EARLY-LIFE EXPOSURES AND ADULT CANCER RISK: A LIFE COURSE APPROACH TO CANCER PREVENTION

by

Megan A. Clarke, MHS

A dissertation submitted to Johns Hopkins University in conformity with the requirements for the degree of Doctor of Philosophy.

Baltimore, Maryland
December, 2015

Abstract

Background

Despite emerging evidence indicating the potential importance of early-life exposures for adult cancer risk, there is limited research investigating cancer risk factors in early-life. The goals of this dissertation are to 1) elucidate whether maternal adiposity influences epigenetic processes in the offspring relevant to obesity and carcinogenesis and 2) inform primary cancer prevention strategies by addressing two modifiable, early-life risk factors: human papillomavirus (HPV) in males and unhealthy diet in postpartum teens.

Methods

Study 1: We evaluated the association of maternal pre-pregnancy body mass index (BMI) and gestational weight gain (GWG) with umbilical cord blood DNA methylation in a prospective study of 112 black and white mothers and infants, enrolled in Baltimore, MD, 2006-2007. Study 2: We identified predictors of HPV vaccination using electronic medical record data from 14,688 males aged 11-26 years in Maryland, 2012-2013. Study 3: We examined associations of perceived school and home food environments with dietary behaviors using baseline data from 853 postpartum teens enrolled in a weight-loss intervention study across 27 states, 2007-2009. Questionnaire items measuring perceived access to healthful items were used to categorize environments as "positive" or "negative".

Results

Study 1: Maternal pre-pregnancy BMI and GWG were significantly associated with DNA methylation in several CpG sites within 17 candidate genes. A majority of these associations were sex-specific.

Study 2: Approximately 15% of males initiated the HPV vaccine. Non-Hispanic black males (vs. non-Hispanic white) and publicly insured males (vs. private), were more likely to initiate the HPV vaccine, but less likely to receive subsequent doses. Frequent clinic visits (>3) were associated with increased uptake of all three doses.

Study 3: A positive school environment was related to healthful eating behaviors such as fruit consumption. In contrast, a positive home environment was associated with frequent consumption of a wider variety of healthful items as well as infrequent consumption of unhealthful food and beverages.

Conclusion

Early-life is an important, yet understudied period with respect to cancer risk. A better understanding of early-life factors from both an etiologic and primary prevention perspective will help to inform interventions that may substantially impact current cancer prevention strategies.

THESIS COMMITTEE ADVISORS AND READERS

Corinne Joshu, PhD, MPH, MA
Advisor, Department of Epidemiology

Elizabeth Platz, ScD, MPH
Thesis Committee and Reader, Department of Epidemiology

Gypsyamber D'Souza, PhD, MPH, MS
Thesis Committee and Alternate Reader, Department of Epidemiology

Srinivasan Yegnasubramanian, MD, PhD
Thesis Committee, School of Medicine

Paul Strickland, PhD, MSc
Thesis Reader, Department of Environmental Health Sciences

Betty Chou, MD
Thesis Reader, School of Medicine

Ana Navas-Acien, MD, PhD, MPH
Alternate Thesis Reader, Department of Environmental Health Sciences

Meghan Moran, PhD, MA
Alternate Thesis Reader, Department of Health, Behavior and Society

Acknowledgments

First and foremost, I would like to express my sincerest gratitude to my advisor, Dr. Corrie Joshu. Dr. Joshu has been a mentor, confidant, and someone whom I respect and look up to a great deal. Despite her myriad of other responsibilities, Dr. Joshu selflessly invested so much of her time and energy in my academic and long-term career development. She gave me the freedom to make decisions about my research, but always made herself available whenever I had questions or needed her guidance. Dr. Joshu had a special talent for devising eloquently simple solutions to things I tended to overcomplicate, and she was exceptionally good at reigning me back me in when I ventured too far off course in my analyses. Her calming presence in an otherwise stressful environment kept me grounded and enabled me to pursue my dissertation research with ongoing enthusiasm and focus.

I would also like to thank Dr. Elizabeth Platz for her mentorship and intellectual guidance, for giving me access to her data, and for all the work that she does to make the Cancer Epidemiology, Prevention, and Control training program such an enriching experience. She is the definition of a leader, and her remarkable dedication, breadth and depth of knowledge, and boundless energy and enthusiasm encompass what I aspire to as a junior scientist. I am extremely grateful to have been able to work alongside both Drs. Joshu and Platz, who have been and will continue to be, significant female role models for me in my career.

I would next like to thank the other members of my thesis advisory committee, Drs. Gypsyamber D'Souza and Vasan Yegnasubramanian, for generously contributing their time, thought, and guidance along the way. Thank you to Drs. Paul Strickland, Betty Chou, Ana Navas-Acien, and Meghan Moran, for your willingness to serve as thesis readers. Thank you to my collaborators Drs. Betty Chou, Darcy Phelan-Emrick, and the rest of the HPV vaccine team who helped me make sense of a crazy dataset, Dr. Debra-Haire Joshu and Cindy Schwarz, for sharing your data with me and helping me break into a completely new field, and finally, Dr. Vasan Yegnasubramanian, Dr. Sarah Wheelan and David Esopi, who were instrumental in generating the methylation data for my first aim. The breadth and wisdom of this multidisciplinary team of collaborators has greatly enhanced my doctoral training experience, and has made me a more well rounded cancer epidemiologist.

I am profoundly grateful for my funding support: the National Cancer Institute, T32 CA009314 and the Doctoral Thesis Research Fund (Department of Epidemiology). I am also deeply appreciative of the support I received in the beginning of my academic career from the National Institute of Allergy and Infectious Disease, T32 AI050056. To Drs. Jacky Jennings and Susan Sherman, thank you for helping me navigate both the academic and institutional environments early on, and for showing me that work-life balance is possible. To Matt Miller, thank you for always having the answer to all my financial questions, and for all that you do to keep students funded.

I would like to recognize the following people whom have I have not yet acknowledged elsewhere: Fran Burman, Ayesha Khan, Allyn Arnold, and the rest of Epi

academic support core, thank you for your administrative support and all that you do to make students' lives a little easier; to Dr. Fannie Fonseca-Becker, I am grateful for your mentorship in the Johnson & Johnson Community Healthcare Scholars' program and for helping to stimulate my interest in childhood obesity; to John Barber, thank you for your statistical consults and advice; and to Mark Schiffman, thank you for your brutal honesty and continued friendship and mentorship along the way.

Finally, I would like to express my gratitude to my friends and family. To my friends, particularly Liza Bronner Murrison, Rebecca Pierce, Kristen Little, and Shaun Truelove, you have made this experience far more enjoyable than it ever would have been without all of your support, friendship, laughter, and commiseration. To Shan Andrews, thank you for being a great friend and even better colleague. To my wonderful parents, Gary Clarke and Donna Viscardi, you have instilled in me the value of an education and hard work. Thank you for your unconditional love and support, and for all that you do to enable me to follow my professional dreams. To my best friend and sister Katie, you are younger, yet so much wiser than I am. Thank you for keeping me calm, for expressing interest in what I do, and for always providing logic and common sense when I need it most. And lastly to my husband, Michael Corso, thank you for making me laugh during the stressful times, for struggling to listen when I talked about things like DNA methylation, and letting me rant about the obesity epidemic and low rates of HPV vaccination. Thank you for tolerating my grocery lists and for keeping our daily lives afloat through the busy times. Your unwavering support and confidence in my

capabilities have carried me through this PhD process, and I never could have accomplished this without you by my side.

Table of Contents

List of Tables

List of Figures

CHAPTER 1

Introduction & Background

Introduction

Current evidence suggests that more than 50% of the approximately 1.6 million incident cancer cases diagnosed annually in the United States (U.S) can be attributed to a small number of key preventable and modifiable risk factors[1]. As shown in Table 1-1, these include tobacco use, risk factors related to energy balance such as obesity, poor diet, and physical inactivity[2,3], alcohol, and exposure to infectious agents, such as the human papillomavirus[4].

Table 1-1. The Most Common Modifiable Causes of Cancer in the United States[a]

Risk Factor	Cancer Sites	Prevalence of Risk Factor	% of Cancer Caused[b]
Tobacco Use	Lung, mouth, larynx, pharynx, esophagus, stomach, colon/rectum, pancreas, kidney, bladder, cervix, ovary, myeloid leukemia	18% of adults and 16% of teens smoke cigarettes	33
Overweight & Obesity	Breast, endometrium, colon/rectum, kidney, pancreas, esophagus, gallbladder, ovary, thyroid, possibly prostate	Nearly 66% of U.S. adults and 33% of teens are overweight or obese	20
Poor Diet	Breast, colon/rectum, possibly pancreas	70% of adults and >75% of teens do not meet daily fruit/vegetable recommendations	5
Physical Inactivity	Breast, endometrium, colon	50% of adults and 75% of teens do not meet daily physical activity recommendations	5
Alcohol	Oral cavity, pharynx, larynx, esophagus, liver, colon/rectum, and breast	5% of adults are heavy drinkers (≥2 per day for men, ≥1 per day for women)	3
Human Papillomavirus	Cervix, Oropharynx (base of tongue, tonsil, larynx), Vulva, Vagina, Penis	69% of females and 58% of males received all 3 doses of HPV vaccine	5

Data Source: [a]American Cancer Society Cancer Prevention & Early Detection Facts &

Figures[1], [b]Adapted from Colditz et al., 2012[5]

In 1998, the American Cancer Society (ACS) set a goal to reduce cancer incidence by 25% by the year 2015[6]. Based on trends estimated in a midpoint assessment in 2009, U.S. cancer incidence rates were projected to fall 50% short of the ACS objective[7]. The most recent Annual Report to the Nation on the Status of Cancer shows cancer mortality rates have declined 1.8 percent per year among men and 1.4 percent per year among women from 2002 to 2011[8]. This decline can be attributed to several successful interventions, most notably, decreased smoking prevalence, as well as increased screening and early detection and decreased use of prostate specific antigen testing for prostate cancer screening[9]. While these data are encouraging, rates for other cancers have increased over the past several years; many of which are associated with modifiable risk factors, including obesity (endometrium, esophagus, pancreas, liver, and thyroid) and infectious agents (liver and HPV-associated oropharyngeal cancer) (Table 1-2). Due to population growth and aging, it is projected that the number of people living with cancer will double to 2.6 million by the year 2050[10].

Table 1-2. Estimated Number of New Cancer Cases and Deaths in 2015, Incidence Trends from 2007-2011 for Select Cancers

Cancer Site	Estimated Number of New Cases, 2015	Estimated Number of Deaths, 2015	Trends in Incidence Rates, 2007-2011
Breast (Invasive)	231,840	40,730	Stable in white women; increase of 0.3% per year in black women
Cervix	12,900	4,100	Decline of 3.4% per year

Colon and Rectum	132,700	49,700	Decline of 4.5% per year in adults ≥50 years; increase of 1.8% per year in adults < 50 years
Endometrium	54,870	10,170	Increase of 2.4% per year
Kidney	61,650	14,080	Stable after increasing over past several years
Liver	35,660	24,550	Increase of 3.4% per year
Lung	221,200	158,040	Decline of 3% per year in men and 2.2% in women
Oropharynx	45,780	8,650	Increase of 1.3% per year in white men (driven by HPV) and stable in white women; decline of 3% per year in black men and 1.4% per year in black women
Ovary	21,290	14,180	Decline of 2% per year in white women; stable in black women
Pancreas	48,960	40,560	Stable after increasing over past 10 years
Prostate	220,800	27,540	Stable in men < 65 years; decline of 2.8% per year in men ≥ 65 years
Thyroid	62,450	1,950	Increase of 4.4% per year
Urinary bladder	74,000	16,000	Decline of 1.6% per year

[a]Estimates adapted from the American Cancer Society's Cancer Facts & Figures, 2015[1]
Data Source: The North American Association of Central Cancer Registries and Surveillance, Epidemiology, and End Results registries.

The growing trend in incidence and mortality for cancers with known links to preventable risk factors underscores the need for new approaches to cancer prevention. These strategies must be informed with a greater understanding of cancer etiology and how to promote lifestyle behaviors that will achieve the greatest benefit. To date, research in cancer epidemiology has primarily focused on adult populations; yet a growing body of evidence suggests exposures in early-life are important for adult cancer risk[5,11]. This growing body of evidence has important implications for current cancer prevention strategies, suggesting a need to shift the focus to risk factors existing in earlier stages in the life course[11]. There are significant challenges to studying early-life exposures including the limited availability, validity, and reliability of early-life exposure measurements. In addition, the long latency periods and lack of intermediate endpoints associated with most cancers make it difficult to assess early-life exposures using traditional epidemiologic methods, particularly given the limited number of long-term biomarkers that have been identified[11]. Because of the obvious need for more evidence, the National Cancer Institute convened several workshops to review opportunities for cancer prevention in early-life, and in 2015, released a Funding Opportunity Announcement (PA-15-126) to stimulate research on early-life exposures and adult cancer development.

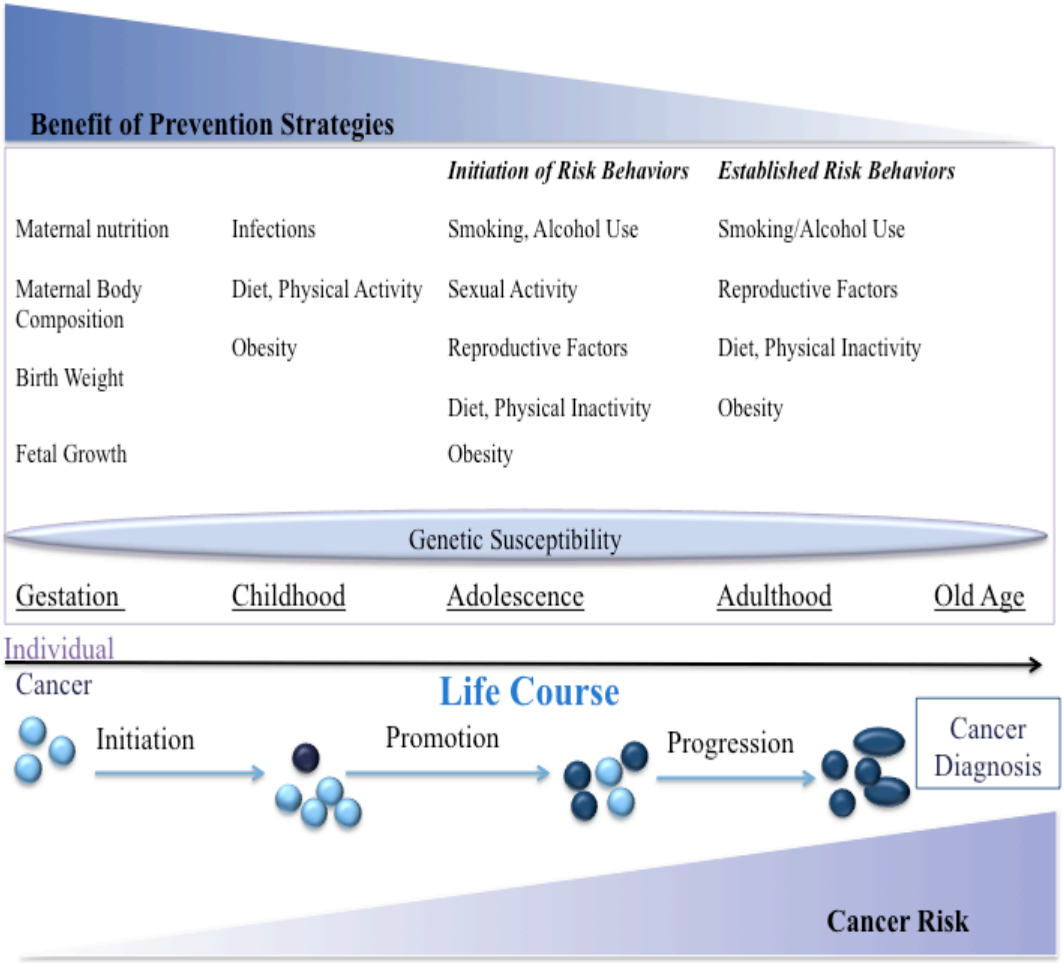

Figure 1-1. Conceptual Framework. Cancer risk factors occur in all stages of development, and are typically initiated in early-life. Similarly, after initiation, cancer typically develops over a period of several years, suggesting most of the underlying biology occurs earlier in life. Thus, targeting cancer prevention efforts in early-life, before behaviors are fully established, is likely to achieve the greatest benefit.

A life course perspective considers modifiable factors operating during different stages of development and the implications for cancer risk later in life[12]. The conceptual framework for this dissertation is shown in Figure 1-1, which depicts the life course as a continuum of cancer risk in the context of the multistage model of carcinogenesis. The underlying biology of the process from initiation to clinical detection typically occurs over a period of several years, with the risk of cancer increasing markedly as people

age[13,14]. Likewise, several modifiable cancer risk factors are initiated during early-life, and become increasingly harder to change as behaviors become established in adulthood. By overlaying the two trajectories, it is clear that prevention strategies targeting earlier stages in life may achieve the greatest benefit.

The purpose of this dissertation is to better understand known modifiable early-life cancer risk factors, from both an etiologic and primary prevention perspective. Even in the absence of cancer outcomes, important gains can be made in informing cancer prevention strategies[15]. Overcoming the methodological challenges associated with studying early-life exposures requires innovative use of existing data. This research leveraged data from three independent studies, each focusing on a modifiable exposure during a different critical period of early-life. The goals were to 1) elucidate how early-life risk factors, such as maternal obesity and weight gain, influence epigenetic processes in offspring relevant to obesity and carcinogenesis (Chapter 2) and 2) generate evidence that will inform primary cancer prevention strategies targeting two modifiable, early-life risk factors: human papillomavirus in males (Chapter 3) and unhealthy diet in a high-risk population of postpartum adolescents (Chapter 4).

Background

In Utero Exposure to Maternal Adiposity: Epigenetic Mechanisms and Implications for Cancer Risk

Obesity is a Modifiable Cancer Risk Factor

It has been estimated that approximately 20% of all cancers are attributed to overweight and obesity, defined as a body mass index (BMI) of 25 –29.9 kg/m^2 and ≥ 30

kg/m^2, respectively[16]. According to the American Institute for Cancer Research and the World Cancer Research Fund's (AICR/WCRF) report, "*Food, Nutrition, Physical Activity and the Prevention of Cancer: A Global Perspective*", obesity is an established risk factor for cancers of the breast (postmenopausal), endometrium, colon and rectum, kidney, pancreas, esophagus (adenocarcinoma), and a probable risk factor for cancer of the gallbladder[17]. Since this publication, there is now convincing evidence supporting an association of obesity with risk of liver cancer and aggressive prostate cancer[18,19], as well as a probable association with increased risk of ovarian cancer, leukemia, non-Hodgkin's lymphoma, and multiple myeloma[1]. Obesity also influences cancer progression and survival after diagnosis, with greater risk of death likely reflecting both biological effects and delays in screening and detection[15]. Based on data from a landmark study of over 900,000 U.S. adults enrolled in the early 1980's, approximately 15–20% of all cancer deaths have been attributed to obesity[15]. It is likely that these data now represent conservative estimates, as the population prevalence of obesity has nearly doubled since 1980[20,21].

Biological Mechanisms Linking Obesity to Cancer Risk

There are several proposed biological mechanisms linking obesity to cancer risk. The most well studied involve pathways related to adipokine production, chronic inflammation, insulin resistance and signaling, and sex hormone metabolism (Figure 1-2)[22–24]. Adipose tissue is an important endocrine and metabolic organ that regulates energy balance and lipid metabolism by releasing free fatty acids (FFAs) and secreting a variety of peptide hormones such as leptin and adiponectin, collectively known as

"adipokines"[23]. Leptin is a potent hormone involved in appetite regulation and tends to be secreted at higher levels in obese individuals[23]. Increased leptin levels have been associated with immune suppression, promotion of angiogenesis and inhibition of apoptosis in cancer cell lines[25]. Additional laboratory evidence suggests that leptin signals through critical pathways involved in cell proliferation and differentiation (e.g., PI3K/Akt, MAPK, JAK/STAT)[19]. Adiponectin production on the other hand, tends to be lower in obese individuals, and has been shown to have anti-proliferative effects[26]. Support for the protective role of adiponectin in carcinogenesis comes from epidemiologic studies showing consistent inverse associations of circulating adiponectin with risk of endometrial, breast, prostate, colorectal, renal and pancreatic cancers[23]. In addition to adipokine regulation, expansion of adipose tissue tends to increase production of pro-inflammatory cytokines including tumor necrosis factor alpha (TNFα), interleukins IL-1β and IL-6, and the pro-inflammatory transcription factor, NF-κB, creating a chronic state of inflammation, which may increase cancer risk[23,24] (Figure 1-2).

In overweight and obese individuals, reduced uptake of fatty acids by adipose tissue combined with increased breakdown of lipids (i.e., lipolysis) leads to higher levels of circulating FFAs and promotion of insulin resistance and hyperinsulinemia[22–24]. Hyperinsulinemia decreases production of insulin-like growth factor binding proteins 1 and 2 (IGFBP-1, IGFBP-2, respectively), resulting in higher levels of unbound, circulating IGF-1[27]. Activation of both the insulin and IGF-1 receptors induces signaling pathways that promote cell proliferation and inhibit apoptosis, two hallmarks of carcinogenesis[27] (Figure 1-2).

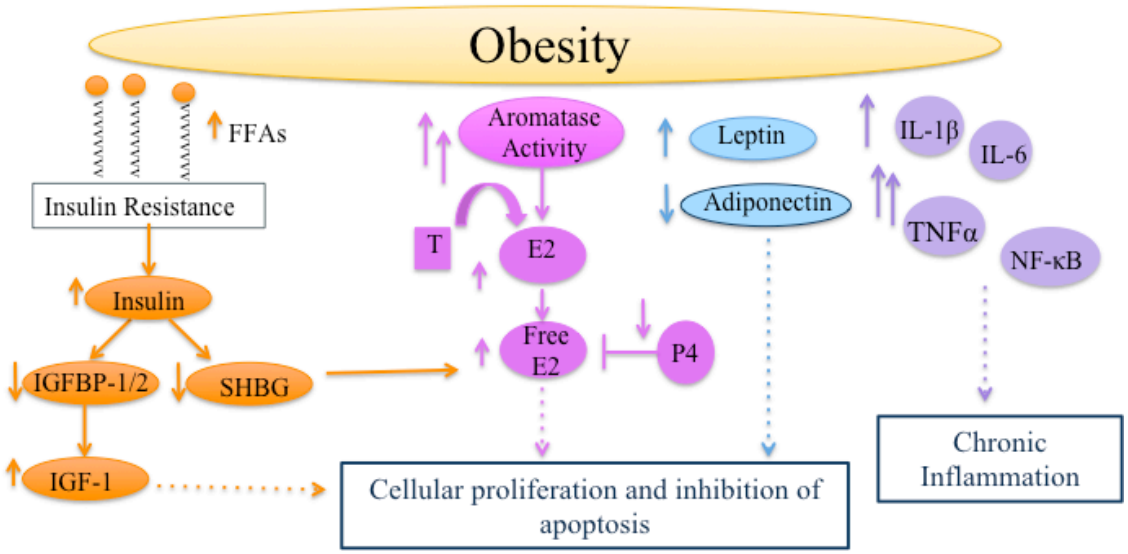

Figure 1-2. Potential Molecular Mechanisms Underlying the Association of Obesity with Cancer Risk. Obesity increases circulation of FFA's, which can lead to insulin resistance and hyperinsulinemia. Hyperinsulinemia decreases concentrations of IGFBP-1 and -2 as well as SHBG. Decreased levels of IGFBP-1 lead to increased IGF-1 levels, which promote cellular proliferation and inhibit apoptosis. Aromatase activity is enhanced in obese individuals. Aromatization of testosterone leads to increased levels of freely circulating E2, particularly when SHBG concentrations are low. Unbound E2 can bind to nuclear receptors to increase cell proliferation and inhibit apoptosis. Progesterone, which typically counteracts effects of estrogen, can be low in obese individuals. Increased leptin and decreased adiponectin concentrations associated with obesity can promote cell proliferation and inhibition of apoptosis. Pro-inflammatory cytokine concentrations are also increased, leading to a chronic state of low-grade inflammation. Abbreviations: FFAs, free fatty acids, IGFBP-1/2, insulin-like growth factor binding protein 1 and 2; IGF-1, insulin-like growth factor-1; SHBG, steroid hormone binding globulin; T, testosterone; E2, estradiol; P4, progesterone; IL, interleukin; TNFα, tumor necrosis factor alpha; NF-κB, nuclear factor kappa-light-chain-enhancer of activated B cells

Another consequence of hyperinsulinemia is reduced synthesis of sex-hormone-binding globulin (SHBG), which binds to sex-steroid hormones, testosterone and estradiol[22]. As an endocrine organ, adipose tissue expresses aromatase, which converts testosterone and Δ^4-androstenedione to estradiol and estrone, respectively[28,29]. The aromatization of androgens increases with obesity and age in both men and women and leads to increased levels of freely circulating estrogen, particularly when SHBG levels

are low as in the case of hyperinsulinemia[22,28]. In certain tissues such as breast and endometrium, bioavailable estrogen can bind to its receptor and stimulate cellular differentiation and proliferation and inhibit apoptosis[30,31]. Data from prospective studies among postmenopausal women have demonstrated that increased levels of freely circulating estrogen largely explain the relationship between obesity and breast and endometrial cancer risk[23]. Among premenopausal women, the presence of progesterone typically acts to counterbalance the proliferative and anti-apoptotic effects of estrogen, in part by increasing synthesis of IGFB-1[32]. However, obesity is associated with anovulation and decreased progesterone levels that are not sufficient to counterbalance the effects of estrogen[31]. This is commonly referred to as the "*Unopposed Estrogen Hypothesis*", and suggests that loss of progesterone may be the most important hormonal risk factor for endometrial cancer in premenopausal obese women[22,23,31].

Evidence for Associations of Obesity in Early-Life with Cancer Risk

Nearly one-third of children aged 2-19 years are overweight or obese[33]. Overweight and obesity during childhood and adolescence have been associated with risk for certain cancers. These associations may be direct, as a result of cumulative exposure to the physiological consequences of obesity over the life course, or indirect, as early-life overweight and obesity increase the risk for obesity in adulthood[22]. The most compelling evidence supporting a direct link between early-life obesity and development of cancer in adulthood comes from numerous prospective studies showing that high BMI in childhood/adolescence is associated with decreased risk of premenopausal breast cancer

later in life[34]. The mechanisms underlying this association are not fully understood, but have been attributed to a greater frequency of anovulatory cycles in overweight and obese females, resulting in lower lifetime estrogen exposure[35]. Aside from breast cancer, there have been few prospective studies assessing early-life obesity or the effect of cumulative obesity exposure across the life course with respect to subsequent cancer risk. Findings from the NIH-AARP Diet and Health Study, a large, prospective cohort study of U.S. adults, suggest a significant association of obesity duration and cumulative exposure to obesity over a lifetime with pancreatic cancer risk[36] and a significant association of overweight and obesity across the life course with increased risk of multiple myeloma[37]. Among women enrolled in the Nurses' Health Study II (NHS II), early-life overweight and obesity and weight gain since age 18 were significantly associated with endometrial cancer risk later in life[38]. Findings from studies on early-life BMI and risk of colorectal cancer have been mixed, with some reporting no association[39–41] and others finding associations of high BMI in adolescence and/or early-adulthood with increased risk of colorectal cancer in men[42–44] as well as a positive association of weight gain since early-adulthood with colorectal cancer risk[41]. In a series of population-based studies linking school health records of over 140,000 children to Danish Cancer Registry data, childhood BMI was associated with increased risk of liver cancer, thyroid cancer, and esophageal adenocarcinoma in adulthood[45–47] while childhood height, but not BMI, was associated with future prostate cancer risk[48,49].

