Design and statistical analysis for the Pathways study[1–3]

CE Davis, Sally Hunsberger, David M Murray, Richard R Fabsitz, John H Himes, Larry K Stephenson, Benjamin Caballero, and Betty Skipper

ABSTRACT We report the design, rationale, and statistical procedures used in Pathways, a randomized, school-based intervention for the primary prevention of obesity in American Indian children. The intervention, which is now being implemented in 7 American Indian communities around the country, includes a health-promotion curriculum, a physical education program, a school meal program, and a family involvement component. Forty-one schools serving American Indian children were randomly assigned to be either intervention or control groups. The intervention will begin in the third grade and continue through the end of the fifth grade. Efficacy of intervention will be assessed by differences in mean percentage body fat, calculated by a prediction equation, between intervention and control schools at the end of the fifth grade. Power computations indicate that the study has power to detect a mean difference of 2.8% in body fat. Data analysis will use intention-to-treat concepts and the mixed linear model. The study will be completed in 2000. *Am J Clin Nutr* 1999;69(suppl):760S–3S.

KEY WORDS American Indians, Native Americans, obesity prevention, study design, school-based studies, children, percentage body fat

INTRODUCTION

Recognizing the prevalence and health consequences of obesity in American Indians (1), the National Heart, Lung, and Blood Institute sponsored a planning and feasibility study for the primary prevention of obesity in American Indian children. Five organizations (the Gila River Indian Community with the University of Arizona, the Johns Hopkins University, the University of Minnesota, the University of New Mexico, and the University of North Carolina), working with 6 American Indian nations (Pima/Maricopa, Tohono O'Odham, Navajo, White Mountain Apache, Oglala Lakota, and Sicangu Lakota) were selected to develop a study design, an intervention, and an evaluation plan. The feasibility study was funded in September 1993 through August 1996. The study investigators and American Indian colleagues selected *Pathways* as the name of the study.

The primary aim of the Pathways study is to develop, implement, and evaluate a culturally appropriate, school-based intervention to promote healthful eating behaviors and increased physical activity to prevent obesity in American Indian children. The secondary aims are to determine whether the intervention will increase the level of physical activity, decrease dietary fat intake, and change knowledge, attitudes, and behaviors related to food choices and physical activity; document the degree to which the intervention was implemented in each intervention school; and evaluate the safety of the intervention.

This article describes the design, rationale, and statistical procedures of the full-scale study, which began after the pilot and feasibility phases were completed and is now underway. The other articles in this supplement provide details concerning the development of the intervention and measurement procedures and present results from the feasibility phase as well.

OVERVIEW OF STUDY DESIGN

A culturally appropriate, school-based intervention involving food service, physical activity, classroom curriculum, and family involvement was designed (2–6) to meet the study aims. This intervention will be evaluated by conducting a randomized study in 41 schools that teach American Indian children, 21 of which were randomly assigned to implement the intervention, with the other 20 serving as control schools (*see* **Figure 1**). The intervention begins in the third grade and lasts for 3 school years. The primary outcome for evaluating the intervention will be mean difference between intervention and control schools in percentage body fat (PBF) at the end of the fifth grade (adjusted for baseline PBF). PBF is estimated by using an equation developed during the pilot study (TG Lohman, B Caballero, JH Himes, et al, personal communication, 1998), using measurements of height, weight, skinfold thickness, and bioelectrical impedence.

The criteria for inclusion of a school in the study are as follows: *1*) all tribal, government, and religious units that control

[1] From the Department of Biostatistics, School of Public Health, University of North Carolina, Chapel Hill; the National Heart, Lung, and Blood Institute of the National Institutes of Health; the Department of Epidemiology, School of Public Health, University of Minnesota; Gila River Indian Community; the Department of Nutrition, Johns Hopkins School of Public Health, Johns Hopkins University; and the Center for Health Promotion and Disease Prevention, University of New Mexico.

[2] Supported by the National Heart, Lung, and Blood Institute of the National Institutes of Health (U01-HL-50869, U01-HL-50867, U01-HL-50905, U01-HL-50885, and U01-HL-50907).

[3] Address reprint requests to CE Davis, University of North Carolina at Chapel Hill, Department of Biostatistics, CB#8030, Chapel Hill, NC 27514. E-mail: Ed_davis@unc.edu.

