Project Viva:

A longitudinal study of health for the next generation

HSC Protocol

07.13.22

Table of Contents

I. Introduction

Project Viva is a prospective cohort study of maternal and child health. In 1999-2002 we recruited 2,670 pregnant women during their first trimester of pregnancy from eight obstetric (OB) offices of a multi-site group practice in eastern Massachusetts, Harvard Vanguard Medical Associates. There were 2128 live singleton births to 2,100 women (28 women enrolled with successive pregnancies) and approximately 1,500 of our original child participants are still involved in the study. We collect data repeatedly from multiple sources, including questionnaires, interviews, medical records, examinations, and biospecimen samples.

Project Viva intends to follow our child participants as long as there is grant funding and interest from the participants. Some of the most beneficial health findings come from long-term follow-up. Project Viva's general study objectives for our child cohort (now Young Adults) are outlined in this protocol. Project Viva data is used in several studies (separate protocols), as well as ancillary studies that fall under our regular data repository, genetic data repository, or epigenetic data repository. This protocol addresses the primary data collection for the child participants originally enrolled in Project Viva, and the main focus areas of research among these now-adults.

II. Project Viva Investigators

Project Viva is reviewed by Harvard Pilgrim Health Care's Institutional Review Board and is led by investigators at the Division of Chronic Disease Research Across the Lifecourse (CoRAL), Department of Population Medicine, Harvard Pilgrim Health Care Institute (HPHCI) and Harvard Medical School (HMS). All Project Viva staff are employees of HPHCI.

A Viva Co-Investigator (Co-I) is anyone listed as a PI or Co-I on the NIH grants that support the majority of Project Viva operations, or the PI of one of the other grants that support Viva science. In addition to being a PI or Co-I, one must also be actively involved with the Co-I meetings and operations. Most co-Investigators are approved under other IRB-approved protocols that contain specific scientific aims but may not support operations.

III. Project Viva Historical Recruitment

This section outlines how and when Project Viva recruited pregnant women and the eligibility criteria.

Project Viva Research Assistants approached pregnant women immediately following their initial prenatal visit (IOB) at one of eight HVMA obstetric offices to determine eligibility. Project Viva recruited women from 1999 – 2002. If a woman was eligible, we asked her to enroll in the study by providing informed consent and completing the early pregnancy visit (V1). The 8 HVMA locations included Kenmore, Copley, Cambridge, Post Office Square, Quincy, Wellesley, Medford, and West Roxbury.

Women were eligible if they met the following criteria:

- Less than 22 weeks pregnant at the time of enrollment, as defined by due date or last menstrual period (LMP), if due date was not available. Gestational age for the inclusion criteria was based on a patient's reported due date. The Research Assistant calculated gestational age from the expected due date using a

pregnancy wheel. If the due date was not known, the Research Assistant calculated gestational age based on LMP. The women had to know either her due date or LMP to be eligible to participate.
- Receive prenatal care at one of the selected HVMA practices.
- Plan on delivering at Brigham and Women's Hospital (BWH) or Beth Israel Deaconess Medical Center (BIDMC).
- Be able to answer questionnaires in English.

Women were ineligible if they met any of the following criteria:
- Planned to terminate the pregnancy.
- Planned to move away from the local area before the end of the initial follow-up period, 6 months after delivery.
- Had multiple gestation (twins, triplets, etc.) since they are likely not be comparable to other women, and the limited number would preclude separate analyses.

IV. Study Objectives

This section outlines the study objectives. This includes outcome focus areas of interest and applicable exposures.

The primary objectives of Project Viva's research with its child cohort fall into five main focus areas.

Focus Area 1: Child Pregnancy and Birth Outcomes

Project Viva is interested in looking at childbirth outcomes, such as fetal growth, birth weight, length and head circumference, and gestational age.

Focus Area 2: Child and Adolescent Obesity and Cardio-Metabolic Outcomes

Project Viva is interested in studying growth, obesity, the development of cardio-metabolic conditions in infancy, childhood and adolescence. Viva has collected extensive data on this topic at multiple time points, including anthropometric measures, blood pressure, cardiovascular fitness, biomarkers of cardiometabolic risk, measures of diet, sleep, and physical activity.

Focus Area 3: Child and Adolescent Asthma and Allergies

Development of asthma and allergies is another focus area. Project Viva has extensive multiple time point data on symptoms and doctor diagnosis of asthma and allergic conditions, biomarkers, lung function and inflammation measurements.

Focus Area 4: Child and Adolescent Development

Project Viva is also interested in learning about child and adolescent development. Project Viva assesses cognition and behavior at each visit, in addition to well-being and other general health questions.

Focus Area 5: Epigenetics and Genetics

Project Viva has collected multiple blood samples from mothers and children, including umbilical cord blood. Project Viva is primarily interested in how genetics plays a role in the development of high blood pressure, asthma, growth, cognitive development, and length of pregnancy, and how these associations are modified by other

lifestyle factors. Project Viva is also interested in epigenetics as both an exposure and an outcome, including DNA methylation, and telomere length.

Project Viva conducts genetic and epigenetic testing only on individuals who have given specific genetics consent for such analyses. Beginning at the "Early Teen" (AG12) visit, consent forms will include language describing the use of biospecimens for epigenetic research. Epigenetic analysis may be performed on biospecimen samples provided by participants at previous visits, provided that the mother signed the updated consent forms containing epigenetics language. If a mother did not sign the updated consent form for either herself or her child, we only perform epigenetic analysis on the mother's biospecimen samples if she previously signed a genetics consent form, and/or on the child's biospecimen samples if the mother previously signed a genetics consent form for her child.

Exposures

In relation to the above outcomes, Project Viva studies a number of maternal, paternal, child, household, and environmental exposures from the pre-pregnancy period through adolescence, that include but are not limited to: body mass index (BMI) and other anthropometric measures; maternal and child/adolescent blood pressure; sleep and physical activity; DNA methylation, epigenetics and genetics; gestational weight gain and post-partum weight loss; post-partum depression; time to pregnancy as measure of fertility; maternal stress, racism and violence; child/adolescent bullying; parenting; sociodemographic variables; breastfeeding and lactation; birth outcomes; maternal pregnancy complications such as gestational diabetes and preeclampsia; blood, urine and hair assays; maternal and child/adolescent diet and eating behaviors; geographic information system (GIS) variables, including distance to roadways and highways and census variables; asthma and allergies; and cognition and behavior. Most of these variables can be exposures, outcomes, or covariates depending on the specific analysis. The lead investigator for each analysis specifies exposures, outcomes, and covariates in an analysis plan.

V. Historical Consent Forms

This section describes all Project Viva consent forms used with our child cohort, including consent forms for main visits, sub-studies, visit addendums, and waivers of consent approvals.

A. Consent forms for enrollment through 6 months old

This section describes all consent forms dealing with data collected between enrollment (the child's birth) and 6 months old. This includes maternal consent for child participation.

i. Original Maternal Consent Covering Visit 1 (1st Trimester) to Visit 4 (6 months)

Project Viva's original consent form for mothers covers maternal data collection for the first four visits. There are several versions of the consent form which reflect changes occurring in the study and from discussions with the IRB. These include adding language about what the blood would be used for, adding the Certificate of Confidentiality, combining data with medical records and adding visit components to the delivery and 6-month visits. The consent provisions are as follows.

I. Purpose: To examine the roles of diet and other factors in maintaining the health of pregnant women and their babies.
II. Early Pregnancy In-person Visit Components (<22 weeks gestation)
    a. Questionnaires
    b. In-Person Interviews
    c. Blood draw at the lab (extra tube drawn at the lab, not a separate draw)
III. Mid-Pregnancy In-person Visit Components (26-28 weeks gestation)

a. In-person interview
b. Questionnaires
c. Blood draw at the lab (extra tube drawn at the lab, not a separate draw)

IV. Delivery In-person Visit
a. In-person interview one to three days post-partum
b. For women delivering at Brigham and Women's, hospital staff to measure the weight of the placenta and collect a sample of umbilical cord blood immediately following delivery in the delivery suite

V. Six-month In-person Visit
a. Questionnaire

VI. HVMA and delivery hospital medical records regarding pregnancy and delivery, including items such as results of laboratory tests, medical procedures, prescription dispensing, medical & reproductive history, and labor & delivery details

VII. Addendums for maternal blood pressure and weight measurements

ii. Maternal Hair Substudy

I. Purpose: To look at the relationship between maternal mercury levels and child growth and development.

II. Participant provided written informed consent for Project Viva to collect a sample of her hair at delivery

b. Child Data Collection

i. Maternal Consent for Child Participation, covering V3 (delivery) & V4 (6 months) for child data

I. Purpose: To examine the effects of diet and other factors during pregnancy on the health of pregnant women and their infants.

II. Post-partum In-person Visit Components
a. Length and circumference measurements
b. Blood pressure

III. Six-month In-person Visit Components
a. VRM visual test and VAT visual test
b. Blood pressure
c. Length and weight measurements
d. Questionnaire about child's diet, environment and habits

IV. HVMA and hospital delivery medical record

V. Addendums for visual tests, blood pressure, and length and weight measurements

ii. Thyroid Function Substudy

I. Purpose: To help understand the relationship of newborn thyroid function with mother's diet, thyroid function, and child development.

II. Participant provided written informed consent (by mail) for Project Viva to obtain newborn thyroid screening results from Mass DPH

B. Consent forms for data collection at the "Early Childhood" visit (~3 years old)

This section describes all consent forms dealing with data collected at 3 years post-partum. This includes maternal consent for child participation.

i. Maternal Consent for Child Participation at Age 3

I. Purpose: To examine the effects of diet and other factors during pregnancy and infancy on child development and health.
II. Visit Components:
    a. Child development
    b. Blood pressure and anthropometric measures
    c. Blood draw
        i. If the child was an HVMA patient, we provided lead and CBC results to the HVMA pediatrician for the child's medical records so an additional needlestick was not needed. This is important to note for IRB reasons related to adding research collected data to clinical medical records.
III. Outpatient and hospital medical records
IV. Age 3 Addendum Components:
    a. Anthropometric measurements
    b. Visual Motor Ability Test
    c. Interview about child's health, environment and development

ii. Maternal Consent for Child Participation in the Immune Substudy at Age 3

I. Purpose: To examine the effects of diet and other factors during pregnancy and infancy on child development and health.
II. Only participants with a cord blood sample were eligible
III. Visit Components:
    a. Child height and weight measurement
    b. Vacuum part of the child's bedroom floor and bed for dust sample
    c. Child blood draw to look at lead level, complete blood counts, and response to allergens
    d. Child medical record review

iii. Maternal Consent for Child Blood Draw at 3 years post-partum

I. Purpose: To examine the effects of diet and other factors during pregnancy and infancy on child development and health.
II. Participant provided written informed consent for a blood draw to look at lead level and complete blood count (CBC).
III. For HVMA patients, results will be entered into the child's medical record and viewed by the child's pediatrician.

C. Consent forms for data collection at the "Mid-Childhood" visit (~7 years old)

This section describes all consent forms dealing with data collected at 7 years post-partum. This includes maternal consent for child participation.

i. Maternal Consent for Child Participation at Age 7

I. Purpose: The purpose of Project Viva is to examine the effects of diet and other factors during pregnancy and infancy on child development and health.

II. Visit Components:

a. General questionnaire, and questionnaires about child's behavior, development and home environment completed by the mother
b. Family medical history interview with the mother
c. Contact information for teacher to send him/her two questionnaires
d. Hospital medical records and insurance claims from birth through the visit date
e. Anthropometric measures and blood pressure
f. Intelligence, memory and motor abilities tests
g. Step test
h. Breathing test
i. DXA scan

ii. Maternal Consent for Child Biospecimen Participation at Age 7

I. Purpose: The purpose of Project Viva is to examine the effects of diet and other factors during pregnancy and infancy on child development and health.

II. Visit Components:

a. Blood collection
b. Urine collection
c. Hair collection

D. Consent and authorization forms for data collection for the "Early Teen" visit

This section describes all consent forms dealing with data collected at the "Early Teen" visit (AG12). This includes maternal consent and child assent for child participation.

i. Maternal and Child Consent and Authorization for Child Participation at "Early Teen" Visit (AG12)

I. Purpose: The purpose of Project Viva is to examine the effects of factors during pregnancy, infancy on childhood on maternal and child health.