In Utero Exposure to Obesity and Implications for Cancer Risk

The idea that *in utero* exposures may increase the risk of disease later in life is commonly referred to as the "*Developmental Origins of Health and Disease"* hypothesis. This hypothesis was originally proposed by David Barker in the late 1980's based on evidence linking low birth weight with increased risk of coronary heart disease and metabolic syndrome in adulthood[50,51]. Barker and colleagues hypothesized that the *in utero* period was a critical window of developmental plasticity during which environmental exposures could lead to adaptive responses that permanently alter offspring development[51]. While the foundation for this hypothesis is largely rooted in cardiovascular disease research, subsequent studies have evaluated this hypothesis in the context of cancer risk. Classic, quasi-experimental studies of exposure to severe caloric restriction during the Dutch Winter Famine in World War II, have demonstrated an association of famine exposure during early gestation with subsequent breast cancer risk among female offspring, possibly due to rapid postnatal catch-up growth[52,53]. The increasing prevalence of obesity among reproductive-aged women as well as children as young as 2 years of age[21,33], has prompted studies investigating the potential harmful effects of maternal overweight and obesity and weight gain on fetal development. Initial sibling studies among Pima Indians, a group with a very high prevalence of obesity and diabetes, provided strong evidence for an association of diabetes during pregnancy with higher offspring BMI and risk of type 2 diabetes later in life[54–56]. Similarly, studies of siblings born before and after maternal bariatric surgery, demonstrated an increased risk of overweight and obesity among children who were born before surgery compared with their siblings born after surgery[57,58]. It is now widely recognized that both maternal

obesity and gestational diabetes are independent risk factors for increased birthweight and body size later in life[59–63]. Moreover, high birthweight, although not a perfect marker for the *in utero* environment, has been linked to adiposity later in life and increased risk of premenopausal breast cancer and testicular cancer[34,64,65]. This cycle of obesity risk has important implications for subsequent cancer development, and warrants further investigation into the biological mechanisms underlying these associations.

Maternal Metabolism During Pregnancy and the Potential Role of DNA Methylation in Fetal Programming

During pregnancy, the maternal metabolism adapts to meet the nutritional needs of the growing fetus[66]. In normal weight women, maternal fat accumulation primarily occurs during early gestation, followed by a switch to lipolysis around the third trimester, resulting in high levels of circulating FFAs and glycerol[67]. Fatty acid transport to the fetus is mediated by the placenta and is maximal in late gestation[68,69]. In obese women, the shift to lipolysis occurs earlier in gestation, thus increasing the availability of circulating triglycerides and FFAs that can be transported to the fetus[70] Interestingly, maternal triglyceride levels have been shown to be independently associated with infant birth weight (Figure 1-3)[71,72]. Lipids have the ability to activate cell-signaling pathways and to bind to nuclear receptors, suggesting that increased fetal lipid exposure may affect gene expression in pathways related to energy storage, oxidation, growth, and inflammation[66]. Support of this hypothesis comes from a recent study of pregnant women with gestational or type 1 diabetes showing selective upregulation of genes involved in fetal-placental lipid metabolism in women with gestational diabetes and

obesity (but not in women with type 1 diabetes)[73]. While the underlying biological mechanisms of these associations are still poorly understood, it is likely that epigenetic processes, such as DNA methylation, may play a role[66].

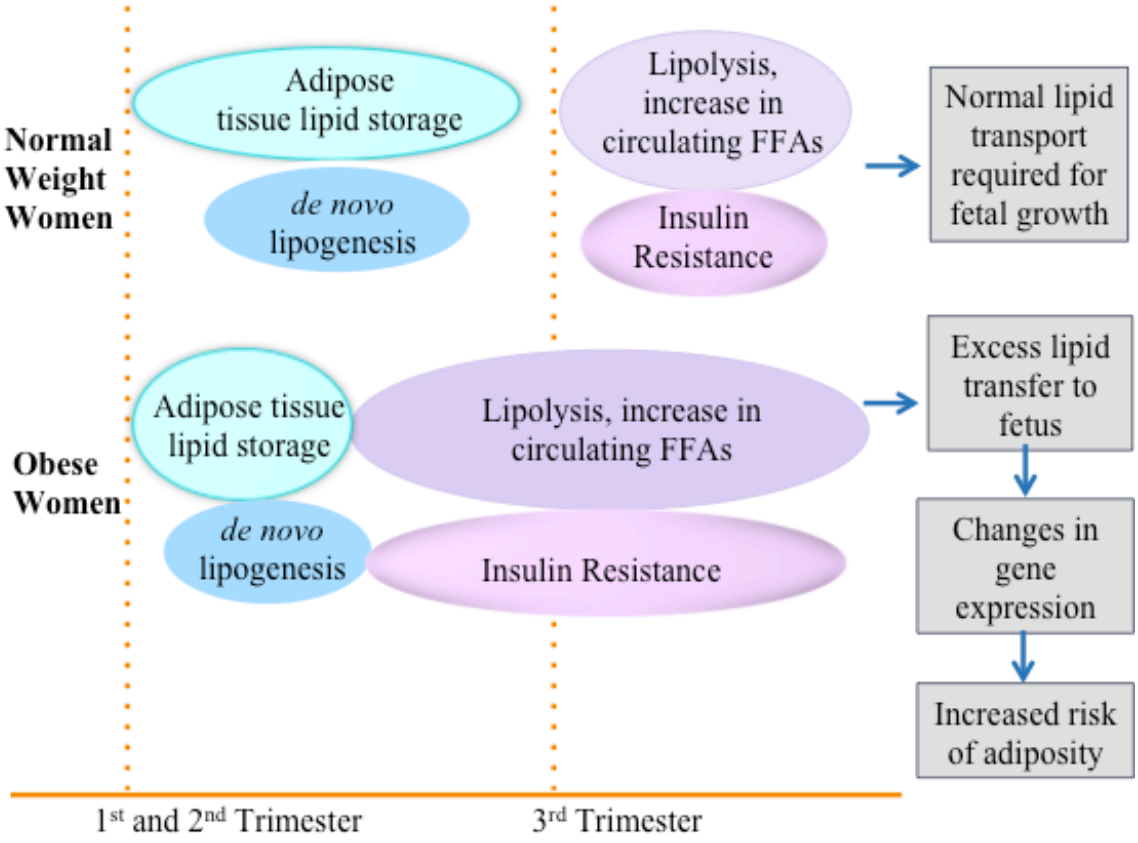

Figure 1-3. Maternal Metabolism During Pregnancy in Normal Weight and Obese Women. In normal weight women early pregnancy (1st and 2nd trimester) is characterized by maternal fat accumulation and lipid storage, followed by a switch to lipolysis and lipid breakdown in late pregnancy (3rd trimester), resulting in adequate levels of circulating FFAs and glycerol molecules transported to the fetus for normal development. In obese women, this switch to lipolysis happens earlier in pregnancy, resulting in increased delivery of triglycerides and FFA's to the fetus. Binding of lipids to nuclear receptors can lead to changes in gene expression, which are hypothesized to increase the risk of offspring adiposity. Abbreviations: FFAs, Free fatty acids

DNA methylation is essential for processes involved in human embryonic development, such as genomic imprinting and X chromosome inactivation[74,75]. In

humans, DNA methylation is facilitated by a family of DNA methyltransferase enzymes that catalyze the addition of a methyl group to cytosines at the 5' position of a cytosine and guanine dinucleotide pair termed "CpG"[76]. In human diploid genomes, any specific CpG site in a single cell can be either methylated, partially ("hemi") methylated, or unmethylated[74]. The quantitative value of methylation at a given CpG site is generally represented as the fraction of sites that are methylated, expressed as a proportion or percent. In promoter regions, binding of a methyl group can alter chromatin conformation and transcriptional binding site affinity, often resulting transcriptional silencing. Methylation of CpG-rich stretches of DNA located in promoter regions of genes, termed "CpG Islands," are essential for regulation of gene expression, while disruption of CpG island methylation has been documented in malignant cellular transformation[76]. Patterns of DNA methylation are established during embryogenesis through a highly dynamic process in response to genetic and environmental cues[77]. Upon fertilization, there is rapid, genome-wide DNA de-methylation, with the exception of imprinted genes, such as *H19/IGF2*, which retain their parent-of-origin marks[75]. Prior to implantation, methylation patterns are re-established *de novo* and are typically stable throughout life[77]. Collectively, this suggests that the *in utero* period may be a critical window for epigenetic programming via DNA methylation, with potential long-term consequences for gene expression.

A critical step in demonstrating a role for DNA methylation as a mediator of adult disease is establishing an association between a prenatal exposure of interest and changes in offspring DNA methylation patterns at birth[78]. Several studies have accomplished this

first step by demonstrating that prenatal conditions, such as maternal stress[79], tobacco smoking[80–82], and arsenic exposure[83–85] can influence methylation patterns in cord blood DNA. A growing body of evidence from both animal and human studies have provided meaningful insights into the role of DNA methylation in mediating the association of maternal nutrition with future offspring risk of obesity later in life[77,86–88]. Classic, quasi-experimental studies of exposure to severe caloric restriction during the Dutch Winter Famine have demonstrated an association with differences in DNA methylation in genes involved in growth and metabolism including *IGF2*, interleukin-10 (*IL-10*) and leptin, up to six decades after exposure[89–91]. In a study conducted in rural Gambia, where nutritional status varies dramatically by season of conception, average birth weight among offspring conceived during the nutritionally challenged rainy season was 200–300 grams lower than the birth weight of offspring conceived during the dry (harvest) season[92]. In subsequent studies, season of conception was related to persistent differences in offspring DNA methylation patterns, evidenced in tissue from all three germ layers[87,88]. More recent data from both animal and human studies exploring the effects of gestational exposure to maternal obesity and excess gestational weight gain (GWG) support the role of DNA methylation in the regulation of key genes involved in adipogenesis, inflammation, growth and signaling[87,88,93–97]. In a very recent study of over 1,000 participants, offspring born to obese mothers had differential methylation patterns in several CpG sites compared with offspring born to normal weight mothers. Using paternal obesity as a negative control, the investigators also demonstrated a maternal-specific association of obesity with DNA methylation patterns, confirming that these

associations were related to intrauterine mechanisms. Interestingly, DNA methylation patterns associated with maternal obesity exposure tended to correlate with offspring adiposity, providing evidence for DNA methylation as potential mediator of the association of maternal obesity with offspring adiposity later in life[95].

HUMAN PAPILLOMAVIRUS VACCINATION IN MALES

Infection with Carcinogenic HPV is a Modifiable Risk Factor for Cancer

Infection with HPV is extremely common. In the U.S., approximately 80 million people are infected with at least one HPV type, and among individuals with at least one opposite sex partner the lifetime probability of acquiring an HPV infection is approximately 85% for females and 91% for males[98]. HPV infections are generally transmitted through sexual contact, or in some cases, through other intimate contact (e.g., oral-genital)[99–101]. Persistent infection with one of the 13 carcinogenic HPV types is causally associated with nearly all cases of cervical cancer, and a substantial proportion of anal, oropharyngeal, vaginal, vulvar, and penile cancers[4,102]. Of these 13 types, HPV16 is the most carcinogenic due to the activity of two oncogenes, E6 and E7, which interfere with tumor suppressor proteins p53 and pRb, respectively[102]. HPV16 causes nearly all cases of cervical cancer and is responsible for a substantial proportion of other HPV-associated cancers in the anogenital tract and oropharynx. HPV18 is the second leading carcinogenic type, and is most commonly associated with cervical adenocarcinomas. Together HPV16 and HPV18 cause approximately 70% of all cervical cancers[103].

Genital HPV Infection and Carcinogenesis

Most of the evidence regarding HPV natural history comes from studies of cervical infection. Genital HPV infections are commonly transmitted by sexual contact, and the likelihood of acquiring an HPV infection is highest within a few years of becoming sexually active. Prevalence estimates calculated before vaccine introduction from National Health and Nutritional Examination Surveys (NHANES, 2003-2006) data suggested an overall HPV prevalence of about 42.5% among U.S. females aged 15 to 59 years. These estimates varied by age group, with prevalence significantly increasing from 32.9% in females aged 14 to 19 years to a peak prevalence of 53.8% in females aged 20 to 24, with subsequent declines to about 38.8% in females aged 50-59 years[104]. Regardless of the age of acquisition however, most genital HPV infections clear within one to two years after exposure; with the rate of clearance decreasing the longer HPV persists[102]. For the 5-10% of carcinogenic HPV infections that persist more than two years, the risk of cervical precancer (i.e., cervical intraepithelial neoplasia grade 3, CIN3) is substantially increased[105–107]. The time course for progression to CIN3 is relatively short, with peak incidence occurring 5 to 15 years after infection, as opposed to most cancers which occur over subsequent decades[102]. Risk factors for progression to CIN3 include smoking, long-term oral contraceptive use, and multiparity[108–110]. Factors that influence the immune system, such as chronic inflammation and co-infection with HIV, are also associated with HPV persistence and progression to CIN3[111,112]. In the U.S., cervical cancer screening with cytology (i.e., Pap smear) is recommended for females 21

to 65 years of age. To lengthen screening intervals, females aged 30 to 65 years may be screened with a combination of cytology and HPV testing every five years[113].

Incidence rates of HPV-associated vaginal and vulvar cancers are much lower than those observed for cervical cancer. In the U.S., HPV has been associated with approximately 69% of invasive vulvar cancers and 97% of vulvar intraepithelial neoplasia (VIN) grade 3, with HPV16 accounting for approximately half of vulvar cancers and 81% of VIN3[114]. Limited data on vaginal cancers suggest HPV is associated with approximately 75% of all invasive cancers and approximately 90% of vaginal intraepithelial neoplasia (VaIN) grade 2/3, with HPV16 accounting for more than 50% of invasive cancers[115,116]. Unlike cervical cancer, most cases of vaginal cancers occur in women over age 60 years[116].

Penile cancer is very rare and most commonly occurs in men aged 50 to 70 years. In the U.S., HPV has been associated with approximately 50% of penile cancers, with HPV16 accounting for approximately 45% of HPV-associated cases[117]. Risk factors for HPV-associated penile cancer include smoking and being uncircumcised[4].

Anal HPV Infection and Carcinogenesis

Data on the natural history of anal HPV infection is limited, particularly among women. The few studies conducted among women have shown comparable prevalence estimates for both cervical and anal HPV infection. In a recent study of over 2,000 women enrolled in the Costa Rican Vaccine Trial, the prevalence of anal HPV infection was 32%, similar to the prevalence of cervical HPV infection in this population (37%)[118]. Studies of anal HPV infection among males are more common, particularly among HIV-

positive men who have sex with men (MSM). Limited data from heterosexual males without HIV infection, suggest an anal HPV prevalence of approximately 12%. In contrast, a pooled meta-analysis of 53 studies estimated a prevalence of anal infection with any HPV type to be approximately 93% among HIV-positive MSM and nearly 64% in HIV-negative MSM[119]. Similar to cervical histopathology, anal intraepithelial neoplasia (AIN) grade 2/3 is the precursor lesion of anal cancer[4]. In the U.S. the prevalence of HPV infection is 91% in anal cancers, with a majority (77%) positive for HPV16[120]. In one of the largest studies involving 34,189 HIV-positive and 114,260 HIV-negative participants, the incidence of anal cancer was 131 per 100,000 person-years among MSM compared with 46 per 100,000 person years among HIV-positive men, and 2 per 100,000 person-years among HIV-negative men[121]. Screening for anal cancer is not currently recommended in the U.S., although some clinics perform anal cytology in high-risk patients such as HIV-infected and uninfected MSM. Rates of AIN3 and invasive anal cancer have been increasing at a steady rate in almost all racial and ethnic groups and particularly among men (see Figure 1-4 below)[122,123].

Oral HPV Infection and Carcinogenesis

The prevalence of oral HPV infection is substantially lower than that of cervical HPV infection. Estimates from the most recent NHANES data (2009-2012) indicate a prevalence of approximately 7% in men and 1.5% in women aged 14-69 years. The prevalence curve of oral HPV infection by age appears to have a bimodal distribution, peaking at ages 25–30 years and again at 55–65 years. Risk factors for oral HPV infection include smoking and lifetime number of sex partners[124]. Interestingly, unlike

HPV infections at other anatomical sites, the rate of oral HPV acquisition among males tends to remain high regardless of age, and infections are less likely to clear as men get older[125,126]. The prevalence of HPV16 is very high with respect to other HPV types, and accounts for over 80% of all oral HPV infections[124] and causes more than 85% of HPV positive oropharyngeal cancers[127]. Limited data on oral HPV natural history suggest similar rates of persistence to genital HPV infections, with most infections clearing within one year[128,129]. Screening for HPV-associated oropharyngeal cancer is not currently recommended. Challenges to implementing a screening program include a lack of clinically validated oral HPV DNA-based tests, a lack of known oral HPV-associated precursor lesions, and the absence of a non-invasive intervention for reducing incidence and mortality[130,131].

Trends in HPV-Associated Cancers

In 2012, the Annual Report to the Nation on the Status of Cancer featured data on HPV-associated cancer incidence rates and short-term trends. Based on the most recent data available (2009), HPV-associated cancers were estimated to account for 3.3% of all cancer cases among females and 2.0% among males[122]. Figure 1-4 has been adapted to show trends in incidence of select HPV-associated cancers in the U.S. from 2000 to 2009. Due in large part to ongoing successful screening programs in the U.S., incidence rates of cervical cancer have significantly decreased among women in all racial and ethnic groups, with the exception of American Indian/Alaskan Native females, who experienced a non-significant increase in cervical cancer of 0.2% (Figure 1-4)[122]. Incidence rates of

HPV-associated anal cancer have increased for all racial and ethnic groups, with significant increases observed among white and black males and females. Interestingly, rates of HPV-associated oropharyngeal cancer were significantly increased among white males and females, but significantly decreased among black males[122]. In countries like the U.S. with ongoing cervical cancer screening programs, the incidence of both anal and oropharyngeal cancers are clearly on the rise, with the number of cases projected to exceed those of cervical cancer in the near future[132].

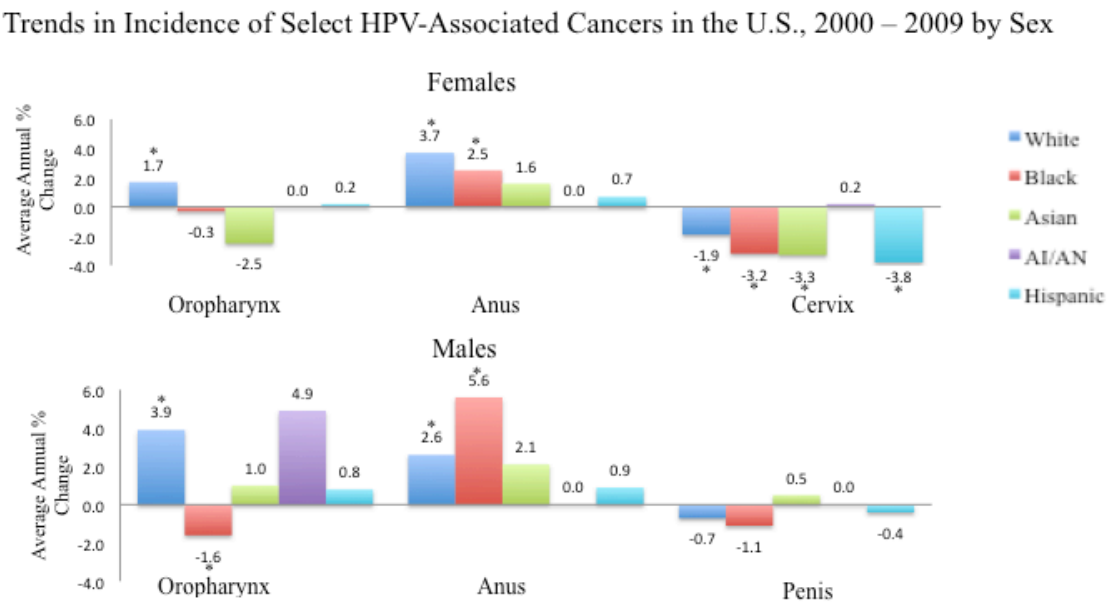

Figure 1-4. Trends in Incidence of Select HPV-Associated Cancers in the U.S., 2000-2009. For classifying HPV-associated cancers, the average annual number of HPV-associated cancers was multiplied by the percentage of each cancer type found attributable to HPV based on genotyping studies. For cervical cancers, all squamous cell carcinomas were selected. *Statistically significant trend in incidence rate ($p<0.05$). Data Source: The National Program of Cancer Registries (NPCR) and/or the NCI's Surveillance, Epidemiology, and End Results (SEER) program. Adapted from Jemal et al., 2012[119]
Abbreviations: HPV, Human Papillomavirus; AI, American Indian; AN, Alaskan Native

Primary Prevention of HPV-Associated Cancers

A bivalent (Cervarix, GlaxoSmithKline, Rixensart, Belgium) and quadrivalent (Gardasil, Merck and Co, Inc., Whitehouse Station, New Jersey) HPV vaccine were originally licensed for the protection against HPV-related disease[4]. The Cervarix vaccine protects against HPV16 and HPV18 and is licensed for use in females aged 9 to 26 years. The Gardasil vaccine protects against types HPV16 and HPV18, plus two low-risk types, HPV6 and HPV11, that are responsible for more than 90% of anogenital warts[133,134]. Data from phase III clinical trials suggest both vaccines are highly efficacious against HPV16- and HPV18-associated CIN 2/3 (primary endpoints) and persistent infection

with HPV16 and HPV18 (secondary endpoints)[135–137]. High efficacy was also observed for protection against HPV16 and HPV18 related VIN 2/3 and VaIN 2/3[138]. Very recently, Merck and Co., Inc released a nonavalent HPV vaccine (Gardasil 9) that protects against the four HPV types included in Gardasil, plus five other carcinogenic types that cause an additional 20% of HPV-associated cancers (HPV types 31, 33, 45, 52, and 58). In a phase III efficacy trial in approximately 14,000 females aged 16 to 26 years, efficacy for the prevention of CIN2 or worse associated with HPV types 31, 33, 45, 52, and 58 was over 95%, and 96% for persistent (6 month) infection with these types[139].

One large clinical trial was conducted to assess the efficacy of Gardasil in over 4,000 males aged 16 to 26 years. This study found high per-protocol efficacy for prevention of HPV6-, HPV11-, HPV16-, and HPV18-associated incident genital warts. In a subanalysis restricted to 602 MSM, per-protocol efficacy for prevention of AIN 2/3 associated with vaccine types was close to 75% and nearly 100% for anogenital warts associated with vaccine types. These findings led to FDA approval of Gardasil for the prevention of AIN and anal cancer for both males and females.

Safety data from both females and males aged 9 to 26 years indicate a larger proportion of injection-site adverse events among HPV-vaccinated groups compared with placebo groups. The most common adverse events included swelling, pain, and erythema (i.e., redness of the skin). There were no significant differences in serious adverse events reported between the HPV vaccine and placebo groups.

The duration of protection for the HPV vaccine has not yet been established, though follow-up data from a phase III Gardasil trial conducted in Nordic countries suggest sustained antibody titers up to 9 years after vaccination[140]. Protection from cervical HPV infection by less than three doses of Cervarix was evaluated in two phase III trials. A very recent combined analysis of these data suggest similar protection against HPV16 and HPV18 infection for one or two doses of the HPV vaccine compared with receiving all three doses in females aged 15 to 25 years[141]. Based on these data and cost-effectiveness analyses, the World Health Organization's Strategic Advisory Group of Experts on Immunization now recommends two doses of the HPV vaccine for females aged 9 to 14 years[142]. Following these recommendations, several countries have switched to two-dose regimens (at 0 and 6 months), including Switzerland, the Netherlands, Mexico, the United Kingdom, and parts of Canada (Quebec)[143].

Due in large part to the limited data on oral HPV natural history and the lack of a precursor lesion associated with oral HPV infection, vaccine efficacy against oral HPV and related disease is unknown. Recognizing this limitation, the WHO recently recommended that incident and persistent HPV infections could be used as surrogate endpoints for risk of HPV-positive oropharyngeal cancers in HPV vaccine trials[124]. Recent data from the Costa Rican Vaccine trial demonstrate high vaccine efficacy against oral HPV16[144–146], and in a recent analysis, vaccine efficacy for multisite infections was observed in females who had been previously exposed to HPV16/HPV18[144]. These data have important implications for future vaccine efficacy trials for the prevention of HPV-associated oropharyngeal cancers.

HPV Vaccine Recommendations in the U.S.

Gardasil was first licensed for females in 2006 and was recommended by the Advisory Committee on Immunization Practices (ACIP) for routine use in females aged 11 to 12 years with catch-up vaccination for females aged 13 to 26 years who had not been previously vaccinated[4]. Cervarix was subsequently approved in 2009 and both vaccines are now recommended by ACIP for use in females. Gardasil was licensed for males in 2009 and subsequently recommended by ACIP in 2011 for routine use in males aged 11 or 12 years with catch-up vaccination for males aged 13 to 21 years who had not been previously vaccinated. Males aged 22 to 26 years may also be vaccinated and HPV vaccination is recommended for MSM up to 26 years of age, including those who are HIV-positive[4]. In 2014, Gardasil 9 was licensed by the FDA for both females and males, and approved for routine use by ACIP in 2015[147]. Each of these vaccines is administered as a 3-dose series, with the second and third doses administered at two and six months after the first dose, respectively[4]. Vaccinating children at age 11 or 12, prior to sexual debut, ensures maximum benefit for prevention of HPV-associated disease.

The HPV vaccine is most commonly administered by primary care providers or by health clinics and should be administered at the same visit as other age-appropriate vaccines such as tetanus, diphtheria, and pertussis (i.e., Tdap) and meningococcal conjugate vaccines. The Vaccines for Children Program (VFC) provides access to the HPV vaccine for Medicaid and underinsured children less than 18 years of age. Under the Patient Protection and Affordable Care Act, most private insurance plans are required

to cover the HPV vaccine at no cost to patients up to 18 years of age[148]. Three jurisdictions currently require HPV vaccines for school attendance, including Rhode Island, Virginia, and Washington D.C. In 2007, lawmakers in the state of Maryland attempted to pass legislation that would require all girls entering the sixth grade to be vaccinated effective in 2008; however this bill was withdrawn and has not since been reintroduced[149].

Barriers to HPV Vaccine Uptake

While the vaccine offers considerable promise for protection against HPV-associated cancers, uptake has been suboptimal in the U.S, with only 39.7% of females and 21.6% of males receiving all three doses in 2014[150]. These estimates fall short of the *Healthy People 2020* goal, which targets 80% coverage for females, and recently extended this goal to include 80% coverage for males[151]. Barriers to HPV vaccination among U.S. adolescents were recently summarized in a systematic review[152]. A majority of studies included in this review focused on vaccine initiation, and were conducted among females. Most studies regarding males were conducted prior to the ACIP recommendation in 2011. Findings from this review suggest that among parents and caregivers, the most important barriers to HPV vaccine initiation included not receiving a recommendation from a healthcare provider, lack of information/knowledge, concerns about side effects, and a perception that their child was too young to get vaccinated for HPV. Among healthcare providers, the most important barriers to HPV vaccine initiation were parents' attitudes and concerns, financial concerns including inadequate insurance coverage and reimbursement, preference for vaccinating older adolescents, and a

preference for vaccinating girls vs. boys[152]. Additional barriers cited for vaccine completion included lack of insurance coverage, lack of a regular "medical home", and infrequent contact with the healthcare system[152]. Important disparities have also been observed with respect to HPV vaccination. Among underserved and minority populations, the available data in both males and females suggest that African Americans and Hispanics as well as those living below the poverty line, are more likely to initiate the vaccine, but less likely to complete the three dose series[150,152]. In one large study of MSM aged 18 to 26 years, a group at high risk for HPV infection and related disease, vaccine uptake in 2011 was only 4.9%. Predictors of HPV vaccination in this study included visiting a healthcare provider in the past year, disclosure of MSM status, being HIV-positive, and receipt of the hepatitis vaccine. Of the 3,000 males who were unvaccinated, 76% had visited a healthcare provider within the past year[153].

