STRONGMINDS MENTAL HEALTH AFRICA

End of Pilot Phase Impact Evaluation for the Treating Depression at Scale in Africa Program in Uganda

May 2014

This Impact Evaluation was completed by StrongMinds.

Dr. Katia Peterson, of In Situ Research LLC, executed the analysis and contributed to the report document.

Table of Contents

Acronyms

DSM-IV	Diagnostic and Statistical Manual of Mental
Disorders GEE	Generalized Estimating Equation
GIPT	Group Interpersonal Psychotherapy
IPT	Interpersonal Psychotherapy
MHF	Mental Health Facilitator
MOH	Ministry of Health
NGO	Non-Governmental Organization
PHQ-9	Patient Health Questionnaire-9
RCT	Randomized Controlled Trial
STATA	Statistical Software by StataCorp.

1. Executive Summary

To follow

2. Synopsis

An Impact Evaluation indicates that the StrongMinds' "Treating Depression at Scale in Africa" Phase II pilot exceeded expectations by successfully reducing the depressive symptoms in 94-97% of the patients treated using Group Interpersonal Psychotherapy (GIPT). This decrease in depression had an impact on the well-being of the participants. Following the 12-week intervention, patients were working and eating

more and required less medical care for chronic physical illness. More families fared better by sleeping under shelter and getting children to school. This suggests that treatment for depression could benefit work being done by other NGOs. Also notable, the patients' support networks improved which could mitigate future depressive episodes. The concept of these support networks is the basis for the StrongMinds Peer Support Group Model that will drive scalability.

Recommendations for successful scale-up include enhancing monitoring and evaluation efforts and employing technology to enable more timely assessment of treatment initiatives; developing and testing a Peer Support Group Model; and partnering with other organizations to improve overall patient well-being.

3. Introduction and Background

During the period of January 2014 to February 2015, StrongMinds implemented a pilot program using Group Interpersonal Psychotherapy (GIPT) to treat 514 depressed women in Kampala, Uganda, a post-conflict and highly impoverished country where 1 out of every 4 adults suffers from depression. Depression in Africa is a pervasive and debilitating mental illness; it is the number one cause of disability for over 60 million African women, over 90% of whom have no access to treatment.

GIPT is a proven model of treating depression. It focuses on the interpersonal relationships of depressed group members and is led by a facilitator who uses a structured model to help group members identify and understand the root causes and triggers of their depression, and then to formulate strategies to overcome those triggers. Since depression is episodic and will continue to recur throughout most people's lives, these newly acquired skills have both an immediate impact and a long-term preventive impact on the lives of those suffering from depression.

The pilot, "Treating Depression at Scale in Africa", tested, assessed and modified key program features over a 14-month period, and allowed StrongMinds to develop a refined model that is currently being implemented at a larger scale.

The pilot treated patients in two cohorts of roughly equal size, over two phases.

StrongMinds concluded Phase I of the pilot by treating 244 depressed women over a 16-week period via 26 Interpersonal Psychotherapy (IPT) groups on September 12, 2014. The 26 groups were led by four Mental Health Facilitators (MHFs) who are employed by StrongMinds. An Impact Evaluation completed in November 2014 indicated that the StrongMinds' Phase I pilot exceeded expectations by successfully reducing the depressive symptoms in 94-97% of the patients treated using Group Interpersonal Psychotherapy (GIPT). This decrease in depression had an impact on the well-being of the participants. During the 16-week intervention, self-employment increased by 22%, unemployment was reduced by 67%, the number of women who were able to save part of their income increased by 63%, and women eating three meals a day increased 245%. Recommendations for Phase II and scale up include changes in the type and quantity of data collected as well as the length of treatment and the severity of cases to include in treatment. Refer to the StrongMinds Impact Evaluation Report dated November 2014 for more details.

Phase II of the pilot treated 270 women with 26 groups and four MHFs. The treatment period was

shortened to 12 weeks based on the Phase I learning that the majority of patients received maximum impact by the end of week twelve. With the completion of Phase II in February 2015, StrongMinds reached the Pilot goal of treating 500 women in total.

4. Impact Evaluation Purpose

The Phase II for the StrongMinds' Treating Depression at Scale in Africa Project, took place in Uganda between November of 2014 and February of 2015. The purpose of this Impact Evaluation is to inform program activities for 2015 and beyond.

The evaluation focused on three major questions:

1. Was the use of Group Interpersonal Psychotherapy, implemented at a scaled approach, effective in treating depression in Uganda?
2. What, if any, were the secondary positive impacts of using GIPT on the depressed patients?
3. What actions are necessary for StrongMinds to improve its programmatic activities in light of the impact evaluation findings?

5. Evaluation Methods and Limitations

The Impact Evaluation primarily used quantitative techniques and was comprised of all the 270 depressed female participants in the treatment intervention group and all the 36 depressed female participants in the control arm. Participants were located in various villages within the Bulenga and Maganjo parishes in suburban Kampala. Basic demographic data was collected at pre-assessment. Map 1 provides a geographical representation; Table 1 provides a specific listing of the participant population groups per village location.

Map 1. Site Location Map (Kampala, Uganda environs)