II. Visit Components:

a. Early Teen questionnaire – for Parents completed by the mother
b. Family medical history interview with the mother
c. Anthropometric measures and blood pressure
d. Dual-energy x-ray absorptiometry (DXA)
e. Biospecimen collection of nasal swabs, urine and hair
f. Step test

g. Early Teen Questionnaire
h. Breathing tests: eNO and Spirometry
i. Actigraphy: physical activity and sleep monitor
j. Outpatient and hospital medical records and insurance claims from birth through the end of the study
k. Blood collection by trained phlebotomist

E. Consent and authorization forms for data collection for the "Mid-Teen" visit

This section describes all consent forms dealing with data collected at the "Mid-Teen" visit. This includes maternal consent and teen assent for teen participation.

i. Maternal and Teen Consent and Authorization for Teen Participation at "Mid-Teen" Visit (AG17)

I. Purpose: The purpose of Project Viva is to examine the associations of factors during pregnancy, infancy, childhood and adolescence with maternal, child, and adolescent health.
II. Visit Components:
   a. Blood collection by trained phlebotomist
   b. Biospecimen collection of urine, hair, and shed baby teeth
   c. Anthropometric measures and blood pressure
   d. Dual-energy x-ray absorptiometry (DXA) – total body
   e. Step test (*removed from consent/operations during COVID)
   f. Breathing tests: eNO and Spirometry (*removed from consent/operations during COVID)
   g. Mid-Teen Questionnaire – for Teens
   h. Internet-based, self-administered dietary recall (1 completed at in-person visit; 2 completed at home)
   i. 7-day sleep and physical activity measurement using the Fitbit Charge 3
   j. Outpatient and hospital medical records and insurance claims from birth through the end of the study
   k. Genetic and epigenetic analysis of biospecimen samples collected at Mid-Teen Visit and previous in-person visits
III. Additional consent opportunity presented:
   a. Data sharing with the ECHO Program: sharing participant data (including genetic and epigenetic data) with the ECHO program, managed by the US National Institutes of Health (NIH).
   b. Collect an additional sample of blood (~1 tablespoon) specifically for the ECHO Program

ii. 18+ Teen Consent and Authorization for 18+ Teen Participation at "Mid-Teen" Visit (AG17)

I. Purpose: The purpose of Project Viva is to examine the associations of factors during pregnancy, infancy, childhood and adolescence with maternal, child, and adolescent health.
II. Visit Components:
   a. Blood collection by trained phlebotomist

b. Biospecimen collection of urine, hair, and shed baby teeth
c. Anthropometric measures and blood pressure
d. Dual-energy x-ray absorptiometry (DXA) – total body
e. Step test (*removed from consent/operations during COVID)
f. Breathing tests: eNO and Spirometry (*removed from consent/operations during COVID)
g. Mid-Teen Questionnaire – for Teens
h. Internet-based, self-administered dietary recall (1 completed at in-person visit; 2 completed at home)
i. 7-day sleep and physical activity measurement using the Fitbit Charge 3
j. Outpatient and hospital medical records and insurance claims from birth through the end of the study
k. Genetic and epigenetic analysis of biospecimen samples collected at Mid-Teen Visit and previous in-person visits

III. Mid-Teen 18+ Teen Addendum Components:
a. Data and biospecimen sharing with the ECHO Program
b. Collect an additional sample of blood (~1 tablespoon) specifically for the ECHO Program

F. Consent form for data collection for the Age 19 survey

This section describes the consent form presented to young adult participants dealing with data collected through the Age 19 survey.

Consent and Authorization Form at Age 19

Purpose: The purpose of Project Viva and ECHO is to examine the associations of factors during pregnancy, infancy, childhood, and adolescence with maternal, child and adolescent health.

Visit Components:

a. Age 19 Questionnaire
b. Genetic and epigenetic analysis of biospecimen samples collected at Mid-Teen Visit and previous in-person visits for the ECHO program
c. Data and biospecimen sharing with the ECHO Program

G. Genetics Study Consent for Children

This section describes when genetics consent forms were presented to participants, the different versions of the consent form, and what is included in the consent forms. **Genetic analyses will remain to be conducted ONLY on participants who specifically provide this consent.**

Project Viva approached mothers for genetics consent at the initial Early Pregnancy visit (V1), the Early Childhood visit (V7), and the Mid-Childhood Visit (~Age 7). The consent forms were only presented at Early and Mid-Childhood visits if a participant choose an undecided option, or no consent form was on file. At the Mid-Teen visit, genetic and epigenetic consent were included on the main consent form, which was presented to all participants who attended an in-person or remote visit.

If participants refused to participate in genetic and/or epigenetic analyses, no future genetic or epigenetic analyses will be run using those participants' samples.

a. Original Genetics Consent Form

Project Viva's original genetics consent form was presented only to women delivering at Brigham and Women's Hospital. This consent form covers both maternal blood and child blood on a single consent form.

Consent form options:

I. Project Viva can use Genetic material from child for future studies. The mother chooses yes or no for each of the following domains:
    a. High blood pressure
    b. Asthma
    c. Growth
    d. Length (duration) of pregnancy (for mothers only)
    e. Other medical conditions identified in the future
II. Project Viva may store, but not use, genetic material from child and can contact the participant in the future as projects arise.
III. Project Viva may not use or store genetic material from child

b. Child Genetics Consent Form

If the original genetics consent form was not completed or option 2 (undecided) was selected, Project Viva Research Assistants administered the child genetics consent form at the ages 3 and/or 7 in-person visits.

Consent form options:

I. Project Viva can use Genetic material from mother and/or child for future studies. The mother chooses yes or no for each of the following domains:
    a. High blood pressure
    b. Asthma
    c. Growth
    d. Other medical conditions identified in the future
II. Project Viva may store, but not use, genetic material from child and can contact the participant in the future as projects arise.
III. Project Viva may not use or store genetic material from child

c. Child Genetics & Epigenetics Ancillary Study Consent Form

Project Viva Research Assistants administered the child genetics & epigenetics consent form at the Early Teen in-person visits.

Consent form options:

I. Epigenetic Study
    a. Biosamples may be used by Project Viva for epigenetic analysis.
    b. Please do not use biosamples for epigenetic analysis at this time.
II. Genetic Study

a. Project Viva can use Genetic material from child for future studies. The mother chooses yes or no for each of the following domains:
    i. High blood pressure
    ii. Asthma
    iii. Growth
    iv. Other medical conditions identified in the future

VI. Data Collection

A. Primary Data Collection

This section outlines Project Viva's primary data collection and provides a description of each visit performed with our child cohort members. In addition to the primary data detailed below, we update contact information at every visit, including alternate contacts and name changes. Appendix B includes a visit flowsheet that offers a graphical depiction of Project Viva's visits.

A1: Completed Visits

Project Viva has completed data collection on all visits through the Mid Teen visit. Each visit is described in detail below. The only foreseeable human subjects risk for completed visits is the risk of a privacy or confidentiality breach.

a. Delivery (V3), in-person

There were 2128 births to 2,100 unique women, of the initial 2,670 pregnancies. We conducted post-delivery follow-up at two hospitals: Beth Israel Deaconess Medical Center and Brigham and Women's Hospital, the receiving hospitals for all deliveries from the HVMA recruitment practices. Each HVMA practice delivered all of their patients at one of the two hospitals. Quincy, Wellesley, Kenmore, West Roxbury and Post Office Square patients delivered at the Brigham and Women's Hospital; Copley, Medford and Cambridge patients delivered at the Beth Israel Deaconess Medical Center. Visits took place 1-3 days after delivery on the post-partum maternity floor, ideally before visiting hours started, and lasted about 30 minutes. Cord blood collection took place in the delivery suite at the time of delivery. We completed 2072 visits.

The child visit components included:

1. Mother's Interview: delivery, child feeding, maternal diet and smoking
2. Supplement (self-administered): growth concerns
3. Child anthropometry (head, chest and abdominal circumferences, and length)
4. Child blood pressure
5. Venous umbilical cord blood collection (Brigham and Women's Hospital only, done in delivery suite at the time of delivery)
    a. Assays completed to date include 16S bacterial microbiome sequencing, 5-OH-mC, Adiponectin, Allele-specific DNA methylation of Igf2 and H19, Bile acids, C peptide, Cortisol/Cortisone, Cytokines (IL6, L10, IL13, INF, TNF), Fatty acids, Genome-wide DNA methylation, hsCRP, hsTnT, IGF-1, IGF-2, IGFBP-3, Insulin, Leptin, LINE1, Lymphocyte proliferation, Metabolomics, Mercury, mtDNA, PGLYPR-1, Pro-BNP, sIL6Rα, sTNFR-II, Vitamin D, TSH

b. Infancy Visit (~Six-Months old (V4)), in-person

Staff completed the infancy visit in the pediatric office at one of the HVMA sites. Ideally visits took place between 5.5 and 8 months after delivery. Child age ranged from 4.9 to 10.6 months, with a mean of 6.4 months. The visit length was approximately one hour, and we completed 1697 visits.

The visit components for the child included:

1. Mother's Interview: child development and child diet
2. Child length and blood pressure measurements
3. Child vision test (VRM and VAT)
4. General questionnaire (mother self-administered): child growth and development, child feeding, social and physical activities, home environment

c. 1 and 2 year (V5, V6), mail

Visit Description – For the Ages 1 and 2 mailed visits, Project Viva mailed participating mothers an annual questionnaire that included child data collection. If the participant was unwilling to complete the questionnaire via mail, Project Viva attempted administer the questionnaire over the phone.

1. For the 1-year visit, Project Viva received 1,256 completed mailed questionnaires and administered 47 questionnaires over the phone. In total, 1,303 Project Viva participants completed the 1-year visit.
2. For the 2-year visit, Project Viva received 1,288 completed mailed questionnaires and administered 127 questionnaires over the phone. In total, 1,415 Project Viva participants completed the 2-year visit.

d. Early Childhood (~3 year (V7)), in-person

This visit ideally took place at the Kenmore HVMA offices, or if that was not feasible, at the participant's home or other convenient location. Participants were eligible for this visit if the mother had completed at least one FFQ during pregnancy and had not disenrolled the child from follow-up. The visit length was approximately 1.5 hours. Children for this visit ranged in age from 2.8 years (2 years, 9 months) to 6.3 years (6 years, 4 months), with a mean of 3.5 years. Participants living too far away, and those unable or unwilling to meet with us in person had the option to complete a "mailed only" visit, which included the same self-administered questionnaires plus an RA-administered family health interview over the phone. We completed 1296 in-person visits and an additional 157 by mail.

The visit components for the child included:

1. General questionnaire completed by the mother about the child's behavior, health, and home environment
2. Child anthropometry measurements, including standing and sitting height, weight, waist and hip circumferences, mid-upper arm circumference, triceps and subscapular skinfolds
3. Child blood pressure
4. Child cognition: Peabody Picture Vocabulary Test, Wide Range Assessment of Visual Motor Abilities (WRAVMA)
5. Dust collection (for immune substudy only)
    a. Project Viva collected 294 dust samples which were analyzed for allergens.
6. Blood collection
    a. Project Viva collected 816 child blood samples at this visit. Project Viva phlebotomists conducted 601 of the blood draws, of which 139 were in the fasting state (>6 hours).
    b. Child blood assays completed to date include plasma leptin, adiponectin, lead, CBC, cytokines, fatty acids, miRNA, tocopherol isoforms (Vitamin E), eosinophil count, allergen-specific and total IgE, and, 25(OH)D; genome-wide DNA methylation, and lymphocyte proliferation.

e. 4, 5 and 6 year (V8 – V10), mail

Visit Description— For the Ages 4, 5 and 6 mailed visits Project Viva mailed participating mothers an annual questionnaire, that included child data collection.

1. For the 4-year visit, Project Viva received 1,244 completed mailed questionnaires.
2. For the 5-year visit, Project Viva received 873 completed mailed questionnaires.
3. For the 6-year visit, Project Viva received 965 completed mailed questionnaires.

f. Mid-Childhood (~7 year (Age7)), in-person

This visit ideally took place at the Kenmore HVMA offices, or if not feasible at the participant's home or other convenient location. All participants still enrolled in Project Viva were eligible for this visit. The visit length was approximately 3 hours, and children ranged in age from 6.6 to 10.9 years, with a mean of 8.0 years. Participants living too far away, and those unable or unwilling to meet with us in person had the option to complete a "mail–only" visit which included the same self-administered questionnaires and an RA-administered family health interview over the phone. We completed 1,116 in-person visits and an additional 163 by mail.

The visit components for the child included:

1. General questionnaire completed by the mother about the child's behavior, health, and home environment
2. Child Behavioral questionnaire completed by the mother and teacher: BRIEF and SDQ
3. Child anthropometric measurements, including height, weight, bioimpedance, waist and hip circumference, middle-upper arm circumference, and triceps and subscapular skinfolds
4. Child blood pressure
5. Child cognitive development: Kaufman Brief Intelligence Test (KBIT-2), Wide Range Assessment of Memory and Learning (WRAML), Wide Range Assessment of Visual Motor Abilities (WRAVMA)
6. Child step test
7. Child spirometry testing: pre- and post- bronchodilator
8. Child body composition: Dual-energy x-ray absorptiometry (DXA)
9. Child urine collection
    a. Project Viva collected 1,045 urine samples.
10. Child hair collection
    a. Project Viva collected 959 hair samples.
11. Child blood collection
    a. Project Viva collected 701 blood samples at this visit, 654 of which were fasting (>6hours).
    b. Child blood assays completed to date include:
        i. IGF-1, IGF-BP3, leptin, adiponectin, insulin and glucose (fasting), IL-6, CRP, sTNFR-II, bile acids, genome-wide DNA methylation, metabolomics, mtDNA, PFAS, serum ALT levels, sTNF-R2 fasting total cholesterol, HDL, triglycerides,25(OH)D, specific IgEs for common allergens and total IgE

i. 8, 9, 10 and 11 year (Age8, Age9, Age10, Age11), mail

For the age 8-, 9-, 10- and 11-year visits, Project Viva mailed or emailed the participating mothers an annual questionnaire that included child data collection. Project Viva implemented online questionnaire completion at Age 9. Mothers of child participants completed online questionnaires through a participant-specific site (protected by entering the child's date of birth) maintained by New England

Research Institute (NERI), our database administrator at the time. Data participants entered into the online version was housed and saved in the same manner as RA entered data. Online questionnaires were temporarily discontinued while Project Viva transitioned from NERI to its current database administrator, REDCap. A portion of Age 11 participants had the option to complete an Age 11 online survey through REDCap.

