



Published in final edited form as:

*Am Heart J.* 2017 May ; 187: 170–181. doi:10.1016/j.ahj.2017.02.020.

## Rationale and Design of *Family-Based Approach in a Minority Community Integrating Systems–Biology for Promotion of Health (FAMILIA)*

Sameer Bansilal, MD, MS, Rajesh Vedanthan, MD, MPH, Jason C. Kovacic, MD, PhD, Ana Victoria Soto, MD, Jacqueline Latina, MD, Johan L. M. Björkegren, MD, Risa Jaslow, MS, RDN, Maribel Santana, BS, Samantha Sartori, PhD, Chiara Giannarelli, MD, PhD, Venkatesh Mani, PhD, Roger Hajjar, MD, Eric Schadt, PhD, Andrew Kasarskis, PhD, Zahi A Fayad, PhD, and Valentin Fuster, MD, PhD

Icahn School of Medicine at Mount Sinai, New York, NY.

### Abstract

**Background**—The 2020 American Heart Association Impact Goal aims to improve cardiovascular health of all Americans by 20% while reducing deaths from cardiovascular disease and stroke by 20%. A large step toward this goal would be to better understand and take advantage of the significant intersection between behavior and biology across the entire life-span. In the proposed FAMILIA studies, we aim to directly address this major knowledge and clinical health gap by implementing an integrated family-centric health promotion intervention and focusing on the intersection of environment and behavior, while understanding the genetic and biologic basis of cardiovascular disease.

**Methods**—We plan to recruit 600 preschool children and their 600 parents or caregivers from 12–15 Head Start schools in Harlem, NY, and perform a 2:1 (2 intervention/1 control) cluster randomization of the schools. The preschool children will receive our intensive 37-hour educational program as the intervention for 4 months. For the adults, those in the “intervention” group will be randomly assigned to 1 of 2 intervention programs: an “individual-focused” or “peer-to-peer based.” The primary outcome in children will be a composite score of knowledge (K), attitudes (A), habits (H), related to body mass index Z score (B), exercise (E), and alimentation (A) (KAH-BEA), using questionnaires and anthropometric measurements. For adults, the primary outcome will be a composite score for behaviors/outcomes related to blood pressure, exercise, weight, alimentation (diet) and tobacco (smoking; Fuster-BEWAT score). Saliva will be collected from the children for SNP genotyping, and blood will be collected from adults for RNA sequencing to identify network models and predictors of primary prevention outcomes.

---

sameer.bansilal@mssm.edu.

Reprint requests: Sameer Bansilal, MD, MS, Icahn School of Medicine at Mount Sinai, One Gustave L. Levy Place, Box 1030 New York, NY 10029.

Conflicts of interest

None of the investigators have any financial conflicts of interest related to this work.

Appendix. Supplementary data

Supplementary data to this article can be found online at <http://dx.doi.org/10.1016/j.ahj.2017.02.020>.

**Conclusion**—The FAMILIA studies seek to demonstrate that targeting a younger age group (3–5 years) and using a family-based approach may be a critical strategy in promoting cardiovascular health across the life-span. (*Am Heart J* 2017;187:170–81.)

---

## Background

Over the last 2 decades, implementation of various primary, secondary, and tertiary preventive strategies have led to a decline in cardiovascular mortality. However, much work remains to be done to accomplish the 2020 American Heart Association (AHA) Impact Goal of improving cardiovascular health of all Americans by 20% while reducing deaths from cardiovascular disease (CVD) and stroke by 20%.<sup>1</sup> We believe that a large step forward would be to better understand and focus on health promotion across the life-span through a family-centric approach, while taking advantage of advances in the understanding of genomics and behavior.

Data from several large-scale, epidemiologic studies have documented the excess burden of important cardiovascular risk factors, among African Americans and Hispanic adults compared with white individuals.<sup>2,3</sup> The National Health and Human Examination Survey III reported that physical inactivity and diabetes were twice as common in black and Mexican American women compared with whites.<sup>4</sup> The Harlem area of New York City (NYC) has among the highest obesity and diabetes prevalence rates in the city. The NYC report on South Bronx and Harlem reported that 96% of the surveyed participants ate less than the recommended daily fruit and vegetable consumption, and nearly 40% did not engage in any form of exercise.<sup>5</sup> Simultaneously, the prevalence of childhood obesity in the United States since the 1960s has been increasing.<sup>6</sup> The 2003–2004 National Health and Nutrition Examination Survey reported that an estimated 17.1% of children and adolescents aged 2 to 19 years are overweight in the United States, compared with 13.9% in 1999 to 2000.<sup>7</sup> Although the prevalence of obesity has declined overall in NYC owing to children-focused efforts, the declines have been blunted in the Hispanic population and negligible in African American children.<sup>8</sup> In Harlem, nearly 1 in 3 children in the preschool Head Start programs are obese, and almost half are overweight or obese. In addition, nearly 1 in 4 children in public elementary schools are obese, and nearly 4 in 10 are overweight or obese.<sup>5</sup>

Furthermore, although genome-wide association study efforts have revealed many reproducible genetic associations with coronary artery disease (CAD), they explain only about 10% of the total estimated variation in CAD/myocardial infarction heritability.<sup>9</sup> Given the complexity of CAD pathogenesis, new analytical approaches are much needed to boost the probability of identifying causal drivers and pathways in this disease.

Over the last few years, we have successfully tested our children's intervention program—SI! (Salud Integral) for preschoolers in Bogota, Colombia, and Madrid, Spain, through cluster randomized trials involving approximately 3,500 children.<sup>10</sup> The Si! Program was designed in partnership with Sesame Street and our academic partners in Colombia (Cardio Infantil) and Spain (Centro Nacional de Investigaciones Cardiovasculares) as a multilevel (school, teachers, families, and children), multicomponent (healthy diet, increased physical activity, understanding the human body, and managing emotions) school-based intervention.