Efforts to Increase HPV Vaccine Uptake

In 2012, the President's Cancer Panel declared increasing HPV vaccine coverage an urgent national priority and defined three critical goals that must be achieved in order to increase uptake, including 1) reduce missed clinical opportunities, 2) increase acceptance of HPV vaccines, and 3) maximize access to HPV vaccination services[154]. In the state of Maryland, the Health Department recently launched a coordinated effort to increase the percent of Maryland children that are fully vaccinated against HPV[155]. Studies have shown that the most successful strategies for improving vaccine uptake include incorporating HPV vaccination in cancer control plans, public communication

campaigns, patient/provider reminder systems, clinician education sessions and assessments, practice-level interventions to educate staff, and educating parents about the importance of HPV vaccination for their children[150].

Environmental Influences of Unhealthy Diet in High Risk Adolescents

Diet is a Modifiable Risk Factor for Cancer

Diet is believed to play an important, yet not fully understood role in cancer risk. The lack of definitive evidence is largely due to the methodological challenges associated with measuring dietary exposures in epidemiologic studies. For example, a majority of observational studies typically rely on self-reported dietary intakes from tools such as the Food Frequency Questionnaire (FFQ), which measures average daily intakes of foods and nutrients over the past several months[156]. Data collected from the FFQ and other self-report dietary assessments are often subject to measurement error due to the inherent difficulty in recalling average dietary intakes over a long period of time, and in estimating consumption patterns and portion sizes. Retrospective assessment of diet in case control studies has been shown to introduce bias in the association of dietary factors with cancer risk[157,158]. Prospective studies of diet and cancer risk tend to mitigate this bias and have been more successful in providing valid associations, particularly when statistical methods such as regression calibration and error correction techniques are applied[156]. Despite these challenges, the body of scientific evidence has yielded some consistent findings relating dietary behaviors and cancer risk. With respect to overall diet

quality and cancer risk, dietary recommendations from agencies such as the AICR/WCRF, ACS, and International Agency for Research on Cancer (IARC), emphasize high intake of fruit, vegetables, and unprocessed whole grains, and low intake of red and processed meats, and alcohol. Dietary guidelines for obesity prevention also have relevance to cancer risk, and emphasize low intake of energy-dense foods, refined sugars, and sugar sweetened beverages[159]. In 2007 the AICR/WCRF published a report titled, "*Food, Nutrition, Physical Activity, and the Prevention of Cancer: a Global Perspective*", based on a thorough review of the existing literature on diet, physical activity and cancer by an international panel of experts[17]. Since the publication of this report, evidence is kept up to date through the Continuous Update Project, which is an ongoing review that synthesizes new research as studies are published[160]. The project collects evidence from randomized controlled trials and cohort studies for 17 cancer types. Probable and convincing evidence relating dietary exposures to selected cancers from this project are summarized below, with particular focus on dietary exposures that are common in the U.S.

Dairy Intake

The 2007 AICR/WCRF report suggested a probable association of milk and calcium intake with decreased risk of colorectal cancer[17]. Since the publication of this report, a meta-analysis including data from six additional studies investigating milk and ten additional studies investigating dietary calcium was published in 2012[161]. The Continuous Project Update meta-analysis estimated a modest 9% decrease in colorectal

cancer risk associated with milk intake that was not statistically significant and a 6% decrease in colorectal cancer risk associated with dietary calcium that was statistically significant. These findings support the original conclusions in the 2007 report, and suggest milk probably protects against colorectal cancer[161]. It is hypothesized that the association of milk intake with reduced colorectal cancer risk is mediated by calcium, which has demonstrated anti-proliferative effects in colon epithelial cells and is involved in processes that promote differentiation of normal cells and apoptosis of transformed cells[162].

Red and Processed Meat

The 2007 AICR/WCRF report cites convincing evidence for an association of red and processed meat (e.g., bacon, sausage, lunch meat, hot dogs) intake with risk of colorectal cancer[17]. Since the report, six additional prospective studies have been conducted. Collectively, these data suggest a 16% increased risk of colorectal cancer for every 100 grams (g) per day of total red and processed meats[161]. Recently, IARC released a summary of the evidence on consumption of red and processed meat and cancer risk. More than 800 studies were reviewed, with the majority focused on colorectal cancer risk. For red meat, a total of 14 cohort and 15 high-quality population-based case-control studies were considered and for processed meat, a total of 18 cohort and nine high-quality, population-based case-controls studies were considered. Based on their review of this evidence, IARC classified red meat as a Group 2A carcinogen (i.e., "probably carcinogenic to humans") and processed meat as a Group 1 carcinogen (i.e.,

"carcinogenic to humans")[163]. The relative risk for colorectal cancer was estimated to be about 18% for every 50 g daily serving of processed meat and about 17% for every 100 g daily serving of red meat[163]. Potential mechanisms underlying this association include the high levels of heme, a component of hemoglobin, found in red meat which promotes the formation of potentially carcinogenic N-nitroso compounds, the production of polycyclic aromatic hydrocarbons and heterocyclic amines when meat is cooked at high temperatures, and the elevated levels of nitrates and nitrites added to preserve processed meat[161,163].

Fruits and Vegetables

Evidence from the 2007 AICR/WCRF report supports a probable risk reduction associated with vegetable and fruit consumption for cancers of the lung, mouth, pharynx, larynx, stomach, and esophagus[17]. An updated systematic literature review for lung cancer was published recently as part of the Continuous Update Project. This meta-analysis included an additional 11 studies assessing the relationship of fruit and vegetable intake with reduced lung cancer risk[164]. Overall, there was a significant association of fruit and vegetable intake with a 14% reduction in lung cancer risk. Relative risk reduction estimates ranged from 8% to 18%, depending on the type of vegetable or fruit assessed, with the highest risk reduction observed for fruit intake[164]. Results from this analysis also provided evidence for a modest dose-response relationship of fruit and vegetable consumption with reduced lung cancer risk (up to 400 g per day)[164]. In general, estimates were attenuated when results were stratified by smoking status, and

tended to only remain significant among current smokers[164]. With regard to other cancers, evidence from the original AICR/WCRF report suggests a probable protective association of consumption of non-starchy vegetables, fruit, and dietary carotenoids with risk of cancers of the mouth, pharynx, and larynx and a probable protective association of consumption of non-starchy vegetables, fruit, foods containing beta-carotene, and foods containing vitamin C with cancer of the esophagus[17]. With respect to stomach cancer, the report suggests a probable protective effect of consumption of non-starchy vegetables, allium vegetables (e.g., garlic, onions), and fruit[17]. For each cancer type, most studies were suggestive of a dose-response relationship; however the majority of those included in the review were case-control studies, which limited the certainty of these findings. The observed protective effect associated with fruit and vegetable intake is likely due to specific nutrients such as carotenoids, polyphenols, and other vitamins and minerals found in fruits and vegetables[17]. Carotenoids and polyphenols for example, can act as potent antioxidants, trapping free radicals produced by oxidative stress and thus protecting against DNA damage[17]. Vitamin C has antioxidant activity, and may also have anti-inflammatory effects. Allium vegetables, such as garlic, have been shown to have antibiotic activity, which may modify the risk of stomach cancers caused by *H. pylori*[17]. In addition to the action of specific nutrients, high fruit and vegetable intake may indirectly influence cancer risk via a relationship with energy intake and/or body weight[159].

Dietary Fiber

The 2007 AICR/WCRF report's conclusion for a protective association of dietary fiber intake with colorectal cancer risk was upgraded from probable to convincing after a systematic review of 11 additional studies was published in 2011[165]. Evidence from this review suggested a significant risk reduction of about 10% for each 10 g daily serving of dietary fiber. This review also considered specific dietary sources of fiber and concluded a protective effect for cereal fiber and whole grains, but not fiber from fruit, vegetable, or legume sources[165]. The mechanisms underlying this protective association are not completely understood, but likely involve fiber's ability to reduce stool transit time in the gastrointestinal tract and to dilute carcinogens present in stool[17,165]. Fermentation of fiber by gut flora produces short chain fatty acids, which may have anti-proliferative and apoptotic effects in the colon[165]. High fiber intake may also indirectly influence colorectal cancer risk by protecting against weight gain[165].

Evidence for Associations of Poor Diet in Early-Life with Cancer Risk

Prospective studies of diet in early-life and cancer risk in adulthood are very limited largely due to the methodological challenges previously discussed. Most studies assessing the relationship between childhood and adolescent diet and cancer rely on data obtained retrospectively from adults, typically before disease onset[166]. Most of the research has been conducted in breast cancer, although evidence for other cancers is emerging. With respect to breast cancer, a study conducted among Chinese immigrant women demonstrated an association of red meat intake in adolescence, but not adulthood, with increased breast density later in life, even after adjusting for acculturation factors[167].

Similarly, findings from an NHS II study revealed a positive association of red meat intake during adolescence, but not in adulthood, with premenopausal breast cancer risk[168,169]. Additionally, among women enrolled in NHS II, consumption of vegetable protein at age 14 was associated with lower risk of benign breast disease in early adulthood, and poultry intake and replacement of red meat with a diet consisting of poultry, fish, legumes and nuts in adolescence was also associated with lower breast cancer risk later in life[168,170]. In the Growing Up Today Study (GUTS), a prospective study of over 9,000 daughters of women enrolled in NHS II, sugar sweetened beverage consumption in adolescence (>1.5 servings per day) was associated with earlier age of menarche, a risk factor for breast cancer, even after adjusting for total energy intake, height and BMI[171].

The relationship between adolescent diet and colorectal cancer risk has also been explored in some studies. In the NHS II for example, recalled adolescent intake of poultry and replacement of red meat with either poultry or fish was associated with reduced risk of colorectal adenomas in adulthood[172]. In the NIH–AARP Diet and Health Study cohort, a lower risk of colon cancer was observed for those with higher intakes of vegetables and vitamin A during adolescence, but not in adulthood, based on recalled adolescent and baseline adult diets[173]. This study also found that for certain foods such as fruit and nutrients such as calcium, the protective effect was strongest when consumed in both adolescence and adulthood. Similarly, the risk of colorectal cancer was strongest for individuals who consumed high amounts of red and processed meat in both adolescence and adulthood, suggesting that consistent or cumulative dietary patterns over the life

course may be important[173]. A very recent analysis from the NIH-AARP study investigated the association of adolescent and mid-life diet and risk of thyroid cancer. Results from this large prospective study suggested greater intake of poultry, tuna (among men), sweet baked goods and vitamin C and reduced intake of butter/margarine were positively associated with risk of thyroid cancer later in life. Mid-life diet appeared to be less important for risk of thyroid cancer, although broccoli intake in mid-life was associated with reduced risk among men[174].

Adolescent Diets in Relation to Cancer Prevention Recommendations

Dietary patterns that are established in childhood and adolescence often persist into adulthood, and thus have important implications for diet- and obesity-related cancers later in life[175]. In general, prominent organizations such as the ACS, AICR, and the WCRF issue similar dietary guidelines with respect to cancer prevention. These guidelines include increasing consumption of fruits, vegetables, and whole grains, and decreasing intake of red and processed meats, as well as high-energy foods with low nutritional value, commonly referred to as "empty calories"[159,176]. The 2010 U.S. Dietary Guidelines, intended for Americans aged 2 years and older, emphasize maintaining energy balance so as to sustain a healthy weight and promote consumption of nutrient-dense foods and beverages. These guidelines also stress the importance of maintaining an appropriate energy balance during each stage of the life course, including childhood, adolescence, adulthood, and pregnancy[177]. To assess adherence to the U.S. Dietary Guidelines, the U.S. Department of Agriculture (USDA) developed the Healthy Eating

Index (HEI), with the most recent guidelines released in 2010 (HEI-2010)[178]. The HEI measures diet quality using 12 different dietary components, each associated with a score ranging from 0 to 20 (higher scores associated with higher quality), with a total score ranging from 0 to 100 representing intake per 1,000 kilocalories[178]. A recent study of NHANES data used HEI scores as a metric to assess diet quality among over 8,000 children and adolescents aged 4 to 18 years in the U.S. from 2005 to 2008[179–181]. Overall diet quality in this study was poor, and fell below U.S. guidelines, which target a total HEI score of 80[179]. Consistent with the current literature, dietary quality seemed to decline from childhood to adolescence: children ages 4 to 8 years had a total HEI score of 52.1, while children ages 9 to 13 years and 14 to 18 years had a total HEI score of 46.9 and 43.6, respectively. In regards to specific foods related to cancer risk, adolescents in both age groups were not meeting recommended intakes for fruits, vegetables, and whole grains, and were consuming excess calories from refined grains and empty calories (Figure 1-5)[179]. These results are in line with recent data from the Youth Behavioral Risk Factor Surveillance Study, which indicate that only 21.9% and 15.7% of adolescents in 9th through 12th grade are consuming three or more fruits and vegetables per day, respectively[182].

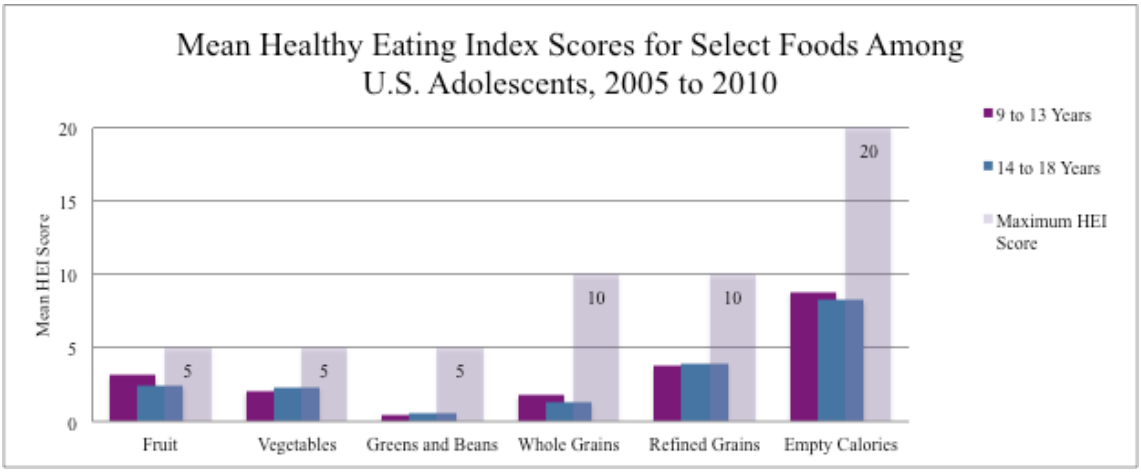

Figure 1-5. Mean Healthy Eating Index Scores for Select Foods Among U.S. Adolescents, 2005 to 2010. Select dietary components were chosen based on relevance to cancer prevention guidelines. Healthy Eating Index Scores for each component were calculated by estimating the intake per 1,000 kilocalories. Each component has a maximum score and higher values correspond to better adherence to U.S. Dietary Guidelines for Americans. Data is shown for male and female adolescents aged 9 to 13 years and 14 to 18 years, respectively. Data Source: National Health and Nutrition Examination Survey (NHANES) 2005-2006, 2007-2008, and 2009-2010. Adapted from Banfield et al., 2015[178] Abbreviations: HEI, Healthy Eating Index

Determinants of Unhealthy Diet Among Adolescents

Dietary behaviors among adolescents are influenced by a number of factors operating at different levels of influence[181]. At the individual level, food preferences, taste and appearance of foods are important determinants of what adolescents choose to eat. Additionally, many adolescents prioritize time and convenience over health and nutrition when it comes to choosing foods, and report taste, availability, convenience, and cost as significant barriers to eating healthy[181]. Although children transition to greater independence and autonomy in making food decisions as they get older, many aspects of the family and home environment are important influences on adolescents' eating behaviors[181]. Engaging in a family meal, for example, is associated with healthy dietary behaviors, including greater fruit and vegetable consumption, and lower consumption of sugar-sweetened beverages and high-fat foods[183]. Other factors associated with the home environment include parent modeling and parental influence over what foods are available and accessible in the home[181]. Demographic characteristics, such as socioeconomic status, influence food choices among adolescents and their families. For individuals living in poverty, the price of fresh fruits and vegetables compared with low cost fast food options creates a significant barrier to healthy eating[184]. Packaged foods that tend to last longer and cost less are often a major source of refined grains, added sugars, and fats. In underserved, low-income neighborhoods, grocery store options are typically very limited; these neighborhoods are instead often populated with convenience stores, liquor stores, and fast food restaurants[185].

In addition to the family and home environment, the school environment plays an important role in shaping dietary behaviors of adolescents. Data suggest that adolescents consume approximately 35 to 50% of their daily calories at school[186]. This is particularly true for the more than 30 million children participating in federally assisted meal programs in the U.S., such as the National School Lunch Program and the School Breakfast Program[187]. Schools participating in these programs are required to offer meals that adhere to nutrition standards implemented by the USDA. Emerging evidence suggests that these policies are having an impact on the nutritional content of school meals and have reduced disparities in healthy food access[187]. However, school meal programs only represent one aspect of the school food environment. Schools also offer a wide range of foods and beverages outside of school meal programs that tend to be high in fat and sugar[188]. These foods are typically sold in school cafeterias, vending machines, school stores, and during fundraisers and are collectively referred to as "competitive foods and beverages" because they displace healthy alternatives[188]. As part of the Healthy, Hunger-Free Kids Act of 2010, the USDA implemented new "Smart Snacks in School" nutrition standards for competitive foods and beverages sold outside of the school meals program during the school day[189]. These standards limit the amount of calories, salt, sugar, and fats in foods and beverages and promote whole grains, low fat dairy, fruits, and vegetables as healthy snack choices[188]. Emerging evidence suggests that these policies may be effective in changing school food environments, but future studies are needed to determine the impact on adolescent eating behaviors and the prevalence of overweight and obesity[188,190].

Postpartum Adolescents: A Group at High Risk for Unhealthy Diets

Nearly 300,000 adolescents become pregnant each year in the U.S., representing a significant, yet understudied population at high risk for obesity and related chronic diseases, such as cancer, later in life[191]. Indeed, adolescents who become pregnant tend to gain excessive weight during pregnancy and retain more weight postpartum compared with their adult counterparts[192–196]. In a recent study of NHANES data, women aged 20 to 59 years who gave birth during adolescence were significantly more likely to be overweight or obese later in life compared with women who gave birth in adulthood[197]. The burden of teenage pregnancy falls disproportionately on minority and socioeconomically disadvantaged populations, underscoring the need for targeted interventions among this particularly high-risk group[198]. The limited evidence on dietary behaviors in this population suggest that postpartum adolescents have diets low in fruits and vegetables, and consume excess calories from high-fat, sugary snacks[199,200]. Given the aforementioned importance of the family, home, and school food environments for shaping dietary behaviors among adolescents, interventions that focus on the food environment may be a particularly effective strategy for preventing unhealthy diets in postpartum adolescents, as they require fewer individual-level resources to be effective[201]. To date, very little is known about how the environment influences dietary patterns among postpartum adolescents. Research addressing this hard-to-reach and

high-risk population will be important for informing intervention strategies and preventing long-term obesity and subsequent cancer risk.

Chapter 2

A targeted, next-generation bisulfite sequencing approach to evaluate the influence of maternal pre-pregnancy body mass index and gestational weight gain on umbilical cord blood DNA methylation levels

Title: A targeted, next-generation bisulfite sequencing approach to evaluate the influence of maternal pre-pregnancy body mass index and gestational weight gain on umbilical cord blood DNA methylation levels

Megan A. Clarke[1], David Esopi[2], Srinivasan Yegnasubramanian[2,3], Sarah J. Wheelan[2,3], Elizabeth A. Platz[1,3], Corinne E. Joshu[1,3]

[1]Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD
[2]Department of Oncology, Johns Hopkins University School of Medicine, Baltimore, MD.
[3]Sidney Kimmel Comprehensive Cancer Center at Johns Hopkins, Baltimore, MD

Keywords: Obesity, cancer, methylation, epigenetics, umbilical cord blood

Financial Support: Megan Clarke was supported by the National Cancer Institute, T32 CA009314

Corresponding Author:
Megan A. Clarke, MHS
Johns Hopkins Bloomberg School of Public Health
615 North Wolfe Street
Baltimore, MD 21205
(p) 443-287-3821
mclark43@jhu.edu

ABSTRACT

Introduction: *In utero* exposure to maternal obesity has been associated with offspring adiposity later in life. This association is hypothesized to involve DNA methylation. Using a candidate gene approach, we assessed maternal pre-pregnancy body mass index (BMI) and gestational weight gain (GWG) in relation to umbilical cord blood DNA methylation.

Methods: We quantified DNA methylation in 112 cord blood leukocyte samples using high-throughput, microfluidic PCR and next-generation, bisulfite sequencing. Generalized linear models with a binomial distribution were fit to assess associations of DNA methylation with pre-pregnancy BMI and GWG. From the sequencing data, we identified patterns of DNA methylation haplotypes and used Fisher's Exact chi-square and logistic regression models with generalized estimating equations to evaluate associations of the most common haplotype patterns with pre-pregnancy BMI and GWG. All analyses were stratified by infant sex and were adjusted for false discovery rate (FDR).

Results: A total of 29 CpG sites within 14 genes were differentially methylated with respect to maternal pre-pregnancy BMI (FDR-adjusted p-value ≤ 0.05). A total of 30 CpG sites within 15 genes were differentially methylated with respect to maternal GWG. A majority of these associations were sex-specific. Common methylation haplotype patterns in the *H19* gene were associated with pre-pregnancy BMI, and females born to

overweight/obese mothers were significantly less likely to have *H19* methylation haplotypes with all CpG sites methylated compared with females born to normal weight mothers.

Discussion: Our findings suggest maternal overweight/obesity and excess GWG may influence DNA methylation within offspring genes, in a sex-specific manner. These findings warrant further replication in large, prospective cohort studies.

INTRODUCTION

Adverse nutritional exposures *in utero* have been shown to influence the risk of chronic diseases, including certain cancers, later in life[51,202]. Although studies have historically focused on maternal nutritional deprivation[52,86–88,203,204], the rising prevalence of obesity among reproductive-aged women over the past several decades has necessitated research investigating the influence of maternal overnutrition and adiposity on offspring health. A growing body of evidence suggests that infants born to obese and overweight mothers are at increased risk of obesity and metabolic disease later in life, both of which are risk factors for certain cancers[59–62]. While the underlying biological mechanisms of these associations are not well understood, evidence from animal models[205] and quasi-experimental human studies of prenatal nutritional deprivation[89,91] suggests that epigenetic processes, such as DNA methylation may play an important role[66].

DNA methylation is essential for processes involved in human development, such as genomic imprinting and X chromosome inactivation[74,75], and aberrant DNA methylation patterns have been associated with diseases such as cancer[206]. In humans, DNA methylation most commonly occurs at cytosines at the 5' position of a cytosine and guanine dinucleotide pair termed "CpG". Clusters of CpG sites in small stretches of DNA termed, "CpG islands", are commonly located in promoter regions of genes[76]. DNA methylation plays a critical role in regulating gene expression by influencing transcription factor binding affinity and recruiting methyl-DNA binding proteins that

regulate transcription[74]. Patterns of DNA methylation are established during embryogenesis through a highly dynamic process[77] and tend to be relatively stable in differentiated cells and tissues over time[207]. Collectively, this suggests that the *in utero* period may be a critical window for epigenetic programming in response to environmental signals, with potential long-term consequences for gene expression.

Very few studies have investigated the association of maternal adiposity and gestational weight gain (GWG) with offspring DNA methylation levels. Findings thus far have been inconsistent, with studies often reporting small differences (~5%) in methylation levels at specific CpG sites within different genes, including sites within peroxisome proliferator-activated receptor gamma (*PPARG),* zinc finger, CCH10 domain containing 10 *(ZCCHC10),* matrix metalloproteinase-7 *(MMP7)* and the retinoic X receptor alpha *(RXRA)*[87,89,90,93–97,208,209]. Possible reasons for discrepancies between studies include technical differences in the array-based platforms used to quantify DNA methylation levels[210], potential misclassification of self-reported pre-pregnancy BMI and variable adjustment for confounders, and the difficulty in obtaining well-characterized samples of cord blood DNA.

Thus, we evaluated the association between maternal obesity, GWG and DNA methylation patterns in previously identified candidate genes in 112 cord blood leukocyte samples with clinically-measured maternal and child characteristics. We prioritized genes with known functional roles, particularly those related to obesity and cancer (i.e., involved in processes such as apoptosis and cell cycle regulation). We quantitatively determined DNA methylation levels using a novel, high-throughput microfluidic PCR

platform for target enrichment and gold standard bisulfite sequencing technology[211,212]. This method provides high-resolution CpG methylation information and allows for the characterization of methylation patterns of multiple, contiguous CpG sites on single DNA molecules (termed "methylation haplotypes")[212]. In addition, we leveraged this technology to assess common patterns of DNA methylation haplotypes and evaluate the distribution of these haplotype patterns with respect to maternal BMI and GWG categories.

METHODS

Study Population

This study consists of black and white neonates who were part of the Hormones in Umbilical Cord Blood Extended Study (eHUB), a prospective study of pregnant women designed to determine how racial differences in the hormonal and growth factor milieu *in utero* contribute to the racial disparity in prostate and breast cancers. The eHUB study was approved by the Institutional Review Board at the Johns Hopkins Bloomberg School of Public Health.

This prospective cohort study included 185 black and white pregnant women enrolled in 2006 – 2007 from a prenatal clinic in Baltimore, MD. Umbilical cord blood samples were collected by the nurse at delivery in EDTA Vacutainer tubes and samples were processed within 24 hours and stored at -70^{o}C. For the current study, inclusion criteria for the infants were full term birth (≥37 weeks), no major birth defects, and singleton birth. Cord blood samples were available for 122 eligible infants, and a total of 112 had sufficient DNA available for next-generation bisulfite sequencing.

Maternal Pre-Pregnancy Body Mass Index and Gestational Weight Gain

Pre-pregnancy body mass index (BMI) (kg/m^2) was calculated using the women's self-reported weight before pregnancy and height at their first questionnaire assessment. Self-reported BMI was highly correlated with clinically measured BMI at the first prenatal visit (~12 weeks) (Pearson correlation coefficient =0.93; see Appendix Figure A5), indicating high accuracy of self-reported weight. Pre-pregnancy BMI was categorized as normal weight (<25.0 kg/m^2), overweight (25.0 to <30.0 kg/m^2) or obese (≥30 kg/m^2). Gestational weight gain (GWG) was calculated as the difference between the measured weight at the last obstetrics visit before delivery and the measured weight at the first obstetrics visit. Within the three BMI categories, GWG was categorized according to the Institute of Medicine's (IOM) recommendations[213] (i.e., normal: 28-40 lbs, overweight: 25-35 lbs, and obese: 11-20 lbs) as less than recommended, within recommendation, or more than recommended.

Other Variables

Women were administered a questionnaire at their first visit (~12 weeks) and at the postpartum visit, with questions pertaining to demographic characteristics and lifestyle behaviors before and during their current pregnancy, respectively. For this study, we categorized education as less than a college degree vs. college or graduate degree, and smoking status as never, former, or current. Additional maternal characteristics including parity (categorized as nulliparous vs. parous), gestational age,

weight at start of pregnancy, weight at each maternity visit and at delivery, and pregnancy complications, and infant characteristics including sex and birth weight, were abstracted from the medical record.