At ages 9, 10 and 11 mailed visits, Project Viva also mailed the child an annual questionnaire. Prior to mailing the questionnaire, the mother had the option to contact us and request we not mail a questionnaire directly to the child. Child participants did not have the option to complete an online questionnaire for these mailed visits.

- For the 8-year visit, Project Viva received 712 completed questionnaires from mothers about their children.
- For the 9-year visit, Project Viva received 927 completed questionnaires from mothers about their children, and 1057 questionnaires from the children themselves.
- For the 10-year visit, Project Viva received 797 completed questionnaires from mothers about their children, and 952 questionnaires from the children themselves.
- For the 11-year visit, Project Viva received 704 completed questionnaires from mothers about their children, and 818 questionnaires from the children themselves.

j. Early Teen (AG12) Visit, in-person

This visit ideally took place at the Kenmore HVMA/Fenway offices, but may have also been conducted at the participant's home or other convenient location if travel to Boston was prohibitive. All participants still enrolled in Project Viva at the start of the Early Teen in-person visit were eligible to participate. The visit length was approximately 2½ hours. Participants had the option of completing a "mailed visit" if they lived too far away, or were unable or unwilling to meet with us in person. The "mailed visit" consisted of the same self-administered questionnaires and an RA-administered family medical history interview. Participants had the option to complete the questionnaires electronically using the REDCap survey option. A unique URL was sent to study participants who wished to complete their questionnaires electronically using the HPHC domain. Starting at the Early Teen "mailed visit", child participants were also given the option to complete the questionnaire electronically. The mother was asked how she wanted her child to be sent the questionnaire before any email communication were initiated with the child. We completed 1,038 in-person visits and 773 blood draws. Our Early Teen Operations Manual served as an up-to-date data collection protocol for the Early Teen Visit.

The visit components for the child included:

1. General questionnaire completed by the mother about the child's behavior, health, and home environment
2. Child blood collection

a. Child blood assays completed to date include: Glucose, hsCRP, hsIL-6, sTNFR-II, HDL Cholesterol, Total Cholesterol, Triglycerides, C peptide, Insulin, IGF-1, IGFBP3, Leptin, Adiponectin, IgE, IgG4, and Metabolomics

3. Child urine collection
4. Child hair collection
5. Child nasal swab collection

   a. We collected 8 swabs for DNA. DNA extraction was used to identify epigenetic markers, specifically methylation. Use of biospecimen samples for epigenetic analyses is included in the genetics & epigenetics ancillary study consents, and we only analyzed nasal swab samples from participants who provided this consent. All child participants who had parental consent and signed their assent were approached to have their nasal swabs collected at the Early Teen visit.

6. Child anthropometric measurements, including height, weight, bioimpedance, waist and hip circumference, middle-upper arm circumference, and triceps and subscapular skinfolds
7. Child body composition: Dual-energy x-ray absorptiometry (DXA)
8. Child blood pressure
9. Child step test
10. Child Exhaled Nitric Oxide testing with NiOX MINO
11. Child spirometry testing: pre- and post- bronchodilator
12. Child 7-day wrist actigraphy and accelerometry and sleep journal
13. Early Teen child questionnaire

k. 14 and 15 year (Age 14, Age 15), mail

For the age 14- and age 15-year visits, Project Viva mailed or emailed the participating mothers an annual questionnaire that included child data collection. Participants had the option to complete electronic versions of the questionnaires through REDCap. A unique URL was sent to study participants who wished to complete their questionnaires electronically using the HPHC domain.

Project Viva also mailed child participants Age 14 and Age 15 annual questionnaires. Prior to mailing the questionnaires, the mother had the option to contact us and request we not mail the questionnaires directly to the child. The mother also had the option to include an email address for her child, so that the child can complete an electronic questionnaire if his or her mother so chooses.

- For the 14-year visit, Project Viva received 602 completed questionnaires from mothers about their children, and 409 questionnaires from the children themselves.
- For the 15-year visit, Project Viva received 679 completed questionnaires from mothers about their children, and 468 questionnaires from the children themselves.

l. 16 year (Age 16), mail

For the age 16-year visits, Project Viva mailed or emailed the participating mothers an annual questionnaire that included child data collection. Participants also had the option to complete electronic versions of the questionnaires through REDCap. A unique URL was sent to study participants who wished to complete their questionnaires electronically using the HPHC domain.

Project Viva also mailed child participants Age 16 annual questionnaires. Prior to mailing the questionnaires, the mother had the option to contact us and request we not mail the questionnaires directly to the child. The mother also had the option to include an email address for her child, so that the child can complete an electronic questionnaire if his or her mother so chooses.

- For the 16-year visit, Project Viva received 813 completed questionnaires from mothers about their children, and 727 questionnaires from the children themselves.

m. Mid-Teen (AG17) Visit, in-person

This visit ideally took place at the HPHCI Fenway office but may have also been conducted at the participant's home or other convenient location if travel to Boston was prohibitive. All participants still enrolled in Project Viva at the start of the Mid-Teen visit were eligible to participate. The visit length was approximately 3 hours. Participants had the option of completing a "remote visit" if they lived too far away or were unable or unwilling to meet with us in person. The "remote visit" consisted of the same self-administered questionnaires, an RA-administered family medical history interview, and a dietary recall. Participants had the option to complete the questionnaires electronically using the REDCap survey option. A unique URL was sent to study participants who wished to complete their questionnaires electronically using the HPHC domain. We completed 705 in-person visits with teens, 101 remote visits with teens, and 528 teen blood draws. Our Mid-Teen Operations Manual served as an up-to-date data collection protocol for the Mid-Teen Visit.

The visit components for the teen included:

1. General questionnaire completed by the mother about her Project teen's health and behavior, and the home environment
2. Teen blood collection
3. Teen urine collection
4. Teen hair collection
5. Teen baby teeth collection
6. Teen anthropometric measurements, including height, weight, bioimpedance, waist and hip circumferences, middle-upper arm circumference, and tricep and subscapular skinfolds
7. Teen body composition: Dual-energy x-ray absorptiometry (DXA)
8. Teen blood pressure
9. Teen step test (ceased data collection March 2020 – August 2021 due to COVID)
10. Teen Exhaled Nitric Oxide testing (ceased data collection March 2020 – August 2021 due to COVID)
11. Teen spirometry testing: pre- and post- bronchodilator (ceased data collection March 2020 – August 2021 due to COVID)
12. Teen questionnaire
13. Internet-based, self-administered dietary recall: 1 completed during in-person visit, 2 completed at home in the 2 weeks following the in-person visit
14. 7-day sleep and physical activity monitoring with the Fitbit Charge 3 device

n. COVID-19 questionnaires, email only

To supplement the data collected during the Mid-Teen visit and the Age 19 questionnaire, Project Viva emailed the participating teens and young adults questionnaires specific to their experiences during the global coronavirus pandemic. The COVID 1.0 Questionnaire was fielded from May 2020 – September 2020, and the COVID 2.0 Questionnaire was fielded from February 2021 – September 2021. Participants completed this electronic questionnaire through REDCap, using unique URL sent to their email, or their mother's email. Father's contact information was also collected within this survey, and if utilized will be done so under a separately approved protocol. We did not send these via mail given the work from home order during the pandemic, that did not allow for mailings to be sent

o. 19 year (Age 19), email/mail

For the age 19 visits, Project Viva mailed or emailed the participating young adults a questionnaire. Participants had the option to complete electronic versions of the questionnaires through REDCap. A unique URL was sent to study participants who wished to complete their questionnaires electronically using the HPHC domain.

A2: Ongoing Visits

a. Annual Surveys, email/mail

Beginning in 2022, Project Viva stopped sending age-based questionnaires and began the practice of sending annual questionnaires. Participants have the option to complete electronic versions of the questionnaires through REDCap. A unique URL is sent to study participants who wish to complete their questionnaires electronically using the HPHC domain.

b. Remote Urine Collection

We are offering our young adult participants who have consented to sharing biospecimen data with ECHO, the opportunity to provide urine specimens remotely. Participants are mailed urine cups directly to their homes, with instructions and pre-paid envelopes in which to return the samples.

c. Young Adult Visit #1, in-person

This visit will ideally take place at the Fenway office, but may also be conducted at the participant's home or remotely. All participants still enrolled in Project Viva at the start of the Young Adult Visit #1 are eligible to participate. This visit will last about 2 hours. Participants have the option of completing a remote visit if they are unable or unwilling to meet with us in person. We expect to complete 1,000 visits during the Young Adult Visit #1 data collection period. Our Young Adult Visit #1 Operations Manual serves as an up-to-date data collection protocol for the Young Adult Visit #1.

The visit components include:

1. Blood collection
2. Urine collection
3. Anthropometric measurements, including height, weight, bioimpedance, waist and hip circumferences, middle-upper arm circumference, and tricep and subscapular skinfolds
4. Body composition: Dual-energy x-ray absorptiometry (DXA)
5. Blood pressure
6. Young Adult Visit Questionnaire
7. Internet-based, self-administered dietary recall: 1 completed during in-person visit, 2 completed at home in the 2 weeks following the in-person visit
8. After-visit summary from recent physician visit (for remote visits only)

B. Secondary Data Collection – Medical Records

Project Viva has obtained medical record data from several sources, including HVMA medical records, hospital birth logs, dental records, and non-HVMA pediatric medical records and newborn screening test results. These arrived in a variety of forms, including electronic and paper. We have derived data from the records and these variables are now part of the existing Project Viva dataset. There is the potential that we could extract additional information from the records for research purpose related to this protocol, or to Project Viva's substudies or ancillary studies. Project Viva will seek additional IRB approval before deriving any variables that may be considered sensitive in nature.

a. Hospital delivery logs

Project Viva received hospital birth delivery logs from Beth Israel Deaconess Medical Center in paper form and from Brigham and Women's Hospital. Data abstracted from these records include gravidity, medical risk factors, anesthesia, delivery method, complications during labor and delivery, obstetric procedures, birthweight, Apgar score, umbilical cord pH, congenital anomalies, abnormal conditions, and delivery location. The paper birth logs are currently archived at Iron Mountain, the Department of Population Medicine's document storage contractor.

b. HVMA/Atrius Medical Records & Insurance Claims

Project Viva collects information from HVMA medical records and insurance claims periodically. The data pulls related to our child cohort are outlined in detail below. These records and datasets are electronic. Project Viva has derived some very important exposure and outcome variables from this information including pre-pregnancy weight and BMI, gestational weight gain, preeclampsia, and gestational diabetes diagnosis.

i. Peri-Pregnancy Period Pull

Project Viva obtained full-text medical records from HVMA on all participants starting 3 months prior to LMP. From these we abstracted the following information onto our Medical Record Abstraction (MRA) form, and entered it into our database. These variables are currently part of Project Viva's data set. The full text medical records still exist in electronic form and remain on HPHC's server. The Medical Record Abstraction forms are archived at Iron Mountain.

1. Alpha-fetal protein and amniocentesis results
2. Ultrasound history and fetal measurements

ii. Early Childhood (~Age 3)

Project Viva obtained HVMA medical records and HPHC insurance claims data on 962 child participants from RSDC during our 3-year visit. This pull included data through December 2004. The data from this pull includes full text medical records, and datasets on growth, prescriptions, immunizations, vital signs and diagnosis codes.

During our 3-year visit, Project Viva worked with HVMA to minimize the number of times the child had to undergo a blood draw. For many participants, we coordinated the regular HVMA 3-year pediatric draw for CBC and lead level with Project Viva's blood draw. If we collected the child's blood, we provided the CBC and lead levels to HVMA to include in the child's medical record. If the child did a separate blood draw at HVMA for CBC and lead, we extracted CBC and lead level from the HVMA medical records.

iii. Mid-Childhood (~Age 7)

Project Viva obtained medical record data from HVMA on 1137 child participants after the 7-year visit. We are unable to collect insurance claims data at this time because of contract agreements between the insurance carriers and HVMA.