In 2014, as part of its long-term vision regarding research in prevention, the AHA funded 4 nationwide Strategically Focused Research Network centers. In the proposed FAMILIA studies, our Strategically Focused Research Network center at the Icahn School of Medicine at Mount Sinai will study a family-based approach to prevention which includes preschool children and their parents or caregivers in simultaneous, parallel intervention studies and leverage genetic materials from children and parents or caregivers to understand the complex relationships between behavior and genomics.

## Methods

### Setting

The study is being conducted through the Head Start program schools (Figure 1) in Harlem, NY. Harlem is located in upper Manhattan and stretches from the East River to the Hudson River, and between 155th Street, where it meets Washington Heights, to an uneven border along the south. Harlem is predominantly a mix of immigrant Hispanic and African American populations. The neighboring Icahn School of Medicine at Mount Sinai has had a long tradition of interaction and works within this community.

**Head Start facilities.**—Head Start is a federally funded program that promotes the school readiness of children from birth to age 5 from low-income families by enhancing their cognitive, social, and emotional development. This program is currently operational at 32 preschool facilities in Harlem, NY. The average Head Start facility in Harlem has approximately 4 preschool classrooms, with each class size ranging from 10 to 23 children and an average of 16 children per class. Head Start places a strong value on promoting adult-child interaction.

### Participants

We plan to recruit approximately 600 preschool children (3–5 years of age) and at least 600 parents or caregivers from approximately 12–15 Head Start schools.

### Study participants

We are targeting preschools with comparable characteristics in socioeconomic level and ethnicity. In addition, we will use the following criteria for inclusion:

1. The preschools must be located in Harlem, NY.
2. The preschools must be public.
3. The preschools must have children between 3 and 5 years of age.
4. The preschools must provide meals for the children.
5. The preschools must make available use of their applicable program operation space.

In addition to the inclusion criteria above, we will ask the participating schools to make the following commitments:

1. Not to participate in any other major structured health intervention program aside from the usual curriculum during the study

All adult caregivers or parents of the recruited school children who consent will be included in the parallel randomized trial for adults. Parents/caregivers are not necessarily blood relatives, but those that interact and significantly impact the child's health.

## Design

The 4-year study period (July 2014–June 2018) is broadly divisible into an initial preparatory pilot phase followed by the randomized trials with active intervention and follow-up. Pilot phase In the 1-year initial pilot period (Figure 2), we will identify the contextual factors, facilitators, and barriers that may impact the implementation of a preschool-based health promotion educational program (Si! Program modified to create the FAMILIA children's intervention program) integrated with adult lifestyle intervention strategies in Harlem, using qualitative research methods such as focus groups for relevant stakeholders such as school directors, teachers, social services/parent engagement staff, Nutrition/Health staff consultants, and parents. We will carry out a pilot intervention of the contextually and culturally adapted program in up to 3 preschools in Harlem. We will assess for acceptability and feasibility, and make any necessary modifications before the implementation of the cluster-randomized trial. Randomized trial: Children's Study: We will use a randomization algorithm that allows for a stratified allocation, so as to ensure that the 2 groups will be balanced for the presence of certain characteristics that may affect the outcomes (primary native language, zip code). Since the intervention program involves minimal if any risk, the control arm participants will receive the intervention program after the initial 4-month randomized trial period, irrespective of the impact of the program. This is in keeping with our responsibility as health care providers to this neighborhood and we believe ensures engagement of the control group. The effectiveness of the FAMILIA children's intervention program for preschool education developed by our team in concert with the Sesame Workshop and our collaborators in Colombia and Spain will be evaluated through a cluster randomized trial of 600 children aged 3 to 5 years enrolled in 12–15 preschools in Harlem through the Head Start program (Figure 2). Children randomized to the intervention arm (n ~ 400) will receive the 4-month, 37-hour health promotion education program along with their parents/caregivers and teachers. The control arm (n ~ 200) will receive the standard curriculum in their schools (Figure 3). Details of the FAMILIA Children's intervention program The FAMILIA Children's intervention program incorporates a comprehensive and integrated vision of cardiovascular health promotion that encompasses 4 basic and interrelated components:

- Body and heart
- Physical activity
- Healthy diet
- Emotional management

The curricular objectives are listed in Table I. These objectives translate into curricular units that align with the child's age-appropriate development phase. Four basic subjects in 5

teaching units have been developed. The specific implementation of the teaching units in the classroom will depend, in large measure, on the pedagogical methodology that each respective school uses and may be spread across a week, done in a crosscutting way, or as a project. There are 4 key activities in each unit to cover its specific contents. In addition to these basic activities, additional ones are provided as a guide to reinforce the learning processes of all the components, encouraging the children to understand the relationship between them. These key activities include short Sesame Street video clip episodes where nutrition, physical activity, body and heart, and social and emotional well-being (Table II).

To encourage the participation of the entire family in the FAMILIA Project, we will invite parents and other family members of the participating children to informational and educational meetings, called FAMILIA days. In addition, the teachers will regularly provide parents a minimum of 11 family health activities during the duration of the child educational intervention. The purpose of the “Family Activities” is for children to share and do with their parents/caregivers to reinforce the home-school connection.

**Adults Study.**—We will recruit approximately 600 parents and caregivers of the preschool children and similarly cluster randomize them in a 2:1 (2 intervention/1 control) fashion, with parent/caregiver assignment aligned to the assignment for the child. Those in the “intervention” group will be randomly assigned to 1 of 2 interventions— an “Intensive Individual Intervention Program” consisting of 8 to 12 individual counseling sessions and a personal activity monitoring device as a motivational agent, or the “Peer-To-Peer Program Intervention,” consisting of monthly group meetings of participants to help everyone in the group collectively improve their cardiovascular risk factor profile. The intervention program will continue for 12 months (Figure 3).

