

B. Collection of Information Employing Statistical Methods

B.1. Respondent Universe

Statistics vary on percent of the population that would identify as lesbian and bisexual female. The National Gay and Lesbian Task Force¹ estimated that 5-8% of the population is lesbian (with 1-3% of the population as bisexual) and the Family Research Report² estimated that 2% of women are homosexual or bisexual. The population in St. Louis and Columbia is 2,975,831. Of that population, 49% are adults, 50% are female, 50% are 40 years old or older, about 5% are lesbian/bisexual and 50-75% are overweight which provides us with a potential recruitment pool of over 13,500. Of those overweight, in a Missouri survey approximately 25% of the lesbian/bisexual population expressed an interest in losing weight. Therefore for the study region, over 3000 individuals might consider enrolling in the project.

For this study, approximately 160 lesbians and bisexual women will be enrolled in the study for a target number of 120 individuals completing all phases of testing: 40 lesbian and bisexual enrolled in the Full Gym intervention group across the two project sites, 40 lesbian and bisexual women in the Smart Pedometer group, and another 40 in an "attention-control" group. We estimate approximate 17% attrition rate. The Center project, one of the project sites, will post announcements at the Center and other LGBT friendly venues and events to recruit potential participants for the intervention groups. They will also advertise the project via the Out, Proud, and Healthy Facebook page, and community listserves. With a larger population, SAGE, the second project site, will utilize their listserve, Vital Voice advertisement (St. Louis' lesbian, gay, bisexual and transgender (LGBT) magazine), Facebook, and other targeted announcements to advertise the project.

B.2. Procedures for the Collection of Information/Limitations of the Study

B.2.1. Statistical Methodology for Stratification and Sample Selection

We have selected two locations in Missouri for our data collection. Columbia, Missouri is a typical mid-western college town while St. Louis, Missouri is a typical large city that reflects an urban lifestyle. Participants will be recruited from these two locations using the following eligibility: 1) 40 years or older; 2) self-identify as lesbian or bisexual female; 3) at risk (or with) heart disease, cancer, diabetes, musculoskeletal problems, asthma, and other morbidities related to overweight and obesity; 4) interested in getting healthy, as self-determined since the promotion will advertise the project's focus on "getting healthy;" 5) medical clearance to participate in an exercise program; and 6) ability to commit to participating in the study and follow-up assessments. In the situation in which eligible partners both express interest in enrolling in the program, one will be randomly selected. The other partner will be allowed to attend the educational class, but will be responsible for the costs of further participation.

Approximately 160 lesbians and bisexual women will be enrolled in the study for a target number of 120 individuals completing all phases of testing (40 lesbians and bisexual women will be enrolled in the Full Gym intervention group across sites, 40 lesbian and bisexual women in the Smart Pedometer group, and another 40 in a “control” group). We estimate an attrition rate of approximately 17%.

B.2.2. Estimation Procedure

For this study, we have decided to conduct a Randomized Control Trial as the method of evaluation. Data will be collected at Baseline, Month 4 and Month 12. We will collect data through direct assessments and surveys of the participants. Every effort will be made to minimize the non-sampling errors in the estimates by maximizing the response rates and taking steps to reduce response errors.

Eligible participants will be randomly assigned across the two intervention groups or the control group so that there are approximately 40 participants in all three groups. Random assignment for these interventions will take place through ongoing rolling enrollment. Rolling enrollment will necessitate randomly assigning women on a regular basis as they are successfully screened for eligibility.

Because this study relies on randomization of study participants, it will produce intervention and control groups that are comparable in every way except for the differences in participation in the fitness program; therefore, we will be able to compare the outcomes for the treatment group against those of the control group, controlling for covariates, and interpret any statistically significant differences in outcomes as impacts attributable to participation in the program. We will develop a basic multivariate model in order to understand the relationship between program characteristics, participant characteristics, environmental characteristics, and other covariates and participant outcomes. Developing a well-specified, parsimonious base model that provides continuity as we examine different characteristics of the participants and the programs is essential. Standard statistical tests such as two-group t-tests (for continuous variables), two-group z-tests (for proportions), and chi-square tests (for categorical measures and distributions) will be used to determine whether estimated effects are statistically significant.