Candidate Gene Selection

Candidate genes were chosen based on literature review (Supplementary Table 2-S1). We prioritized genes with known functional roles, particularly those related to obesity and cancer (i.e., involved in processes such as apoptosis and cell cycle regulation). We selected 20 genes that were previously shown to be associated with maternal pre-pregnancy BMI in a recent study of African-American newborns[96], 3 genes that showed differential methylation patterns in cord blood with respect to body composition in childhood[93], and 2 genes that were found to be differentially methylated with respect to infant birthweight[214]. An additional 4 genes were selected based on their hypothesized role in fetal programming of adiposity, with supporting evidence from animal studies[20–23]. When possible, we designed primers to target specific CpG sites reported in the literature. After testing bisulfite primers for 29 candidate genes, we selected 48 primer pairs that performed robustly, interrogating a total of 24 genes. We will refer to each CpG site by the last four digits of its genomic position (determined by Genome Browser Build 36).

DNA Extraction and Bisulfite Conversion

Cord blood leukocyte DNA was isolated and re-purified to remove potential residual PCR inhibitors using the DNeasy Blood and Tissue kit (Qiagen). DNA concentrations were determined by Qubit (Thermo Scientific, Wilmington, DE). Sodium bisulfite conversion of 120 ng genomic DNA extracted from cord blood was carried out using the EZ DNA Methylation-Gold kit (Zymo, Irvine, CA) according to the manufacture's instruction.

Bisulfite Sequencing

Targeted sequences for each candidate gene were obtained from the UCSC genome browser (https://genome.ucsc.edu). The target-specific bisulfite sequencing primers were manually designed by the laboratory technician (D.E.) and using the online tool, Methprimer (http://www.urogene.org/methprimer/), when necessary. A total of 63 amplicons were designed for the 29 genes and covered an average of 16 CpG sites per amplicon (range 1 to 35) with amplicon sizes of no more than 316 bases. Universal sequencing tags were added to the 5' end of the forward and reverse primers according to the User Guide of the Access Array System for Illumina Sequencing Systems. All primers were ordered from Integrated DNA Technologies (Coralville, IA) and validated in conventional bisulfite PCR reactions using JumpStart Taq DNA Polymerase (Sigma Aldrich, St. Louis, MO) and human male DNA (EMD, Chicago, IL) that was converted with the EZ DNA Methylation-Gold kit. Each 50 µL reaction contained 60 ng of converted DNA, 1X JumpStart Buffer, 200 µM each dNTPs, 1X Access Array Loading Reagent, 200 nM each forward and reverse primers, and 1.25 units JumpStart Taq

Polymerase. Cycling conditions are as follows: 1) 50˚C for 2 minutes; 2) 70˚C for 20 minutes; 3) 94˚C for 1 minute; 4) 5 cycles of 94˚C for 30 seconds, 57˚C for 30 seconds, and 72˚C for 90 seconds, with the annealing temperature decreasing by 1˚C each cycle; 5) 30 cycles of 94˚C for 30 seconds, 50˚C for 30 seconds, and 72˚C for 90 seconds; 6) 72˚C for 5 minutes; 7) a 10˚C hold step until further analysis. Reactions were then run out on a 1% agarose gel and visualized using GelStar (Lonza, Walkersville, MD).

Sequencing libraries were prepared using the microfluidic Access Array system (Fluidigm, San Francisco, CA) as has been described elsewhere[219]. In brief, following bisulfite conversion, cord blood DNA was eluted from the columns in 7 μL water, of which 3 μL eluate was amplified on the Access Array as per the manufacturer's instructions. Cycling conditions were identical to those used for conventional PCR validation of the bisulfite primers. The amplified material was then subjected to a second round of PCR, which incorporated barcoded primers to uniquely label each sample library. For this barcoding PCR, 1 μL of microfluidic PCR reaction was amplified in 20 μL reactions containing 1X NEBNext High-Fidelity PCR Master Mix (NEB, Ipswich, MA) and 4 μL Access Array Barcode primers (Fluidigm). Cycling conditions are as follows: 1) 98˚C for 3 minutes; 2) 15 cycles of 98˚C for 30 seconds, 60˚C for 30 seconds, and 72˚C for 90 seconds; 3) 72˚C for 5 minutes; 4) a 10˚C hold until further processing.

Barcoded PCR products were pooled and purified using AMPure magnetic beads. The quality of PCR products was assessed using the Agilent Bioanalyzer system to confirm expected size distributions. Purified sequencing libraries were quantified with

Qubit dsDNA HS Assay kit (Life Technologies, CA) and sequenced on the Illumina HiSeq 2500 platform using a sequencing-by-synthesis approach.

Methylation status was determined using a bioinformatics pipeline that included demultiplexing based on the unique barcoded sequences. For each CpG site, the methylation ratio was calculated by dividing the number of C reads by the sum of C and T reads at each CpG site. CpG site methylation patterns for each single molecule and counts of each pattern were determined for each amplicon, per individual sample.

Quality Control

A detailed description of the quality control analyses is provided in the Appendix. Briefly, we included 11 technical replicate samples for quality control purposes[220]. Concordance correlation coefficient estimates for each pair of replicates were fair to good, with an average of approximately 0.70 (range 0.42 – 0.82).

Statistical Methods

Descriptive statistics were used to determine means and proportions of maternal and infant characteristics. We fit a generalized linear model with a binomial distribution and robust variance estimation to determine associations of DNA methylation at each CpG site with pre-pregnancy BMI and GWG respectively, adjusting for maternal age, parity, maternal smoking status, infant sex, maternal race, and gestational age. To increase statistical power, we collapsed BMI and GWG categories into binary variables, and compared overweight or obese (n=22 and 21, respectively) vs. normal (n=69) BMI

and more than recommended (n=63, "excess") vs. less than recommended or within recommended (n=12 and 37, respectively). We corrected for multiple testing by controlling for the false-discovery rate (FDR), the expected proportion of false positive findings (0.05), using the Benjamini-Hochberg method[221]. Because early exposures have been shown to influence DNA methylation patterns differentially by sex, we stratified all analyses of significant CpG sites by infant sex.

From the bisulfite sequencing data we determined methylation haplotypes, represented as a combination of the methylation statuses of contiguous CpG sites in a single DNA molecule. For simplicity, we will refer to unmethylated CpG sites as "|" and methylated CpG sites as "M". For example, a gene with 5 CpG sites that were all unmethylated would be represented as "|||||", whereas if all 5 CpG sites were methylated, the haplotype would be represented as "MMMMM". We determined the prevalence of each unique methylation haplotype by counting the number of times a particular haplotype pattern was observed for each gene, for each infant sample. As an exploratory analysis, we assigned each infant the most common haplotype pattern observed for each gene, and evaluated the distribution of methylation haplotype patterns by pre-pregnancy BMI and GWG. For genes with an adequate distribution of different haplotype patterns across infants, we categorized each infant as having all CpG sites methylated ("completely methylated", e.g., MMMMM), some CpG sites methylated ("partially methylated", e.g., |MM||) or none of the CpG sites methylated ("completely unmethylated", e.g., |||||) for their most common haplotype pattern. We used Fisher's Exact chi-square test to determine whether there were significant differences in the

distribution of the most common haplotype pattern by categories of pre-pregnancy BMI and GWG and where appropriate. We also used logistic regression models with generalized estimating equations (GEE) to determine if methylation haplotype categories were significantly associated with pre-pregnancy BMI and/or GWG, accounting for clustering within infant. All significant associations were further stratified by infant sex, and we tested for statistical interaction using GEE logistic regression models. All analyses were conducted using Stata v.13.1 (StataCorp, College Station, TX).

RESULTS

Characteristics of Study Population

The characteristics of 112 mothers and infants included in this study are shown in Table 2-1. The mean age of mothers was 29.5 years (range 18 to 40 years) and about 60% were white. Approximately 60% of mothers had a normal BMI before pregnancy and about 20% were overweight or obese, respectively. A majority of mothers had a college or graduate degree (67%) and were never smokers (63%). More than 50% of mothers gained more weight than recommended during pregnancy according to the IOM guidelines. With respect to infant characteristics, 50% were female and mean birth weight was approximately 3,411 grams (± 470 grams).

Maternal Pre-Pregnancy BMI and Cord Blood DNA Methylation Levels

Maternal pre-pregnancy BMI was significantly associated with cord blood DNA methylation levels in 29 CpG sites within 14 candidate genes (FDR adjusted p-value ≤0.05) (Table 2-2). Maternal pre-pregnancy overweight/obesity (vs. normal weight) was

associated with decreased DNA methylation levels in 12 CpG sites within 9 genes, including Adenylate cyclase 3 (ADCY3, 5895); Anaphase promoting complex subunit 7 (ANAPC7, 5793, 5833); Butyrophilin, subfamily 3, member A1 (BTN3A1, 0680); Calcium binding protein 39 (CAB39, 5747, 5820, 5831); G protein-coupled receptor, family C, group 5, member B (GPRC5B, 3915); Insulin-like growth factor 2 (IGF2, 9059); Proline rich 16 (PRR16, 7885); X-ray repair complementing defective repair in Chinese hamster cells 3 (XRCC3, 1137, 1148); Zinc finger with KRAB and SCAN domains 5 (ZKSCAN5, 0185). Maternal pre-pregnancy overweight/obesity (vs. normal weight) was associated with increased DNA methylation levels in the remaining 17 CpG sites within 11 genes, including ADCY3 (6027), ANAPC7 (5902, 5905), Hes-Related Family BHLH Transcription Factor With YRPW Motif-Like (HEYL, 7781); IGF2, Insulin Receptor (INSR, 4660); Peroxisome proliferator-activated receptor gamma (PPARG, 5208); PRR16 (7848); Suppressor APC domain containing 2 (SAPCD2, 4655, 4747); XRCC3 (1162, 1326, 1335), Zinc finger, CCHC domain containing 10 (ZCCHC10, 0150); ZKSCAN5 (0240, 0280).

To determine whether the observed DNA methylation differences were sex-dependent with respect to maternal pre-pregnancy BMI, we stratified our analyses by infant sex (Supplementary Table 2-S2). Only 3 CpG sites remained significantly associated with pre-pregnancy BMI in both sexes (FDR adjusted p-value ≤ 0.05), with 2 sites showing decreased methylation with respect to pre-pregnancy overweight/obese BMI (CAB39 5820, 5831) and 1 site showing increased methylation with respect to pre-pregnancy overweight/obesity (HEYL, 7781). Among females only, 7 CpG sites

remained significantly associated with maternal pre-pregnancy BMI (FDR adjusted p-value ≤0.05), with 3 CpG sites showing decreased methylation (CAB39 5747, IGF2 9059, and PRR16 7885), and 4 CpG sites showing increased methylation with respect to maternal overweight/obese pre-pregnancy BMI (INSR 4660, SAPCD2 4747, XRCC3 1162, and ZCCHC10 0150). Among males only, 9 CpG sites remained significantly associated with maternal pre-pregnancy BMI (FDR adjusted p-value ≤0.05), with a majority of sites (n=8) showing increased methylation with respect to maternal overweight/obese pre-pregnancy BMI (ADCY3 6027, ANAPC7 5902, IGF2 8982, PPARG 5208, SAPCD2 4655, XRCC3 1335, and ZKSCAN5 0240, 0280) and one CpG site showing decreased methylation with respect to maternal overweight/obese pre-pregnancy BMI (ANAPC7 5833).

Maternal Gestational Weight Gain and Cord Blood DNA Methylation Levels

Maternal GWG was significantly associated with cord blood DNA methylation levels in 30 CpG sites within 15 candidate genes (FDR adjusted p-value ≤0.05) (Table 2-3). Excess maternal GWG (more than recommended vs. within/below recommended) was associated with decreased DNA methylation levels in 12 CpG sites within 9 genes, including ANAPC7 (5833), CAB39 (5693, 5747, 5762, 5780); Docking protein 2 (DOK2, 7072); GPRC5B (4040); PRR16 (7763); XRCC3 (1162, 1272); ZCCHC10 (0146); ZKSCAN5 (0157). Excess maternal GWG (vs. within/below recommended) was associated with increased DNA methylation levels in the remaining 18 CpG sites within 12 genes, including 2-Phosphoxylase phosphatase 1 (PXYLP1, 2624), ANAPC7 (5831);

BTN3A1 (0498); CAB39 (5733, 5743, 5863); HEYL (7781); Integrin, alpha E (ITGAE, 3436); INSR (4466, 4560); PPARG (5246); PRR16 (7721); SAPCD2 (4655, 4708, 4780); XRCC3 (1242, 1276); ZKSCAN5 (0185).

To determine whether the observed DNA methylation differences were sex-dependent with respect to maternal GWG, we stratified our analyses by infant sex (Supplementary Table 2-S3). No CpG sites remained significantly associated with GWG in both sexes. Among females only, 4 CpG sites remained significantly associated with maternal GWG (FDR adjusted p-value ≤0.05), with 2 CpG sites showing decreased methylation (GPRC5B 4040 and PRR16 7763), and 2 CpG sites showing increased methylation with respect to excess maternal GWG (ANAPC7 5831 and PPARG 5246). Among males only, 8 CpG sites remained significantly associated with maternal GWG (FDR adjusted p-value ≤0.05), with a majority of sites (n=6) showing increased methylation with respect to maternal excess GWG (PXYLP1 2624, CAB39 5733, HEYL 7781, INSR 4466, 4560, and SAPCD2 4780) and 2 sites showing decreased methylation with respect to excess maternal GWG (ANAPC7 5833 and CAB39 5762).

Methylation Haplotypes

As an exploratory analysis, we assessed the distribution of methylation haplotypes for each gene and assigned each infant the most common haplotype pattern observed for each gene. The unmethylated haplotype (i.e., no methylation at all contiguous CpG sites for a given amplicon) was observed as the most common haplotype (≥ 50% of infants) for most genes, with the exception of ANGPTL2, CAB39, CDKN1C, DOK2, HEYL (2

amplicons), H19 (4 amplicons), IGF2, INSR, and PLAC1. For these genes, we categorized each infant's most common haplotype pattern as completely methylated, partially methylated, or completely unmethylated. For H19 (amplicon 3), infants born to mothers who were overweight/obese pre-pregnancy were significantly less likely to have haplotypes with all CpG sites methylated compared with infants born to mothers who were normal weight pre-pregnancy (Fisher's exact p-value = 0.023; Table 4). Across all H19 amplicons, taking into account clustering by infant and adjusting for maternal age, we observed a significant decrease in odds of maternal pre-pregnancy overweight/obesity for infants with completely methylated H19 haplotype patterns (GEE OR 0.62; 95% CI 0.39 – 0.97). We did not observe any significant associations between common haplotype patterns and maternal GWG.

To determine whether the associations between H19 methylation haplotype categories and maternal pre-pregnancy BMI were sex-specific, we stratified our analyses by sex and tested for statistical interaction using logistic regression models. Among females, we observed statistically significant associations for H19 (amplicons 1 and 3) such that female infants born to mothers who were overweight/obese pre-pregnancy were significantly less likely to have haplotypes with all sites methylated compared with infants born to mothers who were normal weight pre-pregnancy (Fisher's exact p-value = 0.001 and 0.018, respectively). Among males, methylation haplotype categories were not significantly associated with maternal pre-pregnancy BMI. Across all H19 amplicons, taking into account clustering by infant and adjusting for maternal age, we observed a statistically significant interaction between infant sex and methylation haplotype

categories in H19 such that overweight/obese pre-pregnancy BMI was significantly inversely associated with having a haplotype with all sites methylated in females, but not in males (p-value for interaction = 0.0162).

DISCUSSION

We studied the methylation levels of CpG sites within 24 candidate genes, using a novel next-generation, targeted bisulfite sequencing approach. Overall we observed significant associations of DNA methylation levels with maternal pre-pregnancy BMI and with GWG in several CpG sites within 17 candidate genes. A majority of these CpG sites exhibited sex-specific associations. Differences in methylation included both increases and decreases with respect to maternal exposures, and in some cases within the same gene. Of note, 4 CpG sites were significantly differentially methylated in the same direction with respect to both maternal pre-pregnancy BMI and GWG (*ANAPC7* 5833 and *CAB39* 5747, *HEYL* 7781 and *SAPCD2* 4655). Our findings suggest exposure to maternal adiposity and excess weight gain may influence DNA methylation levels within offspring genes, in particular *ANAPC7*, *CAB39*, *HEYL* and *SAPCD2*. The functional significance of these changes is unknown; however these genes have important roles in cell signaling and cell division processes. Taken together, our analyses support the hypothesis that associations of *in utero* exposures and DNA methylation levels may be different with respect to infant sex[85,89]. In general, the influence of maternal adiposity and excess weight gain on DNA methylation levels appeared to be more pronounced for males, with increases in DNA methylation levels observed for a majority of the significant CpG sites. These findings are in line with other studies, suggesting that males

may be more susceptible to *in utero exposures* such as maternal adiposity and weight gain, and may provide insights into understanding sex-specific differences in obesity and cancer risk observed later in life[222].

In exploratory analyses, we utilized the high-resolution bisulfite sequencing data to investigate the distribution of DNA methylation haplotype patterns within each gene, with respect to maternal pre-pregnancy BMI and GWG. Interestingly, we found an association between methylation haplotypes and pre-pregnancy BMI within the *H19* gene overall, and specifically among females, such that female infants born to mothers with an overweight/obese pre-pregnancy BMI were significantly less likely to have *H19* methylation haplotypes with all CpG sites methylated, compared with female infants born to mothers with normal pre-pregnancy BMI. To our knowledge, this is one of the first studies to explore methylation haplotype patterns in cord blood DNA. *H19* is a paternally imprinted gene, known to have important roles embryogenesis and fetal growth[223]. Previous studies have shown increased methylation at *H19* to be associated with upregulation of paternally-expressed *IGF2*, an imprinted gene that also plays a critical role in fetal growth and development[208]. Our findings are in line with evidence linking maternal adiposity to methylation patterns at the *IGF2/H19* imprinting region, and suggest in studies of cord blood DNA, patterns of contiguous CpG site methylation in single DNA molecules might be more informative than individual CpG site methylation levels.

Although we did not conduct a formal replication analysis, we quantitatively compared our results to those previously published in the literature. Similar to other

studies, we found small differences in DNA methylation levels with respect to maternal pre-pregnancy BMI and GWG, though our findings did not consistently replicate the direction or magnitude of all previously reported associations. We observed similar direction of effect for several CpG sites within certain genes (*ADCY3, ANAPC7, BTN3A1, IGF2, PRR16, SAPCD2,* and *ZKSCAN5*)[96]; however, we also found sites within some of these genes with significant effects in the opposite direction. Possible reasons for these discrepancies include differences in study population, variability in adjustment for potentially important confounders, and the different assays used to assess DNA methylation levels. Indeed, the majority of previous studies in the literature have used array-based approaches to measure associations of maternal pre-pregnancy BMI and GWG with DNA methylation[90,94–97,208,209,214], therefore technical differences in DNA methylation assessment (array versus next-generation sequencing) could contribute to the observed discrepancies. In addition, next-generation sequencing tends to have higher sensitivity for detecting very low levels of methylation, when coverage is adequately high (i.e., >30x)[224,225]

To our knowledge, this study is the first to use targeted, next-generation bisulfite sequencing to quantify DNA methylation patterns in cord blood DNA. Next-generation bisulfite sequencing often considered the "gold standard" in DNA methylation studies as it yields quantitative methylation data with single base pair resolution and produces unambiguous methylation information for haplotypes of DNA molecules in a qualitative and quantitative manner. Additional strengths of our study include the use of a

prospective study design with well-characterized cord blood DNA samples that enabled us to take important confounders into account.

Some important limitations of our study should be noted. First, we measured methylation levels in DNA from cord blood leukocytes, which may not reflect patterns in relevant tissue. Further, we did not have information on cell count distributions so we were not able to account for cellular heterogeneity in our analyses. Although we observed small differences in DNA methylation levels, there is evidence to suggest that even small changes in DNA methylation can influence gene expression[74]. Finally the limited samples size and lack of replication with previous studies limits our ability to draw conclusions, and suggests more research investigating changes in DNA methylation patterns with respect to maternal adiposity and gestational weight gain exposures.

In summary we have identified several CpG sites within obesity-and cancer-related genes that are differentially methylated in cord blood DNA of offspring with respect to maternal overweight/obesity and GWG. *In utero* exposure to maternal adiposity has been shown to increase risk of obesity and metabolic disease later in life, both of which are risk factors for certain cancers[59–62]. Our findings provide evidence to support the hypothesis that DNA methylation may underlie this risk. These findings should be considered preliminary, and further replication studies in large, prospective cohorts are warranted.

Table 2-1. Characteristics of 112 Mothers and Infants in the eHUB Study, 2006-2007

Age (years), mean (SD)		29.5 (5.1)
Race, n (%)		
	Black	43 (38.4)
	White	69 (61.6)
Pre-Pregnancy BMI, n (%)		
	Normal	69 (61.6)
	Overweight	22 (19.6)
	Obese	21 (18.8)
Education, n (%)		
	Less than College degree	37 (33.3)
	College or Graduate degree	74 (66.7)
Smoking Status, n (%)		
	Never	68 (63.0)
	Former	34 (31.5)
	Current	6 (5.6)
Parity, n (%)		
	Nulliparous	62 (55.4)
	Parous	50 (44.6)
IOM Categories for Gestational Weight Gain, n (%)		
	Less than recommended range	12 (10.7)
	Within recommended range	37 (33.0)
	More than recommended range	63 (56.2)
Gestational Age (weeks), n (%)		39.10 (7.1)
Infant Sex, n (%)		
	Female	57 (50.9)
	Male	55 (49.1)
Infant Birth Weight (grams), mean (SD)		3,411.3 (470.9)

Abbreviations: eHUB, Hormones in Umbilical Cord Blood Extended Study; BMI, Body Mass Index; IOM, Institute of Medicine

Table 2-2. Significant Associations of Methylation Proportions at CpG Sites with Pre-Pregnancy BMI in Cord Blood DNA in the eHUB Study (n=112)

Chr.	Gene Symbol	CpG Site Position[a]	Adjusted Mean (95%CI)[b]		Adjusted Mean Difference[b] (95% CI)	FDR Adjusted p-value[c]
			Normal Weight	Overweight/Obese		
2	*ADCY3*	24995895	0.2232 (0.150, 0.297)	0.0781 (0.033, 0.123)	-0.1452 (-0.232, -0.058)	0.032
		24996027	0.0040 (0.000, 0.008)	0.0213 (0.011, 0.031)	0.0173 (0.007, 0.028)	0.043
12	*ANAPC7*	109325793	0.0031 (0.001, 0.005)	0.0009 (0.000, 0.002)	-0.0023 (-0.004, -0.000)	0.032
		109325833	0.0197 (-0.003, 0.043)	0.0005 (0.000, 0.001)	-0.0192 (-0.043, 0.004)	0.011
		109325902	0.0012 (0.001, 0.002)	0.0061 (0.001, 0.011)	0.0049 (-0.000, 0.010)	0.021
		109325905	0.0008 (0.000, 0.001)	0.0066 (0.002, 0.011)	0.0057 (0.001, 0.011)	0.011
6	*BTN3A1*	26510680	0.0139 (-0.005, 0.032)	0.0002 (0.000, 0.000)	-0.0137 (-0.032, 0.005)	7.7×10^{-4}
2	*CAB39*	231285747	0.0076 (-0.005, 0.020)	0.0002 (0.000, 0.000)	-0.0074 (-0.020, 0.005)	0.031
		231285820	0.0310 (-0.004, 0.066)	0.0013 (0.000, 0.002)	-0.0297 (-0.065, 0.005)	0.004
		231285831	0.0776 (0.015, 0.140)	0.0006 (0.000, 0.001)	-0.0770 (-0.140, -0.014)	1.6×10^{-5}
16	*GPRC5B*	19803915	0.0028 (0.000, 0.005)	0.0004 (0.000, 0.001)	-0.0025 (-0.005, -0.000)	0.025
1	*HEYL*	39877781	0.0012 (0.001, 0.002)	0.0764 (-0.013, 0.166)	0.0752 (-0.015, 0.165)	1.1×10^{-4}
11	*IGF2*	2118982	0.0021 (0.000, 0.004)	0.0336 (0.005, 0.063)	0.0316 (0.002, 0.061)	0.003
		2119059	0.0307 (-0.001, 0.062)	0.0008 (0.000, 0.001)	-0.0299 (-0.062, 0.002)	2.7×10^{-4}
19	*INSR*	7244660	0.0012 (0.001, 0.001)	0.0132 (0.002, 0.025)	0.0120 (0.000, 0.024)	4.3×10^{-4}
3	*PPARG*	12305208	0.0014 (0.001, 0.002)	0.0097 (0.001, 0.018)	0.0083 (0.000, 0.016)	0.010
5	*PRR16*	119827848	0.0018 (0.001, 0.003)	0.0323 (-0.004, 0.069)	0.0305 (-0.006, 0.067)	0.011
		119827885	0.0048 (0.002, 0.008)	0.0014 (0.001, 0.002)	-0.0034 (-0.007, -0.000)	0.030
9	*SAPCD2*	139084655	0.0031 (0.001, 0.005)	0.0273 (0.006, 0.048)	0.0242 (0.003, 0.045)	0.004
		139084747	0.0021 (0.000, 0.004)	0.0304 (-0.011, 0.072)	0.0283 (-0.013, 0.070)	0.033
14	*XRCC3*	103251137	0.0358 (0.004, 0.067)	0.0040 (-0.000, 0.008)	-0.0319 (-0.063, -0.001)	0.024
		103251148	0.0619 (-0.004, 0.128)	0.0022 (-0.000, 0.005)	-0.0597 (-0.126, 0.007)	0.010
		103251162	0.0015 (0.000, 0.003)	0.0118 (0.003, 0.021)	0.0103 (0.001, 0.020)	0.011
		103251326	0.0011 (0.001, 0.002)	0.0140 (-0.004, 0.032)	0.0128 (-0.005, 0.031)	0.022
		103251335	0.0011 (0.000, 0.002)	0.0265 (0.011, 0.042)	0.0254 (0.010, 0.041)	2.3×10^{-7}
5	*ZCCHC10*	132390150	0.0016 (0.001, 0.002)	0.0084 (0.002, 0.015)	0.0068 (0.000, 0.014)	0.012
7	*ZKSCAN5*	98940185	0.0083 (-0.004, 0.021)	0.0002 (-0.000, 0.001)	-0.0081 (-0.021, 0.004)	0.050
		98940240	0.0020 (0.000, 0.004)	0.0129 (0.004, 0.022)	0.0110 (0.002, 0.020)	0.032
		98940280	0.0014 (0.001, 0.002)	0.0145 (-0.001, 0.030)	0.0131 (-0.002, 0.029)	0.025

Abbreviations: BMI, Body Mass Index; Chr, Chromosome; CI, Confidence Interval; FDR, False Discovery Rate; [a]Based on NCBI Build 36; [b]Adjusted for maternal age, race, smoking status, parity, gestational age, and infant sex; [c]Adjusted using Benjamini-Hochberg False Discovery Method, alpha =0.05

Table 2-3. Significant Associations of Methylation Proportions at CpG Sites with Maternal GWG in Cord Blood DNA in the eHUB Study (n=112)