We requested medical records only for the participants we saw in person at the 7-year visit, and who provided written informed consent for medical record review. For children this pull included data from date of birth through the visit date. The medical record data from this pull included full-text medical records, and datasets on growth, prescriptions, immunizations, vital signs and diagnosis codes.

iv. Early Teen

For the "Early Teen" visit, we requested medical records and insurance claims for teen participants who provided written informed consent and authorization (or had consent provided on their behalf by their mother) for medical record review from 3 months prior to their birth through the end of the study. The data from this medical record pull included full-text medical records, data sets on growth, prescriptions, immunizations, vital signs and diagnosis codes. Claims data included inpatient, outpatient, and emergency room data (including ICD codes, prescriptions, and billing information). We obtained this data for 946 teen participants.

c. Pediatric medical records, non-HVMA

Project Viva contacted 573 pediatricians to obtain growth data on non-HVMA Project Viva participants. We received 422 charts (74%) and we abstracted weight, height, head circumferences, vaccine information, hemoglobin and lead levels. From height and weight data we calculated each child's age- and sex-specific weight-for-age z-score, height-for-age z-score, height-for-weight z-score, and BMI z-score (for children >2) by use of US national reference data. These data are now included in Project Viva's data set. Project Viva's consent forms covered the collection of this information, but in some cases pediatricians required the mother to sign an additional release form.

d. Newborn screening test for thyroxine

In 2006 Project Viva collected newborn screening records from the New England Newborn Screening Program (NENSP) of University of Massachusetts Medical School. These records included the child's laboratory result for thyroxine. 783 participants were eligible to participate in this substudy of Viva. Project Viva received results for 512 participants. No data exist on paper. The results forwarded to us by NENSP are stored with Viva's dataset on the HPHC server.

VII. Data Management

This section outlines how Project Viva stores data, both in electronic and paper form. It also outlines Project Viva's data access and transmission practices, and Project Viva's Certificate of Confidentiality.

A. Data Storage

a. Electronic Records – REDCap

REDCap is a secure, web-based research database and survey system developed by a multi-institutional consortium and initiated at Vanderbilt University. The system is fully HIPAA compliant provided that security features are utilized, and all study personnel will be trained on use of these features. REDCap may only be accessed by authorized users who must log in to the system and can only view databases that they have been given access to by the project administrator. Additionally, users can be given varying levels of data access (for example, the ability to view de-identified data only, or rights to view but not edit data). REDCap also uses Secure Sockets Layer (SSL) encryption of data, and provides complete audit reports documenting details of all changes that are made to forms and individual fields. HPHCI IT has reviewed and approved the software for use in our department.

Use of REDCap will allow us to increase the security of the data that we are collecting for Project Viva. We will no longer have to rely on an outside institution to house and maintain our database. The database will be hosted on the HPHCI server and users must be logged in to the HPHCI network before they are able to access the HPHCI REDCap system. Authorized Project Viva users will have the ability to export data directly to statistical software programs for analysis, and we can restrict data export by other users. Additionally, REDCap can export completely de-identified datasets, which will allow us to eliminate the possibility of unauthorized individuals inadvertently receiving PHI.

b. Electronic Records – Harvard Pilgrim Health Care Institute (HPHCI) Servers

Project Viva electronic data is also stored on an access restricted HPHCI server. Project Viva data on this server can be accessed only by approved Project Viva study staff and investigators. HPHC machines are password protected, and passwords must be changed at least every 90 days. Access to HPHC's computer systems and file folders are managed by HPHCI, and Project Viva's Project Managers and Data Manager.

c. Electronic Records – Box

A cloud-based file storage system, Box is approved by HPHC's Office of Information Security for use by HPHCI staff for storing non-sensitive *and* sensitive data and PHI. Box and HPHC have a BAA for Box to provide a secure, approved, storage solution for HPHC, and the platform has been vetted by OIS for staff to utilize. Access to Box is controlled via HPHC's account and access validation system, AIM.

d. Electronic Records – Ripple Science Inc.

Ripple™ is a secure web application designed for the storing and management of personally identifying information of research participants. Project Viva will be using Ripple to store meta-data about our participants (contact information, consent status, participation status, etc.). Use of the Ripple platform allows us to more efficiently recruit and track cohort members for our surveys and visits, as it is a platform specifically intended for participants recruitment and tracking, and is designed to house this type of frequently updated meta-data (participant contact information, communication attempt details, communication preferences, etc.).

Authorized Project Viva staff will have the ability to export and import data from/to the Ripple™ platform, and we can restrict access to different sections and functionality of the platform by user account as needed. HPHCI IT has reviewed and approved the software for use in our department. Ripple will only be accessed by approved Project Viva on HPHC-owned computers while on the HPHC server (either in the office or through the VPN). Ripple Science, Inc. staff have access to subject ID numbers and PHI solely for

the purpose of administering our use of their platform and providing technical support, and HPHC and Ripple have jointly executed a Data Confidentiality Agreement.

Ripple was initially developed at the University of Michigan to provide a user-friendly, web-based secure interface where research teams can centralize the storage and management of research participants' personal information, including name, participant ID, demographics, and study workflow (e.g., appointments). Participant information managed with ripple is private and secure. This information is kept in fully encrypted format inside dedicated databases that are segregated from other Ripple accounts and thus only authorized study staff will have access to the study data. Likewise, Ripple infrastructure complies with the privacy and security guidelines of the Health Insurance Portability and Accountability Act (HIPAA), including 2048-bit data encryption in transit and at rest, automatic logoff, audit trail, daily backups in triplicate dedicated servers, firewall, custom access permission for lab members, zxcvbn password strength estimation, and enterprise administrative safeguards to prevent unauthorized staff from accessing participant information.

e. Paper Records

Paper records of Project Viva are stored in our offices at 401 Park Drive, Suite 401, Boston, MA 02215 in locked file cabinets, behind access-restricted doors. Only Project Viva staff knows where the keys are to access these cabinets. Identifiable information is stored in separate locked cabinets from the other information we collect. Cabinets are locked nightly by Research Assistants. Project Viva has a rotating schedule of who will lock cabinets, including a back-up person. Old or archived paper forms are stored at Iron Mountain.

B. Data Access & Transmission

a. Data Access

Project Viva staff have access to all data as a result of their engagement with participants. Project Viva's Principal Investigator, Dr. Emily Oken, and Co-Principal Investigator, Dr. Marie-France Hivert, also have participant contact at times and may see identifiable information. All other investigators have access only to de-identified data (unless otherwise approved) and do not have access to the linking codes. Project Viva policies prohibit staff and investigators from disclosing the linking code under any circumstances.

b. Data Transmission & Types

For specific analyses led by non-DPM investigators, Project Viva's Senior Programmer or Data Manager emails data sets to investigators using HPHC's PGP send feature. These emails include a disclosure statement indicating the recipient is expected to abide by our policies and only use the data for approved purposes.

Project Viva may release the following types of data in the following circumstances:

Personal Health Information (PHI)

PHI is defined by federal law as a data set that includes one of the 18 HIPAA identifiers. Project Viva will release PHI to a non-HPHC investigator only if a Data Security Agreement has been executed between HPHC and the investigator's institution, and HPHC's and the investigator's IRBs have reviewed and approved the research. PHI is sometimes, but very seldom, required for Project Viva analyses. PHI by outside investigators is most likely to be required for substudies or ancillary studies.

LIMITED DATA SETS (LDS)

An LDS is also defined by federal law. It is a data set that includes PHI, but is limited to dates and city, state or zip code. Project Viva will release an LDS to a non-HPHC investigator only if a Data Use Agreement has been signed, and HPHC's and the investigator's IRBs have reviewed and approved the research. LDS are sometimes, but very seldom, required for Project Viva analyses.

DE-IDENTIFIED DATA SETS OR DE-IDENTIFIED AGGREGATE DATA

De-identified data sets are the most common type of data used for Project Viva analyses both for internal and external investigators. The Senior Programmer or Data Manager may provide de-identified data sets to Co-Is at their request for Viva work approved at a Co-Investigator Meeting. De-identified data sets may be provided for substudies, ancillary studies or data repository studies after the appropriate HSC/IRB approvals are in place.

De-identified aggregate data sets are often needed for proposals, including grant proposals or analysis proposals. Project Viva also sometimes provides them to investigators who do not have extensive programming experience and need help completing their analysis. They will be released at the discretion of Viva's Senior Programmer for approved Project Viva work. If the investigator is to use the data for substudies or ancillary studies, they need to obtain appropriate HSC/IRB approvals prior to data release.

C. Certificate of Confidentiality

NIH-funded research is automatically covered by a Certificate of Confidentiality from the Department of Health and Human Services (DHHS), that further protects the privacy of our participants. With this certificate, the investigators cannot be forced to disclose information that may identify participants in any federal, state, or local civil, criminal, administrative, legislative, or other proceedings. Disclosure will be necessary, however, upon request of DHHS for audit or program evaluation purposes.

Appendix A: Current Grants Supporting Project Viva Child Data Collection and Specific Grant Aims

An active grant list is found in Table 3. These grants support ongoing or upcoming data collection related to this protocol. Additionally, the aims for the main Project Viva grant (*4R01HD034568-15: Pre- and Perinatal Predictors of Childhood Obesity)* are given below.

Pre- and peri-natal predictors of childhood obesity (continuation)

(4R01HD034568-15) Project Period: 2/6/2017 – 1/31/2022, PI: Emily Oken

In this continuation of R01 HD34568, our overall goals are to characterize how children's trajectories of growth and adiposity that appear set from early life may improve – or worsen – as they traverse adolescence; to understand underlying metabolic changes, and to identify and quantify modifiable determinants of these changes, especially those that can reverse earlier trajectories of dysmetabolism.

The specific aims are:

I. Characterize trajectories of body mass index, skinfold thickness, DXA fat and fat-free mass, and components of the metabolic syndrome through mid/late adolescence.

II. Use metabolomics profiling in plasma from mid-childhood and mid/late adolescence to refine characterization of trajectories defines only by size, adiposity, or risk factors.

III. Examine modifiable determinants of these trajectory changes at several levels from physiology through behavioral and social factors.

IV. Use single nucleotide polymorphisms as instrumental variables to determine the causal sequences between metabolites and adiposity phenotypes, which may be bi-directional.

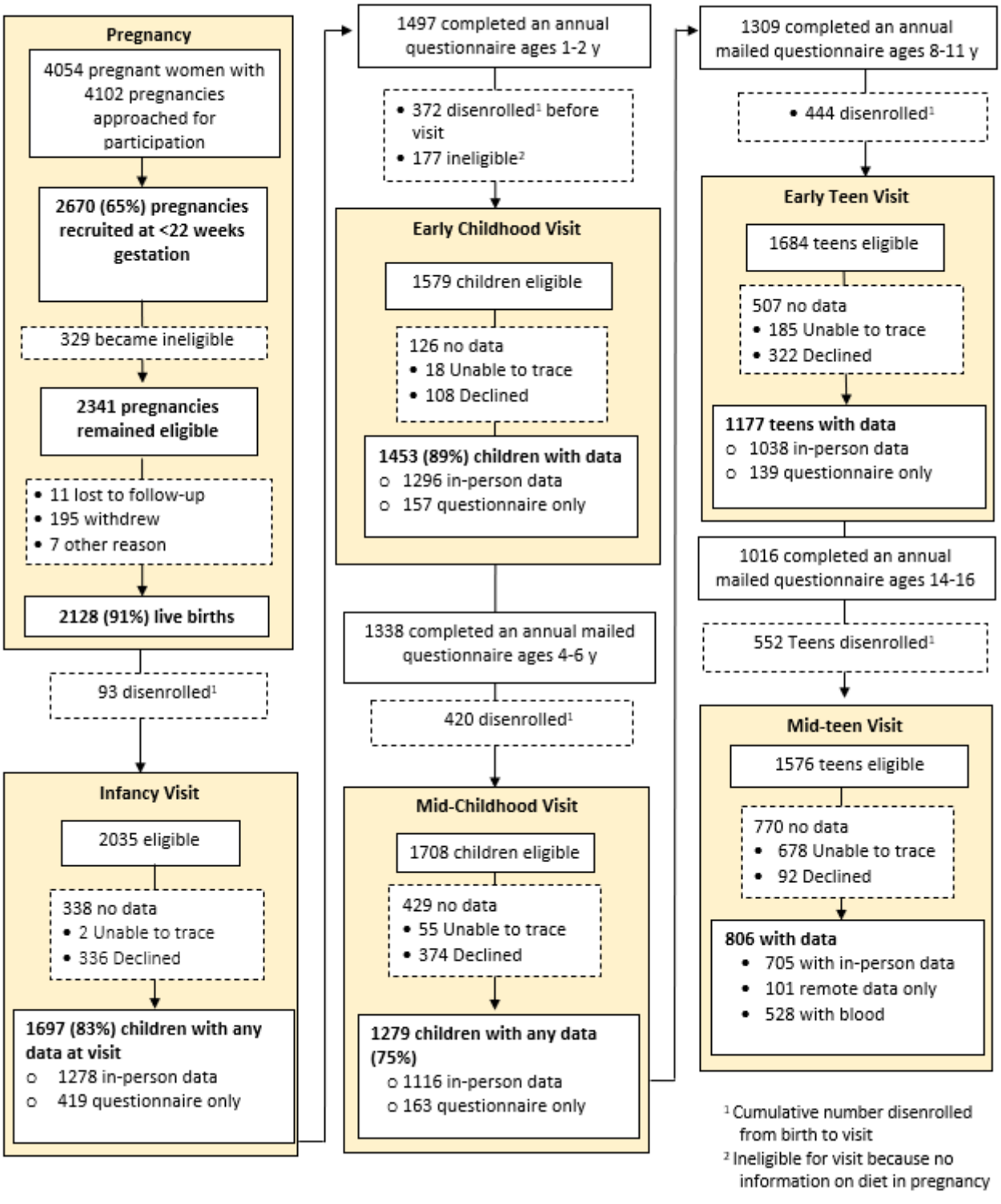
V. With the metabolomics and causal sequences in hand, examine the most promising determinants from Aim 2 as predictors of metabolic trajectory changes in adolescence.