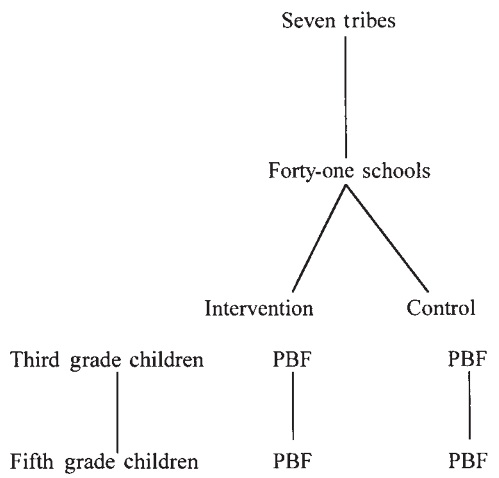

FIGURE 1. Diagram of Pathways study design. PBF, percentage body fat.

access to the school must give consent; *2*) the projected third grade enrollment in the fall of 1997 must be ≥15 children; *3*) ≥90% of children in the third grade are American Indian; *4*) the average retention from third to fifth grade over the preceding 3 y is ≥70%; *5*) the school administers the meal program on site, *6*) the school has the minimum facilities with which to implement the physical activity program; and *7*) there is no indication that the school will close or merge with another school in the 3 y of the study.

Forty-one schools agreed to participate in the full-scale study. They are divided among the field centers and tribes as follows: the Gila River Indian Community with the University of Arizona (Pima, Maricopa, and Tohono O'Odham), 8 schools; the Johns Hopkins University (White Mountain Apache), 6 schools; the University of Minnesota (Oglala Lakota and Sicangu Lakota), 13 schools; and the University of New Mexico (Navajo), 14 schools. One additional community, the San Carlos Apache, will be added to the full-scale phase of the study.

The intervention begins in the fall of the third-grade year and continues until the end of the spring semester in the fifth-grade year. Baseline measurements of body composition, physical activity, menu composition, and dietary intake (by direct observation) are collected in the spring of the second-grade year. This time was chosen to reduce the burden on the schools and children in the fall of the third grade, to collect baseline data before randomization of schools, and to provide adequate time to train teachers for the intervention before the beginning of school in the fall of third grade. Other baseline measures (5, 6) are obtained in the fall of the third grade school year. Measurements are taken according to the schedule in **Table 1**. Baseline 24-h dietary recalls were not obtained because there was no funding for them. The end-of-study 24-h dietary recall will provide an unbiased comparison of intervention and control children because the study is randomized. The baseline and intermediate diet observations provide a less expensive method of monitoring changes in diet for the 2 groups.

The study has been reviewed and approved by institutional review boards at the participating universities and by tribal authorities representing each participating American Indian nation. The National Heart, Lung, and Blood Institute of the National Institutes of Health appointed a protocol review committee that recommended approval of the study protocol and a data and safety monitoring committee to review the progress of the study periodically to ensure the safety of the children participating in the study.

RATIONALE FOR STUDY DESIGN

Randomization

The unit of randomization is the school. Schools were stratified within each American Indian community on the basis of median PBF of students at the end of second grade by using data from the baseline survey. The schools at each participating American Indian community were ranked according to median PBF. The half of the schools with the larger median PBF constitutes one stratum; the half with the smaller median PBF is the second stratum. The Coordinating Center (University of North Carolina) randomly assigned half of the schools within each stratum to be intervention schools. The randomization was constrained to ensure that one-half of each field center's schools was assigned to the intervention.

It is recognized that individual randomization usually yields more information per subject and thus is more efficient than group randomization (7). Nevertheless, schools are the natural unit of randomization for this study because much of the intervention is delivered school-wide. For example, the food service intervention will result in changes in the breakfast and lunch menus for the entire school. It would be very difficult to implement an individually tailored school meal program that gives different foods to different children.

The investigators considered 3 forms for the randomization: *1*) simple randomization without stratification; *2*) pair-matched randomization, in which schools would be paired by median PBF and then randomly assigned to intervention or control; and *3*) the stratified randomization described in the first paragraph. Simple randomization was not used because of the sizable probability of imbalance of baseline PBFs (8).

Both stratification and pair matching are intended to reduce the interschool heterogeneity and thus reduce the error variance used to test for intervention effects (9, 10). Because there are 41 schools in Pathways and the data analysis uses field center as a blocking variable, there would be 19 df for testing intervention effects for the pair-matched study and 36 df for the stratified study, adjusting for field center effects. If the stratified study successfully reduced interschool heterogeneity, it would thus be more powerful than the pair-matched study. However, pair matching provides more stringent control for the intervention comparison, thereby reducing the error variance, and thus it would be more powerful than the stratified design. Insufficient data on PBF were available to make judgments concerning the relative merits of these 2 approaches, and the choice made by the investigators was under the assumption that stratification would provide adequate control for the differences in PBF among schools.