Since our sample sizes are small, it will probably not be possible to examine subgroups within the three treatment and control groups.

B.2.3. Degree of Accuracy Needed for the Purpose Described in the Justification

Health Measurements

We are interested in learning whether the interventions of a gym membership with additional nutrition information or use of a smart pedometer result in a significantly reduced body mass index, waist-to-height ratio, or significantly increased maximal aerobic capacity. We compare the before-and-after measurements for each separate group. We can also compare the before-

and-after differences in body mass index, waist-to-height ratio, and maximal aerobic capacity between the control group and the two intervention groups. Based on Yancey et al (2006), Zhu et al (2002), and Lynch et al (2002), we assume that the baseline mean body mass index is 30.0 (with a standard deviation of 6.5), the baseline mean waist-to-height ratio is 0.51 (with a standard deviation of 0.05), and the baseline maximal aerobic capacity is 2.00 L/min (with a standard deviation of 0.30) for all three groups and that the correlation between the before and after measurements are 0.79 for all three measurements (based upon base year and one year later body mass index measurements from the National Longitudinal Study of Youth, 1997 cohort).³⁻⁵ As described above, our proposed sample sizes are 40 for each of the two intervention groups and 40 for the control group.

Since all three groups have an equal sample size of 40, the before and after comparison within each group has the same power. A sample size of 40 provides 80 percent power to detect a mean body mass index drop of 6 percent (a drop from 30.0 to 28.2). For waist-to-height ratio, which is less variable, a sample size of 40 provides 80 percent power to detect a mean waist-to-height ratio drop of 4 percent (a drop from 0.51 to 0.49). For maximal aerobic capacity, a sample size of 40 also provides 80 percent power to detect a mean aerobic capacity drop of 4 percent (a drop from 2.00 to 1.92). These are all one-sided paired t-tests with $\alpha=0.05$ (type I error rate).

Now, to compare the drops between either of the intervention groups and the control group, we need two-sided two-sample t-tests (we keep $\alpha=0.05$). For randomized control samples, the correct analysis is to calculate the differences (subtract the baseline measurement from the post-treatment measurement) for each for each participant and compare these differences (with the appropriately smaller standard errors due to the high correlation between the before and after measurements). For body mass index, sample sizes of 40 and 40 provide 80 percent power to detect a mean body mass index difference of 1.5 (5 percent). For example, we have sufficient power to detect post-treatment mean body mass indexes between 30.0 in the control group (no change) and 28.5 in the treatment group (a drop of 1.5). For waist-to-height ratio, sample sizes of 40 and 40 provide 80 percent power to detect a waist-to-height ratio difference of 0.015 (3 percent). For example, we have sufficient power to detect post-treatment waist-to-height ratios between 0.510 in the control group (no change) and 0.495 in the treatment group (a drop of 0.015). For maximal aerobic capacity, sample sizes of 40 and 40 provide 80 percent power to detect a maximal aerobic capacity difference of 0.07 (3.5 percent). For example, we have sufficient power to detect post-treatment maximal aerobic capacities between 2.00 in the control group (no change) and 1.93 in the treatment group (a drop of 0.07).

Food and Drink Consumption

We are interested in learning whether the interventions of a gym membership with additional nutrition information or use of a smart pedometer result in significantly reduced sugar-sweetened beverage consumption, alcohol consumption, or significantly increased fruit and vegetable consumption. We compare the before-and-after measurements for each separate group. We can also compare the before-and-after differences in sugar-sweetened beverage consumption, alcohol

consumption, and fruit and vegetable consumption between the control group and the two intervention groups. Based on Chen et al (2010), Bloomfield (1993), and Boehmer and Bowen (2009), we assume that the baseline mean sugar-sweetened beverage consumption is 10.5 fluid ounces per day (with a standard deviation of 11.9), the baseline mean alcohol consumption is 28.7 drinks per month (with a standard deviation of 53.2), and the baseline fruit and vegetable consumption is 3.20 servings per day (with a standard deviation of 2.35) for all three groups and that the correlation between the before and after measurements are again 0.79 for all three measurements.⁶⁻⁸ As described above, our proposed sample sizes are 40 for each of the two intervention groups and 40 for the control group.