**Control arm (n ~ 200 parents/caregivers).**—Parents and caregivers randomized to the control arm will not receive any structured program for the first 4 months (coinciding with their children also not receiving the educational intervention program in the children’s study). After 4 months, the parents and caregivers will receive the relevant components of the FAMILIA curriculum for the rest of the school year (approximately 4 months). During their 4-month FAMILIA Project participation only, the parents and caregivers will be invited for regular biweekly informational and educational meetings to be organized by staff. At all other times through the duration of the intervention, the teacher will regularly send educational material with the students to work on with their families at home. To keep this group engaged, we plan to conduct seminars dealing with issues not related to the outcome of interest, but useful to the participants, such as related to immigration, childcare, financial and tax advice, life and disability insurance, and vocational courses. The control group will have access to the study Web site, which will have additional health information.

## Details of intervention programs

### **Intensive Individual Intervention Program (n ~ 200 parents/caregivers)**

**Personalized lifestyle counseling.:** These will be one-on-one counseling sessions, lasting approximately 45 minutes, with a lifestyle coach who is versed in physical activity, nutrition, motivational interviewing, weight loss, psychology, or social work. The underlying

technique in all these sessions will be motivational interviewing. Motivational interviewing is a counseling style that stimulates behavioral change by focusing on exploring and resolving ambivalence.<sup>11</sup> These 45-minute sessions will occur every 3 to 4 weeks during the first 8 months, with 4 additional complimentary sessions offered to the participants for up to approximately 12 months for this intervention.

The first session (baseline) will capture the participant's lifestyle behavior patterns. A mutually agreed plan of action will be formulated that fits well with the participant's personal circumstances and preferences. Sessions 2 to 5 will be used to educate, monitor, and steer the participant's progress and usage of intervention components. Each of the sessions will be structured around more specific lifestyle and behavioral change modules. Sessions will also focus on maintenance of the behavioral lifestyle changes. Session 8 will provide an overview of the previous sessions as it pertains to the individual's goals as well as a discussion of the culmination of the process and promotion of continuing healthy habits. Each participant will be offered 4 complimentary sessions that they can redeem with a lifestyle coach during 4 subsequent months within the 12-month intervention period.

**Activity monitor.:** Pedometers that self-monitor physical activity were shown in a recent Lancet review to be the most effective intervention tool to improve physical activity.<sup>12</sup> We will be using the Garmin ViVoFit 2 device, Garmin Ltd, Switzerland. This is a next-generation pedometer that also allow for self-monitoring of diet and sleep in addition to physical activity. The device can measure (steps taken, calories burned, floors climbed, and provide goal setting functions). The devices link wirelessly to information communication technology platforms in the form of Web site, smart phone, and tablet applications.

**Peer-To-Peer Program Intervention (n ~ 200 parents/caregivers).**—Similar to a Support Group model, participants will meet in groups of up to 20 and support each other in self-control of CV risk factors. We have found that the collective will and support generated within these groups can amplify motivation to change at the level of the individual.

**Selection of peer leaders.:** In peer education, the leader figure is essential. Each group, using a preset group dynamics and based on established leader characteristics provided by a psychologist and expert on leadership, will pick 2 participants, including those with voluntary will, clear objectives, and availability, who will assume the role of group leaders (leader and co-leader).

**Training of the leaders.:** Chosen leaders will receive specific leadership training. This leadership training will be a 2-day training each lasting 4 hours each. Leaders have the following functions: (i) raise the awareness of other members of the importance of learning to listen and encourage mutual support, (ii) mark the tone of the discussion, (iii) addressing issues that arise during the meeting, (iv) monitor the evolution of the changes, and(v) encourage and maintain member involvement.

**Parent Ambassadors.:** Parents/caregivers of children that have attended a preschool program (Head Start Center) in previous years will work with the participating Head Start

programs, peer leaders, and parents/caregivers to support the adult participant engagement and peer group process.

**Group meetings:** Lasting approximately 12 months on a monthly basis, participants will gather to share their own experiences, problems, and knowledge associated with health habits; evaluate changes, and offer mutual support. Peer groups within a school are formed based on convenience related to participants' schedule, language preference, and so on. The meetings will occur at the schools and last 60 to 90 minutes each. A local school coordinator working with Parent Ambassadors will be available to ensure space and logistics, and one study representative/Parent Ambassadors representative will attend the first 3 meetings. Thereafter, the study representative/Parent Ambassadors will meet with Peer Leaders to check-in on the status of the peer group sessions approximately every quarter. In addition, the Peer Leaders will be requested to complete Peer Leader feedback forms for each session held. Intervention group members have access to the Web site, where they will find information on health habits and risk factors as well as other health-related links.

**Point system incentive structure.:** An established point-based incentive structure will be used in which participants will earn points based on their attendance/participation in the assessments (FAMILIA Days), interventions, educational, and informational meetings. All study participants can accrue points in which they can redeem for items such as sporting event tickets, sports equipment, and gift cards for food/or sports equipment.

## Genomics

We will collect saliva from all children with assent at enrollment. Blood will be collected from all adults who provide consent at enrollment and at the end of the intervention period (Figures 2 and 3). Saliva from children will be used to isolate DNA. Blood from adults will be used to isolate DNA and plasma. In addition from the adults, by in vitro culture of blood-derived mononuclear cells, we will derive macrophages, which will then be driven to become foam cells. As disease-relevant models, foam cells are generated by loading participants' macrophages with acetylated low-density lipoproteins to drive gene expression toward an atherosclerosis-relevant profile. Because macrophages are equally important to atherosclerosis, RNA will be isolated from both macrophages and foam cells from adults. We will bank all samples, but our main analysis will be performed on the samples from 450 children and 600 of their related adult caregivers.