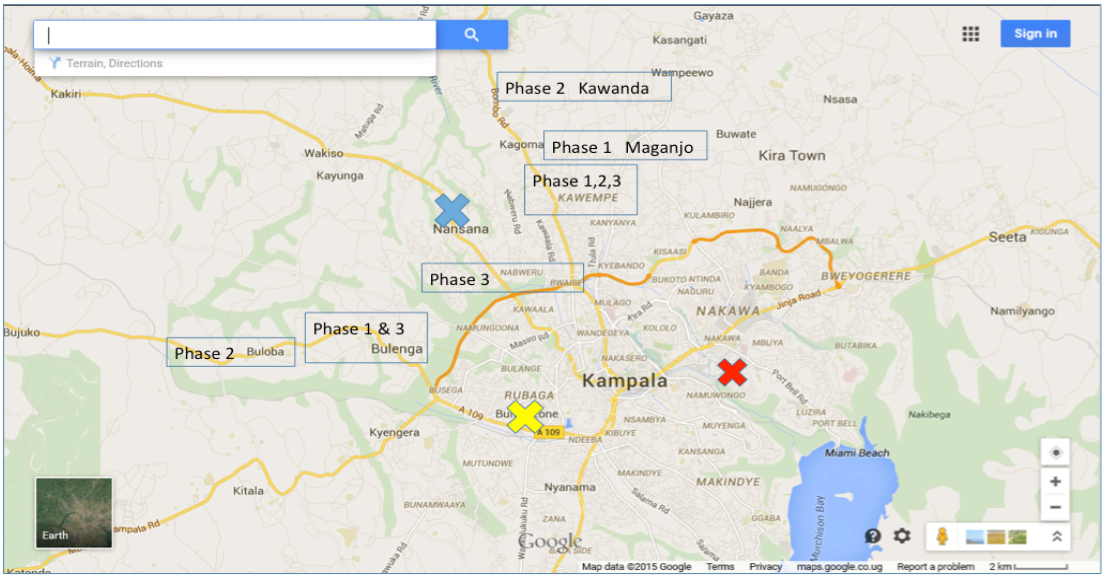


Table 1. Study Participant Location

Location	Intervention Size N (%)	Control Size N (%)
Bulenga Parish		
Nakabugo Rashida	15 (5.6%)	5 (13.9%)
Kisisira 1	14 (5.2%)	1 (2.8%)
Nakabugo Kiwanuka	14 (5.2%)	5 (13.9%)
Kalambi 2	13 (4.8%)	
Kalambi 1	12 (4.4%)	2 (5.6%)
Kirimamboga	12 (4.4%)	
Mayumba	10 (37%)	1 (2.8%)
Bbira	10 (3.7%)	1 (2.8%)
Nnsale'A	10 (3.7%)	
Kisisia 2	9 (3.3%)	
Ntongo	6 (2.2%)	3 (8.3%)
Maganjo Parish		
Wabitembe	27 (10.0%)	
Brac	16 (5.9%)	
Kamanya	15 (5.6%)	3 (8.3%)
Nakyesanja Katende	15 (5.6%)	3 (8.3%)
Nakyesanja	14 (5.2%)	3 (8.3%)
Jinja Karoli Kasatiro	13 (4.8%)	4 (11.1%)
Kubiyinja	10 (3.7%)	
Lukadde	8 (3.0%)	
Jinja Karoli	8 (3.0%)	4 (11.1%)
Nakyesanja Saaka	6 (2.2%)	
Kayii	6 (2.2%)	
Wabitembe 3	6 (2.2%)	
Nakyesanja Norah	1 (0.4%)	1 (2.8%)

Participants in the treatment intervention group and control group were screened using the Patient Health Questionnaire-9 (PHQ-9, a quantitatively based depression diagnostic tool) and diagnosed with depression in October and November of 2014. Participants in the treatment intervention group agreed to join IPT groups. Participants in the control arm were not able to join IPT groups operated by StrongMinds due to capacity constraints. Thus, the control arm participants did not receive any official treatment for depression during this 12-week intervention period, although they were offered the opportunity to join an IPT group with StrongMinds during 2015.

For the treatment intervention group, raw scores from the Patient Health Questionnaire-9 were recorded at pre-assessment ("baseline"), again at every other IPT group meeting from weeks 2-12, and once more as a post-assessment ("endline") in week 13. Data on patient functionality, for both the treatment intervention and control group was collected at baseline and endline only.

Limitations of the evaluation included the possibility of patient response bias; the subjectivity of self-reported data; some missing data, and logistical and time constraints to organize and analyze the

sizeable amount of data.
As background, the PHQ-9 is contained within the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV) published by the American Psychiatric Association and supported by the WHO for use in the developing world. (Please also see Appendix B: Depressed Patient Diagnosis Using the PHQ-9 and PHQ-9 Form.) The PHQ-9 is a series of 9 questions which score the severity of depressive symptoms for a patient, and provides an overall Raw Score for a patient's level of depression, as seen in Table 2. The Raw Score, in turn, equates to a level of depression, or Depression Severity Score.

Minimal Depression, which equates to a PHQ-9 Raw Score of between 1-4, can also be referred to as a normative level of depression. Minimal Depression, in lay terms, means that the individual is not able to be diagnosed with any significant level of depression, and is thus, "depression-free." The goal of the StrongMinds' pilot intervention was to reach the Minimal Depression state for its patients.

Table 2: Raw Score to Depression Score Conversion

PHQ-9 Raw Score	Depression Severity Score
1-4	0 = Minimal Depression
5-9	1 = Mild Depression
10-14	2 = Moderate Depression
15-19	3 = Moderately Severe Depression
20-27	4 = Severe Depression

Respondents with Minimal or Mild Depression (anyone with total raw scores between 1-9) at baseline in both the treatment intervention and control groups were dropped from the analysis. In typical practice around the world, individuals with Minimal/Mild Depression are not considered for inclusion in group therapy, because their depressive symptoms are relatively insignificant. While StrongMinds consciously included these Minimal/Mild cases in this Phase I of its pilot to gain experience with the patient population, their removal from the analysis serves to ensure that the Impact Evaluation is not artificially inflated, since reducing the depressive symptoms of Minimal/Mild Depressive cases is admittedly easier to do.

The result was longitudinal data collected from 270 patients in the treatment intervention group and 36 people in the control population. As previously indicated, all the participants in the study were female.

6. Findings

Demographic Descriptive Statistics (all participants)

Descriptive statistics were generated for both the treatment intervention and control groups. Demographic data consisted of age, location (as noted above), marital status, and parity. The treatment

intervention group consisted of 270 women with an average age of 36 years. The control group consisted of 36 women with an average age of 29 years. Approximately 65% of respondents in the treatment intervention group reported being married at the beginning of the study while nearly 91% reported having at least one child. Control group findings were similar. Tables 3a and 3b below illustrates demographic characteristics of the treatment intervention and control groups.

Table 3a. Baseline Demographic Characteristics

Characteristics	Intervention Size N (%)	Control Size N (%)
Age (years)	270	36
15-19	5 (1.9%)	5 (13.9%)
20-24	28 (10.4%)	8 (22.2%)
25-29	56 (20.7%)	7 (19.4%)
30-34	58 (21.5%)	9 (25.0%)
35-39	40 (14.8%)	3 (8.3%)
40-44	22 (8.1%)	2(5.6%)
45-49	21 (7.8%)	1 (2.8%)
50-54	15 (5.5%)	1 (2.8%)
55+	25 (9.3%)	

Table 3b. Baseline Demographic Characteristics

Characteristics	Intervention Size N (%)	Control Size N (%)
Marital Status	270	36
Married	175 (64.8%)	23 (63.9%)
Divorced	51 (18.9%)	5 (13.9%)
Widowed	31 (11.5%)	7 (19.4%)
Single	13 (4.8%)	1 (2.8%)
Have Children		
Yes	245 (90.7%)	32 (88.9%)
No	25 (9.3%)	4 (11.1%)

Treatment Effects

Tables 4 outlines descriptive statistical data on depression scores and prevalence for both the treatment intervention and control groups at pre and post-assessment. Figures 1 and 2 represent those scores graphically.