Chr.	Gene Symbol	CpG Site Position[a]	Adjusted Mean (95% CI)		Adjusted Mean Difference (95% CI)	FDR-Corrected p-value[c]
			Less Than/Within Recommended	More than Recommended		
3	*PXYLP1*	142432624	0.0009 (0.001, 0.001)	0.0150 (0.005, 0.025)	0.0141 (0.004, 0.024)	1.2×10^{-8}
12	*ANAPC7*	109325831	0.0031 (-0.000, 0.006)	0.0295 (0.005, 0.054)	0.0264 (0.003, 0.050)	0.014
		109325833	0.0292 (-0.006, 0.064)	0.0006 (0.000, 0.001)	-0.0286 (-0.063, 0.006)	0.009
6	*BTN3A1*	26510498	0.0011 (-0.000, 0.002)	0.0191 (0.006, 0.033)	0.0180 (0.004, 0.032)	0.009
2	*CAB39*	231285693	0.0080 (-0.001, 0.017)	0.0008 (0.000, 0.001)	-0.0072 (-0.017, 0.002)	0.038
		231285733	0.0014 (0.001, 0.002)	0.0274 (0.010, 0.045)	0.0260 (0.008, 0.044)	6.9×10^{-7}
		231285743	0.0012 (0.000, 0.002)	0.0152 (-0.003, 0.034)	0.0140 (-0.005, 0.033)	0.040
		231285747	0.0084 (-0.003, 0.020)	0.0004 (0.000, 0.001)	-0.0079 (-0.019, 0.004)	0.016
		231285762	0.0247 (0.008, 0.041)	0.0042 (0.001, 0.007)	-0.0205 (-0.038, -0.003)	0.034
		231285780	0.0170 (0.008, 0.027)	0.0028 (0.000, 0.006)	-0.0142 (-0.024, -0.004)	0.037
		231285863	0.0012 (0.000, 0.002)	0.0062 (0.001, 0.012)	0.0050 (-0.001, 0.011)	0.038
8	*DOK2*	21827072	0.0366 (0.025, 0.049)	0.0029 (0.001, 0.005)	-0.0337 (-0.047, -0.021)	8.3×10^{-5}
16	*GPRC5B*	19804040	0.0450 (0.032, 0.058)	0.0006 (-0.000, 0.002)	-0.0444 (-0.058, -0.031)	0.010
1	*HEYL*	39877781	0.0015 (0.001, 0.002)	0.0488 (-0.006, 0.103)	0.0474 (-0.007, 0.102)	1.4×10^{-4}
17	*ITGAE*	3573436	0.0014 (0.001, 0.002)	0.0231 (0.008, 0.039)	0.0217 (0.006, 0.037)	4.2×10^{-7}
19	*INSR*	7244466	0.0050 (0.001, 0.009)	0.0384 (0.007, 0.070)	0.0334 (0.001, 0.065)	0.038
		7244560	0.0003 (0.000, 0.001)	0.0792 (0.054, 0.104)	0.0790 (0.054, 0.104)	1.2×10^{-6}
3	*PPARG*	12305246	0.0041 (0.000, 0.008)	0.0486 (0.008, 0.089)	0.0445 (0.003, 0.086)	0.017
5	*PRR16*	119827721	0.0016 (0.000, 0.003)	0.0099 (0.001, 0.019)	0.0083 (-0.001, 0.018)	2.7×10^{-3}
		119827763	0.0108 (-0.001, 0.023)	0.0009 (0.000, 0.001)	-0.0099 (-0.022, 0.002)	0.050
9	*SAPCD2*	139084655	0.0017 (0.000, 0.003)	0.0204 (0.002, 0.039)	0.0186 (0.000, 0.037)	0.014
		139084708	0.0005 (0.000, 0.001)	0.0044 (-0.000, 0.009)	0.0039 (-0.001, 0.009)	0.015
		139084780	0.0011 (0.000, 0.002)	0.0110 (-0.001, 0.023)	0.0099 (-0.002, 0.022)	0.051
14	*XRCC3*	103251162	0.0118 (0.003, 0.021)	0.0018 (-0.000, 0.004)	-0.0100 (-0.019, -0.001)	0.052
		103251242	0.0017 (-0.000, 0.004)	0.0171 (-0.008, 0.042)	0.0154 (-0.009, 0.039)	0.047
		103251272	0.0428 (-0.005, 0.091)	0.0015 (0.000, 0.003)	-0.0413 (-0.090, 0.007)	0.023
		103251276	0.0017 (0.001, 0.003)	0.0123 (-0.002, 0.027)	0.0106 (-0.004, 0.025)	0.050
5	*ZCCHC10*	132390146	0.0183 (0.012, 0.024)	0.0065 (0.003, 0.010)	-0.0118 (-0.018, -0.005)	0.011
7	*ZKSCAN5*	98940157	0.0056 (0.001, 0.010)	0.0009 (0.001, 0.001)	-0.0047 (-0.009, -0.000)	0.007
		98940185	0.0012 (0.000, 0.002)	0.0099 (0.002, 0.018)	0.0087 (0.001, 0.017)	0.013

Abbreviations: GWG, Gestational Weight Gain; eHUB, Hormones in Umbilical Cord Blood Extended Study; Chr, Chromosome; CI, Confidence Interval; FDR, False Discovery Rate; [a]Based on NCBI Build 36
[b]Adjusted for maternal age, race, smoking status, parity, gestational age, and infant sex; [c]Adjusted for multiple testing using the Benjamini-Hochberg False Discovery Method, alpha =0.05

Table 2-4. Distribution of Common Methylation Haplotypes for H19 by Pre-Pregnancy BMI in the eHUB Study (n=112)

Amplicon 1 (chr11: 1977937 – 1977819)		
Haplotype, n (%)	Normal (n=63)	Overweight/Obese (n=39)
MMMMMM	42 (66.7)	18 (46.2)
MMMMM |	0 (0.0)	2 (5.1)
MM | MMM	0 (0.0)	1 (2.6)
M | MMMM	1 (1.6)	0 (0.0)
| | | MMM	1 (1.6)	0 (0.0)
| | | | | |	19 (30.2)	18 (46.1)
Completely Methylated	42 (66.7)	18 (46.15)
Partially Methylated	2 (3.2)	3 (7.7)
Completely Unmethylated	19 (30.2)	18 (46.15)
Fisher's Exact p-value	0.10	
Amplicon 2 (chr11: 1977841 – 1977679)		
Haplotype, n (%)	Normal (n=62)	Overweight/Obese (n=40)
MMMMMMMMMMMMMMMMM	11 (17.7)	4 (10.3)
MMMMMMMMMMMMMMMMM |	0 (0.0)	1 (2.6)
MMMMMMMMMMMMMMM | MM	1 (1.6)	0 (0.0)
MMMMMMMMMMMMMM | MM |	0 (0.0)	1 (2.6)
MMMMMMMM | MMMMMMMMM	1 (1.6)	0 (0.0)
MMMMMMMM | | | | | | | | |	1 (1.6)	0 (0.0)
MMMMMM | MMMMMMMMMM	11 (17.7)	4 (10.3)
MMMMMM | MMMMMMMMM |	0 (0.0)	1 (2.6)
MMMMMM | MMMM | MMMMMM	0 (0.0)	1 (2.6)
MMMMMM | MM | MMMMMMMM	0 (0.0)	1 (2.6)
MMMMM | MMM | MMMMMMMM	0 (0.0)	1 (2.6)
MMMM | MMM | MMMMMMMM	0 (0.0)	1 (2.6)
M | MMMM | MMMMMMMMMM	1 (1.6)	0 (0.0)
| | | | | | | | | | | | | | | | |	36 (58.1)	24 (61.5)
Completely Methylated	11 (17.7)	4 (10.3)
Partially Methylated	15 (24.2)	11 (28.2)
Completely Unmethylated	36 (58.1)	24 (61.5)
Fisher's Exact p-value	0.60	
Amplicon 3 (chr11: 1977750 – 1977587)		
Haplotype, n (%)	Normal (n=63)	Overweight/Obese (n=40)
MMMMMMMMMMMMMM	33 (50.8)	14 (35.0)
MMMMMMMMMMMMMM |	1 (1.6)	0 (0.0)
MMMMMMMMMM | MMM	0 (0.0)	1 (2.5)
MMMMMMMMMM | | MM	1 (1.6)	0 (0.0)
MMMMMMMMM | MMMM	0 (0.0)	1 (2.5)
MMMMMMMMM | | | |	0 (0.0)	1 (2.5)
MMMMMMMM | MMMMM	2 (3.2)	1 (2.5)
MMMMMMMM | MMM | M	0 (0.0)	1 (2.5)
MMMMM | | | MMMMM	1 (1.6)	0 (0.0)
| M | | M | | | | | | | |	1 (1.6)	0 (0.0)
| | | M | | | | | | | | |	0 (0.0)	1 (2.5)
| | | | | | | | | | | MMM	2 (3.2)	2 (5.0)
| | | | | | | | | | | M | |	0 (0.0)	1 (2.5)
| | | | | | | | | | | | MM	3 (4.8)	7 (17.5)

||||||||||||M|	2 (3.2)	0 (0.0)
|||||||||||||M	3 (4.8)	5 (12.5)
||||||||||||||	14 (22.2)	5 (12.5)
Completely Methylated	33 (52.4)	14 (35.0)
Partially Methylated	16 (25.4)	21 (52.5)
Completely Unmethylated	14 (22.2)	5 (12.5)
Fisher's Exact p-value	**0.023**	
Amplicon 4 (chr11: 1977679 – 1977841)		
Haplotype, n (%)	Normal (n=57)	Overweight/Obese (n=34)
MMMMMMMMMMMMMMMMM	1 (1.7)	0 (0.0)
MMMMMMMMMMMMMM | MMM	1 (1.7)	2 (5.9)
MMMMMMMMMMM | MMMMMM	15 (26.3)	4 (11.8)
MMMMMMMM | MMMMMMMMM	1 (1.7)	0 (0.0)
MMMMM | MMMMM | MMMMMM	0 (0.0)	1 (2.9)
MMMM | MMMMMM | MMMMMM	1 (1.7)	0 (0.0)
M||| MMMMMMMMM | MM |	0 (0.0)	1 (2.9)
| MMMMMMMMMMMMM | MMM	0 (0.0)	1 (2.9)
| MMMMMMMMMM | MM ||||	1 (1.7)	0 (0.0)
| MMMM | MMMMMMMM | MMM	0 (0.0)	1 (2.8)
||||||||M|||||||	1 (1.7)	0 (0.0)
||||||||||||||||||	36 (63.2)	24 (70.6)
Completely Methylated	1 (1.7)	0 (0.0)
Partially Methylated	20 (35.1)	10 (29.4)
Completely Unmethylated	36 (63.2)	24 (70.6)
Fisher's Exact p-value	0.80	

Abbreviations: BMI, Body Mass Index; eHUB, Hormones in Umbilical Cord Blood Extended study

Table 2-S1. List of Candidate Genes and CpG Sites Included in this Study of Cord Blood DNA Methylation

Chr.	Gene Name	Gene Symbol	Genomic Start Site – End Site[a]	# of Unique CpG Sites	Previously Published CpG Site[b] and Relationship with *in Utero* Exposure of Interest, Corresponding CpG Site in Current Study (if applicable)	Relationship to Obesity/Cancer[c]
3	2-phosphoxylose phosphatase 1	*PXYLP1*	142432505 - 142432710	8	Liu et al., 2014: cg00400028 – positive association with maternal pre-pregnancy BMI. Corresponding CpG site position: 142432606	None identified
2	Adenylate Cyclase 3	*ADCY3*	24995866 - 24996048	9	Liu et al., 2014: cg17644208 – positive association with maternal pre-pregnancy BMI Corresponding CpG site position: 24995895	Cell signaling
12	Anaphase promoting complex subunit 7	*ANAPC7*	109325732 - 109325955	20	Liu et al., 2014: cg04062907 – negative association with pre-pregnancy BMI (males only). Corresponding CpG site position: 109325928	Cell cycle processes, cell division
9	Angiopoietin-like 2	*ANGPTL 2*	128924162 - 128924347	5	Liu et al., 2014: cg11213150 – negative association with pre-pregnancy BMI Corresponding CpG site position: 128924278	Member of vascular endothelial growth factor family
6	Butyrophilin, subfamily 3, member A1	*BTN3A1*	Amplicon 1: 26510723 - 26510896 Amplicon 2: 26510465 - 26510710	7	Liu et al., 2014: cg01840268 – negative association with pre-pregnancy BMI (males only). Corresponding CpG site position: 26510755	Inflammation
2	Calcium binding protein 39	*CAB39*	231285655 - 231285888	35	Liu et al., 2014: cg06874144 – positive association with pre-pregnancy BMI (males only).	Related to insulin receptor signaling
2	Caspase 10, Apoptosis-related cysteine	*CASP10*	201756214 - 201756490	4	Relton et al., 2012: cg13782463 – negative association of methylation at birth with childhood BMI at age 9	Apoptosis

	peptidase					
11	Cyclin dependent kinase inhibitor 1C	*CDKN1C*	2861569 - 2861722	13	Relton et al., 2012: cg17511511 – positive association of methylation at birth with childhood BMI at age 9	Negative regulator of cell cycle and cell proliferation
8	Docking protein 2	*DOK2*	21826811 - 21827097	5	Liu et al., 2014: cg06874144 – negative association with pre-pregnancy BMI (males only). Corresponding CpG site position: 21827005	May modulate cellular proliferation; involved in chronic myelogenous leukemia
17	Erb-B2 Receptor Tyrosine Kinase 2	*ERBB2*	35097895 - 35098036	8	Liu et al., 2014: cg19752722 – negative association with pre-pregnancy BMI (males only).	Member of the epidermal growth factor family, involved in numerous cancers
16	G protein-coupled receptor, family C, group 5, member B	*GPRC5B*	Amplicon 1: 19803968 – 19804144 Amplicon 2: 19803832 – 19803995	19	Liu et al., 2014: cg20312475 – negative association with pre-pregnancy BMI (males only). Corresponding CpG site position: 19803962	May modulate insulin secretion, increased expression associated with Type 2 diabetes
11	H19, Imprinted maternally expressed transcript	*H19*	Amplicon 1: 1977797 – 1977969 Amplicon 2: 1977654 – 1977877 Amplicon 3: 1977563 – 1977778 Amplicon 4: 1977649 – 1977868	26	Perkins et al., 2012: - positive association with *H19* DMR (chr11:2109500 – 2109519, NCBI Build 37.1) methylation at birth and weight-for-age at year 1.	Paternally imprinted gene involved in embryogenesis and growth
1	Hes-related family BHLH transcription factor with	*HEYL*	Amplicon 1: 39877571 – 39877822 Amplicon 2:	27	Liu et al., 2014: cg25462291 – negative association with pre-pregnancy BMI (males only). Corresponding CpG site position: 39877643	Involved in Notch signaling and may be a regulator of cell fate decisions

	YRPW motif-like		39877559 – 39877823			
17	Integrin, alpha E	*ITGAE*	Amplicon 1: 3573407 – 3573622 Amplicon 2: 3573338 - 3573597	13	Liu et al., 2014: cg19585196 – negative association with pre-pregnancy BMI (males only). Corresponding CpG site position: 3573509	Inflammation, cell adhesion
11	Insulin-like growth factor 2	*IGF2*	Amplicon 1: 2118822 – 2119111 Amplicon 2: 2110578 – 2110683 Amplicon 3: 2121687 – 2121895	48	Heijmans et al., 2008: - prenatal famine exposure negatively associated with DNA methylation in *IGF2* DMR (chr11:2126035-2126372)	Insulin receptor signaling cascade, growth promoting activity
19	Insulin receptor	*INSR*	Amplicon 1: 7244620 - 7244755	36	Tobi et al., 2014: prenatal famine exposure positively associated with DNA methylation in *INSR* (chr19: 7110011-7111334)	Insulin receptor signaling cascade
X	Placenta-specific 1	*PLAC1*	133619876 - 133620051	3	Liu et al., 2014: cg14674582 – negative association with pre-pregnancy BMI (females only). Corresponding CpG site position: 133619948	Reproductive and cancer biology
3	Peroxisome proliferator-activated receptor gamma	*PPARG*	12305182 - 12305361	15	Gemma et al., 2009: positive association with peroxisome proliferator-activated receptor-gamma co-activator 1alpha gene (*PPARGC1A*) methylation and maternal pre-pregnancy BMI.	Regulates adipocyte differentiation, implicated in obesity, diabetes, and certain cancers
5	Proline rich 16	*PRR16*	119827677 - 119827915	32	Liu et al., 2014: cg25584626 – negative association with pre-pregnancy BMI (males only).	Regulator of cell size and promotes cell size enlargement
9	Suppressor APC domain	*SAPCD2*	139084637 - 139084851	29	Liu et al., 2014: cg15785720 – negative association with pre-pregnancy BMI.	Associated with certain cancers

	containing 2				Corresponding CpG site position: 139084805	
7	Wingless-type MMTV integration site family, member 16	*WNT16*	120751156 - 120751344	1	Liu et al., 2014: cg24849648 – positive association with pre-pregnancy BMI. Corresponding CpG site position: 120751255	Involved in oncogenesis and patterning during embryogenesis.
14	X-ray repair complementing defective repair in Chinese hamster cells 3	*XRCC3*	103251106 - 103251360	34	Engel et al., 2014: cg02194129, cg12798040, cg14172849, cg23369670 – positive association with infant birth weight	Involved in maintaining chromosome stability and repairing DNA damage.
5	Zinc finger, CCHC domain containing 10	*ZCCHC10*	Amplicon 1: 132389960 – 132390173 Amplicon 2: 132389960 – 132390173	26	Liu et al., 2014: cg01422136 – positive association with pre-pregnancy BMI. Corresponding CpG site position: 132390123	None identified
7	Zinc finger with KRAB and SCAN domain 5	*ZKSCAN5*	Amplicon 1: 98940190 – 98940361 Amplicon 2: 98940106 - 98940314	19	Liu et al., 2014: cg01422136 – negative association with pre-pregnancy BMI (males only). Corresponding CpG site position: 98940249	None identified

Abbreviations: Chr, Chromosome; BMI, Body Mass Index; DMR, Differentially methylated region

[a]Based on NCBI Build 36

[b]CpG probe ID according to Illumina HumanMethylation27 BeadChip (Liu et al., 2014), Illumina GoldenGate Methylation Cancer Panel I (Relton et al., 2009), Infinium HumanMethylation450 BeadChip (Engel et al., 2014), otherwise labeled according to genome position based on NCBI Build 36

[c]Determined from www.ncbi.nlm.nih.gov/gene/

Table 2-S2. Significant Associations of Methylation Levels at CpG Sites with Pre-Pregnancy BMI in Cord Blood DNA in the eHUB Study, Stratified by Infant Sex (n=112)

Chr.	Gene Symbol	CpG Site Position[a]	All		Females (n=57)		Males (n=55)	
			β (95% CI)[b]	FDR-Adjusted p-value[c]	β (95% CI)[b]	FDR-Adjusted p-value[c]	β (95% CI)[b]	FDR-Adjusted p-value[c]
2	*ADCY3*	24995895	-1.28 (-2.1, -0.5)	**0.032**	-1.39 (-2.5, -0.3)	0.310	-0.77 (-1.9, 0.4)	0.999
		24996027	1.71 (0.6, 2.8)	**0.043**	0.89 (-0.9, 2.6)	0.999	2.84 (1.9, 3.7)	**1.1×10^{-7}**
12	*ANAPC7*	109325793	-1.3 (-2.1, -0.5)	**0.032**	-1.59 (-2.6, -0.6)	0.067	-0.21 (-1.3, 0.8)	0.210
		109325833	-3.75 (-5.7, -1.8)	**0.011**	-6.23 (-11.1, -1.4)	0.296	-3.19 (-4.7, -1.7)	**0.002**
		109325902	1.66 (0.7, 2.6)	**0.021**	0.30 (-0.5, 1.1)	0.999	2.25 (1.2, 3.3)	**0.002**
		109325905	2.12 (1.0, 3.3)	**0.011**	1.07 (-0.0, 2.1)	0.892	1.84 (0.5, 3.2)	0.210
6	*BTN3A1*	26510680	-4.50 (-6.5, -2.5)	**7.7×10^{-4}**	-3.75 (-5.7, -1.8)	0.73	-2.74 (-4.8, -0.7)	0.275
2	*CAB39*	231285747	-3.64 (-5.9, -1.4)	**0.031**	-9.19 (-14.1, -4.3)	**0.018**	-1.41 (-2.8, -0.1)	0.758
		231285820	-3.29 (-4.9, -1.7)	**0.004**	-5.75 (-7.9, -3.6)	**7.1×10^{-5}**	-4.09 (-6.6, -1.6)	**0.050**
		231285831	-5.34 (-7.3, -3.4)	**1.6×10^{-5}**	-4.05 (-5.7, -2.4)	**3.3×10^{-4}**	-4.01 (-6.0, -2.1)	**0.003**
16	*GPRC5B*	19803915	-2.02 (-3.2, -0.8)	**0.025**	-1.93 (-3.3, -0.5)	0.185	-1.45 (-2.7, -0.2)	0.515
1	*HEYL*	39877781	4.24 (2.6, 5.9)	**1.1×10^{-5}**	4.09 (1.8, 6.4)	**0.022**	8.61 (3.7, 13.5)	**0.025**
11	*IGF2*	2118982	2.89 (1.5, 4.3)	**0.003**	3.86 (1.2, 6.5)	0.144	3.00 (1.6, 4.4)	**0.002**
		2119059	-3.74 (-5.3, -2.2)	**2.7×10^{-4}**	-2.47 (-3.8, -1.1)	**0.022**	-4.19 (-8.3, -0.0)	0.850
19	*INSR*	7244660	2.42 (1.4, 3.4)	**4.3×10^{-4}**	2.98 (1.3, 4.6)	**0.019**	1.24 (0.3, 2.1)	0.216

3	*PPARG*	12305208	2.00 (1.0, 3.0)	**0.010**	-0.53 (-0.9, -0.1)	0.240	2.62 (1.4, 3.8)	**0.002**
5	*PRR16*	119827848	3.00 (1.4, 4.6)	**0.011**	3.97 (2.3, 5.6)	0.086	0.83 (0.0, 1.6)	0.735
		119827885	-1.26 (-2.0, -0.5)	**0.030**	-0.90 (-1.5, -0.3)	**$2.8x10^{-4}$**	-1.06 (-2.0, -0.1)	0.693
9	*SAPCD2*	139084655	2.40 (1.2, 3.6)	**0.004**	-0.38 (-1.1, 0.3)	0.999	3.80 (2.2, 5.4)	**$5.9x10^{-4}$**
		139084747	2.79 (1.1, 4.5)	**0.033**	3.87 (1.9, 5.8)	**0.009**	-1.86 (-3.2, -0.5)	0.230
14	*XRCC3*	103251137	-2.28 (-3.6, -0.9)	**0.024**	-3.70 (-6.8, -0.6)	0.393	-1.90 (-3.7, -0.1)	0.718
		103251148	-3.42 (-5.3, -1.6)	**0.010**	-4.04 (-7.3, -0.8)	0.354	-5.27 (-10.1, -0.4)	0.715
		103251162	2.16 (1.0, 3.3)	**0.011**	2.72 (1.3, 4.2)	**0.017**	-0.52 (-1.0, -0.0)	0.722
		103251326	2.55 (1.1, 4.0)	**0.022**	1.80 (0.2, 3.4)	0.457	2.49 (1.0, 4.0)	0.051
		103251335	3.42 (2.3, 4.5)	**2.3×10^{-7}**	-0.67 (-1.4, 0.0)	0.999	4.41 (3.4, 5.5)	**$1.2x10^{-13}$**
5	*ZCCHC10*	132390150	1.70 (0.8, 2.6)	**0.012**	2.11 (0.9, 3.3)	**0.017**	0.43 (-0.2, 1.0)	0.999
7	*ZKSCAN5*	98940185	-2.61 (-4.3, -0.9)	**0.050**	-2.76 (-4.8, -0.7)	0.197	-1.07 (-1.9, -0.2)	0.301
		98940240	2.39 (0.9, 3.9)	**0.032**	1.55 (0.3, 2.8)	0.352	2.98 (1.3, 4.7)	**0.025**
		98940280	1.96 (0.8, 3.1)	**0.025**	2.23 (0.4, 4.1)	0.367	2.46 (1.3, 3.6)	**0.002**

Abbreviations: BMI, Body, Mass Index; eHUB, Hormones in Umbilical Cord Blood Extended Study; Chr, Chromosome; CI, Confidence Interval; FDR, False Discovery Rate

[a]Based on NCBI Build 36

[b]Adjusted for maternal age, race, smoking status, parity, gestational age, and infant sex

[c]Adjusted for multiple testing using the Benjamini-Hochberg False Discovery Method, alpha =0.05, significant p-values are shown in bold.

Table 2-S3. Significant Association of Methylation Levels at CpG Sites with Maternal GWG in Cord Blood DNA in the eHUB Study, Stratified by Infant Sex (n=112)

			All		Females (n=57)		Males (n=55)	
Chr.	Gene Symbol	CpG Site Position[a]	β (95% CI)[b]	FDR-Adjusted p-value[c]	β (95% CI)[b]	FDR-Adjusted p-value[c]	β (95% CI)[b]	FDR-Adjusted p-value[c]
3	*PXYLP1*	142432624	2.87 (2.0, 3.7)	**1.2×10^{-8}**	1.46 (-0.4, 3.3)	0.999	3.06 (2.1, 4.0)	**2.1×10^{-8}**
12	*ANAPC7*	109325831	2.45 (1.1, 3.8)	**0.014**	4.09 (1.9, 6.3)	**0.034**	1.68 (-0.7, 4.1)	0.999
		109325833	-4.10 (-6.2, -2.0)	**0.009**	-3.44 (-5.7, -1.2)	0.138	-3.06 (-4.3, -1.9)	**1.5×10^{-4}**
6	*BTN3A1*	26510498	3.08 (1.5, 4.7)	**0.009**	0.32 (-1.1, 1.7)	0.999	3.72 (2.0, 5.5)	**0.007**
2	*CAB39*	231285693	-2.28 (-3.7, -0.9)	**0.038**	-2.57 (-4.4, -0.7)	0.252	-1.62 (-3.3, 0.0)	0.999
		231285733	3.62 (2.4, 4.8)	**6.9×10^{-7}**	2.89 (1.0, 4.8)	0.141	3.89 (2.3, 5.5)	**7.1×10^{-4}**
		231285743	2.55 (1.0, 4.1)	**0.040**	1.17 (-0.3, 2.6)	0.999	3.17 (1.3, 5.0)	0.212
		231285747	-2.97 (-4.6, -1.3)	**0.016**	-3.90 (-6.2, -1.6)	0.085	-1.55 (-3.1, -0.0)	0.999
		231285762	-2.05 (-3.3, -0.8)	**0.034**	1.08 (0.2, 1.9)	0.348	-4.15 (-6.1, -2.2)	**0.006**
		231285780	-2.00 (-3.2, -0.8)	**0.037**	-1.12 (-2.3, 0.1)	0.999	-0.43 (-0.9, 0.1)	0.999
		231285863	1.68 (0.6, 2.7)	**0.038**	1.60 (-0.1, 3.3)	0.999	1.65 (-0.1, 3.4)	0.999
8	*DOK2*	21827072	-4.51 (-6.3, -2.7)	**8.3×10^{-5}**	-0.71 (-1.5, 0.1)	0.999	0.03 (-0.8, 0.8)	0.999
16	*GPRC5B*	19804040	-9.37 (-14.3, -4.5)	**0.010**	-11.13 (-16.8, -5.5)	**0.020**	-0.32 (-1.3, 0.7)	0.999
1	*HEYL*	39877781	3.57 (2.1, 5.0)	**4.2×10^{-7}**	3.46 (1.3, 5.7)	0.135	6.48 (3.2, 9.8)	**0.028**
17	*ITGAE*	3573436	2.88 (1.9, 3.8)	**1.4×10^{-4}**	2.35 (1.0, 3.7)	0.057	3.73 (1.6, 5.9)	0.184

19	*INSR*	7244466	2.11 (0.8 – 3.4)	**0.038**	0.22 (-1.3, 1.8)	0.999	2.72 (1.3, 4.1)	**0.045**
		7244560	7.62 (5.0, 10.2)	**1.2 x 10^{-6}**	0.57 (-0.2, 1.3)	0.999	8.55 (6.3, 10.8)	**2.3x10^{-11}**
3	*PPARG*	12305246	2.54 (1.1, 4.0)	**0.017**	4.42 (2.3, 6.5)	**0.008**	2.15 (0.7, 3.6)	0.900
5	*PRR16*	119827721	1.89 (0.7, 3.1)	**2.7 x 10^{-3}**	0.31 (-0.4, 1.0)	0.999	2.34 (0.9, 3.7)	0.286
		119827763	-2.53 (-3.9, -1.1)	**0.050**	-3.53 (-4.9, -2.2)	**1.9x10^{-4}**	0.26 (-0.1, 0.6)	0.999
9	*SAPCD2*	139084655	2.64 (1.2, 4.1)	**0.014**	0.75 (-0.2, 1.7)	0.999	1.97 (0.7, 3.3)	0.886
		139084708	2.15 (0.8, 3.5)	**0.015**	2.32 (0.7, 4.0)	0.236	2.30 (0.6, 4.0)	0.999
		139084780	2.33 (0.8, 3.9)	**0.051**	1.65 (0.4, 2.9)	0.302	3.10 (1.5, 4.7)	**0.039**
14	*XRCC3*	103251162	-1.98 (-3.3, -0.7)	**0.052**	-2.42 (-4.0, -0.9)	0.150	1.87 (0.0, 3.7)	0.999
		103251242	2.37 (0.8, 3.9)	**0.047**	3.03 (0.8, 5.2)	0.252	-1.83 (-3.5, -0.2)	0.999
		103251272	-3.54 (-5.6, -1.5)	**0.023**	-14.26 (-34.9, 6.4)	0.999	0.23 (-0.9, 1.4)	0.999
		103251276	2.01 (0.7, 3.3)	**0.050**	2.73 (0.9, 4.6)	0.164	1.87 (0.0, 3.7)	0.999
5	*ZCCHC10*	132390146	-1.25 (-1.9, -0.6)	**0.011**	0.05 (-0.6, 0.7)	0.999	-1.53 (-2.4, -0.7)	0.165
7	*ZKSCAN5*	98940157	-1.79 (-2.7, -0.9)	**0.007**	-0.86 (-1.9, 0.2)	0.999	-2.58 (-4.0, -1.1)	0.141
		98940185	2.17 (1.0, 3.3)	**0.013**	2.60 (0.8, 4.3)	0.175	1.95 (0.4, 3.5)	0.999

Abbreviations: GWG, Gestational Weight Gain; eHUB, Hormones in Umbilical Cord Blood Extended Study; Chr, Chromosome; CI, Confidence Interval; FDR, False Discovery Rate

[a]Based on NCBI Build 36

[b]Adjusted for maternal age, race, smoking status, parity, gestational age, and infant sex

[c]Adjusted for multiple testing using the Benjamini-Hochberg False Discovery Method, alpha =0.05, significant p-values are shown in bold.