The genetic, genomic, and molecular data isolated from macrophages/foam cells will primarily be used to infer network models of clinical outcomes in response to primary preventive measures in children and adults. In brief, the expression level of macrophage/foam cell genes, derived at baseline and follow-up (2 samples per adult), will be used to infer undirected (coexpression) gene networks/modules. Next, each coexpression module will be linked to the clinical characteristics of each adult participant, including the Fuster-BEWAT score. Directed (Bayesian) networks (including information from participants' DNA) will be integrated into the analysis of coexpression modules to identify those causally linked to the clinical characteristics. To improve the predictive power of these causal networks, we will integrate existing RNA sequence data sets from CAD patients and

directed and undirected network models from other in-house data sets of other complex diseases related to CAD, such as type 2 diabetes.<sup>13,14</sup> Finally, expression quantitative trait loci (eQTLs—loci in the genome that control gene expression levels) of the causal network models of prevention outcomes will be examined for their association with CAD using the CARDIo-GRAM meta-analysis genome-wide association data set of nearly 4 million DNA variants, of which approximately 200 are associated with CAD.<sup>15,16</sup>

We will then seek to define key gene network modules operative in CVD prevention. We will examine the network models and eQTLs (isolated from macrophage/foam cell RNA sequence data), including their association with molecular readouts both before and after the intervention period in adults, to identify changes in genetic-molecular interactions that result from the lifestyle interventions in adults. We will link prevention-induced changes in the network models to the clinical read-outs, including the Fuster-BEWAT score. Finally, we will generate models to predict the molecular state of networks in children based on the appearance of these models in related caregivers and adults. These models will be validated by using them to predict, in a blinded fashion, the clinical response of children to lifestyle intervention using KAH-BEA score and body mass index.

To identify the genetic, genomic, and molecular signature of favorable vs poor responders to lifestyle intervention, thus permitting future tailored approaches, we will derive risk scores to rank participants and examine whether these risk scores are associated with the Fuster-BEWAT score and other clinical responses following different intervention approaches. In this fashion, we will define the impact of genetic, genomic, and molecular effects on the efficacy of the CVD prevention approaches. Using the network models on foam cell formation/macrophages, the adult study participants will be scored in 3 ways: (1) genetically, based on the degree of inherited risk carried by the network eQTLs (considering both the number and the statistical significance of the CAD-associated eQTLs); (2) genomically, based on macrophage/foam cell network models; and (3) biochemically, based on the cell phenotype (eg, foam cell cholesterol-ester accumulation *in vitro*) glucose levels and lipid parameters. Next, adults will be ranked according to the combined genetic, genomic, and molecular risk scores. This ranking will be used to examine subgroups of participants in different prevention groups, to determine how this score affects CVD risk reduction in the intensive and control intervention groups.

The final component of these studies will be to identify novel therapeutic and diagnostic targets in network models of early atherothrombotic disease. We will mine the models created above to identify novel biomarkers of nascent atherothrombotic disease and therapeutic targets for early intervention by joint analysis of DNA genotypes and foam cell RNA sequence profiles. We will examine the network models for association to phenotypic data collected at baseline, after prevention, and change between time points. The phenotypic data will include the Fuster-BEWAT score, biochemistry (eg, glucose and lipid levels), and demographic clinical data. We will deploy a large collection of standard in-house reference data sets to determine the extent to which molecular processes previously linked to CAD (inflammation, metabolism, lipid trafficking) are represented in the identified networks. We will seek key drivers of disease and response to intervention in all network models. These key drivers are central to the overall function of the molecular network and hence are



promising drug targets because they frequently have high-hierarchy roles in disease networks.<sup>17</sup>

The parents/caregivers of the children will provide consent for the children to participate in the study, before conducting any trial-specific procedures. This consent form will include relevant details of the collection of a saliva sample from the children which will be used to provide DNA samples. Parents/caregivers will sign a separate consent form for their own participation. This consent form will include a separate section regarding collection of blood samples for genomic analyses. The study is approved by the Icahn School of Medicine Institutional review board.

### Assessments and outcomes

**Children's study.**—The primary outcome of the preschool-based health promotion educational program will be a change in composite KAH-BEA score at approximately 4 months (at the end of the intervention period): knowledge (K), attitudes (A), habits (H), related to body mass index Z score (B), exercise (E) and alimentation(A) (KAH-BEA) of the preschool children will be measured at approximately 4 months and the change in this score from baseline will be the primary outcome. To evaluate the sustainability of the impact of a preschool-based health promotion educational program on the KAH-BEA of preschool children, we will assess the children 2 additional times over the ensuing 2 school years.

In addition, anthropometric measures such as height, weight, and waist circumference of the participating children will be obtained, both in the intervention group schools and in the control group schools, at each assessment. To evaluate the fourth component of the program, management of emotions in the children, we will use the Test of Emotional Comprehension. Details of the assessment methods are provided in Supplementary Appendix.

**Adults' study.**—For the parents/caregivers study, the primary outcome will be a composite score consisting of a 0–15 scale for behaviors/outcomes related to blood pressure, exercise, weight, alimentation (diet), and tobacco (smoking; Fuster-BEWAT score, 0–3 for each domain). Assessments will be performed at baseline, 4 months (early-effect), and 8 months (interim-effect) at the end of the intervention (12-month peak-effect). The 12 month assessment will be used to calculate the between group difference for the change in the Fuster-BEWAT score and will be the primary outcome measure for the adults. A final 24-month assessment will be performed to evaluate the sustainability of the impact of the 2 intensive lifestyle interventions on healthy behaviors and biological parameters 12 months after the intervention program ends. For the assessment of secondary outcome measures, participants will undergo point-of-care testing for a lipid panel and blood sugar assessment, and a 3-dimensional ultrasound imaging (carotid, iliofemoral) to assess for presence of atherosclerosis at baseline and after the intervention period.

Secondary outcome measures will include a comparison of the following:

- Mean final scores for each of the individual domains assimilated in the composite score and change in the mean score

- Mean blood pressure at final visit and change in blood pressure (systolic and diastolic)
- Mean weight at final visit and change in mean weight
- Proportion of participants who lost 5% or 10% of initial body weight
- Mean International Physical activity Questionnaire score at final visit and change in mean score on Physical activity Questionnaire scale
- Proportion of participants with any positive change in behavior
- Interaction testing for presence or absence of atherosclerosis for above-mentioned measures
- Change in plaque volume (carotids and ileofemoral separately)

Details of the assessment methods are provided in Supplementary Appendix and the flow is shown in Figure 4.