Table 3. Active grants supporting Project Viva Child Cohort Data Collection

Grant name	HPHCI PI	Prime Institution (Prime PI)	IRBNet Project #
Pre- and Peri-natal Predictors of Childhood Obesity	Emily Oken	HPHCI	235301
Common and Distinct Early Environmental Influences on Cardiometabolic and Respiratory Health: Mechanisms and Methods	Emily Oken	HPHCI	951581
Long-term health consequences of birth by cesarian section	Emily Oken	HPSH	1366044
Environmental Chemicals, Adiposity, and Bone Accrual across Adolescence	Emily Oken	Maine Medical Center	1369580
Exposure to Air Pollutants and Upper Airway Microbial Communities in Project Viva, A Pre-Birth Cohort Study	Joanne Sordillo	HPHCI	1473502
Epigenetics, Air Pollution & Childhood Mental Health	Marie-France Hivert	University of Cincinnati	1671551
Pre-and Perinatal predictors of childhood obesity: Nutrition-stress interactions	Emily Oken	HPHCI	235301

Appendix B: Visit Completion Breakdown and Summaries

Figure 1: Flow of Participant Involvement in Project Viva through the Mid-Teen visit



Appendix D: Administrative Supplement to Project Viva: Pre- and perinatal predictors of childhood obesity: Nutrition-Stress Interactions

I. SUMMARY OF PARENT GRANT (5R01HD034568-16 (Oken, PI))

A now-vast animal experimental literature and a growing human counterpart demonstrate that factors operating at the earliest stages of human development-even before birth-can have lifelong consequences for obesity and cardiometabolic outcomes. Yet major questions still exist regarding how pre- and peri-natal factors operate to influence these outcomes. After the prenatal period and infancy, adolescence represents the next "critical period" in the life course as it is characterized by significant changes (auxologic, physiologic, behavioral, and psychosocial) implicated in life-long cardiometabolic health. Adolescence thus provides opportunities to reverse risk trajectories set in motion by pre- and perinatal determinants. Project Viva's "kids," followed with their mothers since before birth, are now 15-18 years of age and the large majority have reached peak height and completed puberty. **The primary goal** of the current cycle of the parent grant is to characterize how children's trajectories of growth and adiposity that appear set from early life may improve—or worsen—as they traverse adolescence; to understand underlying metabolic changes, and to identify and quantify modifiable determinants of these changes, especially those that can reverse earlier trajectories of dysmetabolism. Determinants may range from physiologic (e.g., hormone levels), to auxologic (height growth), behavioral (diet, activity, sleep), psychological (depressive symptoms, stress), and social (relationships with peers and parents).

The aims of the parent award are as follows:
1. Characterize trajectories of adiposity measures, and components of the metabolic syndrome through adolescence.
2. Use metabolomics profiling in plasma from to refine characterization of trajectories defined only by size, adiposity, or risk factors.
3. Examine modifiable determinants of these trajectory changes at several levels from physiology through behavioral and social factors.
4. Use single nucleotide polymorphisms as instrumental variables to determine the causal sequences between metabolites and adiposity phenotypes, which may be bi-directional.
5. With the metabolomics and causal sequences in hand, examine the most promising determinants from Aim 3 as predictors of metabolic trajectory changes in adolescence

The impact of the proposed work lies in identifying and quantifying sets of modifiable factors that can precisely guide intervention studies in early adolescence to preserve or re-calibrate health trajectories, just as we continue to do in the pre- and perinatal period, in the area of obesity and cardiovascular diseases.

II. PROPOSED SUPPLEMENT TO PARENT GRANT

1. SPECIFIC AIMS OF SUPPLEMENT:

Programming of child obesity is multifactorial with a nuanced interplay between pre-, peri-, and post-natal factors. We have made tremendous gains towards achieving our overall study goal to identify and quantify prenatal1 (maternal nutrient intakes, air pollutants), and postnatal (early childhood hormonal patterns such as leptin and androgens) factors contributing to cardiometabolic health in adolescent-aged participants in Project Viva. We have also characterized predictors of body mass index (BMI) trajectories through mid-childhood.

However, **trajectories** characterizing the development of adiposity **through late adolescence,** and *modifiable factors* that improve – or worsen –these trajectories, remain unclear. Maternal nutrition is a key factor in the fetal programming of obesity. For example, higher pre-pregnancy BMI, greater dietary intake of saturated fats, and dietary inflammation are all associated with offspring adiposity based on results from animal studies, observational cohorts, and some more sophisticated study designs such as sibling comparisons and pre/post maternal bariatric surgery follow-up. However, limited longitudinal data exists on the association of maternal

nutrition in pregnancy with offspring adiposity trajectories through adolescence. Furthermore, emerging animal and human data suggest that fetal exposure to excessive glucocorticoids or stress can also program adiposity later in life. However, little is known about the moderating effect of prenatal stress – the prevalence of which is high in the US– on the associations of maternal nutrition with offspring adiposity trajectories. Exposure to psychosocial stress influences dietary intake and metabolic response to food, and conversely, specific nutrients may modulate mood and perceived stress, furthering a vicious cycle. In light of this bidirectional relationship that exists between prenatal nutrition and stress, it is important to address whether prenatal stress may exacerbate the effects of diet alone on offspring adiposity. Finally, a deeper understanding of the mechanisms underlying these associations is warranted to better inform intervention studies. Animal data suggest that maternal high-fat diet induces alterations in stress-related pathways (hypothalamic-pituitary axis (HPA)) in the fetal brain, which may alter the programming of homeostasis and metabolism associated with later obesity and cardiometabolic risk. However, the role of offspring HPA function as a mediating factor in the associations of maternal nutrition with offspring adiposity is unknown.

With support from this supplement, we propose to address this knowledge gap via investigations that are based within Aims 1 and 3 of the parent grant. Specifically, we will investigate the role of nutrition in pregnancy as predictor of childhood adiposity trajectories and characterize the role of maternal and child stress in these associations (Figure 1). Via the following supplemental aims, we will examine:

i) Aim A (1): The associations of nutrition in pregnancy with offspring adiposity trajectories across early childhood (median 3.3y), mid-childhood (median 7.9y), early adolescence (median 13.2y) and late adolescence (15-18y).
***Aim A (2)*:** The extent to which these associations vary by child's age and sex.

ii) ***Aim B (1)*:** The relationship between maternal stress and nutrition during pregnancy
Aim B (2): The moderating effect of maternal stress on the associations of nutrition in pregnancy with offspring adiposity trajectories

iii) ***Aim C*:** The mediating effect of offspring HPA function, as measured by cortisol, on the associations of nutrition in pregnancy with adiposity in adolescence.

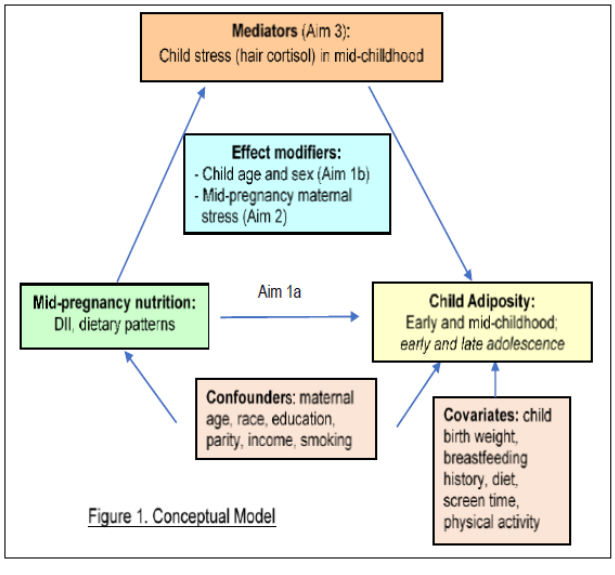


Figure 1. Conceptual Model

Our overall hypothesis is that from childhood to teenage years, children born to mothers with a highly inflammatory diet in pregnancy will gain weight and adiposity more rapidly compared to children born to mothers with a less inflammatory diet, and that both maternal stress and offspring stress will play a critical role in this association. Specifically, maternal stress will potentiate the effect of an inflammatory diet in pregnancy on offspring adiposity trajectories. Furthermore, offspring cortisol (a marker of chronic stress) will mediate associations of maternal nutrition with adiposity in adolescence.

The impact of the proposed supplemental work lies in identifying and quantifying modifiable factors that may improve or worsen growth and adiposity trajectories from early childhood through adolescence. Results of this work could lead to new scientific knowledge on the role stress in pregnancy and childhood on the development of child obesity. Given the burden of both psychosocial stress and childhood obesity, findings from this study could inform intervention studies targeting this intersection.

2. RESEARCH ACTIVITY

Project Viva is a prospective longitudinal pre-birth cohort that recruited pregnant women at in-person visits from April 1999 to July 2002 during the first trimester (median 9.9 weeks of gestation) of pregnancy from 8 obstetric offices of Atrius Harvard Vanguard Medical Associations, a multi-site group practice in Eastern Massachusetts since 1999. Exclusion criteria included multiple gestation, inability to answer questions in English, gestational age of ≥ at least 22 weeks at the initial prenatal care appointment, or plans to move away from the area before delivery. We conducted follow-up in-person research visits at late 2nd trimester (26–28 weeks gestation), delivery, early childhood (mean 3.3 ± standard deviation 0.3 years), mid-childhood (7.9 ± 0.8 years), and early adolescence (13.2 ± 0.9 years) and are currently conducting in-person visits in late adolescence (15-18 years). For this analysis, we will include the 1666 participants with available maternal dietary data in mid-pregnancy and any growth outcomes in childhood. Specifically, we will aim to:

i) Examine associations of maternal nutrition in pregnancy with offspring growth trajectories from childhood through adolescence and examine the extent to which these associations vary by child age and sex. **Our exposures are mid-pregnancy measures of nutritional status, which include (1) dietary inflammatory index (DII) scores and (2) dietary pattern scores**. We obtained data on maternal diet from mid-pregnancy (~28 weeks) food frequency questionnaires (FFQ) which assessed diet intake during the previous 3 months. We calculated the DII scores based on these dietary data. The DII is a population-based aggregate measure of the inflammatory potential of an individual's diet that has been validated with various inflammatory markers and validated in pregnant and non-pregnant populations. We chose DII as an exposure as it is a comprehensive approach to quantify and classify dietary pattern based on inflammatory potential, and inflammation has been postulated as a mechanism in the fetal programming of obesity. Furthermore, we and others have also shown that DII may predict neonatal and mid-childhood adiposity; but the association of this index with trajectories through adolescence is unknown. Using the FFQ results we also determined two dietary pattern scores: (a) a modified version of the Mediterranean diet score (excluding alcohol consumption), and (b) the Alternate Healthy Eating Index (AHEI-P) score, a measure of diet quality based on modified recommendations from the US Department of Agriculture and modified for pregnancy. Specific dietary patterns in pregnancy have been associated with early child growth but longitudinal evidence through adolescence is limited. **Our outcomes are (1) anthropometrics (BMI-z scores) and (2) adiposity measures using a) Waist: Hip circumference ratio and b) skin fold thicknesses (Sum of Subscapular (SS)+Triceps (TR) skinfold thicknesses for total adiposity, and the SS:TR ratio for central adiposity) obtained from early childhood through adolescence; and c) DXA-derived fat mass obtained from mid-childhood through adolescence**. We will model exposure measures categorically (quartiles) and outcome measures continuously. We will use mixed-effects regression for repeated measures to model each continuous outcome as a function of categories of maternal nutrition. We will also adjust for relevant confounders. This method accounts for the correlation between repeated measures on the same individual at different ages, estimates the overall pattern of growth in the cohort, and generates a unique growth trajectory for each child. We will include cubic or linear spline functions for age to capture the trend in adiposity outcomes for each period of childhood and to assess our main hypothesis that maternal nutrition is associated with offspring measures and that such effects are age-dependent. We will include interaction terms for sex to determine the extent to which adiposity trajectories differ between girls and boys. The mixed model will provide estimated model coefficients with 95% confidence intervals.

