important? Why should it be targeted and studied apart from clients' reproductive goal attainment? The theory states that the quality of care improves family planning continuation and thus increases the prevalence of contraception, which in turn results in diminished fertility (Jain, 1989). But this theory referred to the concept of continuation basically to distinguish between clients already using contraception from clients still to become users. If we already have a cohort of clients that have made the decision to use contraception and, further, have chosen a method, why not directly target the cohort's prevalence of family planning use over the following months, as we did in this study? In an experimental framework, the study of family planning continuation forces the researcher to ignore instance of failure to initiate the use of the method chosen and its consequent negative contribution to prevalence.

Moreover, by labeling as continuation the subject matter, a specific methodology is implicitly selected. The study of contraceptive continuation by means of life tables was invented to deal with virtual client cohorts in demographic studies but is dysfunctional in experimental studies with real cohorts. It displaces segments of reinstituted use to month zero and creates confusion between proportions of segments of continuous use and proportions of clients protected. The use rate is a more simple indicator that should be prioritized in studies of impact of the quality of care on the use of contraception.

Practical Issues

Consider in context the practical value of the amount of change achieved

The standard for judging the scientific value of an experimental finding is the statistical significance and size of the effects achieved, satisfactory in the case of this study. On the other hand, increasing the use of family planning in the target population by 3.1 points (quality-related use rate) may be regarded as a limited practical achievement. Increasing contemporary reproductive goal satisfaction from 94.71 percent to 96.39 percent can also be called into question as a meaningful programmatic accomplishment.

Such practical judgments, however, will be prone to error if they take the findings of this study independently of their historical context. The *Balanced Counseling Strategy* was not yet fully developed when the study started and what was developed was not opportunely and/or fully implemented. The method pamphlets were not ready for the first round of provider workshops. The intervention was limited to three days of provider training of which the first two days introduced uncomfortable method checklists instead of the method pamphlets for clients. The supervisory component remained inert due to a fiscal crisis in Peru. Probably as a consequence of these limitations, only 37 percent of treated providers were found using the *Strategy's* job aids in their daily interactions with clients.

That provider compliance with the behavioral requirements of the new counseling model can be substantially improved was shown in a subsequent study in Guatemala in which over 70 percent of treated providers were found using the job aids in their daily

interactions with clients (León *et al.*, 2003a). Therefore, the practical question to be asked is, what is the amount of change attained when a fully developed *Strategy* is implemented effectively? This points to the need for further research to respond to a practical question.

Establish the impact of the intervention under typical provider performance

A related need stems from the fact that the client cohorts were recruited after receiving services from providers observed by a third party during the counseling. The quality of care is enhanced when providers know that they are under observation (Miller *et al.*, 1991; Ndhlovu, 1999; León *et al.*, 2003c). Thus, the measurement of client knowledge, use of needed contraception, and attainment of contemporaneous reproductive goals can be assumed to have occurred under conditions of maximal provider performance in the experimental group. Whereas this probably affected control and treated providers alike and cannot be regarded as a biasing factor, the study cannot claim the observed client outcomes as products of typical provider behavior.

To maximize practical knowledge, replication studies should focus on the short- and long-term impacts of typical provider performance rather than on the maximal provider performance that in all probability was promoted by the observation of the client-provider interactions. What programs need to know is what the level of impact is under typical everyday circumstances, not when the provider is especially motivated to show her or his best counseling behavior.

Identify the niche that maximizes impacts of the Balanced Counseling Strategy

So far, the *Strategy* has been implemented in two types of settings and the results have been markedly different. One encompassed the primary health centers or posts of the Ministries of Health of Peru and Guatemala at the national or regional level. Immediate effects of the intervention were unambiguously established in terms of provider behavior in these settings. The quality of care was clearly improved in each study (León *et al.*, 2002b, 2003a).

The second setting is the social security system, represented by EsSalud in Peru and the Guatemalan Institute of Social Security (IGSS) in Guatemala. The studies took place in the capital cities of Lima and Guatemala City and involved hospitals and other

facilities that were considerably larger than the average facility covered in the MOH studies. In both countries the evidence in favor of the new counseling model were weak and some of the results ambiguous or plainly negative (León *et al.*, 2002a, 2003b).

Methodological differences may help explain the contrast. However, a more parsimonious interpretation would be one that uses findings from the Peru follow-up study reported here to account for the different level of success of the quality-of-care interventions in the MOH and social security settings. The negative correlations between DISA size and client use of family planning after method choice that were found separately in each cohort suggest that contextual factors are important. The greater negative correlation observed in the experimental than the control cohort of this study further suggested – not proved - that the quality of care intervention was more effective in the smaller DISAs. This factor could explain the greater level of success attained in MOH settings than in social security settings. It may be worth pursuing this lead.