Chapter 3

Predictors of Human Papillomavirus Vaccination in a Large Clinical Population of Males Aged 11 to 26 years in Maryland, 2012 – 2013

Title: Predictors of Human Papillomavirus Vaccination in a Large Clinical Population of Males Aged 11 to 26 years in Maryland, 2012 – 2013

Megan A. Clarke[1], Francesca Coutinho[1], Darcy F. Phelan-Emrick[1], MaryAnn Wilbur[2], Betty Chou[2], Corinne E. Joshu[1,3]

[1]Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD
[2]Department of Gynecology and Obstetrics, Johns Hopkins University School of Medicine, Baltimore, MD.
[3]Sidney Kimmel Comprehensive Cancer Center at Johns Hopkins, Baltimore, MD

Keywords: Human papillomavirus, vaccine initiation, vaccine completion, vaccine doses, males

Financial Support: This work was supported by P30 CA006973. Megan Clarke was additionally supported by the National Cancer Institute, T32 CA009314.

Corresponding Author: Megan A. Clarke, MHS
Johns Hopkins Bloomberg School of Public Health
615 North Wolfe Street
Baltimore, MD 21205
(p) 443-287-3821
mclark43@jhu.edu

This chapter is currently in press at *Cancer Epidemiology Biomarkers and Prevention.*

ABSTRACT

Background: Despite the recommendation for routine human papillomavirus (HPV) vaccination in males, coverage estimates remain low. We sought to identify predictors of receiving each HPV vaccine dose among a large clinical population of males.

Methods: We conducted a cross-sectional analysis of electronic medical records for 14,688 males aged 11-26 years attending 26 outpatient clinics (January 2012 –April 2013) in Maryland to identify predictors of each HPV vaccine dose using multivariate logistic regression models with generalized estimating equations. All analyses were stratified in accordance with vaccine age recommendations: 11-12 years, 13-21 years, and 22-26 years. Analyses of predictors of receipt of subsequent HPV doses were also stratified by number of clinic visits (≤3, >3).

Results: Approximately 15% of males initiated the HPV vaccine. Less than half of males eligible received the second and third dose, 49% and 47%, respectively. Non-Hispanic black males (vs. non-Hispanic white) aged 11-12 and 13-21 years and males with public insurance (vs. private) aged 13-21 years, had significantly greater odds of vaccine initiation, but significantly decreased odds of receiving subsequent doses, respectively. Attendance to >3 clinic visits attenuated the inverse association between public insurance and receipt of subsequent doses.

Conclusion: Overall, rates of HPV vaccine initiation and of subsequent doses were low. While non-Hispanic black and publicly insured males were more likely to initiate the HPV vaccine, they were less likely to receive subsequent doses.

Impact: Tailoring different intervention strategies for increasing HPV vaccine initiation versus increasing rates of subsequent doses among males may be warranted.

INTRODUCTION

Vaccination against human papillomavirus (HPV) in males is substantially lower compared to other adolescent vaccines. As of 2014, approximately 42% of males aged 13-17 years in the U.S. initiated the HPV vaccine series, as compared with approximately 79% coverage for meningococcal conjugate and 88% coverage for tetanus, diphtheria, and pertussis[150]. Completion rates for the HPV vaccine are even lower, with 21.6% of males receiving all three doses in 2014[150]. The HPV vaccine was originally licensed for males in 2009[226]. In October 2011, the Advisory Committee on Immunization Practices (ACIP) recommended routine HPV vaccination for males aged 11-12 years, with catch-up vaccination for males aged 13-21 years, and permissive vaccination up to 26 years of age[227]. Although HPV vaccination coverage among U.S. males has increased, more than half of the target population still remains unvaccinated[150].

The quadrivalent HPV vaccine (Gardasil, Merck and Co, Inc.) is administered as a 3-dose series, with the second and third doses administered at 2 and 6 months after the first dose, respectively[227]. The vaccine protects against high-risk HPV types 16 and 18 and low-risk HPV types 6 and 11, and is most effective when administered prior to HPV exposure before sexual debut[135,228]. Persistent infection with HPV types 16 and 18 is causally associated with a significant proportion of anal, penile and oropharyngeal cancers in males[103] and infection with low-risk HPV types 6 and 11 are responsible for

nearly all cases of genital warts[133]. In females, nearly 70% of all cervical cancers are caused by HPV types 16 and 18[229].

Data on determinants of HPV vaccination among males are limited, but suggest a health care provider's recommendation as one of the most important predictors of HPV vaccine initiation[230]. Additionally, rates of vaccine initiation are generally higher among non-Hispanic black and Hispanic males (vs. non-Hispanic whites) and among males living below the poverty level (vs. at or above the poverty level)[150,152,230]. Less is known about factors related to HPV vaccine completion among males; however a few studies have shown that rates are lower among non-Hispanic black and Hispanic males and among uninsured/underinsured male adolescents[150,152,230]. Frequent contact with the healthcare system has also been cited as an important predictor of completion, particularly among low-income and minority males[152,231]. Most of the evidence on predictors of HPV vaccine initiation and completion among males was generated before the ACIP recommendation in 2011 with a majority of studies focusing on vaccine acceptability[232–239]. More recent studies, such as the National Immunization Survey-Teen (NIS-Teen), include limited age ranges (13 – 17 years) and do not include males in the target age range (11-12 years)[150,230,237,240]. To our knowledge, no recent studies have examined predictors of the second dose of the HPV vaccine, despite growing interest in reduced dosing schedules of the HPV vaccine series[141,241,242]. To this end, the purpose of our study was to identify predictors of each dose of the HPV vaccine series among a large, clinical population of males aged 11 to 26 years after the ACIP recommendation, January 2012 through April 2013.

METHODS

Study Population

We evaluated EMR data from 15,996 males aged 11 to 26 years attending Johns Hopkins Community Physicians (JHCP) clinics from January 2012 through April 2013. JHCP is a university-affiliated practice comprised of 26 primary care outpatient sites in 11 counties in Maryland. Our study population was drawn from the Family Practice, Internal Medicine/Pediatrics (IM/Peds), Internal Medicine, and Pediatrics practice specialties at these facilities. Males who received an HPV vaccine dose outside of a JHCP clinic (n=101) and those who initiated the HPV vaccine series prior to the start of our study (n=1,207, 7.6%) were excluded. We created three analytic cohorts to evaluate HPV vaccine initiation, receipt of the second dose of the HPV vaccine, and HPV vaccine completion. Therefore, the analytic cohort for HPV vaccine initiation included 14,688 males who had not received an HPV vaccine dose as of 2012. Dates in the EMR data included only visit year (vs. month and year). As such, we could not determine whether males who initiated in 2013 (n=346) or those who received the second dose in 2013 (n=202) had enough time (i.e., 6 months) to complete the series; these males were excluded from the second dose and completion analytic cohorts, respectively. Thus, the analytic cohort for the second dose of the HPV vaccine included the 1,834 males who initiated the HPV vaccine in 2012, and the analytic cohort for the completion analysis included the 702 males who received the second dose in 2012. This study protocol was approved by the Johns Hopkins Medicine Institutional Review Board.

HPV Vaccination Outcome Definitions

Information on HPV vaccination status was available from the EMR. HPV vaccine "initiation" was defined as receipt of at least one dose of the HPV vaccine, the second dose was defined as receipt of two doses of the HPV vaccine, and HPV vaccine "completion" was defined as receipt of all three doses of the HPV vaccine series.

Demographic and Clinical Predictors of HPV Vaccination

Demographic and clinical characteristics were available from the EMR. We evaluated age at the first clinic visit during the study period (i.e., "baseline") as a continuous variable and also categorized baseline age according to the ACIP recommendations: 11-12 years (target age range for vaccination, "Target"), 13-21 years (catch-up age range for vaccination, "Catch-Up", and 22-26 years (permissive age for vaccination, "Permissive"). Race/ethnicity was self-identified in the registration files of the EMR and defined as non-Hispanic white, non-Hispanic black, Hispanic, Asian, or other race/ethnicity. Insurance was categorized as private, public, or military. The number of clinic visits during the study period was categorized as ≤3 visits (the minimum number of visits required to complete the HPV vaccine series) vs. >3 visits. JHCP clinic location was defined as urban or suburban using U.S. census data and JHCP practice specialty was categorized as Family Practice, IM/Peds, Internal Medicine, or Pediatrics. Because males could visit more than one practice specialty type during the study period, we assigned each male the most common practice specialty observed. When we were

unable to identify the most common practice specialty because a male attended an equal numbers of different specialties (n=602, 3.8%), we used the practice specialty at the male's first visit. Among males who were vaccinated, agreement between the assigned practice specialty and the specialty associated with the vaccine visit was 95%.

Statistical Analysis

In this cross-sectional analysis, we calculated means and proportions for demographic and clinical predictors, using descriptive statistics with t-tests and Pearson's chi-square tests to assess differences by uptake of each HPV vaccine dose. Multivariable logistic regression models using generalized estimating equations (GEE) were used to calculate adjusted odds ratios (aORs) and 95% confidence intervals of associations of demographic and clinical predictors with each HPV vaccine dose, accounting for clustering within JHCP clinics. All models were stratified by baseline age group and mutually adjusted for continuous baseline age, race/ethnicity, insurance type, number of clinic visits, JHCP clinic location, and JHCP practice specialty. Since additional clinic visits are required for receipt of subsequent doses of the HPV vaccine, we conducted a sub-analysis to explore whether the number of clinic visits modifies any potential association of race/ethnicity and insurance type with receipt of the second and third dose of the HPV vaccine, respectively. In this analysis we focused on race/ethnicity and insurance because these factors are known to be differentially associated with healthcare utilization patterns (14). To increase statistical power and adequately test for interaction, we combined the target and catch-up age groups and re-categorized race/ethnicity as non-

Hispanic white, non-Hispanic black and other (Hispanic, Asian/Pacific Islander, and other race/ethnicity). We stratified our models by number of clinic visits (≤3 and >3), and tested for statistical interaction using the Wald test. All analyses were conducted using Stata v.13.1 (StataCorp, College Station, TX). All tests were 2-sided and results were considered statistically significant if $p<0.05$.

RESULTS

HPV Vaccine Initiation

Of the 14,688 males eligible for the first dose of the HPV vaccine, a total of 2,180 (14.8%) initiated the series. The average baseline age of males eligible for the first dose of the HPV vaccine was 18.0 ± 4.7 years, and the majority were non-Hispanic white (50.6%), followed by non-Hispanic black (35.1%; Table 3-1). More than half of all males were privately insured (Table 3-1). The majority of males attended ≤3 clinic visits during the study period, visited JHCP clinics in suburban locations and practices with a Family Practice or Pediatrics specialty (Table 3-1).

Multivariable aORs for HPV vaccine initiation by age group are shown in Table 3-2. In the target age group, non-Hispanic black males had 39% greater odds of initiating the HPV vaccine compared with non-Hispanic white males ($p=0.02$); and males with public insurance had 45% greater odds of HPV vaccine initiation compared with males with private insurance ($p=0.02$). Attending >3 clinic visits during the study period was associated with over a two-fold increase in odds of HPV vaccine initiation compared with ≤3 visits during the study period ($p<0.001$), and visiting a clinic in an urban location was associated with over a three-fold increase in odds of HPV vaccine initiation compared

with clinics in a suburban location ($p<0.01$). Baseline age and JHCP practice specialty were not significantly associated with odds of HPV vaccine initiation in this target age group.

In the catch-up age group, similar to males in the target age group, non-Hispanic black race/ethnicity, public insurance, attending >3 clinic visits during the study period, and urban clinic location were significantly associated with increased odds of HPV vaccine initiation ($p<0.01$, respectively). Additionally in the catch-up age group, older age at baseline was significantly associated with a 14% decrease in odds of HPV vaccine initiation ($p<0.001$) and Internal Medicine practice specialty was significantly associated with a 73% decrease in odds of HPV vaccine initiation compared with Family Medicine practice specialty ($p<0.001$).

Similar to males in both the target and catch-up age groups, non-Hispanic black race/ethnicity in the permissive age group and attending >3 clinic visits during the study period were significantly associated with increased odds of HPV vaccine initiation ($p<0.001$, respectively). Like males in the catch-up age group, older age at baseline and Internal Medicine practice specialty were significantly associated with decreased odds of HPV vaccine initiation ($p=0.05$ and $p=0.02$, respectively). Additionally in the permissive age group, males with military insurance had nearly two-and-a-half times the odds of initiating the HPV vaccine compared to males with private insurance ($p=0.04$). JCHP clinic location was not significantly associated with odds of HPV vaccine initiation in this permissive age group.

Second Dose of the HPV Vaccine

Of the 1,834 eligible males (those who received their first HPV vaccine dose in 2012), a total of 904 (49.1%) received the second dose of the HPV vaccine. Males eligible for the second dose of the HPV vaccine tended to be younger than those eligible for initiation, with a mean baseline age of 14.9 ± 3.2 years (Table 3-1). A majority of these males were non-Hispanic black (57.9%), publicly insured (43.5%), and attended ≤3 clinic visits during the study period (70.7%; Table 3-1). Approximately half attended JHCP clinics in suburban locations and the most common JHCP practice specialty was Pediatrics (62.9%; Table 3-1).

Multivariable aORs for the second dose of the HPV vaccine by age group are shown in Table 3-3. Results for the permissive age group are not shown due to limited statistical power. In the target age group, non-Hispanic black males, Hispanic males, and males who identified as other race/ethnicity had 27%, 61%, and 74% decreased odds, respectively, of receiving the second dose of the HPV vaccine compared with non-Hispanic white males ($p \leq 0.05$, respectively). Attending >3 clinic visits during the study period was significantly associated with a four-fold increase in odds of receiving the second dose compared with ≤3 visits ($p < 0.001$). Baseline age, insurance type, JHCP clinic location, and JHCP practice specialty were not significantly associated with odds of receiving the second dose of the HPV vaccine in this age group.

In the catch-up age group, similar to males in the target age group, non-Hispanic black race/ethnicity was significantly associated with decreased odds of receiving the second dose of the HPV vaccine ($p = 0.02$), and attending >3 clinic visits during the study

period was associated with nearly a six-fold increase in odds of receiving the second dose of the HPV vaccine in the catch-up age group ($p<0.001$). Additionally, males with public insurance had 27% decreased odds of receiving the second dose of the HPV vaccine compared with males with private insurance ($p=0.02$) and IM/Peds practice specialty was significantly associated with a 48% decrease in odds of receiving the second dose of the HPV vaccine compared with Family Practice specialty ($p=0.04$). Baseline age and JHCP clinic location were not significantly associated with odds of receiving the second dose of the HPV vaccine in this age group.

In the target and catch-up age groups, non-Hispanic black males with ≤3 visits had lower odds (aOR 0.62, 95% CI 0.4 – 0.9) of receiving the second dose of the HPV vaccine compared to their non-Hispanic white male counterparts. This association was nearly equivalent for non-Hispanic black males with >3 visits (aOR 0.66, 95% CI 0.4 – 1.1). Similar patterns were observed for males in the combined "Other" race/ethnicity category (data not shown). Number of clinic visits did not modify the association between race/ethnicity and receipt of the second dose of the HPV vaccine (p-interaction = 0.07). In contrast, in the target and catch-up age groups, publicly insured males with ≤3 visits had significantly lower odds of receiving the second dose of the HPV vaccine (aOR 0.67, 95% CI 0.5 – 0.9) compared to their privately insured counterparts; however, this association was attenuated for publicly insured males with >3 visits (aOR 1.05, 95% 0.7 – 1.6). There was no significant difference in the odds of receipt of the second dose of the HPV vaccine when comparing males with military insurance to their privately insured counterparts, irrespective of number of clinic visits (≤3 visits: aOR 0.85, 95% CI 0.6 –

1.2 vs. >3 visits: aOR 0.81, 95% CI 0.3 – 1.9). Number of clinic visits modified the association between insurance type and receipt of the second dose of the HPV vaccine (p-interaction = 0.001).

HPV Vaccine Completion

Of the 702 eligible males (those who received their second HPV vaccine dose in 2012), a total of 331 (47.2%) completed the series during the study timeframe. Males eligible for the third dose of the HPV vaccine tended to be younger than those eligible for initiation, with a mean baseline age of 14.7 ± 3.2 years (Table 3-1). A majority of these males were non-Hispanic black (46.7%) and privately insured (44.2%; Table 3-1). About half attended ≤3 clinic visits during the study period, the majority attended JHCP clinics in suburban locations (61.1%) and the most common JHCP practice specialty was Pediatrics (57.3%; Table 3-1).

Multivariable aORs for HPV vaccine completion by age group are shown in Table 3-4. Results for the permissive age group are not shown due to limited statistical power. In the target age group, attending >3 clinic visits during the study period was significantly associated with over a three-and-a-half-fold increase in odds of HPV vaccine completion compared with ≤3 visits ($p<0.001$) and visiting a clinic in an urban location was significantly associated with a 43% decrease in odds of HPV vaccine completion compared with suburban locations ($p<0.01$). Baseline age, race/ethnicity, insurance type, and JHCP practice specialty were not significantly associated with odds of HPV vaccine completion in this target age group.

In the catch-up age group, similar to males in the target age group, attending >3 clinic visits during the study period was significantly associated with over a three-and-a-half-fold increase in odds of HPV vaccine completion ($p<0.001$). Additionally, males with public insurance had 50% decreased odds of HPV vaccine completion compared with males with private insurance ($p=0.05$). Baseline age, race/ethnicity, JHCP clinic location, and JCHP practice specialty were not significantly associated with odds of HPV vaccine completion in this age group.

In the target and catch-up age groups, non-Hispanic black males with ≤3 visits had lower odds (aOR 0.53, 95% CI 0.2 – 1.3) of completing the HPV vaccine compared to their non-Hispanic white male counterparts. This association was nearly equivalent for non-Hispanic black males with >3 visits (aOR 0.69, 95% CI 0.3 – 1.5). Similar patterns were observed for males in the combined "Other" race/ethnicity category (data not shown). Number of clinic visits did not modify the association between race/ethnicity and completing the HPV vaccine (p-interaction = 0.14). In contrast, in the target and catch-up age groups, publicly insured males with ≤3 visits had significantly lower odds of completing the HPV vaccine (aOR 0.46, 95% CI 0.3 – 0.6) compared to their privately insured counterparts; however, this association was attenuated for publicly insured males with >3 visits (aOR 0.72, 95% 0.3 – 1.7). There was no significant difference in the odds of completing the HPV vaccine when comparing males with military insurance to their privately insured counterparts, irrespective of number of clinic visits (≤3 visits: aOR 0.76, 95% CI 0.3 – 1.7 vs. >3 visits: aOR 0.89, 95% CI 0.5 – 1.5). Number of clinic

visits modified the association between insurance type and completing the HPV vaccine (p-interaction = 0.0001).

DISCUSSION

In this large clinical population of over 14,500 males aged 11 to 26 years, the overall proportion of HPV vaccine initiation was low, with approximately 15% of males receiving at least one dose of the vaccine between January 2012 and April 2013. We observed differences in rates of initiation by age group; approximately 25% of males in the target age group initiated the HPV vaccine, while 18.5% and 2% of males in the catch-up and permissive age groups initiated the HPV vaccine, respectively. Our rates of initiation were lower than those reported from the NIS-Teen, which estimated that 35% of males aged 13-17 years initiated the HPV vaccine in the U.S. in 2013 (21% in 2012[230]). In our study, among all males who initiated the HPV vaccine in 2012, 49% received the second dose, and among those who received the second dose in 2012, 47% completed the HPV vaccine series. Our rates of completion were comparable with those reported for the general U.S. male population in the NIS-Teen study, which estimated that 48% of males who initiated the HPV vaccine (45.1% in 2012[230]) completed the series in 2013.

Among all age groups, we found that non-Hispanic black males were more likely to initiate the HPV vaccine compared with non-Hispanic white males. Irrespective of race/ethnicity, males in the target and catch-up age groups who were publicly insured were also more likely to initiate the HPV vaccine. These findings are in line with

previous studies among both males and females suggesting higher HPV vaccine initiation rates among non-Hispanic black and publicly insured populations[150,152,230]. Although cost has been previously cited as a barrier to HPV vaccination[243–245], efforts over the past several years have focused on improving HPV vaccine reimbursement[246]. For low-income children, the Vaccine For Children (VFC) program provides access to the HPV vaccine for Medicaid and underinsured children less than 18 years of age[247]. In the private sector, the Affordable Care Act requires most private insurance plans to cover the HPV vaccine at no cost to patients up to 18 years of age[248]. We also observed that males in the permissive age group with military insurance were more likely to initiate the HPV vaccine compared with males with private insurance. HPV vaccination is a covered benefit for males aged 11 to 26 years under military insurance plans[249]. Given that cost should not be a barrier going forward, interventions targeting parents and/or providers to increase HPV vaccine initiation may be warranted.

In contrast to our findings for HPV vaccine initiation, we found that non-Hispanic black males (vs. non-Hispanic white) in both the target and catch-up age groups and males with public insurance (vs. private insurance) in the catch-up age group were less likely to receive subsequent doses of the HPV vaccine. These findings are comparable with previous studies reporting lower completion rates among non-Hispanic black and publicly insured/underinsured populations[152]. It is unclear why the same males who are more likely to initiate the vaccine series are less likely to receive subsequent doses. Our data indicated that for non-Hispanic black males, returning for additional clinic visits did not explain this disparity; however for publicly insured males, those who attended >3

clinic visits during the study period were equally likely to complete the HPV vaccine series compared to males with private insurance. These findings suggest provider alerts and/or patient reminder systems may facilitate HPV series completion for all males, and could be particularly effective among minority and publicly insured male patients.

We also found important clinical predictors associated with HPV vaccination in our study. Among all age groups, attending >3 clinic visits was associated with increased odds of HPV vaccine initiation and with receipt of subsequent doses. These findings are similar to other studies reporting that males require more primary care visits to complete HPV vaccine series[231,250]. We also found that males in the catch-up and permissive age groups who primarily attended Internal Medicine clinics (vs. Family Practice) were less likely to initiate the HPV vaccine, however once they initiated, they were equally likely to receive subsequent doses. Together these findings have important implications for clinical intervention strategies. For example, broad interventions encouraging routine healthcare visits may promote HPV vaccine initiation and completion among all age-eligible males, whereas more targeted interventions focused on increasing vaccine initiation among patients of Internal Medicine physicians may be needed for increasing coverage in males who require catch-up HPV vaccination.

To our knowledge, this is one of the first and largest studies of demographic and clinical predictors of HPV vaccination among age-eligible males after the ACIP began routinely recommending the vaccine in 2011[152]. Our study is unique in that we assessed independent predictors of each dose of the HPV vaccine, and contributes to the literature by identifying predictors of the second dose of the vaccine. Additional strengths include

our diverse study population in terms of patient age, race/ethnicity, insurance, and practice specialties. However, some important limitations are worth noting. First, our study was limited to clinics affiliated with a single academic-institution in Maryland, and therefore our results might not be generalizable to non-academic practice settings or other geographic regions. Second, because we assigned each male a practice specialty type based on his most common visit or first visit (if most common was not available), it is possible that we misclassified practice specialty type; however we would expect such misclassification to be non-differential by vaccine status. Third, we did not have exact visit date, and therefore were limited in our ability to assess timing of each vaccine dose. Finally, we used data obtained from the medical record, which is subject to the limitations of databases that were not designed for research purposes (e.g., lack of data on potential confounders such as parent perceptions, provider recommendation to vaccinate, etc.).

In conclusion, our study indicates that a substantial proportion of age-eligible males attending primary care clinics did not receive the HPV vaccine during visits with their healthcare provider. Consistent with the literature, we found important disparities in HPV vaccine completion by race/ethnicity and insurance status. Moreover, we provide new evidence demonstrating that these disparities are as equally important for receipt of the second dose of the HPV vaccine. These findings point toward a need for understanding barriers to receiving subsequent doses of the HPV vaccine and focused interventions among minority and publicly insured males to ensure HPV vaccine series completion. Further, our data suggest that interventions may need to be targeted by

provider specialty, and warrant future research on provider-level factors associated with HPV vaccination.

Acknowledgements:

We wish to thank the Optimizing Cancer Screening Working Group, also known as the Johns Hopkins Program to Individualize Cancer Screening (*in*CAS) for their support of this study. We would also like to thank the Research Staff at Johns Hopkins Community Physicians for their expert assistance with this project.