ii) Examine the relationship between maternal stress and nutrition during pregnancy, as well as the moderating effect of maternal stress on the associations of nutrition in pregnancy with offspring adiposity trajectories. We hypothesize that maternal stress correlates with nutrition during pregnancy. We also hypothesize that among mothers with high psychosocial and biological stress, maternal nutrition will have the strongest association with offspring adiposity. **We will consider both self-reported and biological markers of mid-pregnancy stress.** We will evaluate psychological stress in the following 4 domains: (1) anxiety, assessed using the Pregnancy Related Anxiety Questionnaire (2) prenatal depression, assessed using the Edinburgh Postnatal Depression Scale (3) history of abuse, assessed using the Personal Safety Questionnaire and (4) social support, using Partners Support Scale. We

will evaluate biomarkers of stress using (1) maternal mid-pregnancy plasma corticotropin-releasing hormone (CRH) and (2) cord blood cortisol/cortisone ratio. This ratio is a marker of fetal exposure to prenatal stress and excess cortisol. Maternal stress has been shown to inhibit placental activity of 11-beta-hydroxysteroid-dehydrogenase type 2 (11B-HSD2), an enzyme which converts cortisol to its inactive form, cortisone, to limit fetal exposure to glucocorticoids in-utero. The cortisol/cortisone ratio in cord blood inversely correlates with measures of this enzyme activity in placental tissue. A higher ratio indicates higher exposure to fetal glucocorticoids and therefore prenatal stress. We will first examine bivariate relationship between maternal stress and dietary patterns using univariate analysis. We will then determine whether maternal stress (analyzed categorically) modifies the associations of maternal nutrition with offspring adiposity by performing interaction analysis. We will perform stratified analysis for those interactions with pint<0.05.

iii) Investigate the mediating effect of offspring HPA function on the associations of maternal nutrition in pregnancy with adiposity in adolescence. **Measures of HPA function will include hair cortisol levels measured in mid-childhood and early adolescence**. We will use multivariable linear regression models to first investigate nutrition in pregnancy as predictor of hair cortisol levels in adolescence. Then, we will examine cortisol as predictor of adiposity in adolescence. Finally, we will conduct formal mediation analyses using methods by VanderWeele. Mediation analyses will include those covariates that are associated with both the mediator (hair cortisol) as well as the outcome (mediator-outcome confounding) to control for collider bias.

The candidate's role in the proposed research activities during the supplement award period will consist of three components. First, Dr. Monthé-Drèze will enhance her research skills by planning, designing and conducting the analysis for all the supplemental proposed specific aims (Table 1). Secondly, she will have the opportunity to work closely with an established senior investigator and highly productive and collaborative research team thereby acquiring important experience in research analysis as well as project leadership. Finally, Dr. Monthé-Drèze will submit an K23 mentored patient oriented research award to NICHD at the end of this supplement period, using results this supplement as preliminary data to propose a novel prospective study on the role of HPA dysregulation on reward pathways, feeding behaviors, and child obesity. Details on the proposed research and career development activities are presented below.

Table 1.Timeline of benchmarks for proposed research activities

Year	*2019*	*2020*		*2021*
Quarter	*3 4*	*1 2*	*3 4*	*1 2*
Aim 1	* Meetings with biostatistician and review of pertinent datasets * Present research and analysis plan at Project Viva Co-Investigator Meeting	* Complete data analysis (Aim 1) * **Manuscript #1:** Preparation and submission: Maternal Nutrition and Offspring Growth Trajectories Through Adolescence		
Aim 2	* Preliminary data analysis (Aim 1 and 2) * Present at Work-in-Progress (WIP) * Submit abstract (Aim 1)	* Meetings with biostatistician * Complete data analysis (Aim 2) * Present at WIP (Aim 2) * Submit abstract (Aim 2)	* **Manuscript #2:** Preparation and submission: Does Pregnancy Stress Potentiate the Effect of Maternal Nutrition on Offspring Adiposity? * **Manuscript #3**: Preparation and submission: Cord blood cortisol/cortisone ratio: a better marker of fetal exposure to prenatal stress?	
Aim 3			* Present research and analysis plan at Project Viva Co-Investigator Meeting * Meetings with biostatistician * Preliminary data analysis (Aim 3) * Present at WIP (Aim 3) * Submit abstract (Aim 3)	* Complete data analysis (Aim 3) * **Manuscript #4:** Preparation and submission: Chronic Stress and Adiposity in Childhood, Is There a Link with Maternal Nutrition * K23 grant preparation and submission

3. CAREER DEVELOPMENT PLAN

The candidate's **long-term overall goal is to become an independent researcher in the field of child obesity prevention. Her objectives are to 1) characterize novel pre- and postnatal determinants of childhood obesity 2) devise innovative behavioral and nutritional interventions based on these determinants 3) conduct longitudinal trials to determine the effectiveness of these interventions for the prevention of childhood obesity. Her specific interest is in exploring psychosocial, hormonal and behavioral pathways via which maternal nutrition programs child obesity.** She will collaborate with researchers in the Harvard community and beyond to design and evaluate lifestyle interventions in the prenatal and early childhood periods. With the training, skills, and experience she will receive through the proposed supplement and her career development grant, Dr. Monthé-Drèze intends to become a nationally recognized expert in the field of the fetal origins of child obesity. **Her short-term goal over the next two years is to develop the specific skills she needs to independently analyze longitudinal data using advanced biostatistical techniques and to additionally acquire hands-on training** with epidemiologists, biostatisticians, and clinical researchers within and beyond the Brigham and Women's Hospital (BWH) Pediatric Newborn Medicine Department.

While Dr. Monthé-Drèze's education and experience thus far have provided a strong foundation for her ongoing work, the proposed research studies, mentoring and career developmental activities under this Diversity Supplement will provide her with the additional experience, skills, and background data necessary to put together a unique tool-kit to support a successful K23 award application. She will avail herself of the superior environmental resources of the Harvard Community. In addition, she will continue to receive mentorship from a team of experienced and outstanding mentors and collaborators who are leaders in their field. To meet her short-term objectives over the next two years, the mentoring team has developed a program that builds and extends on Dr. Monthé-Drèze's prior experience and training, and which focuses on the following 3 goals **(Table 2):**

1) To develop expertise in the methods of assessing nutritional intake in the context of disease. She will take the Nutrition *"ID 214"* course at the Harvard School of Public Health (HSPH), which reviews methods for assessing the dietary intake of populations and individuals and reviews specific diet/disease relationships.

2) To develop expertise in advanced statistical methods. She will pursue further *formal education* in advanced statistical analysis; The HSPH *Biostatistics (BST) 210* and 215 courses cover complex regression and longitudinal analysis, and will allow her to successfully complete the analysis of this and future projects. These courses will be essential towards her goal of evaluating long-term child obesity outcomes.

3) To acquire scientific enrichment and career guidance via formal and informal meetings and training programs. To gain additional opportunities for education and networking with experts and researchers in the parent grant, she will participate in the following meetings with other investigators in the parent grant: ***i)*** Monthly co-Investigator meetings, where updates on parent study and new research proposals are presented ***ii)*** Biweekly CoRAL (Division of Chronic Disease Research Across the Lifecourse) 'Methods/trajectory modelling' meetings, where advanced statistical methods used in trajectory analyses are explored ***iii)*** Biweekly Project Viva data/operations meetings, where she will acquire expertise in study management and leadership. In addition to team/laboratory meetings with primary mentor and co-mentor, journal clubs, and formal mentoring as detailed below, her plan also includes bi-annual presentations of her work at the Dept. of Newborn Medicine work-in-progress meetings, at CoRAL's research-in-progress meetings, and at national meetings. To further facilitate her career development, she also plans to engage in formal leadership training, as well as take advantage of Grant writing workshops offered through the department and the Harvard Medical School Clinical and Translational Science Center (CTSA) (Table 2).

Table 2. Timeline for career development (CD) goals

	Activity type	*Description*	*2019*	*2020*			*2021*	
Quarter			*3*	*4*	*1*	*2*	*3*	*4*	*1*	*2*
CD #1	Coursework at Harvard T.H Chan School of Public Health (HSPH)	- ID 214- Nutritional epidemiology			x					
CD #2	Coursework at HSPH	- BST 210- Applied Regression Analysis	x							
		- BST 215- Linear and Longitudinal Regression			x					
	Summer Program in Clinical Effectiveness at HSPH	- Includes training in Biostatistics including: BST 215-Linear and longitudinal regression					x			
CD #3	Scientific Enrichment	- Monthly meetings with primary mentor	x		x		x		x	
		- Twice weekly meetings with co-mentor	x		x		x		x	
		- Weekly lab meetings with Dr. Sen	x		x		x		x	
		- Monthly lab journal clubs	x		x		x		x	
		- Monthly Project Viva co-investigators meetings	x		x		x		x	
		- Bi-weekly CoRAL's research-in-progress meetings	x		x		x		x	
		- Bi-weekly CoRAL's 'methods/trajectory modelling' meetings	x		x		x		x	
		- Bi-weekly Project Viva data & operations meetings	x		x		x		x	
		- Annual American Academy of Pediatrics National Meetings	x				x			
		- American Societies of Nutrition Meeting			x					
		- Annual Pediatric Academic Societies National Meetings			x				x	
	Professional Development	- Bi-annual meetings with Department Chair	x		x		x		x	
		- BWH monthly K-club meetings	x		x		x		x	
		- AAMC Minority Faculty Leadership Development Seminar	x							
		- Perinatal Research Society Grant Writing Workshop for Young Investigators	x							
		- Leadership Strategies for the Researcher course (CTSA)							x	
	Grant Application	- KL2/Harvard Catalyst Mentored-Medical Research Training Award Application			x					
		- Foundation Grant Application	x		x					
		- NIH K23 application							x	
	Dissemination of Research	- Abstracts submissions to scientific meetings	x		x		x		x	
		- Manuscripts submissions/publications			x (1)		x (2)		x (1)	

The supplement award will allow Dr. Monthé-Drèze to focus 75% of her effort on child obesity research activities, career development activities, publishing manuscripts, disseminating results, enrolling in short-term courses, developing her independent research, and applying for the K23 career development award.

4. MENTORING PLAN

Dr. Emily Oken is the **Principal Investigator** of Project Viva, the cohort in which this supplement is based. As detailed in her biosketch, Dr. Oken is a well-trained and well-funded investigator with relevant experience in internal medicine, pediatrics, and nutritional epidemiology. Moreover, she is an experienced and committed mentor with an established, productive relationship with the candidate. She was a member of the candidate's scholarship oversight committee during her fellowship. Her mentorship has been invaluable and has resulted in a publication in *Pediatric Research*. Dr. Oken has mentored over 3 dozen trainees, several of whom have gone on to receive K awards. Her commitment to mentorship has been recognized nationally with a K24 mid-career development grant from NICHD, and locally with the Young Mentor Award and A. Clifford Barger Award for Excellence in Mentoring at HMS, as well as the Harold Amos Faculty Diversity Award at HMS. She serves on the HMS Council of Mentors. She and co-mentor Dr. Sen have worked closely for several years and published several high impact manuscripts together. As a mentorship team, they have been instrumental to Dr. Monthé-Drèze's academic success so far and will continue this successful mentorship that began during her fellowship.

The candidate will meet with Dr. Oken monthly to review her progress. Dr. Oken will also share grantsmanship strategies that have contributed to her funding success. The candidate will continue to attend Project Viva and CoRAL meetings, which will include an interdisciplinary team of experts in obesity, growth, epidemiology, and biostatistics. The candidate will present her work and seek feedback at these meetings. She will engage with individual members and experts in the Project Viva team to develop expertise in the pre-and post-natal determinants of childhood obesity and their complex interactions, and to learn sophisticated biostatistical research skills to address this complex problem of public health importance.

Dr. Sarbattama Sen (Co-mentor): For the past 3 years, the candidate has had a very successful mentor-mentee relationship with Dr. Sarbattama Sen, who is an Assistant Professor in Pediatrics at HMS and Neonatologist at BWH. She is committed to ensuring that Dr. Monthé-Drèze becomes a successful investigator and has continued to provide her with the necessary tools. Her research expertise, which aligns closely with the candidate's proposed project, is in maternal obesity-related metabolic and nutritional derangements and their impact on offspring health, including child obesity. Dr. Sen has examined these topics with funding from a K23 award from NICHD (on

which Dr. Oken is a mentor) and numerous foundation grants. In the past 3 years, under the mentorship of Dr. Sen, the candidate has published three manuscripts, and has a fourth currently under review, and has co-authored two others in preparation. Dr. Sen has a strong track record in mentorship, having mentored or co-mentored several neonatology and obstetrics-gynecology fellows and junior faculty members who have successfully completed and published their research, and who currently hold academic faculty positions.

Dr. Sen will provide close oversight as the candidate pursues her research project. The candidate will continue to have twice weekly in-person meetings with her, including lab meetings, during which recent pertinent literature will be reviewed and statistical methods explored with the input of the biostatistician in the team. At the one-on-one meetings, they will discuss the progress of the candidate's research, and identify opportunities for new collaboration, for grant funding and to present the candidate's research. Dr. Sen has always been available as needed, whether in-person, by phone or emails, and will continue to be so.