COHORT SELECTION

In group-based randomization studies, a choice must be made between longitudinal cohorts and serial cross-sectional surveys (11, 12). Because the loss of children from the cohorts was expected to be low and the correlation between baseline and

TABLE 1
Schedule of measurements of participating children in the Pathways study[1]

	Grade and session						
	2	3		4		5	
	Spring 1997	Fall 1997	Spring 1998	Fall 1998	Spring 1999	Fall 1999	Spring 2000
Height and weight	X		X		X		X
Body composition	X						X
Diet observation	X		X		X		X
24-h diet recall							X
Physical activity questionnaire	X		X		X		X
Accelerometer	X						X
KAB questionnaire		X	X		X		X
Process evaluation	---------- Continuous ----------						

[1] KAB, knowledge, attitudes, and behaviors.

ending PBF was likely to be strong, the use of a cohort in Pathways was chosen.

The Pathways investigators considered 3 possible cohort types: *1*) a single cohort followed from second through fourth grade, *2*) a single cohort followed from third through fifth grade, and *3*) 2 sequential cohorts followed from third through fifth grade. Because of the perceived difficulty in developing and implementing a health-promotion curriculum as well as in obtaining baseline measures through use of questionnaires in younger children, the first cohort type was not selected. The third cohort type would have the major advantages of reducing the number of schools required for the study and increasing the strength of the intervention because teachers and other staff members would have experience in implementing it. The disadvantages of this 2-cohort approach, however, are that the study would last 1 y longer and the baseline measures on the second cohort would not be free of the school-based intervention. For example, the second cohort would have been exposed to the food service intervention for 1 y before the baseline PBF measures are made because the changes in the school meals apply to all children. The Pathways investigators concluded that the disadvantages of the 2-cohort design outweigh the advantages and thus adopted the second cohort type considered, a single cohort followed in grades 3–5.

DATA ANALYSIS

The statistical analyses will use a mixed linear model (13). PBF at the end of the fifth grade will be the primary response variable. Fixed effects in the model will be baseline PBF (end of second grade) and treatment group (control or intervention). Random effects will be field center, school within stratum, and student within school. The test of treatment effect will be a two-sided test, with a significance level of 0.05. A preliminary test of center × intervention interaction will be conducted. If this test does not indicate that the effect of the intervention is heterogeneous across centers, no claim of different effects by center (or tribe) will be made. If this test indicates that there is a center × intervention effect, results will be reported separately for each center. Assuming no center × intervention interaction, the analysis of variance table will be as shown in **Table 2** (14). If a school should withdraw before completing the study, we will impute values in a conservative manner by reducing the numerator df in the F test by the number of schools for which we have imputed a mean.

A secondary analysis will use only data from students with both second and fifth grade PBFs (ie, no imputed values) in the mixed model, as above. Secondary statistical analyses will use body mass index (in kg/m^2) during grades 3, 4, and 5, and use the mixed model for longitudinal data. This analysis is considered a secondary analysis because it could show a short-term effect (for example, at the end of grade 4) that could disappear before the end of grade 5. Other secondary analyses will include evaluation of the intervention effects on physical activity (as measured by the TriTrac accelerometer; Professional Products, Reining International), positive differences in diet (as measured by 1-d dietary recall) and knowledge, attitudes, and behaviors [KAB; as measured by the Pathways KAB questionnaire (5)]. For these variables (KAB, dietary behavior, and physical activity assessment) mixed linear models, as described above for the primary response variable, will be used.

POWER OF THE STUDY

Power calculations are based on a t test where $n - 2 - 3 = n - 5$ df, and n is the total number of schools. The formula for calculating the detectable difference in cluster randomization studies is

$$\Delta = (t_{\alpha/2} + t_{1-\beta})\ \sigma\ \{2[1 + (m - 1)\ \rho]\ (1-R^2)/mn\}^{1/2} \qquad (1)$$

where $t_{\alpha/2}$ and $t_{1-\beta}$ are the upper $\alpha/2$ and $1 - \beta$ percentiles of the t distribution with $n - 5$ degrees of freedom, m is the number of children per school, n is the number of schools, ρ is the intraschool correlation and R is the correlation between the baseline and final measurements of a child.

The following estimates from a pilot study conducted in the spring of 1995 on 11-y-old American Indian children were used to calculate an estimate of Δ: SD (σ) 7.8, intraclass correlation (ρ) 0.13, and correlation between baseline and end measurement

TABLE 2
Analysis of variance in the Pathways study[1]

Source	df
Intervention	1
Stratum	7
School within stratum	31
Pupil	

[1] Test of treatment effect is $F = MS$(intervention)/MS(school within stratum), which is approximately equivalent to a t statistic with 31 df.

(R), 0.7. With 40 schools, 15 children per school, $\alpha = 0.05$, and $1 - \beta = 0.8$, the detectable mean difference (Δ) is estimated to be 2.2%.