Table 4. Treatment Intervention Pre and Post-Assessment Depression Scores

Characteristics	Intervention N (%)		Control N (%)	
	Pre	Post	Pre	Post
Median PHQ-9 Raw Score	12	0	11	10
Median Depression Score	2	0	2	2
Prevalence of Depression by Type				
(0) Minimal	2 (0.7%)	267 (94-97%)	2 (5.6%)	4 (11.1%)
(1) Mild	44 (16.3%)	3 (1.1%)	8 (22.2%)	11 (30.6%)
(2) Moderate	152 (56.3%)	0 (0%)	19 (52.8%)	17 (47.2%)
(3) Moderately Severe	54 (20.0%)	0 (0%)	4 (11.1%)	4 (11.1%)
(4) Severe	18 (6.7%)	0 (0%)	3 (8.3%)	0 (0%)

Figure 1. Treatment Intervention Population Pre and Post-Assessment Depression Scores[4]

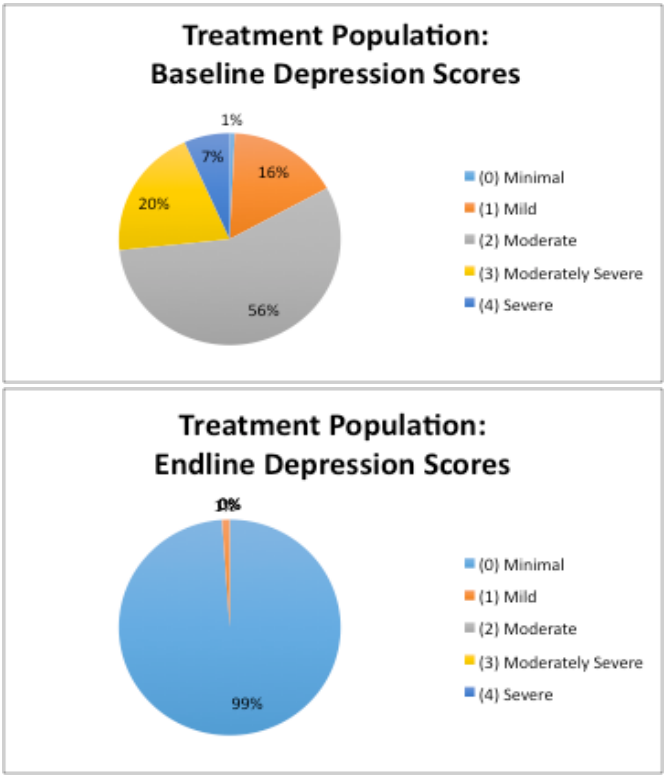
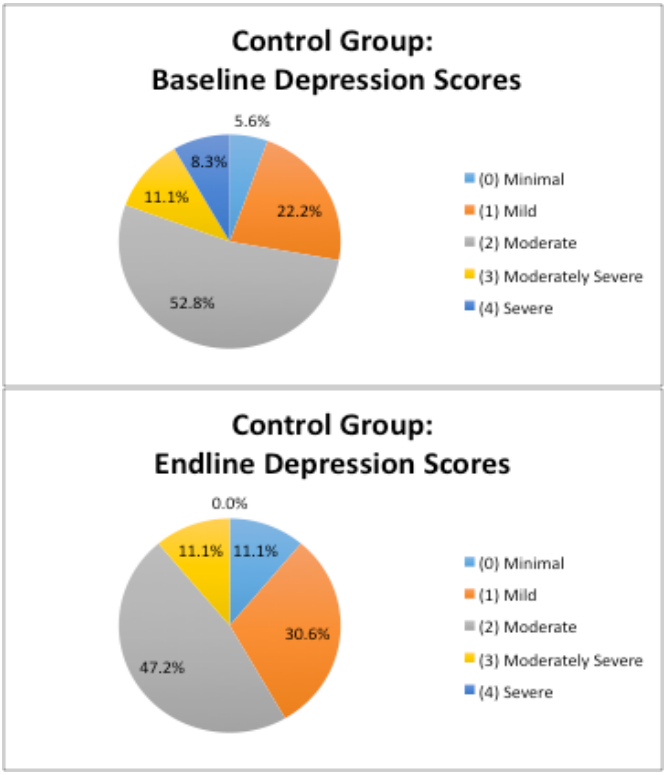


Figure 2. Control Group Population Pre and Post-Assessment Depression Scores[4]



The number of patients in the treatment intervention group responding to Minimal Depression (or "depression-free") at post-assessment (Week 13) ranges from 94% to 97%. As previously explained, all of the respondents who reported Minimal or Mild depression at the pre-assessment (or baseline) were excluded from the analysis (45 patients in total). As such we removed these people not only from the numerator, but also from the denominator for post-assessment calculations and attained the 97% figure. This could result in an overestimation for treatment effect in the post-assessment. If these individuals are removed from the numerator but included in the denominator, the treatment drops to 94% reporting Minimal Depression.

GEE Analysis

A generalized estimating equation (GEE) model was utilized by In Situ Research LLC to analyze correlated data and intra-subject changes in raw PHQ-9 score changes over the 12-week study period and at post-assessment (week 13) among both treatment intervention and control groups. The GEE model was selected in order to determine the overall impact of the intervention on the average scores of individuals in the treatment intervention group relative to those in the control group, and to determine the effect size in score changes. After performing exploratory data analysis with a continuous response variable, it was hypothesized that it may be reasonable that the correlation structure would be autoregressive and thus fit an AR(1) correlation structure to our GEE model.

Data was analyzed using STATA /SE version 12.1 using the xtgee commands. Missing data was not imputed as we found them to be missing at random (and that the probability of drop out may be

related to covariates and pre-drop out responses). Further, GEE models use the "all available pairs" method, in which all non-missing pairs of data are used in the estimating the working correlation parameters. In our case the GEE model only lost the observations that the subject is missing, not all measurements. Some groups had fewer than two observations such that it was not possible to estimate correlations for those groups. Fifty-six groups were omitted from estimation.