The candidate will meet with Drs. Oken and Sen monthly to review progress, abstracts and manuscripts prior to submission. Together, they will provide guidance and support on grant applications. They will evaluate the candidate's progress in meeting her career development and research goals as defined by milestones in Table 1 and 2. Together, they will guide the candidate's transition to become an independent investigator.
Other instrumental members of the candidate's research committee will consist of **Sheryl Rifas-Shiman, MPH,** Senior Analyst for Project Viva and Research Associate in the Department of Population Medicine; and **Sara Cherkerzian, ScD,** Senior Biostatistician and Associate Epidemiologist in the Department of Pediatric Newborn Medicine at BWH. They will oversee and guide the planning and implementation of the complex biostatistical methods used in this project. The candidate has worked with both Sheryl and Sara in the past, and they will continue their well-established and productive relationship in the next two years. The candidate will have monthly scheduled meetings with Sara, and will meet with Sheryl as needed for additional biostatistical oversight.

During the period of this award, markers of success for the candidate will include the following milestones, which will aid towards her transition to becoming an independent investigator: **(1)** Four first-authored and 3 co-authored publications submitted and/or accepted **(2)** Four oral or poster presentations at scientific meetings **(3)** Receipt of external grant funding for resources toward the collection of pilot data for a career development award application **(4)** Submission of a K23 career development award application and **(5)** Participation in at least 4 professional development activities via the CTSA or other programs, as detailed above in Table 2.

Overall, we believe that the combination of a committed, well trained candidate, rigorous training program, outstanding mentorship team, and exceptional research environment will ensure the successful completion of the proposed scientific and career development aims proposed here, and will provide a strong foundation for the candidate's continued success as a physician-scientist.

Appendix E: Administrative Supplement to Project Viva: Early vitamin D status and supplement use and later body composition and bone health

Summary of Parent Project (R01 HD034568)
Project Viva, primarily funded by R01 HD034568, "Pre and peri-natal predictors of childhood obesity", is an ongoing prospective cohort study examining early life factors in relation to pregnancy outcomes and child health. Women were enrolled during pregnancy; inclusion criteria included < 22 weeks gestation, ability to answer questions in English, singleton pregnancy, and plans to remain in the study area until delivery. The Project Viva cohort includes 2,128 singleton infants born 1999-2002 and their mothers; 1,558 mother-child pairs are still enrolled. We regularly collect information, including data on diet and supplement use, from mothers and children via interviews and questionnaires at both in-person and remote (mailed or electronic surveys) visits. We also perform anthropometric measurements and DXA scans and collect biosamples at in-person visits (**Table 1**).

The goals of the current renewal of R01 HD034568 are to characterize how unhealthful trajectories of growth and adiposity that appear set from early life may improve or worsen during adolescence; to employ modern technologies to understand underlying metabolic changes and capture behaviors, and to investigate the modifiable determinants of these changes, especially those that can reverse earlier trajectories of dysmetabolism. With this application we are requesting supplemental funds to examine one modifiable potential determinant of trajectories of body composition and bone health – early vitamin D status, with a focus on mitigation of insufficient dietary intake of vitamin D through use of vitamin D-containing supplements in early life. The aims of the parent R01 are as follows; this supplement will contribute to Aims 1 and 3:

1. Characterize trajectories of body mass index, skinfold thicknesses, DXA fat and fat-free mass, and components of the metabolic syndrome through mid/late adolescence. We will take advantage of approaches to estimating trajectories that our research team has advanced using data from Project Viva and other cohorts. We are especially interested in characterizing Viva participants who change their trajectories as they traverse adolescence.
2. Use metabolomics profiling in plasma from mid-childhood and mid/late adolescence to refine characterization of trajectories defined only by size, adiposity, or risk factors. Our team has been a leader in accurate measurement, data flow, and interpretation of metabolomics readouts in developmental origins studies, including previous Viva analyses.
3. Examine modifiable determinants of these trajectory changes at several levels from physiology through behavioral and social factors. We will use sophisticated multi-level longitudinal modeling to identify the most promising candidates.
4. Use single nucleotide polymorphisms as instrumental variables to determine the causal sequences between metabolites and adiposity phenotypes, which may be bi-directional.
5. With the metabolomics and causal sequences in hand, examine the most promising determinants from Aim 2 as predictors of metabolic trajectory changes in adolescence.

Adequate vitamin D in early life is important for healthy bone formation, and reduces risk of fracture in childhood through adulthood. In addition, early-life vitamin D deficiency has been associated with greater adiposity and other adverse cardio-metabolic outcomes. Research indicates that many U.S. children do not meet the minimum daily vitamin D intake of 400 IU recommended by the American Academy of Pediatrics, although vitamin D supplements are affordable, widely available, and easy to administer to infants and young children. We propose to use data previously collected from Project Viva children to study the impact of vitamin D supplementation in early life on plasma 25(OH)D concentrations in early childhood and on body composition and bone health in mid-childhood and early adolescence.

Although we have stored plasma samples available for analysis of early childhood vitamin D status, we currently do not have funding for 25(OH)D assays and therefore will not be able to complete this study without the supplemental funds we are requesting here. We also require funding to support personnel time to calculate vitamin D supplementation levels from the available data (see **Table 2** for a description of the assessment of vitamin D-containing supplement use in Project Viva).

Research Strategy

2. Significance
Breastfeeding, avoidance of direct sun exposure, and sunscreen use are commonly recommended as best practices for infant/child health and safety yet can contribute to vitamin D deficiency. Vitamin D is essential for healthy bone formation and may have additional cardiometabolic, respiratory, and immune system benefits. Deficiency in early life contributes to higher fracture risk in childhood and through adulthood, and has also been associated with poor muscle tone, greater adiposity, insulin resistance, and higher risk of type I diabetes in childhood. Vitamin D deficiency can lead to clinically observable skeletal abnormalities such as rickets. Subclinical vitamin D insufficiency, which does not manifest in skeletal or metabolic abnormalities, is also common in pediatric populations with unclear long-term consequences. Therefore, the American Academy of Pediatrics (AAP) recommends a minimum daily intake of 400 IU of vitamin D for all infants and children. Fully and partially breastfed infants, infants who do not consume enough formula to meet the minimum requirement, and toddlers and young children who drink less than 1 quart of milk per day, typically require vitamin D supplementation to meet these requirements. However, in a 2010 study of U.S. infants, fewer than half of infants overall met the AAP recommendation for daily vitamin D intake, and fewer than 15% of fully or partially breastfed infants met the recommendation. Additionally, more than 85% of infants in that study did not receive any vitamin D supplement. The impact of supplementation on reducing prevalence of deficiency and risk of adverse outcomes has not been well-studied in U.S. children.

As many parents follow recommendations to use sun protection for their young children, and because body composition and bone density trajectories may be established in early life, the period from infancy to early childhood is an optimal age range to study the use of supplements in correcting dietary vitamin D deficiencies, without interference of endogenous vitamin D synthesis in response to sun exposure. We will conduct this study within Project Viva, a well-established, NIH-supported cohort of children enrolled in eastern Massachusetts, where the prevalence of vitamin D deficiency is high, particularly in the winter months.

3. Approach:
3.1. Study Design: We will perform a longitudinal prospective cohort analysis leveraging existing data and stored biosamples. See **Table 1** for a description of completed study visits and corresponding data, samples, and measurements collected from Project Viva participants.

Table 1: Completed Project Viva Visits and Child Measures

Visit (Median age), Type	Child Data/Samples/Measurements
Birth (39.7 weeks of gestation), in person	Anthropometry, blood pressure, cord blood sample
Infancy (6.3 months of age), in person	Blood pressure, anthropometry, questionnaire and interview (completed by mother)
1 and 2 years, remote	Questionnaire and diet assessment via FFQ (completed by mother)
Early childhood (3.2 years of age), in person	Anthropometry, blood pressure, cognitive assessment, blood sample, questionnaire, interview, and diet assessment via FFQ (completed by mother)
4, 5 and 6 years, remote	Questionnaire (completed by mother)
Mid-childhood (7.7 years of age), in person	Anthropometry; DXA scan; step test; blood pressure; spirometry; cognitive assessment; fasting blood sample; hair and urine; questionnaire and interview (completed by mother)
9, 10 and 11 years, remote	Questionnaires (completed by mother and child, interview (completed by mother)
Early teen (13.0 years of age), in person	Anthropometry; DXA scan; step test; blood pressure; spirometry; exhaled nitrous oxide; nasal swab; fasting blood; hair and urine; interview (completed by mother); questionnaires (completed by mother and child)

3.2. Exposures: Early-life diet and supplement use and plasma 25(OH)D concentrations

Assessment of vitamin D supplement use: As shown in **Table 2**, mothers reported their child's supplement use in detail at 6 months of age. We propose to use some of the funding from this proposal to support time for a research assistant to research formulations of each supplement at the time our data were collected and to conduct a thorough review of the "other" vitamins and supplements reported at each time point to determine which contained vitamin D and how much. We will use the reported information on the timing and frequency of supplement use and the vitamin D content of each of the infant dietary supplements to create a continuous measure of vitamin D supplement use in infancy. Additionally, mothers reported their child's use of multi-vitamin and other supplements on annual questionnaires by indicating whether their child had used any supplements in the past month. We will use the responses from infancy and early childhood (1, 2, and 3y) to create a composite measure of vitamin D supplement use at multiple time points (see Statistical Analysis section for Aim 1 below for additional detail). In early adolescence, teen participants indicated whether they were currently taking a multivitamin and how often.

Table 2: Assessment of Child Vitamin D Supplement Use in Project Viva

Time	Reported By	Question				
6 months	Mother	In the past month, has your baby taken any… a. Fluoride drops (e.g. Tri-Vi-Flor or Poly-Vi-Flor b. Tri-Vit (Tri-Vi-Sol) c. Poly-Vi-Sol d. Other vitamins or supplements:____ Yes ➡ / No	When did your baby start taking them? __ __ months OR __ __ weeks	Is your baby still taking them? Yes No ➡	When did your baby stop taking them? __ __ months OR __ __ weeks	On average, how many times per week has your baby taken this DTTP? __ __ / week
1 year	Mother	In the past month, has your child taken any of the following vitamins or supplements? a. Fluoride drops, for example, Tri-Vi-Flor or Poly-Vi-Flor b. Multi-vitamin drops such as Tri-Vit (Tri-Vi-Sol) or Poly-Vi-Sol c. Other vitamins or supplements: _____ Yes / No				
2 year	Mother	In the past month, has your child taken any of the following vitamins or supplements? a. Fluoride drops, for example, Tri-Vi-Flor or Poly-Vi-Flor b. Multi-vitamin drops such as Tri-Vit (Tri-Vi-Sol) or Poly-Vi-Sol c. Other vitamins or supplements: _____ Yes / No				
3 year	Mother	In the past month, has your child taken any of the following vitamins or supplements? a. Chewable multi-vitamin (for example Flintstones, Sesame Street, Centrum) b. Multi-vitamin drops such as Tri-Vit (Tri-Vi-Sol) or Poly-Vi-Sol c. Other vitamins or supplements: _____ Yes / No				

Assessment of infant and child diet: We will use existing data on infant and child diet for this project. Mothers reported infant feeding mode (exclusive breastfeeding, exclusive formula feeding, or mixed feeding) at 6 months, as well as child supplement use at 6 months, 1 year, 2 years and 3 years. See Table 2 for a description of the assessment of vitamin D-containing supplement use among Project Viva children at each time point. At the Project Viva 2-year and 3-year visits, mothers reported their child's diet via a semi-quantitative food frequency questionnaire (FFQ) validated for use in preschool-aged children. The FFQs assessed the child's diet during the past month, and we estimated usual intake of vitamin D from foods using the Harvard nutrient composition database, which includes food composition values from the US Department of Agriculture and is supplemented by other sources. We will calculate dietary vitamin D intake by multiplying the amount of vitamin D in the specific portion size of each vitamin D-containing food by the consumption frequency of each food and summing across all food items. We will adjust individual nutrient estimates for total energy intake using the nutrient residual method.

Measurement of 25(OH)D in early childhood plasma samples: We have previously collected and stored plasma samples from early childhood for approximately 800 Project Viva participants. We will measure 25(OH)D2/D3 in

these samples at the Clinical and Epidemiological Research Laboratory (CERLab) at Boston Children's Hospital. The laboratory uses an ELISA (Immunodiagnostic Systems Inc.), which is standardized using internal standards which are traceable to the ID LC MS/MS 25 hydroxyvitamin D Reference Measurement Procedure - traceable to the National Institute of Standards and Technology Standard Reference Material 2972. This assay measures total 25(OH)D (D2+D3) and is approved by the FDA for clinical use. The assay is sensitive down to 5.0 nmol/L, and day-to-day variabilities of the assay at concentrations of 40.3, 72.0 and 132.0 nmol/L are 4.6, 6.4 and 8.7%, respectively. We have an established, long-term collaboration with the CERLab, which conducts many assays on Project Viva samples and consistently produces high-quality data.