**Genomics study.**—The primary objectives of the genomics component of the FAMILIA study are as follows: (1) to identify network models and predictors of primary prevention outcomes (KAH-BEA score in children and Fuster-BEWAT score in adults) using genetic, genomic, and molecular data from children and adults before and after lifestyle intervention; (2) to identify the genetic, genomic, and molecular signature of response to lifestyle intervention, thus permitting future-tailored approaches; (3) to identify novel therapeutic and diagnostic targets in network models of early atherothrombotic disease.

## Statistical plan

### Power and sample size

The present study is designed as a cluster-randomized trial with randomization occurring at the level of each school, and the primary power calculation is derived from the children's study.

### Children

Twelve-Fifteen schools will be randomized in a 2:1 fashion to intervention and control (ie, 2 intervention schools for every control school). We assumed an intraclass correlation (school-group effect) coefficient of 0.15 based on our previous work using a similar methodology.<sup>10</sup> In addition, we anticipated a mean difference of 3.5 units in the overall score between groups based on earlier work.<sup>10</sup> We calculated that 12–15 schools (40–50 children per school year) would provide greater than 80% power to detect a difference of this magnitude with a type I error rate of 0.05. Assuming an approximate 10% loss to follow-up, we plan on enrolling a total of 600 children aged 3 to 5 years from 15 separate schools. Among these, 8–10 schools will be allocated to intervention, whereas 4–6 will receive control (2:1).

### Adults

To observe the mean difference actually observed in the Fifty-Fifty study (0.7) for the Fuster-BEWAT score, and taking into account the SD (2.3), we need about 338 participants

in a 1:1 randomization. With a 2:1 randomization as in our case, we require closer to 400 participants. However, because we randomize participants in clusters, the total number required increases. We computed a range of number of clusters depending on the cluster size. With an intraclass correlation of 0.01, we need to recruit 12–15 cluster with a minimum of 40–50 participants each (total  $n = 600$ ).

For the genomics studies, with a sample size of 400, for QTL explaining 10% of the trait variation, the power is greater than 80%. Therefore, we would be well powered to pick up a very significant fraction of the gene expression QTL, and given that the cell-based assay traits will have larger effect sizes, closer to the expression traits, we have reasonable power to detect QTL for those traits as well, especially for those in which the additive genetic variance is N5%. Because we are planning on a sample of 600 adults, we anticipate that our analysis will be quite robust.

### Statistical analyses

Summary statistics describing baseline parameters of all study participants will be presented as means (SD) or frequencies for continuous and categorical variables, respectively. Differences in baseline parameters between control and intervention groups will be compared using the Student  $t$  test or  $\chi^2$  test for continuous and categorical variables.

### Children's intervention

To evaluate changes in KAH after intervention, analyses will be conducted in those children who have data for the primary outcome (overall score) at baseline and at 4 months. Mixed linear models that take account of the cluster-randomized design will be used to test for intervention effects. Fixed effects in each model include the corresponding baseline score, the class year, and the treatment group. Schools will be entered as random effects. We will also look for interaction by school and within that by classroom to tease out specific confounding by quality of instruction. Similar mixed models will be generated for the subcomponents of the primary end point in children and also for the secondary outcome variables in parents, teachers, and the school environment. Additional secondary comparisons will include changes in the proportion of eutrophic children between groups at the end of the study intervention.

### Adult intervention

To evaluate the results of the adult intervention program, the primary end point of the adult study is the between-group difference for the change in the BEWAT score at 12 months (baseline to end of active intervention). The Fuster-BEWAT score includes several dimensions, including blood pressure, exercise, weight, alimentation, and tobacco use. The prevalence of overweight, obesity, hypertension, smoking, and physical inactivity will be calculated in both groups and compared using the  $\chi^2$  test. In all individuals, we will calculate the value of the total score from the 5 dimensions of the same, and calculate its mean, SD, median, and range. To account for the cluster-randomized design, mixed-effects logistic regression will be used to test for intervention effects. The dependent outcome in these models will be the proportion of individuals with any improvement in overall score from baseline. Fixed effects in each model include the corresponding baseline score and the

treatment group. Schools will be entered as random effects. Similarly, linear mixed models will be generated for the score expressed as a continuous measure and for the different dimensions of the total score. To evaluate the impact of subclinical atherosclerosis measured using 3-dimensional carotid or ileofemoral ultrasound on the primary end point, all analyses will be repeated after stratifying by the presence or absence of any subclinical atherosclerosis. A formal test of interaction between any subclinical atherosclerosis and treatment allocation on the primary end point will be performed. In addition, presence of subclinical atherosclerosis will be entered as an additional covariate in the mixed-effects models to assess significance after accounting for the cluster-randomized design.

For all outcomes, the first analysis will compare the aggregate intervention group vs the control. Next-level analysis will evaluate pairwise analyses.

To evaluate intergenerational associations, in response to lifestyle interventions, we will perform additional analyses that take into account changes in the primary end point metrics for both children (KAH scores) and adults (Fuster-BEWAT score). First, we will examine the strength of correlation between change in adult and change in children's scores over time using the Pearson correlation coefficient. Using linear regression, we will also estimate the amount of change in children's scores that may be explained by change in adult scores. Models will be adjusted for baseline scores with school entered as a random effect. To further assess the degree to which change in the children's scores depend on the presence or absence of adult intervention, we will compare the magnitude of change in the KAH scores among those children whose parents received aggressive intervention and those children whose parents were allocated to control. This comparison is possible because all children eventually receive an intervention, including those children whose parents remain in the control group. Sensitivity analyses will be performed taking into account if the parent/caregiver lives in the same household as the child. Using linear regression, we will also examine whether or not adult treatment allocation (intervention vs control) is an independent correlate of change in children KAH scores. Models will be adjusted for baseline KAH scores and school will be included as a random effect. Differences in scores between additional time points within and between groups will be calculated in an analogous fashion. All analyses will be performed on an intention-to-treat basis using Stata version 12.1 (College Station, TX).