KATIA: "The GEE analyses (excluding those with <9 at baseline). I ended up re-running parts of the analysis for various reasons and came up with some different results than what I had told you. First off, the model is unstable because so many people are excluded because they have baseline scores <9. Forty-five people were removed which is about 17% of the data. That is a big chunk. I was trying to make the GEE model more stable but just couldn't. In the future, we may have to do something a bit more different than a GEE and it will be a bit more complicated and require more time. "

As evidenced in Table 5, the intervention does indeed have an impact on depression score, when comparing control vs. intervention groups. Those in the intervention group, on average, had a 4.52 point reduction in their total depression score as compared to control populations.

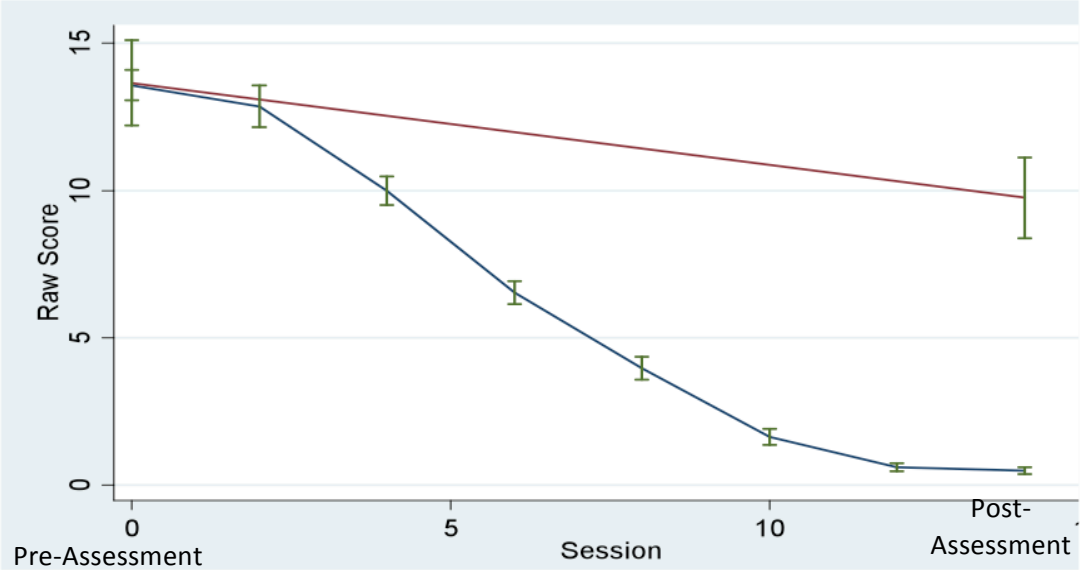
The rates for depression scores of 4 or less at week 12 are 252 out of 253 patients or 99.6%. For the post assessment, 98.8% or 267 out of 270 patients reported scores of 4 or less.

Table 5. Results of the GEE Analysis

Treatment Session	Coefficient	Standard Error	Z	P> |z|	95% Confidence Interval	
group	4.526957	0.4828243	9.38	0.000	3.580639	5.473276
session	-.8634286	0.0209841	-41.15	0.000	-0.9045566	-.8223005
_cons	8.586239	0 .572214	15.01	0.000	7.46472	9.707758

The results from the GEE model demonstrate that the intervention did have a decreasing effect on the final PHQ-9 Raw depression scores. Although both the control and treatment intervention group did have an overall decrease in depression scores, the treatment intervention group had a greater drop in PHQ-9 Raw depression scores and by the 12th session 99% of patients had PHQ-9 Raw Scores of 1-4 in which they would be classified as Minimally Depressed (or "depression-free"). Figure 3 displays the average PHQ-9 scores for both the treatment intervention and control groups throughout the intervention period.

Figure 3: Line Graph Displaying Bi-weekly Average PHQ-9 Raw Scores Treatment Intervention vs. Control Group



Functionality

Functionality was measured on 46 factors and collected from 268 treatment intervention patients at pre- and post-assessment. Due to missing values and the questionability of some indicators as appropriate measures of functionality, 11 of the 46 indicators were analyzed for results. They cover employment, physical health, family and support network status.
The sample size was too small to conduct similar functionality analysis of the control group.

McNemar's test was used to analyze dependent categorical variables (i.e. yes/no answers) from paired matches. The purpose of McNemar's test is to determine whether row and column frequencies are different. The null hypothesis (H_0) was established as: the probability of the event at baseline is equal to probability of the event at end line. The alternative hypothesis (H_1): the baseline probability does not equal the end line probability.

$$H_0: p_{bl} = p_{el}$$
$$H_1: p_{bl} \neq p_{el}$$

McNemar's test statistic χ^2 was calculated[1] (and in some instances the exact binomial distribution was calculated when $b+c$ <25). The p-value was set at 0.05 (two-tailed test).

The Wilcoxon Signed Rank Test was used to analyze dependent ordinal variables (i.e. pain scale). The null and alterative hypothesis were:

H_0: $p_{bl} = p_{el}$ (median difference between the pairs is zero)
H_1: $p_{bl} \neq p_{el}$ (median difference is not zero)

Since many of the cell frequencies exceeded '10' (n>10) the z-test was used instead of the Wilcoxon test statistic. The critical value for the two-tailed test ($p<0.05$) was assigned at -1.96 and +1.96.

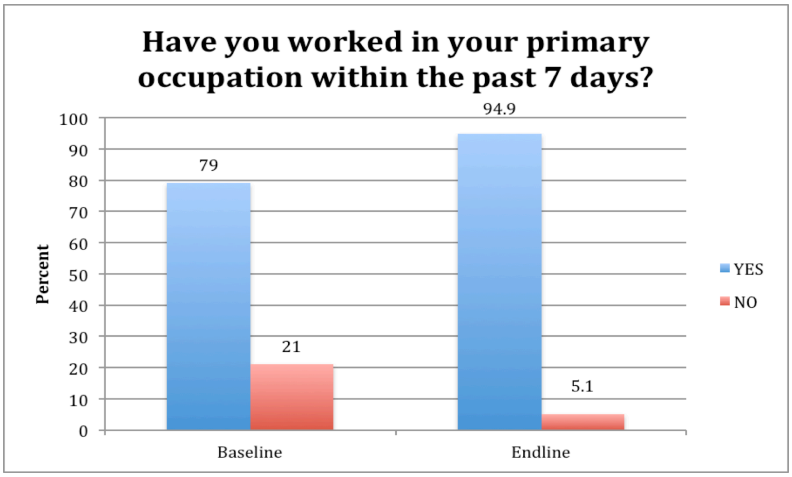
Given the nature of the analysis—before vs. after comparison of the same individual—only complete pairs of data were included. A complete pair has no missing values or discordant values between pairs. Discordance refers to the situation when there is one response and one missing or N/A value[2]. Paired responses coded as 'N/A' were also removed from the analysis.