Table 3-1. Demographic and Clinical Characteristics of Males Aged 11-26 Years Attending JHCP Clinics from 2012-2013 by HPV Vaccine Dose Eligibility

	HPV Vaccine Dose Eligibility		
	HPV Vaccine Initiation	HPV Vaccine 2nd Dose	HPV Vaccine Completion
Total N	14,688	1,834	702
Total Vaccinated n (%)	2,180 (14.8)	904 (49.1)	331 (47.2)
Mean Baseline Age (SD)	18.0 (4.7)	14.9 (3.2)	14.7 (3.2)
Age Group n (%)			
Target	2,471 (16.8)	493 (26.9)	197 (28.1)
Catch-Up	8,011 (54.5)	1,270 (69.2)	472 (67.2)
Permissive	4,206 (28.6)	71 (3.9)	33 (4.7)
Race/Ethnicity n (%)			
White	7,432 (50.6)	585 (31.9)	291 (41.4)
Black	5,163 (35.1)	1,062 (57.9)	328 (46.7)
Hispanic	684 (4.7)	66 (3.6)	32 (4.6)
Asian/Pacific Is.	454 (3.1)	29 (1.6)	13 (1.9)
Other	955 (6.5)	92 (5.0)	38 (5.4)
Insurance Type n (%)			
Private	8,688 (59.1)	709 (38.7)	310 (44.2)
Public	3,114 (21.2)	797 (43.5)	226 (32.2)
Military	2,494 (17.0)	318 (17.3)	162 (23.1)
Missing	392 (2.7)	10 (0.5)	4 (0.5)
Number of Clinic Visits n (%)			
1-3 Visits	12,325 (83.9)	1,296 (70.7)	346 (49.3)
>3 Visits	2,363 (16.1)	538 (29.3)	356 (50.7)
JHCP Clinic Location n (%)			
Suburban	10,955 (74.6)	939 (51.2)	436 (61.1)
Urban	3,730 (25.4)	895 (48.8)	266 (37.9)
JHCP Practice Specialty n (%)			
Family Practice	5,643 (38.4)	494 (26.9)	230 (32.8)
IM/Peds	1,021 (7.0)	115 (6.3)	41 (5.8)
Internal Med	3,553 (24.2)	72 (3.9)	29 (4.1)
Pediatrics	4,471 (30.4)	1,153 (62.9)	402 (57.3)

Abbreviations: Asian/Pacific Is., Asian/Pacific Islander; IM/Peds, Internal Medicine/Pediatrics; Internal Med, Internal Medicine

Table 3-2. Associations of Demographic and Clinical Characteristics with HPV Vaccine Initiation by Age Group Among 14,688 Males Attending JHCP Clinics from 2012-2013

	Target (11-12 years)			Catch-Up (13-21 years)			Permissive (22-26 years)		
	HPV Vaccine Initiation			HPV Vaccine Initiation			HPV Vaccine Initiation		
	Yes (n=617)	No (n=1,854)	aOR (95% CI)	Yes (n=1,483)	No (n=6,528)	aOR (95% CI)	Yes (n=80)	No (n=4,126)	aOR (95% CI)
Mean Baseline Age (SD)	11.43 (0.5)	11.42 (0.5)	1.07 (0.82 – 1.40)	15.80 (2.2)	17.15 (2.6)	0.86 (0.84 – 0.89)	23.53 (1.3)	24.03 (1.4)	0.84 (0.72 – 0.97)
Race/Ethnicity n									
White	187	943	1.0 (Reference)	469	3,400	1.0 (Reference)	39	2,394	1.0 (Reference)
Black	368	660	1.39 (1.06 - 1.84)	866	2,165	1.39 (1.12 - 1.74)	32	1,072	1.93 (1.20 - 3.09)
Hispanic	20	65	1.36 (0.79 - 2.35)	54	312	1.04 (0.76 - 1.42)	5	228	1.58 (0.42 - 5.89)
Asian	11	57	1.15 (0.69 - 1.90)	21	198	0.85 (0.46 - 1.57)	2	165	0.76 (0.18 - 3.26)
Other	31	129	1.16 (0.77 - 1.75)	73	453	1.01 (0.76 - 1.33)	2	267	0.50 (0.12 - 2.11)
Insurance Type n									
Private	193	896	1.0 (Reference)	588	3,692	1.0 (Reference)	59	3,260	1.0 (Reference)
Public	312	457	1.45 (1.07 - 1.96)	625	1,237	1.21 (1.05 - 1.39)	9	474	0.61 (0.28 - 1.37)
Military	112	493	1.25 (1.03 - 1.51)	261	1,466	1.06 (0.91 - 1.25)	10	152	2.46 (1.06 - 5.71)
Number of Clinic Visits n									
1-3 Visits	429	1,528	1.0 (Reference)	1,103	5,562	1.0 (Reference)	51	3,652	1.0 (Reference)
>3 Visits	188	326	2.39 (1.84 - 3.11)	380	966	2.19 (1.76 – 2.73)	29	474	4.08 (2.20 –7.56)
JHCP Clinic Location n									
Suburban	299	1,474	1.0	765	5,298	1.0	63	3,056	1.0

			(Reference)			(Reference)			(Reference)
Urban	318	380	3.27 (1.53 - 6.99)	718	1,229	3.84 (2.23 – 6.61)	17	1,068	0.98 (0.35 – 2.76)
JHCP Specialty n									
Family Practice	120	613	1.0	423	2,873	1.0	43	1,571	1.0
			(Reference)			(Reference)			(Reference)
IM/Peds	39	125	1.17 (0.73 - 1.88)	86	515	1.00 (0.74 - 1.34)	7	249	0.97 (0.35 - 2.66)
Internal Med	5	4	1.43 (0.78 - 2.63)	55	1,187	0.27 (0.15 - 0.47)	29	2,273	0.47 (0.22 - 0.98)
Pediatrics	455	1,112	0.81 (0.51 - 1.28)	919	1,953	1.28 (0.95 - 1.73)	1	33	1.77 (0.23 - 13.37)

All models mutually adjusted for all variables listed in the table. Abbreviations: aOR – Adjusted Odds Ratio; CI, Confidence Interval; Asian/Pacific Is., Asian/Pacific Islander; IM/Peds, Internal Medicine/Pediatrics; Internal Med, Internal Medicine

Table 3-3. Associations of Demographic and Clinical Characteristics with 2nd Dose of the HPV Vaccine by Age Group Among 1,834 Males Attending JHCP Clinics from 2012-2013

	Target (11-12 years)			Catch-Up (13-21 years)		
	2nd Dose HPV Vaccine			2nd Dose HPV Vaccine		
	Yes (n=261)	No (n=232)	aOR (95% CI)	Yes (n=603)	No (n=667)	aOR (95% CI)
Mean Baseline Age (SD)	11.43 (0.5)	11.46 (0.5)	0.86 (0.65 - 1.13)	15.55 (2.1)	15.97 (2.2)	0.95 (0.88 - 1.02)
Race/Ethnicity n						
White	104	43	1.0 (Reference)	257	147	1.0 (Reference)
Black	132	160	0.63 (0.40 - 1.00)	283	458	0.65 (0.46 - 0.92)
Hispanic	8	10	0.39 (0.17 - 0.89)	21	23	0.72 (0.33 - 1.54)
Asian	6	4	0.50 (0.17 - 1.46)	10	7	1.11 (0.52 - 2.40)
Other	11	15	0.26 (0.09 - 0.81)	32	32	0.71 (0.37 - 1.37)
Insurance Type n						
Private	100	61	1.0 (Reference)	266	230	1.0 (Reference)
Public	104	142	0.77 (0.43 - 1.39)	195	347	0.73 (0.56 - 0.96)
Military	57	29	1.03 (0.60 - 1.79)	138	86	0.78 (0.52 - 1.16)
Number of Clinic Visits n						
1-3 Visits	136	194	1.0 (Reference)	326	596	1.0 (Reference)
>3 Visits	125	38	4.01 (3.22 - 4.99)	277	71	5.82 (4.13 - 8.20)
JHCP Clinic Location n						
Suburban	149	89	1.0 (Reference)	374	273	1.0 (Reference)
Urban	112	143	1.08 (0.35 - 3.34)	229	394	0.77 (0.46 - 1.31)
JHCP Specialty n						
Family Practice	64	37	1.0 (Reference)	205	149	1.0 (Reference)
IM/Peds	19	13	0.63 (0.34 - 1.19)	32	44	0.52 (0.28 - 0.97)
Internal Med	0	3	1.00 (1.00 - 1.00)	17	28	0.65 (0.31 - 1.40)
Pediatrics	178	179	0.57 (0.30 - 1.06)	349	446	1.07 (0.74 - 1.54)

All models mutually adjusted for all variables listed in the table. Results for the permissive age group are not shown due to limited statistical power

Abbreviations: aOR – Adjusted Odds Ratio; CI, Confidence Interval; Asian/Pacific Is., Asian/Pacific Islander; IM/Peds, Internal Medicine/Pediatrics; Internal Med, Internal Medicine

Table 3-4. Associations of Demographic and Clinical Characteristics with HPV Vaccine Completion by Age Group Among 702 Males Attending JHCP Clinics from 2012-2013

	Target (11-12 years)			Catch-Up (13-21 years)		
	HPV Vaccine Completion			HPV Vaccine Completion		
	Yes (n=86)	No (n=111)	aOR (95% CI)	Yes (n=228)	No (n=244)	aOR (95% CI)
Mean Baseline Age (SD)	11.45 (0.5)	11.41 (0.5)	1.27 (0.72 - 2.23)	15.47 (2.1)	15.55 (2.0)	1.01 (0.91 - 1.12)
Race/Ethnicity n						
White	40	38	1.0 (Reference)	116	83	1.0 (Reference)
Black	36	60	0.63 (0.31 - 1.29)	87	130	0.54 (0.25 - 1.19)
Hispanic	4	4	1.01 (0.38 - 2.73)	10	11	0.57 (0.17 - 1.88)
Asian	3	2	1.95 (0.17 - 23.02)	2	6	0.24 (0.06 - 0.99)
Other	3	7	0.36 (0.11 - 1.21)	13	14	0.62 (0.31 - 1.26)
Insurance Type n						
Private	35	42	1.0 (Reference)	112	95	1.0 (Reference)
Public	29	45	1.10 (0.74 - 1.63)	55	94	0.50 (0.25 - 1.01)
Military	22	23	1.14 (0.69 - 1.86)	58	55	0.67 (0.43 - 1.04)
Number of Clinic Visits n						
1-3 Visits	29	72	1.0 (Reference)	78	157	1.0 (Reference)
>3 Visits	57	39	3.69 (2.00 - 6.82)	150	87	3.56 (2.34 - 5.40)
JHCP Clinic Location n						
Suburban	54	62	1.0 (Reference)	151	144	1.0 (Reference)
Urban	32	49	0.57 (0.38 - 0.85)	77	100	1.42 (0.64 - 3.12)
JHCP Specialty n						
Family Practice	18	32	1.0 (Reference)	82	83	1.0 (Reference)
IM/Peds	7	7	1.51 (0.49 - 4.69)	11	13	0.82 (0.37 - 1.80)
Internal Med	0	0	-- --	7	7	1.10 (0.32 - 3.83)
Pediatrics	61	72	2.15 (0.93 - 5.00)	128	141	1.34 (0.81 - 2.20)

All models mutually adjusted for all variables listed in the table. Results for the permissive age group are not shown due to limited statistical power

Abbreviations: aOR – Adjusted Odds Ratio; CI, Confidence Interval; Asian/Pacific Is., Asian/Pacific Islander; IM/Peds, Internal Medicine/Pediatrics; Internal Med, Internal Medicine

Chapter 4

Influence of Home and School Environments on Specific Dietary Behaviors Among Postpartum, High-Risk Teens, 27 States, 2007-2009

Title: Influence of Home and School Environments on Specific Dietary Behaviors Among Postpartum, High-Risk Teens, 27 States, 2007-2009

Megan A. Clarke, Debra L Haire-Joshu, Cynthia D. Schwarz, Rachel G. Tabak, Corinne E. Joshu

Author Affiliations: Debra L. Haire-Joshu, Cynthia D. Schwarz, Rachel G. Tabak, Washington University in St Louis, St Louis, Missouri; Corinne E. Joshu, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland.

Keywords: Obesity, environment, adolescent, postpartum, diet

Corresponding Author:
Megan A. Clarke, MHS
Johns Hopkins Bloomberg School of Public Health
615 N Wolfe St, Baltimore, MD 21205.
Telephone: 443-287-3821.
Email: mclark43@jhu.edu.

This chapter has been published:

Clarke MA, Haire-Joshu DL, Schwarz CD, Tabak RG, Joshu CE. Relative influence of home and school environments on specific dietary behaviors among postpartum, high-risk teens, 2007 – 2009. 2015;12:140437. PMCID: PMC4436050

ABSTRACT

Introduction: The objective of this study was to determine whether perceptions of the home and school food environments are related to food and beverage intakes of postpartum teens.

Methods: Our study was a baseline, cross-sectional analysis of 853 postpartum teens enrolled in a weight-loss intervention study across 27 states from 2007 through 2009. Eight-item scales assessed perceived accessibility and availability of foods and beverages in school and home environments. Associations between environments and intakes were assessed by using $\chi 2$ and using logistic regression with generalized estimating equations (GEE), respectively.

Results: Overall, 52% of teens perceived their school food environment as positive, and 68% of teens perceived their home food environment as positive. A positive school environment was independently associated with fruit consumption and 100% fruit juice consumption. A positive home environment was independently associated with fruit, vegetable, and water consumption and infrequent consumption of soda and chips ($\chi 2$ $P < .05$). Having only a positive school environment was associated with fruit consumption (GEE odds ratio [OR], 3.1; 95% confidence interval [CI], 1.5–6.5), and having only a positive home environment was associated with fruit (GEE OR, 2.9; 95% CI, 1.6–5.6), vegetable (GEE OR, 3.1; 95% CI, 1.5–6.2), and water (GEE OR, 2.6; 95% CI, 1.7–4.0) consumption and infrequent consumption of soda (GEE OR, 0.5; 95% CI, 0.3–0.7).

Results for positive home and school environments were similar to those for positive home only.

Conclusion: Home and school environments are related to dietary behaviors among postpartum teens, with a positive home environment more strongly associated with healthful behaviors.

INTRODUCTION

Nearly one-third of adolescents are overweight or obese and thus are at greater risk for obesity and its long-term health consequences, such as diabetes, in adulthood[251,252]. This risk is significantly heightened for postpartum, teenaged mothers who have sociodemographic and behavioral risk factors for overweight and obesity, such as low socioeconomic status and poor diet[197]. Both the school and home environments influence dietary behaviors of teenagers, particularly in low-income and racial/ethnic minority populations[185,253]. Aspects of food environments that may be particularly important include availability and accessibility of healthful foods such as fruits and vegetables, low-fat snacks, and low-calorie beverages[254–256].

More recent evidence suggests that school-based interventions and policies may not be sufficient to overcome risks posed in other settings[257,258]. Reports from the Institute of Medicine suggest that although the school environment is a key target for obesity prevention programs, emphasis is also needed on the role of parents or caregivers in shaping dietary behaviors in the home[255,259].

Little is known about how postpartum teens perceive their food environments and whether those perceptions are related to their dietary behaviors[185,260]. In previous work with high-risk, postpartum teens, we found a stronger relationship between the perceived home food environment (vs school) and healthful dietary behaviors[261]. Here we aim to build on these findings by examining the associations between perceived school and home food environments and consumption of specific food and beverage items and

examining whether relationships vary by body mass index (BMI) and participation in nutrition assistance programs. We hypothesize that positive perceptions of food environments will be associated with healthful food and beverage intakes, and that these associations will differ by type of environment.

METHODS

Study population

This cross-sectional study includes baseline data from postpartum teens enrolled in the Moms for a Healthy Balance Weight-loss Intervention Study (BALANCE), a group-randomized, nested cohort study with an intervention component designed to reduce postpartum weight retention in young mothers[200]. BALANCE was developed in partnership with Parents as Teachers (PAT), a child development–parent education program supported by federal and state funds and delivered free of charge to over 200,000 families in all 50 states[262]. For this study, we selected 27 states on the basis of the number of adolescent parents expected in the state.

Detailed methods on the BALANCE intervention have been described elsewhere[200]. Briefly, trained PAT parent educators delivered an evidence-based curriculum via home visits, group activities, and online resources. Adolescents were eligible to participate if they were enrolled in the PAT Teen Program, were less than 1 year postpartum, and were not pregnant or planning to become pregnant. We enrolled 1,325 eligible adolescent mothers from 2007 through 2009, and the study concluded in 2010. A total of 141 of the 1,325 teen participants randomized did not complete the baseline assessment, and 45 were missing baseline data for the calculation of BMI,

leaving a total of 1,139 with complete data. For this analysis, teens who were underweight at baseline (n = 19) as well as those who reported they were not currently in school (n = 221) were excluded. An additional 46 teen participants did not have information on food environments, leaving a total of 853 included in this analysis. The institutional review board of Saint Louis University and Washington University in St Louis approved this study, and informed consent was obtained from all participants.

Measures

Teen mothers self-reported characteristics including age, race/ethnicity, current education level, length of time since giving birth (postpartum status), breastfeeding status at baseline, and participation in the Supplemental Nutrition Assistance Program (SNAP) and the National School Lunch Program (NSLP).

Trained staff measured height and weight at baseline in accordance with the National Health and Nutrition Examination Survey (NHANES) study procedures[263]. Weight, height, and age data were used to calculate age-appropriate BMI categories, following the Centers for Disease Control and Prevention algorithm[264]. BMI was dichotomized as normal (<85th percentile) and overweight/obese (≥85th percentile).

Questionnaire items measuring perceived access of 4 healthful items (fruits and vegetables, low-fat products, low-calorie beverages, and low-calorie snacks) were used to characterize the home and school food environments[265,266]. For each environment, 8 statements assessed the availability and selection of healthful items at home (eg, "it is easy to find/there is a large selection of low-fat products in my home") and ease of

purchase and selection of healthful items at school (eg, "it is easy to purchase/there is a large selection of low-fat products in school"). Ratings were scored on a 5-point Likert scale (1 = "strongly agree" to 5 = "strongly disagree"). A mean score of the 8 items was created for the school and home food environments (Cronbach's $\alpha = 0.897$ and 0.902, respectively) and dichotomized as less than 3.0 being a positive environment and 3.0 or higher being a negative environment.

Dietary behaviors were assessed using the Snack and Beverage Food Frequency Questionnaire (SBFFQ) developed from our previous work[267,268]. A validation study and pilot testing were completed with 60 teen participants. The SBFFQ examined the young mothers' intake of 31 items during the prior 7 days by asking on how many days, how many times per week, and how much of the item she consumed. Items that were consumed by less than 25% of teens were excluded. Because of the nature and distribution of the data, data on the frequency of specific food and beverage items were collapsed into binary categories of infrequent (0–3 d/wk) and frequent (4–7 d/wk) consumption as a more conservative approach[269].

Statistical analyses

Descriptive statistics were calculated to evaluate baseline characteristics of all postpartum teens and by positive and negative school and home food environments. Differences in baseline characteristics by environment were assessed by using Pearson $\chi 2$ tests and t tests. Relationships between environments and frequency of food and beverage consumption were assessed by using Pearson $\chi 2$ tests. To evaluate the relative

strength of association between home and school environments and dietary behaviors, we created the following categories: “negative school and home,” “positive school only,” “positive home only,” and “positive school and home.” We used multiple logistic regression with generalized estimating equations (GEE) to account for clustering within a state. Potential confounders including NSLP and SNAP participation, race/ethnicity, age, and postpartum status, were identified on the basis of a priori knowledge and assessed by using a backward selection procedure. Final regression models were adjusted for race/ethnicity, age, and postpartum status, and results were calculated as GEE odds ratios (ORs) and 95% confidence intervals (CIs). To determine whether there were any differences by baseline weight status or participation in nutrition assistance programs, all models were stratified by BMI (ie, normal weight vs overweight/obese) and NSLP or SNAP participation. Data were analyzed by using Stata (Stata Intercooled, version 13; Stata Corp LP).

RESULTS

The mean age of the postpartum teens was 17 years (range, 12–20) and there were no significant age differences by perceived school or home environment (Table 4-1). Most teens identified as white (44%), black (29%), or Hispanic (20%). Racial distribution varied significantly by home environment, with a greater proportion of white teens reporting a positive home environment ($\chi 2$ $P < .05$). Slightly more than half of the teens had a normal BMI, and no significant differences were observed between home or school environment and BMI. Participation in SNAP and NSLP was common (30% and 40%, respectively) and varied significantly by home environment, with a greater

proportion of postpartum teens reporting a negative home environment also reporting receiving SNAP and/or NSLP benefits ($\chi 2$ $P < .05$). Most teens were from neighborhoods in rural or suburban settings, and neighborhood location varied significantly by school environment; teens living in a suburban neighborhood were more likely to perceive a negative school environment ($\chi 2$ $P < .05$). Approximately 75% of teens were 3 months or more postpartum and 12% reported that they were currently breastfeeding.

Overall, the item most likely to be consumed more than 3 days per week was chips, followed by cereal (Table 4-2). A positive school environment was significantly associated with eating fruit more than 3 days per week, while a positive home environment was significantly associated with eating cereal, fruit, and vegetables on more than 3 days per week and chips and chocolate on 0 to 3 days per week ($\chi 2$ $P < .05$). When we stratified by baseline BMI, the relationships between a positive home environment and frequency of chips and chocolate consumption were significant only among normal-weight teens ($\chi 2$ $P < .05$). When we stratified by NSLP and SNAP participation, patterns of frequency of intake of food items were similar to the patterns observed for all teens except 1) the relationship between positive school environment and frequency of fruit consumption was significant only for teens participating in NSLP ($\chi 2$ $P < .01$), and 2) the relationship between a positive home environment and frequency of fruit consumption was significant only among teens not receiving SNAP benefits ($\chi 2$ $P < .01$).

Overall, the beverage item most likely to be consumed more than 3 times per week was water, followed by regular soda (Table 4-2). A positive school environment was significantly associated with frequent consumption of 100% fruit juice as well as 2 types of sugar-sweetened beverages: fruit punch and sports drinks ($\chi 2$ $P < .05$). A positive home environment was significantly associated with frequent water, 100% fruit juice, and whole or 2% milk consumption, and infrequent regular soda consumption ($\chi 2$ $P < .05$). We found similar results when we stratified by baseline BMI; however, the significant relationship between a positive home environment and whole or 2% milk consumption was observed only for overweight/obese teens ($\chi 2$ $P < .05$). When we stratified by NSLP and SNAP participation, patterns of frequency of intake of beverage items were similar to the patterns observed for all teens except that a positive school environment was significantly associated only with drinking 100% fruit juice more than 3 days per week among teens who did not participate in NSLP ($\chi 2$ $P < .05$).

When compared with teens reporting negative school and home environments, a positive school environment only was significantly associated with increased odds of frequent fruit consumption (GEE OR, 3.1; 95% CI, 1.5–6.5) (Table 4-3). Compared with teens reporting negative school and home environments, a positive home environment only was significantly associated with frequent consumption of cereal (GEE OR, 2.3; 95% CI, 1.4–3.7), fruit (GEE OR, 2.9; 95% CI, 1.6–5.6), and vegetables (GEE OR, 3.1; 95% CI, 1.5–6.2) and infrequent consumption of chips (GEE OR, 0.5; 95% CI, 0.3–0.8), and a positive home and school environment was associated with increased odds of

frequent cereal (GEE OR, 1.7; 95% CI, 1.1–2.8), fruit (GEE OR, 2.9; 95% CI, 1.6–5.4), and vegetable (GEE OR, 3.2; 95% CI, 1.7–6.2) consumption.

Reporting only a positive school environment was not significantly associated with frequent consumption of any beverage items. Compared with teens reporting negative school and home environments, teens reporting a positive home environment only had increased odds of frequent water (GEE OR, 2.6; 95% CI, 1.7–4.0) and 100% fruit juice (GEE OR, 1.9; 95% CI, 1.2–2.9) consumption and infrequent consumption of regular soda (GEE OR, 0.5; 95% CI, 0.3–0.7). Compared with teens reporting negative school and home environments, teens reporting both positive home and school environments had similar results to those reporting only a positive home environment. Teens reporting both a positive home and school environment had significantly greater odds of frequent 100% fruit juice (GEE OR, 2.3; 95% CI, 1.5–3.6) and water consumption (GEE OR, 1.8; 95% CI, 1.2–2.6) and infrequent consumption of regular soda (GEE OR, 0.7; 95% CI, 0.5–1.0) than those reporting both negative home and school environments. Relative relationships between school and home food environments and food and beverage item consumption did not vary by baseline BMI. Significant associations between the positive school food environment and frequent consumption of healthful items such as fruit (GEE OR, 4.8; 95% CI, 1.6–14.6) and 100% fruit juice (GEE OR, 2.4; 95% CI, 1.1–5.6) were observed only among teens participating in the NSLP. The relationships between a positive home environment and both positive home and school environments did not differ substantially by NSLP participation. The relationship between the positive school food environment and dietary intake did not

differ by SNAP participation, but significant associations between a positive home environment and infrequent consumption of unhealthful items such as chips (GEE OR, 0.4; 95% CI, 0.2–0.8) and soda (GEE OR, 0.2; 95% CI, 0.1–0.5) were observed only among teens who received SNAP benefits. The same patterns were generally observed for both positive home and school environments.

DISCUSSION

Our findings indicate that the school and home food environments have differential relationships with food and beverage intakes. Our findings were similar to those from other studies: we found that a perceived positive school environment was primarily related to healthful eating behaviors such as frequent fruit or 100% fruit juice intake but not unhealthful eating behaviors[270,271]. In contrast, a perceived positive home environment was associated with frequent consumption of a wider variety of healthful items as well as infrequent consumption of unhealthful food and beverage items such as soda and chips. Our findings regarding the impact of positive school and home food environments suggest that for certain items consumed by teens, the major benefit lies within the home environment. This study contributes to our understanding of the relationship between both the home and school food environment and dietary behaviors of this understudied population of postpartum teens.

Numerous studies have documented the impact of policy and behavioral interventions promoting healthful school food environments on positive dietary change in youth[256,272,273]. Increased access to fruit and various juices may be a result of enhanced school wellness and nutrition policies, which promote access to and availability of select

foods[273,274]. In addition, school meal programs such as NSLP that promote fruit and vegetable intake in school environments provide opportunities for increased fruit and vegetable consumption among low-income teens[275]. However, easy access to and availability of high-calorie and high-fat snacks and sugar-sweetened beverages (ie, "competitive foods") that had been commonly sold in vending machines and at after-school fundraisers may have limited the effectiveness of school food policies and the influence of a positive school environment on teens' eating behaviors[272,276]. Our results as well as findings from other studies indicate that while a positive school environment may be related to frequent intake of certain healthful food and beverage items, it was not associated with infrequent intake of unhealthful items such as sweet and salty snacks and sugar-sweetened beverages[185,270,271,277]. These findings support the importance of recent changes in school food policies that limit access to unhealthful snacks by requiring improvements in the nutrition content of vending machine foods.

Unlike childhood obesity interventions in the school setting, interventions conducted in the home have not been common. Many of these interventions have focused on individual behavior change without addressing the home food environment, limiting their impact on dietary intake and other obesity-related outcomes[255,258]. Results from our study are consistent with the literature suggesting the home environment has an important relationship with dietary intake among adolescents[257,278]. The home food environment represents a substantial part of the full environmental context in which a postpartum teen grows, develops, eats, and behaves and is guided by "family policies" informed by tradition and culture as well as neighborhood and economic

environment[257,278]. Additionally, new mothers may be particularly aware of and sensitive to the health quality of their home setting[273]. Our findings suggest the multiple and variable influences of a positive home environment have the added benefit of reducing unhealthful behaviors among postpartum teens.

To our knowledge, ours is the first study to examine whether associations between the school and home environments and food and beverage intake differ by participation in nutrition assistance programs. Other studies have shown mixed associations between SNAP and NSLP participation and dietary behaviors[257,279]. Our findings suggest that the relationship between the food environment and frequency of consumption of certain items may be stronger among postpartum teens receiving nutrition assistance than those who did not receive assistance. Future research is needed to determine whether there are differences in the relationship between the environment and dietary behavior among teens that do and do not participate in nutrition assistance programs.