## Discussion

The FAMILIA studies target parents/caregivers and preschool-aged children, to demonstrate the impact of aggressive family-based lifestyle intervention on behavioral risk factors. The studies will optimally leverage existing infrastructure by partnering with the federally funded Head Start Program for access to low-income, underserved, at-risk families in Harlem, NY, and use existing materials from the Sesame Workshop used successfully in previous research programs in Colombia and Spain tailored to the proposed study population. If successful, the FAMILIA studies could potentially impact a generation of children and their caregivers to build healthy communities. We believe that the studies will serve as a model for evaluating the impact of early childhood, family-based interventions on long-term risk for the development of CVD and CVD risk factors, throughout the United

States and other parts of the world. In addition, we will leverage powerful genetic and molecular tools to study the environmental vs genetic basis of the mechanisms of improvement in CVD-related health arising from these interventions and other aspects of CVD.

In designing the proposed studies, we channeled our previous successes with the kids program in Colombia and Spain, and we attempted to address many of the shortcomings of studies conducted before FAMILIA. Upon review of studies conducted in children and adolescents, we realized that solely attempting to study biological outcomes led to “negative” studies. In keeping with recent consensus within the research and clinical community, we have included pragmatic intermediate objectives related to changes in knowledge, attitude, dietary patterns, or levels of physical activity as early “surrogate” measures that improve cardiovascular health. Studies that failed at achieving these objectives demonstrated some common themes, which we attempted to avoid in FAMILIA, by design. These included programs that were conducted with older children (8 years and older), those developed in response to a specific perceived crisis, participation of schools with low levels of coordination and preparation, programs based on external speakers and resources with little involvement from school staff, and little or no investment in training teachers and providing resources for support.

A major strength of our study is the conduct of parallel randomized trials for the parents or caregivers overlapping with the children’s study. Among others, we relied on the sweeping review of health disparities literature by Peek and colleagues<sup>18</sup> which concluded that “smart,” “culturally tailored” interventions that target participants through culturally/linguistically tailored interventions, peer support, and 1-on-1 interventions, technology-based decision support reminders/prompts, in-person feedback, and problem-based learning had the best odds of achieving meaningful improvement in health care outcomes in minority communities.

In designing the genomics component of this study, we leveraged our recent findings regarding the molecular interactome structure that is reflected in genotypic and RNA sequence data to uncover biologically meaningful gene modules/networks involved in the development of complex disease.<sup>19</sup> Targeting such causal networks to restore them to a normal state has been proposed as a path to treat disease,<sup>20, 21</sup> but this potential is only just starting to be realized. The FAMILIA study will be a big step in the direction of tailored behavioral genomics and nonpharmacologic approaches to precision and preventive medicine.

## Conclusion

The FAMILIA studies will aim to evaluate a family-based approach targeting preschool-aged (3–5 years) children and their parents or caregivers through a combination of a structured educational and behavior programs. This may prove to be a critical strategy in promoting cardiovascular health with positive and long-lasting effects.

## Acknowledgements

The study investigators would like to thank Claire Kofler, Julie Lambert, Carles Peyra, Jessica Bustamante, Peter Warburton, and Inna Gurewitz for their administrative and operational support in the conduct of these studies. Special acknowledgement is due to Rita Prats, Dianne Spann, and Andrea Hayes, school directors of our Vanguard Head Start sites in Harlem, for their contributions.

Clinical Trial Registration: NCT02343341 and NCT02481401.

The FAMILIA study is funded by the American Heart Association (14SFRN20490315). R.V. is supported by the Fogarty International Center of the National Institutes of Health (K01TW009218). J.C.K. is supported by the National Heart, Lung, and Blood Institute (K08HL111330), a Foundation Leducq Transatlantic Network of Excellence Award and AstraZeneca. V.F. is supported by the National Heart, Lung, and Blood Institute (5U01HL114200). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health or the American Heart Association.

## References

1. Lloyd-Jones DM, Hong Y, Labarthe D, et al. Defining and setting national goals for cardiovascular health promotion and disease reduction: the American Heart Association's strategic Impact Goal through 2020 and beyond. *Circulation* 2010;121:586–613. [PubMed: 20089546]
2. Mensah GA, Mokdad AH, Ford ES, et al. State of disparities in cardiovascular health in the United States. *Circulation* 2005;111: 1233–41. [PubMed: 15769763]
3. Bonow RO, Grant AO, Jacobs AK. The cardiovascular state of the union: confronting healthcare disparities. *Circulation* 2005;111:1205–7. [PubMed: 15769758]
4. Winkleby MA, Kraemer HC, Ahn DK, et al. Ethnic and socioeconomic differences in cardiovascular disease risk factors: findings for women from the Third National Health and Nutrition Examination Survey, 1988–1994. *JAMA* 1998;280:356–62. [PubMed: 9686553]
5. Matte TEJ, Bedell J, Selenic D, et al. Obesity in the South Bronx: a look across generations. New York City: Department of Health and Mental Hygiene 2007.
6. Ogden CL, Fryar CD, Carroll MD, et al. Mean body weight, height, and body mass index, United States 1960–2002. *Adv Data* 2004:1–17.
7. (NHANES) Nhanes. [http://www.cdc.gov/nchs/about/major/nhanes/nhanes2003-2004/nhanes03\\_04.htm](http://www.cdc.gov/nchs/about/major/nhanes/nhanes2003-2004/nhanes03_04.htm). National Center for Health Statistics, CDC 2004.
8. Obesity prevalence among low-income, preschool-aged children— New York City and Los Angeles County, 2003–2011 *MMWR Morb Mortal Wkly Rep* 2013;62:17–22. [PubMed: 23325351]
9. Hindorff LA, Sethupathy P, Junkins HA, et al. Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proc Natl Acad Sci U S A* 2009;106:9362–7. [PubMed: 19474294]
10. Penalvo JL, Santos-Beneit G, Sotos-Prieto M, et al. The SI! Program for cardiovascular health promotion in early childhood: a cluster-randomized trial. *J Am Coll Cardiol* 2015;66:1525–34. [PubMed: 26429075]
11. Rollnick S, Butler CC, Kinnersley P, et al. Motivational interviewing. *BMJ* 2010;340:c1900. [PubMed: 20423957]
12. Heath GW, Parra DC, Sarmiento OL, et al. Evidence-based intervention in physical activity: lessons from around the world. *Lancet* 2012;380(9838):272–81. [PubMed: 22818939]
13. Emilsson V, Thorleifsson G, Zhang B, et al. Genetics of gene expression and its effect on disease. *Nature* 2008;452:423–8. [PubMed: 18344981]
14. Zhong H, Yang X, Kaplan LM, et al. Integrating pathway analysis and genetics of gene expression for genome-wide association studies. *Am J Hum Genet* 2010;86:581–91. [PubMed: 20346437]
15. Clarke L, Zheng-Bradley X, Smith R, et al. The 1000 Genomes Project: data management and community access. *Nat Methods* 2012;9: 459–62. [PubMed: 22543379]
16. Consortium CAD, Deloukas P, Kanoni S, et al. Large-scale association analysis identifies new risk loci for coronary artery disease. *Nat Genet* 2013;45:25–33. [PubMed: 23202125]