Pre and post functionality data was descriptively analyzed and specific indicators of functionality were compared for the treatment effect using a z-test.[3]

Employment Status

Respondents were queried about their employment status at pre and post-assessment. In Phase I, there were notable improvements, specifically, there were more women who identified as 'self employed' and fewer people who identified as 'unemployed' post intervention. In Phase II, there was a significant increase from a baseline of 79.0% to end line of 94.9% in those who reported working in their primary occupation within the most recent seven days.

Figure 4. Percentage of Respondents Who Have Worked In Their Primary Occupation Within The Last Seven Days



There was a significant increase in job satisfaction scores over the previous 30 days. The median score increased from 3 ('neither satisfied or dissatisfied') to 2 ('satisfied'). A Z-score of -8.345 and p-value < 0.05 were evident.

Table 6: Job Satisfaction Over The Last 30 Days

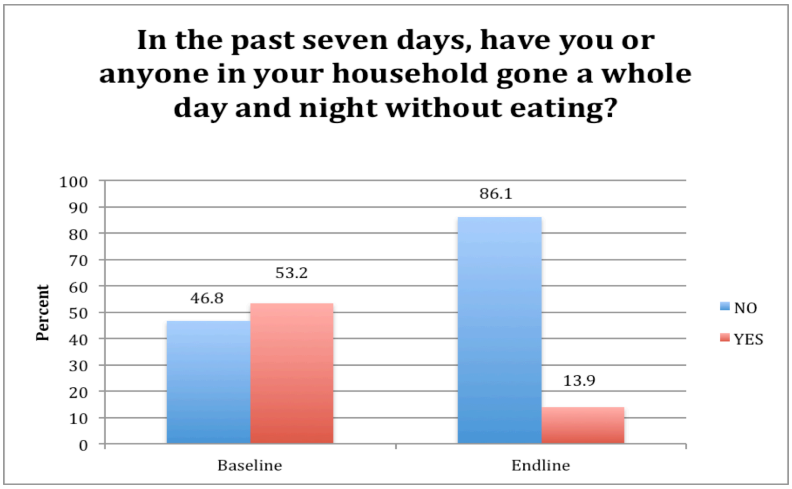
	N	Mean	Std. Deviation	Minimum	Maximum	Percentiles		
						25th	50th (Median)	75th
ES12BL	154	3.39	.959	2	5	3.00	3.00	4.00
ES12EL	154	2.42	.711	1	5	2.00	2.00	3.00

Physical Health Status

The nutritional status of participants was captured in Phase I using the number of meals eaten within the previous 24 hours. The difference between pre and post assessment for three meals a day posted the largest change in the number of meals eaten between the two assessments.

In Phase II, we asked how many times in the past seven days has anyone in your household gone a whole day and night without eating. There was a significant decrease in the number of people who reported going a whole day and night, or 24 hours, without a meal: 53.2% at baseline and 13.9% and end line.

Figure 5. Percentage of Respondents That Have Gone With Out Eating for 24 hours Within The Past Seven Days

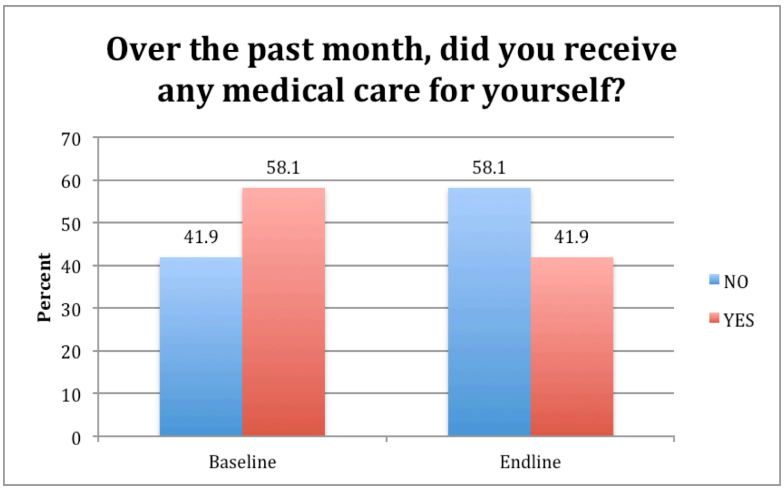


Further, there was a significant increase in the number of meals eaten within a 24 hour period from baseline to end line. The score increased from 3 (two meals per day) to a score of 4 (three meals per day). Here there was a Z-score of -11.282 and p-value < 0.05.

	N	Mean	Std. Deviation	Minimum	Maximum	Percentiles		
						25th	50th (Median)	75th
NS1BL	269	2.83	.898	1	6	2.00	3.00	3.00
NS1EL	269	3.74	.833	1	6	3.00	4.00	4.00

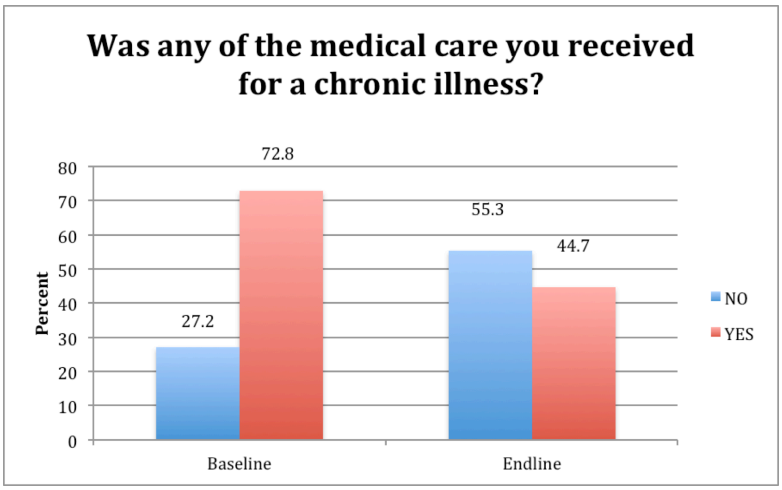
There was a significant decrease in the number of people who sought medical care at baseline (58.1%) compared with end line (41.9%).