3.3. Outcomes: body composition and bone density:
We performed dual-energy X-ray absorptiometry (DXA) scans (Hologic, Discovery A) on Project Viva participants at the mid-childhood and early adolescence visits. We used the same DXA machine on all participants at all time points and calibrated it daily with a standard synthetic phantom to check for machine drift. We analyzed data with pediatric software (Hologic, QDR versions 12.6 and 13.4 for the mid-childhood and early adolescence visits, respectively). At each visit, a single, trained investigator checked all scans for positioning, movement, and artifact, and defined body regions for analysis. We derived data on total, truncal, and non-truncal fat mass, fat-free mass, and bone mineral density (BMD) from the scans at each time point. Intra-rater reliability on duplicate measurements was high (ICC > 0.91) for all variables.

Consistent with our other research on predictors of BMD 14, we will use total body less head aBMD (kg/cm2) as the bone health outcome. BMD measures bone surface density (bone mass (g)/cm2) and is used clinically. As compared to regional BMD measures used in adults, total body BMD has less variation during skeletal development and greater reproducibility, so is preferred for pediatric evaluations. We exclude the skull because it comprises a large proportion of the skeleton in growing children and is not responsive to physical activity or other environmental influences. For this analysis, we will evaluate BMD rather than bone mineral content (BMC), a marker of bone mass (g) which does not track as strongly as BMD with Z-score at skeletal maturity and is not typically used clinically. We will use sex-, age-, race-, and height-standardized Z-scores, which we have derived for BMD using U.S. national reference data.

3.4. Data collection on potential covariates: We have extensive information on many factors that we plan to adjust for or otherwise consider in our statistical analyses, including maternal and child race/ethnicity; maternal and child weight, height, and adiposity; maternal and household socioeconomic status; season of birth and of 3-year blood draw; and child physical activity, screen time, and dietary behaviors.

3.5. Statistical Analysis:
Before beginning our formal analysis, we will examine univariate statistics, graphs, and bivariate analyses to identify outliers and implausible values. In addition to reviewing the literature and using knowledge of biological mechanisms, we will identify potential confounders by examining bivariate associations of other factors with exposures and outcomes. We will check for linearity of associations between exposures and outcomes and normality of residuals and apply transformations if necessary to meet assumptions for linear regression. We will use multiple imputation for missing data to increase the available sample size for all analyses. We will compare results to those obtained from complete-case analysis in a sensitivity analysis.

Aim D: To assess plasma 25(OH)D concentrations in early childhood (~age 3 years) among children who were given vitamin D supplements at ages 6 months, 1 year, 2 years and/or 3 years, and to determine key time points for supplementation to attain higher 25(OH)D levels in early childhood.
For our primary analysis, we will examine 25 (OH)D levels in relation to vitamin D supplementation reported at the 6-month visit. From our detailed data on the timing and frequency of supplement use, we will create a continuous measure of vitamin D supplementation in infancy. We will use multivariable linear regression models to examine associations between vitamin D supplement use and 25(OH)D levels in early childhood. We will examine models adjusted for other factors that could potentially confound the association, including child race/ethnicity, family socioeconomic status (household income, maternal education level), child screen time and physical activity at 3y

(as a proxy for sun exposure), and reported dietary intake of vitamin D at 2 years and 3 years. To reduce variability in the outcome, we will also adjust our models for child age at blood draw and season of blood draw. We will consider child BMI at 3 years as a potential confounder and mediator.

As a secondary analysis, we will explore the impact of additional vitamin D supplementation beyond infancy on 25(OH)D levels in early childhood. To do this, we will first use Analysis of Variance (ANOVA) to compare mean 25(OH)D levels in early childhood in participants who were given vitamin D supplements at each of the 4 time points to those who were not given supplements at each time point. Next, we will create a composite supplement score ranging from 0-4, indicating the number of time points that the mother reported giving her child a vitamin D-containing supplement from 6 months to 3 years of age. Our preliminary data (**Table 3**) indicate that there is considerable variability in supplement use among the Project Viva children. Among the 935 participants who responded to the child supplement questions at all time points, 33% reported no supplement use at any time point, 52% used vitamin D-containing supplements at 1 or 2 time points and 15% reported using these supplements at 3 or all 4 time points. The proportions were nearly identical among the subset of 555 participants who have an early childhood blood sample. We will examine associations of the vitamin D supplement score with early childhood 25(OH)D levels using multivariable linear regression models adjusted for potential confounders.

Table 3: Tracking of Vitamin D Supplement Use in Project Viva Participants (Infancy-Early Childhood)

# of time points with reported vitamin D supplement use	Participants (full cohort), N (%)	Participants (with 3-year blood), N (%)
0	312 (33)	196 (35)
1	314 (34)	187 (34)
2	168 (18)	97 (17)
3	112 (12)	60 (11)
4	28 (3)	15 (3)

Aim E: To determine the relationship of dietary vitamin D intake in infancy and at age 2 years with plasma 25(OH)D levels at age ~3 years, and to examine the impact of concurrent vitamin D supplementation on these relationships. In our primary analysis, we will compare mean early childhood 25(OH)D levels in participants who were exclusively breastfed, exclusively formula-fed, or both breast- and formula-fed before 6 months of age using ANOVA. Additionally, we will analyze FFQ-derived data on vitamin D intake reported at 2 years of age in relation to 25(OH)D levels measured at 3 years using multivariable linear regression models. In the second part of this analysis, we will use multivariate linear regression models with interaction terms to determine whether vitamin D supplement use reported at the time of diet assessment (i.e. supplement use reported at 6 months when infant feeding method is the exposure of interest and vitamin D supplement use reported at 2 years when diet at 2 years is the exposure of interest) modifies the relationship between dietary vitamin D intake and 25(OH)D levels at 3 years. As shown in **Table 4**, 24% of fully breastfed infants, 13% of infants fed both breast milk and formula, and 9% of fully formula-fed infants were also given vitamin D supplements in infancy. We will consider adjusting the models for child race/ethnicity, family socioeconomic status (household income, maternal education level), child age at blood draw and season of blood draw. We will also examine potential confounding by child BMI at the time of exposure measurement and will consider BMI as a potential effect modifier using stratified models and interaction terms. Finally, we will examine BMI as a potential mediator of the association between dietary vitamin D intake at 2 years and 25(OH)D levels at 3 years by adding the change in BMI from 2-3 years to our models and looking for attenuation in the effect estimates.

In a secondary analysis, we will repeat our models using vitamin D intake and supplement use assessed at the 3-year visit. Although these assessments were done at the same time that the blood sample used for 25(OH)D measurement was taken, the FFQ and questionnaire assess dietary and supplemental vitamin D intake in the preceding month which, in combination with the results of the models using 2-year vitamin D intake, will help us to assess how quickly children's 25(OH)D levels respond to vitamin D intake from their diet and supplements

Table 4: Reported Supplement Use at 6 months – 3 years by Infant Feeding Method among Project Viva Children with a 3-year Blood Sample

Infant Feeding Method at 6 months	Reported Vitamin D Supplement Use			
	6 months	1 year	2 years	3 years
Fully breastfed	52/213 (24%)	45/192 (23%)	111/207 (54%)	115/208 (55%)
Mixed (both breast- and formula-fed)	63/473 (13%)	93/413 (23%)	222/457 (49%)	236/464 (51%)
Fully formula-fed from birth	6/64 (9%)	16/60 (27%)	32/71 (45%)	22/62 (35%)
Total	121/750 (16%)	154/665 (23%)	365/735 (50%)	373/734 (51%)

Aim F: *To examine body composition (fat mass and lean mass) and BMD in mid-childhood and early adolescence and trajectories of adiposity through early adolescence among children with varying levels of vitamin D supplementation in infancy and early childhood, and to examine the interaction of vitamin D supplementation with dietary vitamin D intake in relation to these outcomes.*
Our primary analyses will examine body composition and BMD in both mid-childhood and early adolescence in relation to vitamin D supplement use in infancy and the composite score (0-4) for supplement use in infancy through early childhood. We will first look at body composition and bone health at each time point using separate multivariable linear regression models, and if the associations are in the same direction, we will use mixed models to incorporate the longitudinal repeated measurement of these outcomes. We will adjust our models for potential confounders that may influence body composition and BMD, including child race/ethnicity, family socioeconomic status (household income, maternal education level), child dietary intake of milk and dairy products, and child screen time and physical activity. We will also consider examining mediation of the associations by factors that may be on the causal pathway, including lean mass, pubertal status, inflammatory cytokines, and insulin resistance. Finally, because there may be some fetal programming of body composition and bone health by gestational vitamin D status we will examine correlations between maternal and cord blood 25(OH)D (existing data) and early childhood 25(OH)D (to be assayed as part of this project). If there is a moderate to strong relationship, we will adjust our models for maternal and/or cord blood 25(OH)D in order to isolate the impact of postnatal vitamin D supplementation on these outcomes.

As a secondary analysis, we will examine whether vitamin D supplement use interacts with dietary vitamin D intake to impact with body composition and BMD in mid-childhood and early adolescence. We will examine whether vitamin D supplementation in infancy modifies any association of infant feeding method with body composition and BMD, and whether vitamin D supplementation in early childhood modifies associations of dietary vitamin D intake assessed at 2 and 3 years with these outcomes. This will help us to understand whether early vitamin D supplementation can help to correct dietary deficiencies that may contribute to adverse outcomes.

The overall aim of this supplemental study is to isolate the effect of early vitamin D status on later health outcomes. Therefore, we will consider adjusting the models for supplement use at the time of outcome measurement (reported by the mother in mid-childhood and the child in early adolescence as frequency of multivitamin use in the past month), as well as 25(OH)D levels in mid-childhood, which we already measured on 655 Project Viva children in 2011. (These earlier 25(OH)D assays were done at Massachusetts General Hospital by liquid chromatography-tandem mass spectrometry [LC-MS]. The method used was an isotope dilution, LC-MS assay optimized based on published procedures [18]. The limit of detection was 5 nmol/L for D2 and 7.5 nmol/L for D3. The between-run CV for a quality control serum containing a total 25(OH) D concentration of 57 nmol/L was 7.5%.) We will evaluate the correlation between 25(OH)D levels in early and mid-childhood once the proposed early childhood assays are completed by the CERLab.

4. Limitations
While we have detailed information on vitamin D supplement use in infancy and can estimate a continuous measure of vitamin D supplement use at 6 months of age, we have more limited information on supplement

use at the later time points. We will focus our research on the impact of varying levels of vitamin D supplementation as a complement to the infant diet (breast milk vs. formula) on plasma 25(OH)D concentrations measured in early childhood and later body composition and bone health. In addition, we will use our available data to explore the impact of continued and/or later use of vitamin D-containing supplements reported at 1, 2, and 3y, on the same outcomes and as modifiers of associations of dietary vitamin D intake with these outcomes.

We also acknowledge that vitamin D obtained from the diet and supplements is only one determinant of 25(OH)D concentrations and that endogenous vitamin D synthesis can also be a major contributor to vitamin D status. Since all participants were enrolled in the Boston area, their ambient UV exposure will tend to be similar, but individual exposure will depend on other factors. To address this issue, we will consider adjusting our models for individual factors that may affect UV exposure and consequent vitamin D synthesis, including skin type, screen time, and physical activity.

Finally, we recognize that we are proposing a substantial amount of work for a 1-year project. The funds from this supplement will support the 25(OH)D assays and data cleaning and calculation of infant supplemental vitamin D intake, and we are confident that we can complete these activities within the 1-year supplement period. We also expect to be able to complete most of the analysis during this year, making use of previously collected data and drawing on our extensive experience working with the Project Viva data. However, if data analysis and manuscript preparation extend past the end of the funding period, the parent R01, which is funded through January 2022, will support our effort in completing this supplemental project.

5. Anticipated Project Timeline

If this project were funded in April/May 2019, we are prepared to begin 25(OH)D assays on the early childhood plasma samples immediately. While the lab conducts the assays, we will simultaneously perform data cleaning to estimate a more precise measure of vitamin D supplement intake at each time point, especially in infancy as we have additional detail on the timing of supplement use (see Table 2). We can also begin preliminary data analysis. After receiving the 25(OH) assay results, we will allow approximately six months to complete data analysis for all three aims and an additional two months for manuscript preparation and submission. It is thus feasible to complete this project within the 1-year funding period.

Table 5: Project Timeline

	April 2019	May 2019	June 2019	July 2019	Aug 2019	Sept 2019	Oct 2019	Nov 2019	Dec 2019	Jan 2019	Feb 2019	Mar 2019
25(OH)D assays												
Data cleaning (calculate vitamin D supplement intake)												
Data analysis												
Manuscript preparation												

Summary: The proposed supplemental project will provide a cost-effective opportunity to study the impact of using widely accessible dietary supplements of vitamin D, which is emerging as a potential key determinant of many health outcomes across the life-course, to mitigate the impact of the inadequate vitamin D levels that are common in U.S. children.