17. Zhong H, Beaulaurier J, Lum PY, et al. Liver and adipose expression associated SNPs are enriched for association to type 2 diabetes. *PLoS Genet* 2010;6:e1000932. [PubMed: 20463879]
18. Peek ME, Cargill A, Huang ES. Diabetes health disparities: a systematic review of health care interventions. *Med Care Res Rev* 2007;64(5 Suppl):101S–56S. [PubMed: 17881626]
19. Chen Y, Zhu J, Lum PY, et al. Variations in DNA elucidate molecular networks that cause disease. *Nature* 2008;452:429–35. [PubMed: 18344982]
20. Schadt EE, Björkegren LJ. NEW (Network-Enabled Wisdom) in Biology, Medicine and Healthcare. *Science Translational Medicine* 2012; 2012 [Planned publication January 2012].
21. Schadt EE. Molecular networks as sensors and drivers of common human diseases. *Nature* 2009;461:218–23. [PubMed: 19741703]

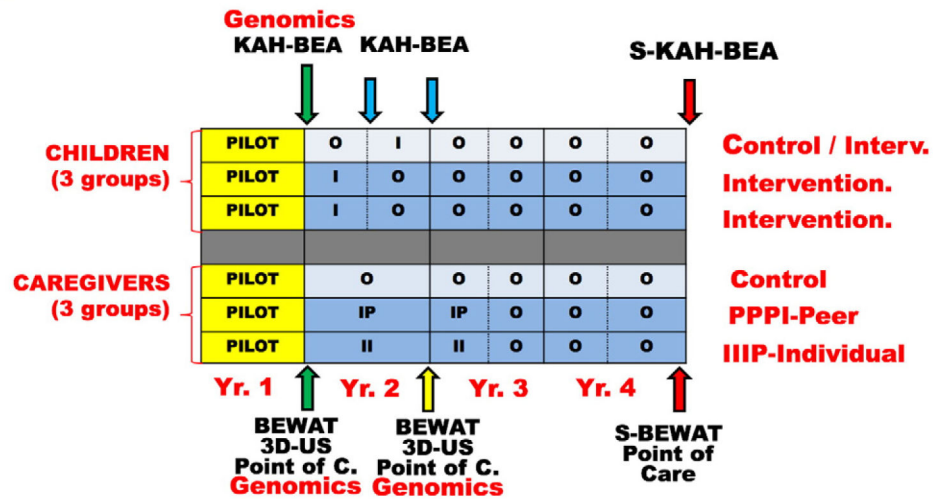


>Figure 1.  
Study schools.





# Summary Of The Proposal



**KAH-BEA = Knowledge, Attitude, Habit, -BMI, Exercise, Alimentation ;**  
**BEWAT = BP, Exercise, Weight, Alimentation, Tobacco; S=Sustained**

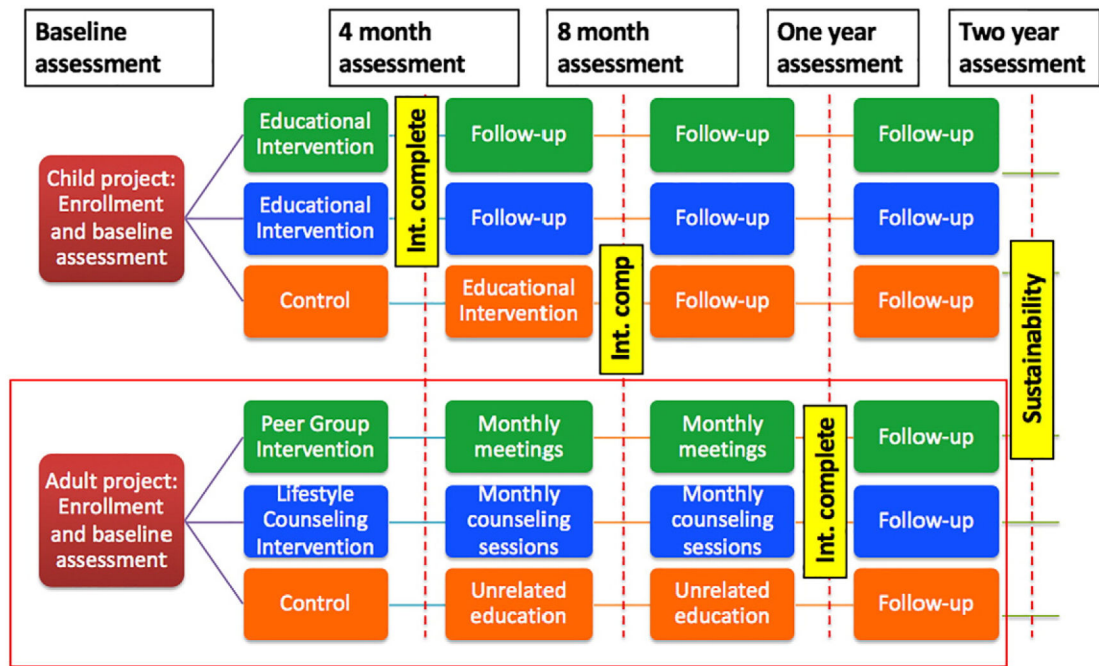
Figure 2. Proposal summary.