Figure 6. Percentage of Respondents Who Received Medical Care Over The Past Month



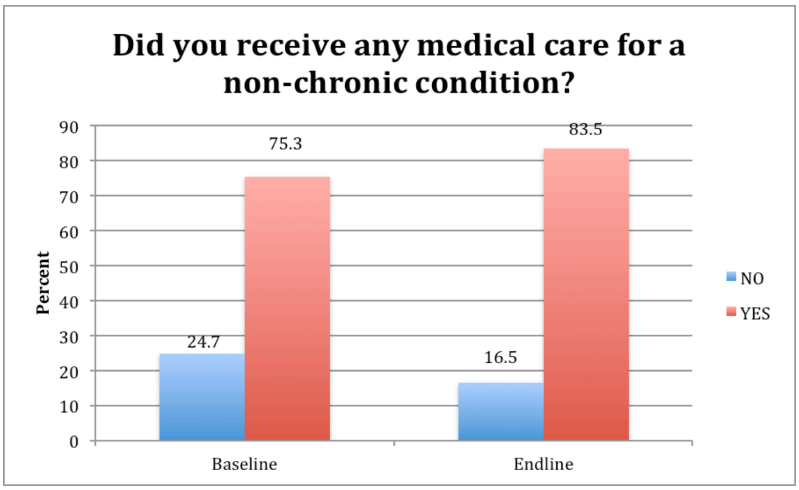
There was also a significant decrease in the number of people who sought medical care for a chronic condition from baseline (72.8%) to end line (44.7%).

Figure 7. Percentage of Respondents Who Received Medical Care for a Chronic Illness



There was no significant difference in the number of people who sought medical care at baseline and end line for a non-chronic condition.

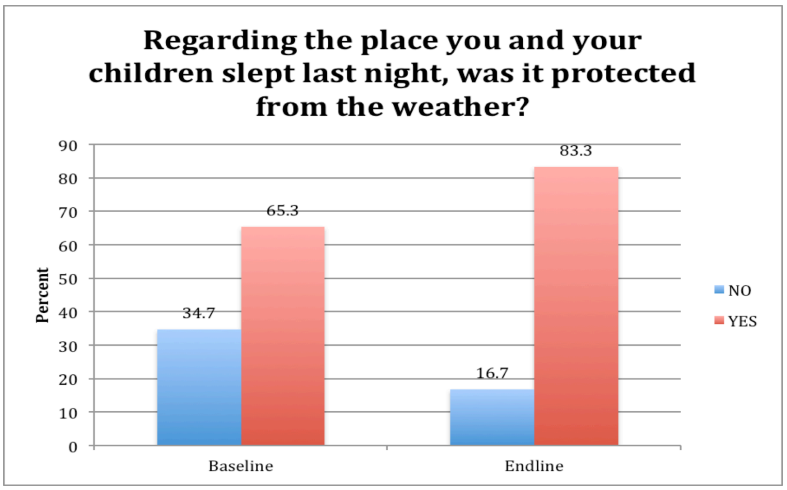
Figure 8. Percentage of Respondents Who Received Medical Care For A Non-Chronic Condition



Family Status

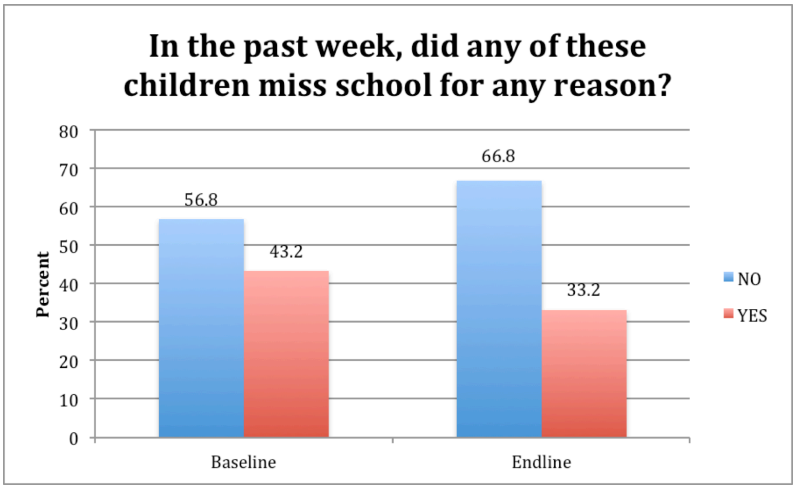
There was a significant increase in people reporting that the place where they and their children slept the previous night was protected from the weather (65.3% to 83.3%).

Figure 9. Percentage of Respondents Who Were Able To Sleep With Their Children In A Protect Shelter The Preceding Night



There was a significant decrease from baseline to end line (43.2 to 33.2%) in those who reported a child absent from school (for any reason excluding holidays) within the most recent 7 days.

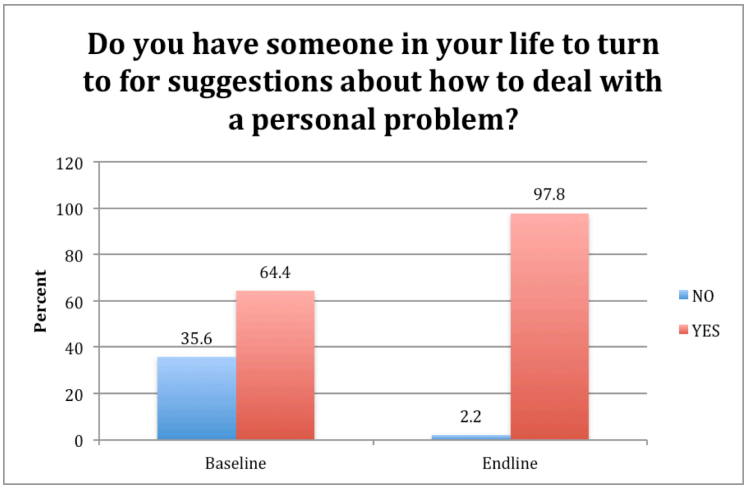
Figure 10. Percentage of Respondents Who Reported That Their Children Missed School In The Preceding Week