Since all three groups have an equal sample size of 40, the before and after comparison within each group has the same power. A sample size of 40 provides 80 percent power to detect a mean sugar-sweetened beverage consumption drop of 28 percent (a drop from 10.5 to 7.5). For alcohol consumption, which is less variable, a sample size of 40 provides 80 percent power to detect a mean alcohol consumption drop of 50 percent (a drop from 28.7 to 14.4). For fruit and vegetable consumption, a sample size of 40 also provides 80 percent power to detect a mean aerobic capacity drop of 18 percent (a drop from 3.2 to 2.6). These are all one-sided paired t-tests with $\alpha=0.05$ (type I error rate).

Now, to compare the drops between either of the intervention groups and the control group, we need two-sided two-sample t-tests (we keep $\alpha=0.05$). The correct analysis is again to calculate the differences (subtract the baseline measurement from the post-treatment measurement) for each for each participant and compare these differences (with the appropriately smaller standard errors due to the high correlation between the before and after measurements). For sugar-sweetened beverage consumption, sample sizes of 40 and 40 provide 80 percent power to detect a mean sugar-sweetened beverage consumption difference of 2.4 fluid ounces per day (23 percent). For example, we have sufficient power to detect post-treatment mean sugar-sweetened beverage consumption between 10.5 in the control group (no change) and 8.1 in the treatment group (a drop of 2.4). For alcohol consumption, sample sizes of 40 and 40 provide 80 percent power to detect an alcohol consumption difference of 10.3 drinks per month (36 percent). For example, we have sufficient power to detect post-treatment alcohol consumptions between 28.7 in the control group (no change) and 18.4 in the treatment group (a drop of 10.3). For fruit and vegetable consumption, sample sizes of 40 and 40 provide 80 percent power to detect a fruit and vegetable consumption difference of 0.5 servings per day (16 percent). For example, we have sufficient power to detect post-treatment maximal aerobic capacities between 3.2 in the control group (no change) and 2.7 in the treatment group (a drop of 0.5).

B.2.4. Unusual Problems Requiring Specialized Sampling Procedures

There are no unusual problems requiring specialized sampling procedures.

B.2.5. Use of Periodic (Less Frequent Than Annual) Data Collection Cycles

Our study will collect data four times within a year. Data will be collected at Enrollment, Baseline, Month 4 and Month 12.

B.3. Methods to Maximize Response Rates and Deal with Issues of Non-response

Yancey et al provided a review of effective retention strategies for ethnic-racial minority populations, and many of these same successful strategies will be used to enhance retention in our project.³ One strategy identified as successful was using culturally appropriate messaging.^{9,10} We plan to provide messages in poster format with culturally appropriate images at the gym, such as older women working out at the gym, as well as providing all study documents using culturally appropriate language and images. Community involvement by project staff is critical for retention and the structure of this intervention which engages community LB centers supports retention of participants.^{11,12} Similarly, providing social support (through the support group study design) will also assist in retaining participants.³ For intensive follow-up which aids in retention, we plan to have a social media presence, updated regularly by a project assistant.¹³ In addition, we will send out reminder notices to participants for their appointments and data entry times throughout the intervention, periodic cards, e.g., birthdays, announce coming-out day events, and community festival dates to increase contact after the conclusion of the 16-week intervention. We will also be in monthly contact with the participants post invention to the end of the study. We will also obtain preferred contact information for reminders as well as alternative contact, such as best friend or partner to help located any who may have moved.

Finally, we will be providing incentives for the follow-up data collections. The use of incentives as positive reinforcement has been found to be successful in retaining participants.¹⁴ These numerous implicit and explicit strategies integrated into the study design increases our likelihood of retaining participants for the study period.