Detectable difference estimates were also calculated under assumptions that some schools may drop out of the study and that not all children will complete the study. If 10% of the original schools do not complete the study, the estimated detectable mean difference increases to 2.3%. If 10% of the schools drop out and only 10 of the 15 children in a school complete the study, the detectable mean difference is 2.5%.

We believe that differences as large as this will be important and are attainable with our intervention. For example, consider the effect of changing the milk served from 2% to 1% fat for 2 meals/d for an entire school year. If 80% of schools change completely, 20% of schools change for half of the year, and within a school 60% of students change for the full year, 30% change for half of the year, and 10% do not change, the estimated difference in body fat would be 0.526 kg in 3 y. Similarly, if physical activity is increased by an additional 30 min/wk in one-third of the schools, by 15 min/wk in one-third of the schools, and not at all in one-third of the schools, the difference would be 0.30 kg in 3 y. If these 2 interventions are additive, the total difference in fat between control and intervention schools would be 1.43 kg. For children who start with a PBF of 37%, this would be a 2.55% reduction. Therefore, given the other interventions (such as rinsing cooked ground meat and reducing consumption of sugar-containing drinks), the ≈2.8% reduction we can detect is attainable.

Moreover, we believe that a reduction of this magnitude in mean PBF will lead to a sizable reduction in the proportion of children who are obese. Although we are unaware of any defined value of PBF that leads to a classification of obese, we assume that most persons would agree that a child whose PBF is >45% is obese. Assuming that PBF is approximately normally distributed with a mean of 37.1% and a SD of 7.8% (the values observed in our pilot study), one would expect that 15.6% of children would have a PBF >45%. If the mean PBF in the population were reduced by 2.8–34.3%, with no change in the SD, one would expect that 8.5% of children would have a PBF > 45%, a reduction of >7% in the prevalence of obesity.

CONCLUSION

Pathways was designed as a randomized, school-based study and detailed statistical analysis plans have been developed. The feasibility phase of the study resulted in the development of an intervention program, data collection instruments, and preliminary data sufficient to support and justify the study. Implementation began in spring 1997 and will be completed in the year 2000.

REFERENCES

1. Story M, Evans M, Fabsitz RR, Clay TE, Holy Rock B, Broussard B. The epidemic of obesity in American Indian communities and the need for childhood obesity-prevention programs. Am J Clin Nutr 1999;69(suppl):747S–54S.
2. Davis SM, Going SB, Helitzer D, et al. Pathways: a culturally appropriate obesity-prevention program for American Indian schoolchildren. Am J Clin Nutr 1999;69(suppl):796S–802S.
3. Lohman TG, Caballero B, Himes J, et al. Body-composition assessment in American Indian children. Am J Clin Nutr 1999;69(suppl):764S–6S.
4. Going SB, Levin S, Harrell J, et al. Physical activity assessment in school-age American Indian children in the Pathways study. Am J Clin Nutr 1999;69(suppl):788S–95S.
5. Stevens J, Cornell CE, Story M, et al. Development of a questionnaire to assess knowledge, attitudes, and behaviors in American Indian children. Am J Clin Nutr 1999;69(suppl):773S–81S.
6. Helitzer DL, Davis SM, Gittelsohn J, et al. Process evaluation in a multisite primary obesity-prevention trial for American Indian schoolchildren. Am J Clin Nutr 1999;69(suppl):816S–24S.
7. Cornfield J. Randomization by group: a formal analysis. Am J Epidemiol 1978;108:100–2.
8. Grizzle JE. A note on stratifying versus complete random assignments in clinical trials. Control Clin Trials 1982;2:365–8.
9. Gail MH, Byar DP, Pechacek TF, Corle DK. Aspects of statistical design for the community intervention trial for smoking cessation (COMMIT). Control Clin Trials 1992;13:16–21.
10. Martin DC, Diehr P, Perrin EB, and Koepsell TD. The effect of matching on the power of randomized community intervention studies. Stat Med 1993;12:329–38.
11. Duncan GC and Kalton G. Issues of design and analysis of surveys across time. Int Stat Rev 1987;55:97–117.
12. Dwyer J, Feinleib M. Introduction to statistical models for longitudinal observation. In: Dwyer JH, Feinleib M, Lippert P, Hoffmeister H, eds. Statistical models for longitudinal studies of health. New York: Oxford University Press, 1992:3–47.
13. Bryk AS, Raudenbush SW. Hierarchical linear models. Newbury Park, CA: Sage Publications, 1992.
14. Murray DM, Hannan PJ, Jacobs DR. Assessing intervention effects in the Minnesota Heart Health Program. Am J Epidemiol 1994;139:19–103.