Support Network Status

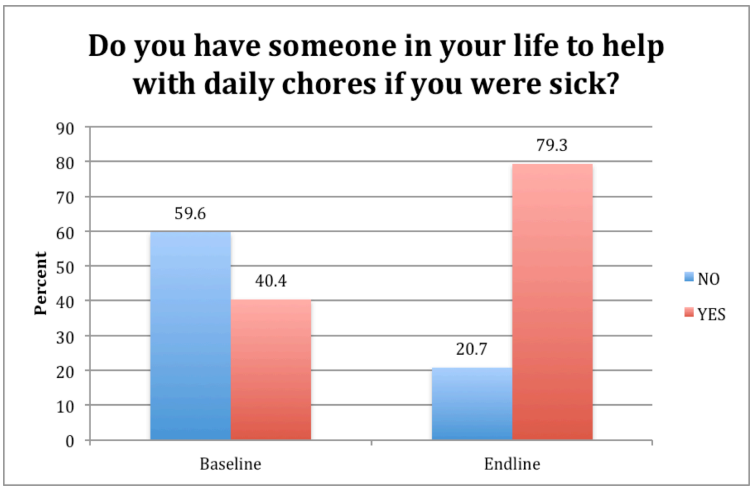
There is a significant increase in those who report having someone to turn to for help dealing with a personal problem (64.4% to 97.8%).

Figure 11. Percentage of Respondents Who Have Someone In Their Lives To Help With Problems



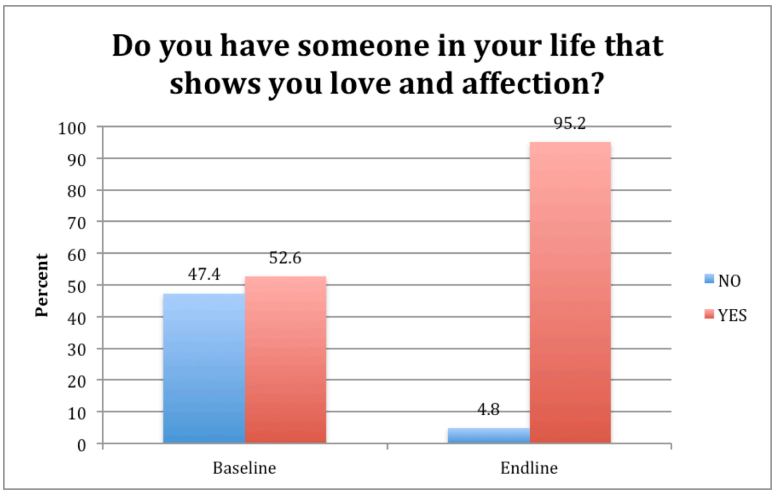
There is a significant increase in those who report having someone to help with chores if they were sick (40.4% to 79.3%).

Figure 12. Percentage of Respondents Who Have Someone Who Could Help At Home Should They Fall Ill



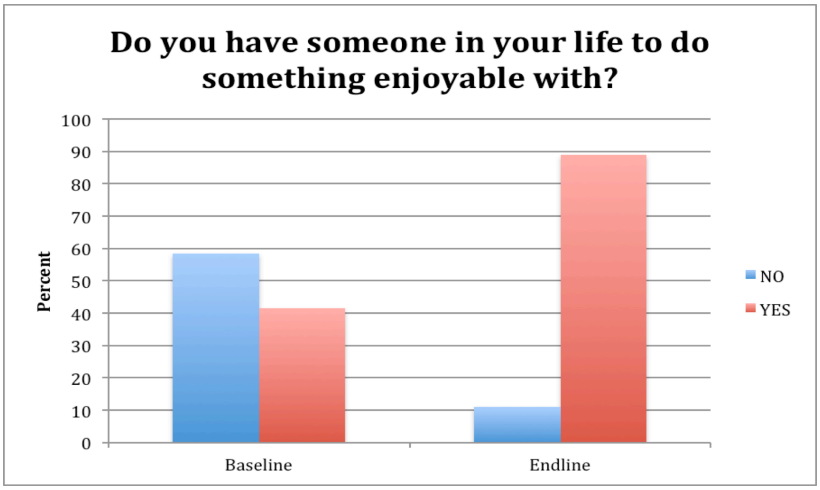
There is an increase in the number of people who report they have someone in their life who shows them love and affection (52.6% to 95.2%).

Figure 13. Percentage of Respondents Who Have Someone Who Shows Them Love And Affection



There was an increase in the number of people reporting they had someone in their life to do something enjoyable with (41.5% to 88.9%).

Figure 14. Percentage of Respondents Who Have Someone With Whom They Can Do Something Enjoyable



7. Discussion

The first question posted in this Impact Evaluation was:

1. Was the use of Group Interpersonal Psychotherapy (GIPT), implemented at a scaled approach, effective in treating depression in Uganda?

The GIPT intervention results for Phase II were consistent with the findings in Phase I. **The Evaluation finds a range of 94-99% of the patients treated by StrongMinds were depression-free after the 12-week GIPT intervention.** As in the Phase I, these results were in the range of the 92% success rate achieved by a Randomized Controlled Trial for GIPT conducted in Uganda in 2002.[1] The 2002 RCT used 9 MHFs to treat 224 individuals whereas StrongMinds, using a scaled approach, and in Phase II used 4 MHFs to treat 270 people. The RCT staff MHFs were lay individuals with high school training only; the StrongMinds MHFs consisted of two nurses and two women with degrees in community psychology. In this pilot, StrongMinds set a goal of reaching depression-free status for 75% of its patients.

Furthermore, the analysis determined that **depressed female patients who completed the GIPT intervention, on average, experienced a 4.5 point reduction** in their total PHQ-9 Raw Score[2][3] over the entire 12-week intervention period, compared to the control group. Additionally, for PHQ-9 taken (every other week), **these women experienced an average 1.7 reduction in their PHQ-9 Raw Score for depression.** These findings were both statistically significant, and indicate the magnitudes by which their depressive symptoms were reduced.

The above findings were verified in several manners. First, after the conclusion of the 12-week intervention, a post-assessment evaluation was conducted within the following week, in which group

members were re-evaluated by a different MHF, using the PHQ-9 diagnosis tool. The reason for this visit was to correct for any patient bias or allegiance to the MHF, and also to correct for any MHF quality issues in administering the diagnosis tool. In addition, outside mental health experts have reviewed the rates of improvements of GIPT group members, and found the information to be reasonable.