References

Bauman, KF, Viadro CI, Tsui AO, Use of true experimental designs for family planning program evaluations: Merits, problems, and solutions. *International Family Planning Perspectives*, 1994; 20, 3: 108-113.

Blanc AK, Curtis S, Croft T, Does contraceptive discontinuation matter? Quality of care and fertility consequences. *MEASURE Evaluation Technical Report Series, No. 3*. Chapel Hill: Carolina Population Center, University of North Carolina at Chapel Hill, 1999.

Bruce J, Fundamental elements of the quality of care: a simple framework. *Studies in Family Planning*, 1990; 21(2): 61-91.

Cook, TD, Campbell DT, *Quasi-Experimentation: Design and analysis issues for field settings*. Chicago: Rand-McNally, 1979.

Cotten N, Stanback J, Maidouka H, Taylor-Thomas JT, Turk T, Early discontinuation of contraceptive use in Niger and The Gambia, *International Family Planning Perspectives*, 1992, 18: 145-149.

Jain AK, Fertility reduction and the quality of family planning services. *Studies in Family Planning*, 1989, 20, 1: 1-16.

Jain AK, Should eliminating unmet need for contraception continue to be a program priority? *International Family Planning Perspectives*, 1999, 25(Supplement): S39-S43 & S49.

Jones EE, Kanouse DE, Kelley HH, Nisbett RE, Valins S, and Weiner B, Eds. *Attribution: perceiving the causes of behavior*. Morristown, NJ: General Learning Press, 1972.

Kirk R, *Experimental design: procedures for the behavioral sciences*. Belmont, CA: Brooks/Cole, 1969.

Koenig MA, Hossain MB, and Whittaker M, The influence of quality of care upon contraceptive use in rural Bangladesh, *Studies in Family Planning*, 1997, 28(4): 278-289.

Lacuesta M, RamaRao S, Costello M, Jones H, Pangolibay B. Quality matters: Evidence from the Davao project. Paper presented at the 2001 Annual Meeting of the Population Association of America. Washington, D.C.: March 28-30, 2001.

Lei Z-W, Wu SC, Garceau RJ, Jiang S, Yang Q-Z, Wang W-L, and Vander Meulen TC, Effect of pretreatment counseling on discontinuation rates in Chinese women given Depo-medroxyprogesterone Acetate for contraception, *Contraception*, 1996, 53: 357-361.

León FR, Impacts of the Peru MOH's job aids-assisted *Balanced Counseling Strategy* on client contraceptive behavior and outcomes and program costs. *Frontiers in Reproductive Health Program Project Proposal*. Lima: Population Council, 15 January 2002.

León FR, Brambila C, De la Cruz M, Bratt J, García Colindres J, Vásquez B, Morales C, Improving provider-client interactions at Guatemala MOH clinics: extent and cost. *Frontiers in Reproductive Health Program Final Report*. Guatemala City: Population Council, 31 January 2003(a).

León FR, Brambila C, De la Cruz M, Bratt J, Gutiérrez G, Dávila V, Bonatto C, Ríos A,

Effects of IGSS' Job Aids-Assisted *Balanced Counseling Algorithms* on Quality of Care and Client Outcomes. *Frontiers in Reproductive Health Program Final Report*. Guatemala City: Population Council, 5 September 2003(b).

León FR, Espinoza V, Espinosa A, Meza B, *Estrategia de Consejería Balanceada para Mejorar la Calidad de Atención en Planificación Familiar en EsSalud/Lima*. *Frontiers in Reproductive Health Program Final Report*. Lima: Population Council, 29 April 2002(a).

León FR, Espinoza V, Espinosa A, Meza B, Expectancy of being observed and quality of care offered. *Paper prepared for presentation at the 131st Meeting of the American Public Health Association, Session 4301.0*. San Francisco: November 18, 2003(c).

León FR, Ríos A, Zumarán A. The *Balanced Counseling Strategy*. I. Effects on family planning care, cost, and first-level client outcomes. *Unpublished manuscript*. Lima: 19 September 2003.

León FR, Ríos A, Zumarán A, Bratt J. Improving provider-client interactions in Peru MOH clinics: Extent, benefit, and cost. *Frontiers in Reproductive Health Program Final Report*. Lima: Population Council, 13 September 2002(b).

León FR, Roca S, Ríos A, Zumarán A, Feijoo AR, *Peru PCI-QoC Project Bulletin No. 23*. Lima: Population Council, 10 February 2003(d).

León FR, Roca S, Ríos A, Zumarán A, Adamchak S, Feijoo AR, Impacts of a successful quality-of-care intervention on family planning use six months later. *Paper delivered at the Population Association of America 2003 Annual Meeting*. Minneapolis, MI: 1 May 2003(e).