We expect to achieve at least an 83 percent response rate for the completion of both the baseline and follow-up surveys.

B.4. Tests of Procedures or Methods

The survey instrument has been drafted and has undergone two reviews: (1) an internal review conducted by NORC's Internal Review Board and (2) a pretest with five adults. Revisions were made to the instruments in response to comments received from both of these reviews.

Modifications to the length, content, and structure of the survey have been made based on the results of the survey pre-test interviews. Respondents provided generally positive feedback indicating that they could readily answer the questions and that the time to complete the survey was not onerous (about 30 minutes).

B.5. Names and Telephone Numbers of Individuals Consulted

The information for this study is being collected by NORC at the University of Chicago, an independent research organization, and the University of Missouri - Columbia, on behalf of the Office of Women's Health. With OWH oversight, NORC and MU are responsible of the study design, instrument development, data collection, analysis, and report preparation.

The instrument for this study and the plans for statistical analyses were developed by NORC and MU in conjunction with the other contract awardees. The staff team is composed of Dr. Elizabeth Hair, Project Director; Dr. Jane McElroy, Principal Investigator. Contact information for these individuals is provided below.

Name	Number
Elizabeth Hair	(301) 634-9386
Jane McElroy	(573) 882-4993

References

- (1) Ray, N. "Lesbian, gay, bisexual and transgender youth: An epidemic of homelessness," 2007.
- (2) Valanis, B. G.; Bowen, D. J.; Bassford, T.; Whitlock, E.; Charney, P.; Carter, R. A. *Archives of Family Medicine* **2000**, 9, 843-53.
- (3) Yancey, A. K.; McCarthy, W. J.; Harrison, G. G.; Wong, W. K.; Siegel, J. M.; Leslie, J. *Journal of Womens Health* **2006**, 15, 412-429.
- (4) Zhu, S. K.; Wang, Z. M.; Heshka, S.; Heo, M. S.; Faith, M. S.; Heymsfield, S. B. *American Journal of Clinical Nutrition* **2002**, 76, 743-749.
- (5) Lynch, N. A.; Ryan, A. S.; Berman, D. M.; Sorkin, J. D.; Nicklas, B. J. *Menopause-the Journal of the North American Menopause Society* **2002**, 9, 456-462.
- (6) Chen, L. W.; Caballero, B.; Mitchell, D. C.; Loria, C.; Lin, P. H.; Champagne, C. M.; Elmer, P. J.; Ard, J. D.; Batch, B. C.; Anderson, C. A. M.; Appel, L. J. *Circulation* **2010**, 121, 2398-2406.
- (7) Bloomfield, K. *Drug and Alcohol Dependence* **1993**, 33, 257-269.
- (8) Boehmer, U.; Bowen, D. J. *Preventive medicine* **2009**, 48, 357-361.
- (9) Kumpfer, K.; Alvarado, R.; Smith, P.; Bellamy, N. *Prevention Science* **2002**, 3, 241-46.
- (10) Zayas, L. H.; McKee, M. D.; Jankowski, K. R. B. *Annals of Family Medicine* **2004**, 2, 504-508.
- (11) Gauthier, M. A.; Clarke, W. P. *Alzheimer Disease & Associated Disorders* **1999**, 13, S29-S33.
- (12) Hessol, N. A.; Schneider, M.; Greenblatt, R. M.; Bacon, M.; Barranday, Y.; Holman, S.; Robison, E.; Williams, C.; Cohen, M.; Weber, K. *American Journal of Epidemiology* **2001**, 153, S192-S192.
- (13) Mychasiuk, R.; Benzies, K. *Child Care Health and Development* **2012**, 38, 753-756.
- (14) Parra-Medina, D.; D'Antonio, A.; Smith, S. M.; Levin, S.; Kirkner, G.; Mayer-Davis, E. *Journal of the American Dietetic Association* **2004**, 104, 70-75.