For Phase I, The StrongMinds team of four Ugandan female MHFs who led the groups were all first time GIPT implementers. They were supervised by two experienced Ugandan GIPT experts who were trained by US personnel in 2006 and have implemented many IPT groups. There was concern following Phase I that the MHFs' inexperience may have led to a partially inflated success rate by incorrectly using the depressive diagnostic tools, for example. This appears to be less so the case given the results for Phase II. However, the StrongMinds team continues to test and probe into possible biases, such as social desirability bias.

In Phase I, the weekly administration of the PHQ-9 created a heightened awareness of the patients' depression that resulted in an awareness bias and therapeutic effect. Weekly collection also posed a data collection burden for the MHFs and the patients. For Phase II, PHQ-9s were completed bi-weekly to minimize those effects.

The Impact Evaluation analysis removed all respondents who reported at baseline either Minimal or Mild Depression and is only based upon those women who were either moderately or severely depressed at baseline in order to avoid inflating the success of the intervention. A possibility exists that the inclusion of these Minimal and Mild Depressive patients in the group settings may have had some influence on the progress of the other more severely depressed patients in the same groups. It is not known if this influence was positive or negative.

The control group in both Phase I and II of the pilot experienced a reduction of depressive symptoms. In Phase I, 33% of the control group were depression-free after the intervention period while 47.2% scored as depression free in Phase II. Following Phase I, based on mental health experts and program staff assessment, measures were taken to mitigate programmatic influence on the control group.

- PHQ-9s were administered only at Phase II baseline and endline to the control group compared with every week during Phase I.
- Control group members were more geographically isolated from intervention group members in Phase II to decrease the likelihood of intervention group members influencing the control group members' behavior.

Despite these measures, the Phase II control group had greater improvement than the Phase I control group. The StrongMinds team in Uganda is exploring possible explanations as to why the control group did improve, despite these members not receiving any formal treatment by conducting random interviews? However, given that this Phase II was the final phase of the pilot, control group formation will not be part of the program structure in 2015.

The second question posed in this Impact Evaluation was:

2. **What, if any, were the secondary positive impacts of using GIPT on the depressed patients?**

The GIPT intervention appears to have had an impact in addition to significantly reducing the depressive symptoms of these women patients. The following is a summary of the statistically significant findings with respect to functionality for women who completed the GIPT intervention, comparing baseline to end line data:

- Patients who worked in their primary occupation increased by 15.9%.
- The number of women who reported not going a day without eating improved by 32.9%.
- Respondents who received medical attention for a chronic illness decreased 28.1%, perhaps suggesting that they were in better overall health and better able to manage their chronic illness.
- Women and their children who were able to sleep in a shelter protected from the weather increased by 18.0%.
- The percentage of children who missed school for any reason declined by 10.0%.

The increases in various well-being indicators collectively demonstrate that GIPT appears to have had significant secondary positive impacts on the well-being of these women. Since depression is the number one cause of disability for women in Africa, it is logical that by reducing these women's level of depression, their level of disability is reduced and there are corresponding increases in work and management of physical health which likely drives the meals consumed and shelter acquired.

There were also improvements in patients' support networks:

- At endline, 97.8% of patients reported having someone in their life to turn to for help with personal problems compared to 64.4% at baseline, a 33.4% increase.
- When it comes to having someone to help with daily chores if they were to fall ill, 79.3% reported having that support after the intervention versus 40.4% before they began, a 38.9% increase.
- Over 95.2% of the patients indicated that they have someone in their life that shows them love and affection at endline, compared with 52.6% at baseline, a 42.6% improvement.
- Following treatment, 88.9% of respondents said that they have someone in their life to do something enjoyable with compared with 41.5% prior to treatment, a 47.4% increase.

The support network findings are indicative of the number of groups that continue to meet following intervention and the number of women who have completed the intervention and want to become peer support group facilitators.

There is a possibility that some or all of these positive impacts resulted from some other outside factor(s) non-attributed to this StrongMinds' intervention, and that these positive impacts, in turn, drove the positive improvements in the depressive states of the women patients. Future efforts by StrongMinds will need to address this issue.

The third question posed in the Impact Evaluation was:

1. What actions are necessary for StrongMinds to improve its programmatic activities in light of the Impact Evaluation findings?

The table below lists the recommendations that came out of the Phase I findings and indicates whether those recommendations were implemented in Phase II.

Recommended Actions from Phase I Findings	**Implementation in Phase II**
Revise functionality data collection tools to reflect more appropriate indicators of patient and their family well-being, to include income, health, educational achievement, etc. and link these tools to local Uganda District Health Survey tools to allow for comparisons.	√
Consider the addition of qualitative methods such as in-depth interviews of random participants in order to capture context around quantitative results to better understand cause/effect.	X
Ensure functionality data collection is completed for both patient intervention group and control group.	√
Consider reducing data collection frequency to limit the burden and also the number of missing data values.	√
Consider automation of data collection efforts, using laptops/tablets/etc.	X
Develop new programmatic strategies for addressing depression with males.	X
Consider utilizing independent/external mental health experts to implement PHQ-9 diagnoses to 10% of participants from each intervention group in order to capture/correct any response and MHF bias, and to serve as a quality control check.	X
Consider excluding Minimal or Mild cases of depression in future groups unless there are extenuating circumstances (for example, suicidality). Determine if there are other, non-GIPT methods by which StrongMinds can assist these case types.	√
Consider reducing the length of GIPT interventions to less than 16 weeks, in light of the high degree of success by week 12.	√
Ensure future control groups are not at risk of contamination; limit their contact with StrongMinds' MHFs and patients.	√

Three of the recommendations that were not implemented in Phase II have been or will be addressed in

2015.

- For example, StrongMinds' Country Director in Uganda has been conducting random in-depth interviews with patients and has found his findings helpful in developing MHF training and supervision as well as programs with partners.
- StrongMinds held its first all male group in 2015.
- StrongMinds is in the process of bringing in external evaluators in addition to hiring a monitoring and evaluation professional with the expectation that they will be able to begin work in July.

The fourth recommendation that was not implemented regarding the automation of data collection has been discussed with a provider in East Africa and a proposal from them is forthcoming. As of the writing of this evaluation, it is not clear whether implementation will be feasible in 2015.