Lipsey, MW, *Design sensitivity: Statistical power for experimental research*. Newbury Park: Sage, 1990.

Makhlouf H, El Hadidi Y, Nawar L, Ibrahim M, Kharboush I, Impact of improved client-provider interaction on women's achievement of fertility goals – Egypt, Phase II. *Draft Final Report of Frontiers-sponsored Project*. El Cairo: Cairo Demographic Center, August 2003.

Miller RA, Ndhlovu L, Gachara MM, Fisher A. The situation analysis study of the family planning program in Kenya, *Studies in Family Planning*, 1991; 22(3):131-143.

Ndhlovu L. Quality in the context of Kenya's integrated reproductive health services, dissertation, University of London, London, 1999.

Pariani SD, Heer M, Van Arsdol MDJr. Does choice make a difference to contraceptive use? Evidence from East Java. *Studies in Family Planning*, 1991, 22, 6: 384-390.

Patel D, Patel A, Mehta A, The effects of service quality on IUD continuation among women in rural Gujarat. In MA Koenig & ME Khan, Eds., *Improving Quality of Care in India's Family Welfare Programme: The Challenge Ahead*. New York: Population Council, 1999, pp. 333-345.

Nisbett RE, Wilson TD, Telling more than we can know: verbal reports on mental processes. *Psychological Review*, 1977, 8, 231-259.

Rosenthal R, Rubin DB, The counternull value of an effect size: a new statistic. *Psychological Science*, 1994, 5(6): 329-334.

Sanogo D, RamaRao S, Jones H, N'diaye P, M'bow B, Bamba Diop C, Improving quality and contraceptive use in Senegal. *Unpublished paper of the Frontiers Program, Population Council*. New York: January 21, 2003.

Siegel S, *Nonparametric statistics for the behavioral sciences*. New York: McGraw-Hill, 1956.

Sperber D, Premack D, and Premack AJ, Eds., *Causal cognition: A multidisciplinary debate*. Oxford, England: Oxford University Press, 1995.

Appendix 1. Calendar Module

INFORMACION PARA SER CODIFICADA EN CADA COLUMNA

INSTRUCCIONES GENERALES

SOLO UN CODIGO POR COLUMNA

COLUMNAS 1,3,4,5 Y 6 SE LLENAN CC

COL 1. Nacimientos, Embarazos, Uso de anticonceptivos

- N NACIMIENTOS
- E EMBARAZOS
- A ABORTOS
- 0 NO USO DE MÉTODO
- 1 ESTERILIZACIÓN FEMENINA
- 2 ESTERILIZACIÓN MASCULINA
- 3 PÍLDORAS
- 4 MINIPÍLDORAS
- 5 DIU
- 6 INYECTABLES
- 7 CONDÓN
- 8 TABLETAS VAGINALES
- 9 LACTANCIA MATERNA
- 10 MOCO CERVICAL
- 11 RITMO
- 12 OTROS _____
(especificar)

COL 2. Discontinuidad de anticonceptivos

- 1 EMBARAZO DURANTE EL USO
- 2 QUERÍA QUEDAR EMBARAZADA
- 3 DESAPROBACIÓN DE LA PAREJA
- 4 EFECTOS SECUNDARIOS
- 5 MOTIVOS DE SALUD
- 6 DISPONIBILIDAD
- 7 ACCESO
- 8 QUERÍA UN MÉTODO MÁS EFECTIVO
- 9 SEXO INFRECUENTE / AUSENCIA DE PAREJA
- 10 INCONVENIENTES EN EL USO
- 11 DOCTOR DIO OTRO MÉTODO
- 12 COSTOS
- 13 FATALIDAD
- 14 DIFICULTAD PARA QUEDAR EMBARAZADA
- 15 DISOLUCIÓN MARITAL / SEPARACIÓN
- 16 OTROS _____
(especificar)
- 17 NO SABE

	1	2		3	4	5
			31	Agosto		
			15			
			31	Julio		
			15			
			30	Junio		
2			15			
0			31	Mayo		
0			15			
2			30	Abril		
			15			
			31	Marzo		
			15			
			28	Febrero		
			15			
			31	Enero		
			15			
			31	Diciem.		
			15			
			30	Noviembre		
			15			
			31	Octubre		
			15			
2			30	Setiembre		
0			15			
0			31	Agosto		
1			15			
			31	Julio		
			15			
			30	Junio		
			15			
			31	Mayo		
			15			

COL 4. Marital Status Exposición a relaciones sexuales

- 0 SIN PAREJA SEXUAL
- 1 CON PAREJA SEXUAL

COL 5

Tipos de localidad

- 1 CAPITAL DE DEPARTAMENTO
- 2 CIUDAD
- 3 PUEBLO

COL 3. Fuente de abastecimiento de método

- 0 NO HUBO ABASTECIMIENTO