Author Manuscript

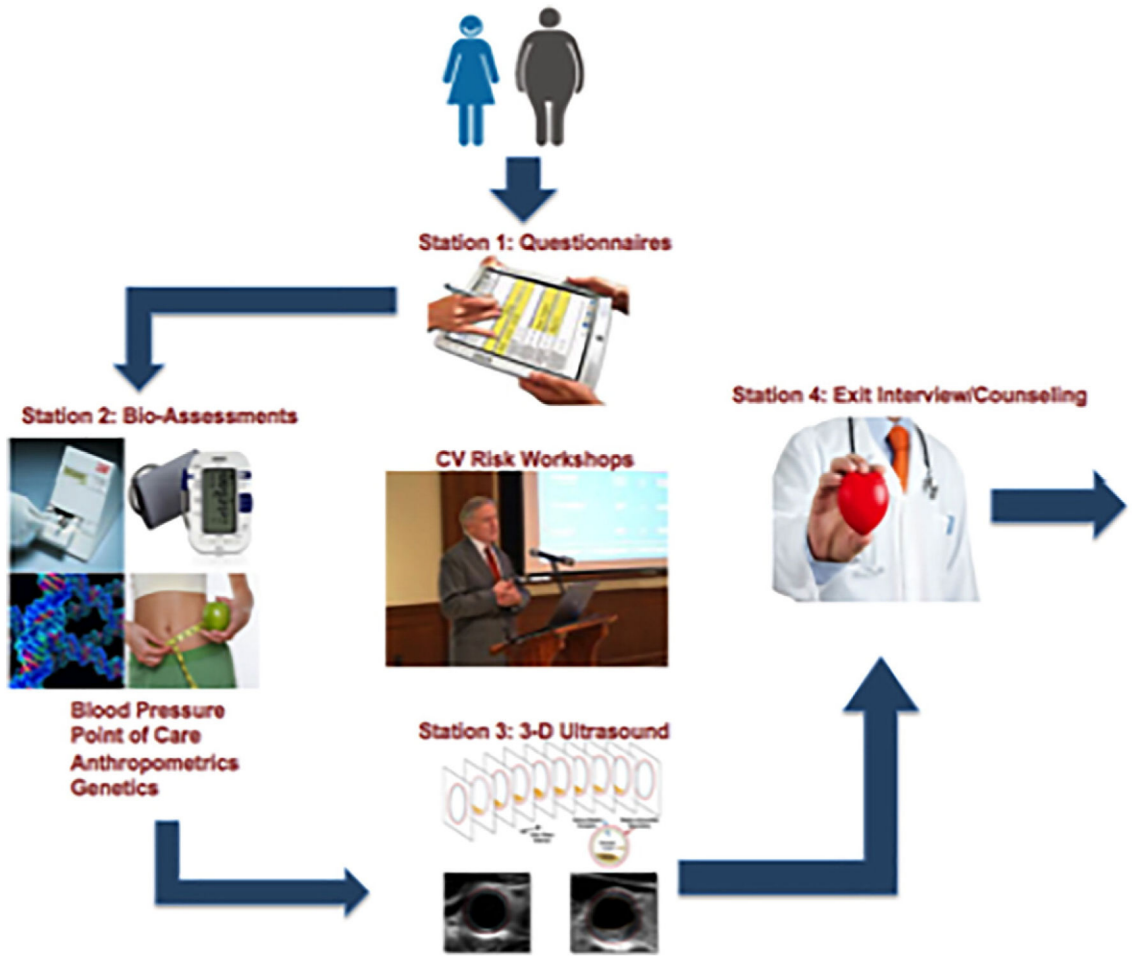
Author Manuscript

Author Manuscript

Author Manuscript



**Figure 3.**  
Study design.



**Figure 4.**  
Structure of assessments.

**Table I.**

Curricular objectives

Domain	Goals	
Diet	Understanding and learning to value health issues regarding a healthy diet, the health characteristics of foods, and their origins, traditions, and attitudes to food hygiene	Observing and exploring the family; natural and social surroundings
Body and heart awareness	Understanding and valuing health issues in the care of one's own body and welfare according to one's own possibilities and limitations	Developing increasingly more independent habits of care and body hygiene and respecting differences among peers
Physical activity Beginning the	Beginning the use of logical/mathematical and reading/writing skills, together with movement, gesture, and rhythm	Practicing, identifying, and developing an interest in cardiovascular physical education and sport as a means of promoting a healthy life in accordance with the values of fair play
Emotion management	Developing emotional comprehension and capabilities	Developing communication skills through word development and forms of expression
	Forming relationships to other people and gradually developing basic tools for coexistence and social relations, as well as developing ways of peacefully resolving conflict	communicative possibilities through different languages and responding to feelings of affection, respecting diversity, and developing helpful and collaborative attitudes

**Table II.**

Si! Program

Level of intervention	Objectives	Intervention activities	Intervention materials	Minimum hours
Children	Acquiring KAH regarding the components of FAMILIA	- Classroom instruction	(3–5 y)—5 didactic units 4 mo. Including 4 key activities per unit, and associated resources. Sesame Street audiovisuals, tales on healthy living, and Family Newsletters	37
Parents	Improving KAH regarding the components FAMILIA	- FAMILIA educational and informational meetings - Family-based (child and parent/caregiver health activities	- Family activities brochures, and videos, - FAMILIA Web site	12