Our study has several limitations. This was a cross-sectional analysis; thus, we cannot evaluate causal relationships. Furthermore, reliance on self-reported data for dietary intake may be subject to recall bias and measurement error such as underreporting of items consumed. We attempted to limit potential misclassification by collapsing food and beverage frequency into dichotomous categories, but misclassification is a concern when using SBFFQ data[254,267]. Although we were not able to compare data on the school and home environments with objective measures, studies have shown that perceived quality of home- and school-based settings independently influences dietary

behavior[185,260]. Therefore, we consider using perceptions of the school and home food environments a strength of this study, particularly because we are among the first to address perceptions of the school and home food environments and how they are related to behavior. Additional strengths of this study include a large and nationally representative sample of postpartum teens, an understudied population with a high risk for overweight and obesity.

Our study highlights the importance of both the school and home food environments and their differential relationships with dietary behaviors among teens at high risk for obesity. Further work targeting interventions across both home and school environments simultaneously are needed. In addition, it is important to understand whether different subpopulations respond differently to environmental influences to tailor effective obesity interventions and policies. Improving the home environment may be particularly important among this population of teen mothers who directly control the food environment of their infants. Environmental interventions in this high-risk and hard-to-reach population may not only be important for reducing the risk of adult-onset obesity in the teenaged mother but may also have substantial impact in minimizing the intergenerational transfer of obesity-related behaviors to offspring[200].

Acknowledgments

This project was funded by the National Cancer Institute of the National Institutes of Health, grant no. USPHS 1 R01 CA1215. This study was also supported in part by grant

no. 1P30DK092950 and cooperative agreement no. U48/DP001903 from the Centers for Disease Control and Prevention (the Prevention Research Centers program). Ms Clarke was supported by the National Cancer Institute (T32CA009314).

Table 4-1. Characteristics of 853 Postpartum Teens and Their School and Home Food Environments[a], 27 States, 2007-2009

		School		Home	
Characteristic	**Total[b]**	**Positive**	**Negative**	**Positive**	**Negative**
Total N (%)	853 (100.0)	442 (51.8)	411 (48.2)	577 (67.6)	276 (32.4)
Age, y, mean (SD)	17.4 (1.1)	17.3 (1.1)	17.4 (1.0)	17.4 (1.0)	17.4 (1.1)
Race/ethnicity, n (%)[c]					
White	379 (44.4)	193 (43.7)	186 (45.3)	264 (45.7)	115 (41.7)
Black	247 (29.0)	131 (29.6)	116 (28.2)	151 (26.2)	96 (34.8)
Hispanic	173 (20.3)	86 (19.5)	87 (21.2)	121 (21.0)	52 (18.8)
Other/missing	54 (6.3)	32 (7.2)	22 (5.3)	41 (7.1)	13 (4.7)
BMI[d], n (%)					
Normal	480 (56.3)	248 (56.1)	232 (56.4)	314 (54.4)	166 (60.1)
Overweight/obese	373 (43.7)	194 (43.9)	179 (43.6)	263 (45.6)	110 (39.9)
Education, n (%)					
9th grade	88 (10.6)	53 (12.4)	35 (8.7)	55 (9.8)	33 (12.3)
10th grade	148 (17.9)	80 (18.7)	68 (17.0)	100 (17.9)	48 (17.9)
11th grade	251 (30.3)	125 (29.2)	126 (31.5)	172 (30.7)	79 (29.5)
12th grade	341 (41.2)	170 (39.7)	171 (42.8)	233 (41.6)	108 (40.3)
SNAP benefits[c], n (%)	254 (30.0)	133 (30.3)	121 (29.6)	155 (27.1)	99 (36.0)
NSLP benefits[c], n (%)	346 (40.8)	188 (42.8)	158 (38.6)	214 (37.4)	132 (48.0)
Neighborhood[e], n (%)					
Rural	345 (40.4)	186 (46.2)	159 (41.7)	237 (44.8)	108 (42.4)
Suburban	260 (33.2)	116 (28.8)	144 (37.8)	176 (33.3)	84 (32.9)
Urban	179 (22.8)	101 (25.1)	78 (20.5)	116 (21.9)	63 (24.7)
Time since giving birth, n (%)					
<3 months	158 (25.1)	81 (25.6)	77 (24.5)	116 (27.0)	42 (20.9)
3–6 months	193 (30.6)	107 (33.9)	86 (27.4)	130 (30.3)	63 (31.3)
>6 months	279 (44.3)	128 (40.5)	151 (48.1)	183 (42.7)	96 (47.8)
Breastfeeding[c], n (%)	96 (11.7)	56 (13.2)	40 (10.1)	81 (14.6)	15 (5.6)

Abbreviations: BMI, body mass index; NSLP, National School Lunch Program; SNAP, Supplemental Nutrition Assistance Program.

[a] See the Methods section for a description of how positive and negative perceptions were determined.

[b] Counts may not sum to overall total because of missing data.

[c] Significantly different for home environment χ^2 $P < .05$.

[d] Weight, height, and age data were used to calculate age-appropriate BMI categories, following the Centers for Disease Control and Prevention algorithm.

[e] Significantly different for school environment χ^2 $P < .05$.

Table 4-2. Associations Between Frequency of Food and Beverage Items Consumed and School and Home Food Environments[a] for 853 Postpartum Teens, 27 States, 2007-2009

		School		Home	
Item Consumed	Total, N (%)	Positive, n (%)	Negative, n (%)	Positive, n (%)	Negative, n (%)
Chips[b]					
0–3 d/wk	624 (73.2)	319 (72.2)	305 (74.2)	434 (75.2)	190 (68.8)
4–7 d/wk	229 (26.8)	123 (27.8)	106 (25.8)	143 (24.8)	86 (31.2)
Crackers					
0–3 d/wk	802 (94.0)	410 (92.8)	392 (95.4)	538 (93.2)	264 (95.7)
4–7 d/wk	51 (6.0)	32 (7.2)	19 (4.6)	39 (6.8)	12 (4.3)
Granola bars					
0–3 d/wk	812 (95.2)	417 (94.3)	395 (96.1)	545 (94.5)	267 (96.7)
4–7 d/wk	41 (4.8)	25 (5.7)	16 (3.9)	32 (5.5)	9 (3.3)
Cakes					
0–3 d/wk	764 (89.6)	394 (89.1)	370 (90.0)	522 (90.5)	242 (87.7)
4–7 d/wk	89 (10.4)	48 (10.9)	41 (10.0)	55 (9.5)	34 (12.3)
Cookies					
0–3 d/wk	785 (92.0)	402 (91.0)	383 (93.2)	531 (92.0)	254 (92.0)
4–7 d/wk	68 (8.0)	40 (9.0)	28 (6.8)	46 (8.0)	22 (8.0)
Chocolate[b]					
0–3 d/wk	750 (87.9)	389 (88.0)	361 (87.8)	520 (90.1)	230 (83.3)
4–7 d/wk	103 (12.1)	53 (12.0)	50 (12.2)	57 (9.9)	46 (16.7)
Hard candy					
0–3 d/wk	794 (93.1)	412 (93.2)	382 (92.9)	542 (93.9)	252 (91.3)
4–7 d/wk	59 (6.9)	30 (6.8)	29 (7.1)	35 (6.1)	24 (8.7)
French fries					
0–3 d/wk	738 (86.5)	381 (86.2)	357 (86.9)	505 (87.5)	233 (84.4)
4–7 d/wk	115 (13.5)	61 (13.8)	54 (13.1)	72 (12.5)	43 (15.6)
Pizza					
0–3 d/wk	811 (95.1)	415 (93.9)	396 (96.4)	551 (95.5)	260 (94.2)
4–7 d/wk	42 (4.9)	27 (6.1)	15 (3.6)	26 (4.5)	16 (5.8)
Cereal[b]					
0–3 d/wk	646 (75.7)	335 (75.8)	311 (75.7)	418 (72.4)	228 (82.6)
4–7 d/wk	207 (24.3)	107 (24.2)	100 (24.3)	159 (27.6)	48 (17.4)
Fruit[b,c]					
0–3 d/wk	712 (83.5)	357 (80.8)	355 (86.4)	468 (81.1)	244 (88.4)
4–7 d/wk	141 (16.5)	85 (19.2)	56 (13.6)	109 (18.9)	32 (11.6)
Vegetables[b]					
0–3 d/wk	722 (84.6)	367 (83.0)	355 (86.4)	468 (81.1)	254 (92.0)
4–7 d/wk	131 (15.4)	75 (17.0)	56 (13.6)	109 (18.9)	22 (8.0)
Water[b]					
0–3 d/wk	251 (29.4)	130 (29.4)	121 (29.4)	144 (25.0)	107 (38.8)
4–7 d/wk	602 (70.6)	312 (70.6)	290 (70.6)	433 (75.0)	169 (61.2)
Regular soda[b]					
0–3 d/wk	456 (53.5)	229 (51.8)	227 (55.2)	337 (58.4)	119 (43.1)
4–7 d/wk	397 (46.5)	213 (48.2)	184 (44.8)	240 (41.6)	157 (56.9)
100% Fruit juice[b,c]					
0–3 d/wk	597 (70.0)	292 (66.1)	305 (74.2)	381 (66.0)	216 (78.3)

4–7 d/wk	256 (30.0)	150 (33.9)	106 (25.8)	196 (34.0)	60 (21.7)
Fruit punch[c]					
0–3 d/wk	712 (83.5)	358 (81.0)	354 (86.1)	477 (82.7)	235 (85.1)
4–7 d/wk	141 (16.5)	84 (19.0)	57 (13.9)	100 (17.3)	41 (14.9)
Sports drinks[c]					
0–3 d/wk	787 (92.3)	397 (89.8)	390 (94.9)	530 (91.9)	257 (93.1)
4–7 d/wk	66 (7.7)	45 (10.2)	21 (5.1)	47 (8.1)	19 (6.9)
Whole or 2% milk[b]					
0–3 d/wk	472 (55.3)	234 (52.9)	238 (57.9)	304 (52.7)	168 (60.9)
4–7 d/wk	381 (44.7)	208 (47.1)	173 (42.1)	273 (47.3)	108 (39.1)
Sweet tea					
0–3 d/wk	711 (83.4)	371 (83.9)	340 (82.7)	483 (83.7)	228 (82.6)
4–7 d/wk	142 (16.6)	71 (16.1)	71 (17.3)	94 (16.3)	48 (17.4)

[a]See the Methods section for a description of how positive and negative perceptions were determined.

[b] Significantly different for home environment, χ^2 $P < .05$.

[c] Significantly different for school environment, χ^2 $P < .05$.

Table 4-3. GEE Logistic Regression[a] Analysis of Food Environments [b] and Frequency of Food and Beverage Consumption Among 853 Postpartum Teens, 27 States, 2007-2009

		GEE OR (95% CI)		
Item Consumed	**Negative School and Home (n = 179)**	**Positive School Only (n = 97)**	**Positive Home Only (n = 232)**	**Positive School and Home (n = 345)**
Food				
Chips	1 [Reference]	0.8 (0.4–1.3)	0.5 (0.3–0.8)[c]	0.8 (0.5–1.1)
Crackers	1 [Reference]	1.9 (0.6–6.1)	1.7 (0.6–4.7)	2.3 (0.9–5.9)[d]
Granola bars	1 [Reference]	3.8 (0.9–17.0)[d]	3.5 (0.9–13.6)[d]	3.4 (0.9–12.8)[d]
Cakes	1 [Reference]	1.0 (0.5–2.2)	0.6 (0.3–1.3)	0.8 (0.5–1.5)
Cookies	1 [Reference]	1.3 (0.5–3.1)	0.9 (0.4–1.9)	1.2 (0.6–2.4)
Chocolate	1 [Reference]	1.6 (0.9-3.1)	0.7 (0.4–1.3)	0.6 (0.3–1.0)
Hard candy	1 [Reference]	1.1 (0.5–2.6)	0.7 (0.3–1.5)	0.7 (0.3–1.3)
Fries	1 [Reference]	1.1 (0.5–2.1)	0.7 (0.4–1.3)	0.8 (0.5–1.4)
Pizza	1 [Reference]	1.5 (0.5–4.1)	0.5 (0.2–1.4)	1.2 (0.5–2.5)
Cereal	1 [Reference]	1.2 (0.7–2.4)	2.3 (1.4–3.7)[e]	1.7 (1.1–2.8)[c]
Fruit	1 [Reference]	3.1 (1.5–6.5)[e]	2.9 (1.6–5.6)[e]	2.9 (1.6–5.4)[b]
Vegetables	1 [Reference]	1.3 (0.5–3.3)	3.1 (1.5–6.2)[e]	3.2 (1.7–6.2)[b]
Beverage				
Water	1 [Reference]	1.3 (0.8–2.1)	2.6 (1.7–4.0)[e]	1.8 (1.2–2.6)[e]
Regular soda	1 [Reference]	1.4 (0.8–2.3)	0.5 (0.3–0.7)[e]	0.7 (0.5–1.0)[c]
100% Fruit juice	1 [Reference]	1.5 (0.8–2.6)	1.9 (1.2–2.9)[e]	2.3 (1.5–3.6)[e]
Fruit punch	1 [Reference]	1.4 (0.7–2.7)	1.1 (0.6–1.9)	1.5 (0.9–2.6)
Sports drinks	1 [Reference]	2.1 (0.8–5.5)	1.1 (0.4–2.6)	2.0 (1.0–4.4)[d]
Whole or 2% milk	1 [Reference]	1.2 (0.8–2.2)	1.5 (1.0–2.2)[e]	1.6 (1.1–2.3)[c]
Sweet tea	1 [Reference]	0.6 (0.3–1.1)	0.7 (0.4–1.2)	0.8 (0.5–1.3)

Abbreviations: GEE, generalized estimating equations; OR, odds ratio; CI, confidence interval.

[a] Adjusted for race, age, and length of time since giving birth.

[b] See the Methods section for a description of how positive and negative perceptions were determined.

[c] $P < .01$

[d] $P < .1$, significant for trend.

[e] $P < .05$.

Chapter 5

Conclusions

The previous chapters suggest a number of conclusions regarding opportunities for cancer prevention in early-life. For example, etiologic evidence linking gestational exposure to maternal obesity with altered DNA methylation patterns in offspring genes may provide additional insights into the *in utero* period as a critical window for adult cancer risk. Likewise, a better understanding of how to prevent exposure to early-life modifiable risk factors, such as HPV infection, and unhealthy diet, have the potential to lead to more effective primary cancer prevention strategies.

SUMMARY OF FINDINGS

In utero exposures, such as maternal adiposity, are recognized as having an important, yet not fully understood influence on fetal growth and later risk of obesity and related diseases, such as cancer. Epigenetic mechanisms, such as DNA methylation, are hypothesized to play an important role in mediating this risk[280]. Although some evidence is emerging, epidemiologic studies are relatively limited due to the inherently complex nature of studying the role of DNA methylation in mediating fetal programming of adult obesity and cancer risk. As an essential first step, it is necessary to establish whether DNA methylation patterns in offspring genes vary with respect to maternal obesity and excess gestational weight gain[280]. Our findings from Chapter 2 provide evidence supporting this critical first step, suggesting *in utero* exposure to maternal overweight/obesity and excess gestational weight gain may influence DNA methylation levels within offspring genes, some of which play a role in cell signaling and cell division processes. Further, our analyses support the hypothesis that associations of *in utero*

exposures with DNA methylation levels may be different with respect to infant sex[89,281]. To our knowledge, this is the first study to evaluate the association of maternal overweight/obesity and gestational weight gain with offspring cord blood DNA methylation patterns in candidate genes using microfluidic PCR and next-generation, bisulfite sequencing technology.

While promising, our findings should be interpreted with caution until replicated in large, independent studies. Similar to other studies, the magnitude of the difference in methylation levels at several CpG sites observed in our study was small (≤5%), thus the biological relevance of these differences remains unclear. Prospective cohort studies assessing whether differences in DNA methylation patterns persist over time and whether these patterns are associated with subsequent offspring BMI, will be important for determining the long-term functional significance of these associations. Future studies investigating the influence of maternal obesity and weight gain on offspring DNA methylation patterns will also need to address whether patterns observed in peripheral blood leukocytes reflect those occurring in disease-relevant tissue, with different cellular compositions[280].

The identification of a persistent biomarker of early-life obesity exposure would be useful for epidemiologic studies of early-life exposures and cancer risk, and could potentially provide more meaningful endpoints for interventions designed to reduce the risk of offspring obesity and metabolic disease later in life. Furthermore, early-life epigenetic modifications that occur in genes related to obesity and cancer may elucidate previously unknown pathways involved in carcinogenesis.

The critical window for HPV vaccination begins in childhood, before the onset of sexual activity[4]. Our study of predictors of HPV vaccination in males suggests that a substantial proportion of age-eligible males attending primary care clinics did not receive the HPV vaccine during visits with their healthcare provider. We were able to corroborate previous findings of higher rates of initiation, but lower rates of completion among non-Hispanic black and publicly insured/underinsured males compared with their non-Hispanic white and privately insured counterparts, respectively[150,152]. Further, we provided novel evidence demonstrating that these disparities are as equally important for receipt of the second dose of the HPV vaccine. We also observed that attending >3 clinic visits was positively associated with receipt of each HPV vaccine dose among all males and that frequent visits mitigated the inverse association between public insurance and receipt of subsequent doses. In regards to provider specialty, attendance to Internal Medicine clinics (vs. Family Practice) was inversely associated with HPV vaccine initiation among older males (i.e., in catch-up and permissive age groups).

Our findings present a number of opportunities for future investigation. Studies designed to better understand barriers to receiving subsequent doses of the HPV vaccine will help to inform interventions among minority and publicly insured males to ensure HPV vaccine series completion. Further, our study underscores the need for future research on provider-level factors associated with HPV vaccination in males, and suggests that interventions may need to be targeted by provider specialty. Indeed, a recent review summarizing the evidence on interventions designed to increase HPV

vaccine uptake suggest that those conducted at the community- and practice-level may be more effective than individual-level educational interventions; although only two of the studies reviewed included males, and few focused on adolescents in the target age range of 11 to 12 years[282]. Based on our current findings and those published in other studies, interventions that incorporate different strategies for increasing HPV vaccine initiation versus increasing rates of subsequent doses among males may be warranted.

Unhealthy dietary behaviors such as decreased fruit and vegetable intake and higher consumption of energy-dense foods and sugar-sweetened beverages are often established during adolescence, a critical period for the development of overweight and obesity[283,284]. This risk is especially heightened for certain subgroups such as postpartum adolescents, who tend to gain excessive weight during pregnancy and retain more weight postpartum compared with their adult counterparts[192–196]. Research addressing environmental influence on dietary patterns among this difficult to reach and high-risk population will be important for informing intervention strategies and preventing long-term obesity and subsequent cancer risk. Numerous studies have investigated the impact of policy and behavioral interventions promoting healthy school environments on dietary behaviors in adolescents. Schools are a unique setting to promote healthy lifestyles and emerging evidence suggests that recent policies are having an impact on the nutritional content of school meals and have reduced disparities in healthy food access[187]. In contrast to school-based interventions and policies, interventions conducted in the home setting are relatively uncommon and understudied. Results from our study are consistent

with the literature suggesting the home environment has an important relationship with dietary intake among adolescents, particularly in reducing unhealthy dietary behaviors[183,257,261,278]. The home environment is an important setting for families, particularly for teen mothers who are in a position to model healthy dietary behaviors for their infants[285].

COMMON THEMES AND IMPLICATIONS FOR CANCER PREVENTION

Although each chapter represents an independent study of a discrete cancer risk factor, it is worth noting a few overarching themes that may be applicable to various aspects of cancer prevention in early-life.

Pregnancy is a Critical Period for Obesity Risk

The high prevalence of overweight and obesity among reproductive-aged women underscores the importance of pregnancy as a critical period for obesity risk for both the mother and infant. Diet and other lifestyle factors during pregnancy can have profound effects on an infant's weight at birth that persist into childhood and adulthood. Pregnancy may potentially be an optimal time for intervention, as women may be more receptive to making lifestyle changes as they contemplate pregnancy, and when they are pregnant, to increase the likelihood of having a healthy baby[286]. During the postpartum period, many women may be willing to make substantial changes to protect the health of their infant. Pregnancy also involves more frequent contact with the healthcare system, providing increased opportunities for patient-provider interaction[286].

Preconception care interventions that provide clinical guidance, screening, and interventions for women of reproductive age, represent an opportunity to reduce the risk of adiposity and weight gain in both the mother and infant[286]. Indeed, the American Congress of Obstetrics and Gynecologists (ACOG) emphasizes the importance of preconception care and strongly recommends antenatal obesity screening and the provision of specific information concerning maternal and fetal risk factors associated with obesity in pregnancy[287,288]. As others and we have demonstrated, maternal pre-pregnancy overweight and obesity can influence cord blood DNA methylation patterns in genes related to cell growth and differentiation, which may have potential long-term consequences for gene expression and disease susceptibility in offspring. Our findings highlight the importance of preconception care in addressing obesity-related risk factors *before* a woman becomes pregnant.

To reduce the risk of postpartum weight retention, ACOG also recommends nutrition and physical activity counseling to all overweight or obese women during pregnancy, through the postpartum period and before attempting another pregnancy[287]. We have shown that the availability and accessibility of foods in the home environment are associated with dietary patterns in postpartum teens. As maternal dietary preferences and behaviors tend to shape the preferences and consumption patterns among offspring, dietary counseling that continues through the postpartum period is critical for mitigating the risk of postpartum weight retention and offspring risk of unhealthy diet and weight gain[285].

Interventions targeting the entire spectrum of pregnancy have the potential to curb the trajectory of weight gain throughout the life course and halt the perpetual cycle of obesity risk. Increased awareness and recognition from cancer institutes and organizations may help to stimulate research and funding opportunities for cancer prevention efforts during this critical period of increased risk and opportune time for intervention.

Cancer Prevention for Children and Adolescents Should be Integrated With Routine Primary Care

Primary care providers should take full advantage of routine clinic visits to provide cancer prevention services to their pediatric and adolescent patients. As our study on determinants of HPV vaccine uptake helps to illustrate, routine encounters with a healthcare provider can result in cancer risk reduction behaviors, such as receipt of age-appropriate immunizations as well as early identification and reduction of modifiable cancer risk factors and screening for early-onset disease. The United States Preventive Services Task Force (USPSTF) issues guidelines for preventive service recommendations for children and adolescents; however, some important topics are not included in the recommendations due to the strict criteria employed by the USPSTF when assessing the evidence. For example, while the USPSTF recommends screening for obesity among children beginning at age 6, they do not find sufficient evidence to support primary care counseling to promote physical activity or nutrition among children and adolescents[289]. In this case, physicians should use clinical judgment and consider alternative sources for evidence-based recommendations for preventive services. One such source, The *Bright*

Futures: Guidelines for Health Supervision for Infants, Children and Adolescents is the primary resource for pediatricians in delivering preventive services[290]. Led by the American Academy of Pediatrics (AAP), *Bright Futures* issues guidelines "to improve the health and well-being of all children through culturally appropriate interventions that address their current and emerging health promotion needs at the family, clinical practice, community, health system, and policy levels". *Bright Futures* covers the spectrum of cancer risk factors, emphasizing the promotion of healthy weight as a significant and growing challenge among children and adolescents, and just recently added an "HPV Champion Toolkit" which provides educational resources to parents and healthcare professionals[291]. *Bright Futures* is a leading example of a comprehensive approach that integrates a wide range of preventive and health promotion services that could have real impact in reducing the burden of cancer risk factors in early-life. Future studies should address how to best engage children and adolescents in primary care settings to achieve maximum benefit from these services. Moreover, research is needed to assess how to best train and incentivize primary care providers and to evaluate clinical support tools and technology that will ensure that cancer prevention services are delivered to pediatric and adolescent patients.

Consideration of Vulnerable Populations

This dissertation focused on vulnerable populations that are typically understudied with respect to cancer research. Federal regulations require additional human subjects' protections for pregnant women, human fetuses, and neonates[292], and these groups are often underrepresented in epidemiologic studies of obesity and cancer

risk[293]. Developing fetuses are extremely vulnerable to maternal exposures that may affect organ development and programming of disease susceptibility. As they have no control over their environments, developing fetuses are completely dependent on the mother for nutrients that are essential for normal growth and development. This vulnerability may manifest as long-term obesity and cancer risk in postnatal life.

Young males represent a population that has been significantly understudied in regard to HPV vaccination, particularly after the ACIP revised their recommendation in October 2011. The rising incidence of HPV-associated anal and oropharyngeal cancers among men underscores the need to vaccinate males before they become sexually active[144]. Relying on female vaccination alone will not provide adequate protection in the form of herd immunity, particularly among MSM who do not stand to benefit from female vaccination[4]. In recognition of the importance of HPV vaccination in males, an objective was added to *Healthy People 2020* in 2014 to increase coverage with all three doses for males by ages 13 to 15 years[151].

Nearly 300,000 adolescents become pregnant each year in the U.S., representing a significant, yet understudied population at high risk for obesity and related diseases, such as cancer, later in life. Teenage mothers have unique needs and competing demands that limit their ability to participate in research studies. Future studies should explore options for home- and school-based interventions to ensure participation[200]. Furthermore, measurement tools that are appropriate for use in postpartum adolescents should be adapted and/or designed to ensure relevant content is being assessed[200]. High rates of both childhood obesity and teenage pregnancy in the U.S. emphasize a critical need for

public health interventions targeting postpartum teens. Addressing the physiological, sociodemographic, and environmental influences on dietary behaviors in this high-risk and understudied population is critical to interrupting the cycle of obesity risk.

More Research on Early-Life Exposures and Adult Cancer Risk is Needed

The purpose of this dissertation was to emphasize early-life as an important, yet understudied period with respect to cancer research. Both the Centers for Disease Control and Prevention and the National Cancer Institute (NCI) have accentuated the need for more research in this area. Recently, the NCI released a Funding Opportunity Announcement (PA-15-126) to simulate research with the goal of better understanding 1) early-life factors that are associated with later cancer development; 2) how early-life factors mediate biological processes relevant to carcinogenesis; and 3) whether predictive markers associated with early-life exposures can be measured and developed for use in cancer prevention strategies. Identifying available resources that can be utilized for research questions addressing early-life and cancer risk should be a priority. The pressing need for more research will require innovative use of existing data from cohorts with relevant intermediate markers or risk factors as well as prospective studies that are designed to capture cancer risk factors in early-life. In the U.S., state-based birth and cancer registries may also be combined to create new opportunities to investigate early-life exposures related to cancer risk[11]. Biomarker studies of early-life exposures and/or intermediate cancer endpoints are also needed.

Appendix

DNA METHYLATION QUALITY CONTROL ANALYSES

Quality Control Samples: Technical Replicates

A total of 11 technical replicate samples were included for quality control (QC) purposes. The replicates were amplified on separate Fluidigm Access Array chips (1, 2 or 3) and sequenced on a single chip. To our knowledge, there is currently no standard approach for quality control analyses of data generated from microfluidic PCR-based target enrichment and next-generation bisulfite sequencing technology[1220]. Traditional quality control estimates such as the coefficient of variation (CV) were calculated; however, because the distribution of methylation values was highly right-skewed (i.e., methylation values were close to 0), the CVs were inflated and very sensitive to small changes in the mean. As an alternative approach, we calculated the concordance correlation coefficient, which combines measures of both precision and accuracy to determine how far the observed data deviate from perfect concordance (i.e., a 45^{o} line)[294]. Concordance correlation coefficient estimates for each pair of replicates were modestly high, with an average of approximately 0.70 (range 0.42 – 0.82). To assess whether concordance varied by specific genes, we calculated the absolute value of the difference in methylation proportions between each pair of replicates for each CpG site and took the average across all replicates to generate a mean difference (CpG site-level mean difference) for each CpG site (Figure A1). We then calculated the mean and standard deviation (SD) of the mean difference across CpG sites, by gene (gene-level mean difference and SD). Any CpG-site level mean difference that was greater than + 2 SD's was flagged as a potential outlier. Out of the 526 total CpG sites, 28 (5.3%) were

identified as outliers using this approach (Table A1). From this analysis, the degree of concordance did not appear to vary by gene.

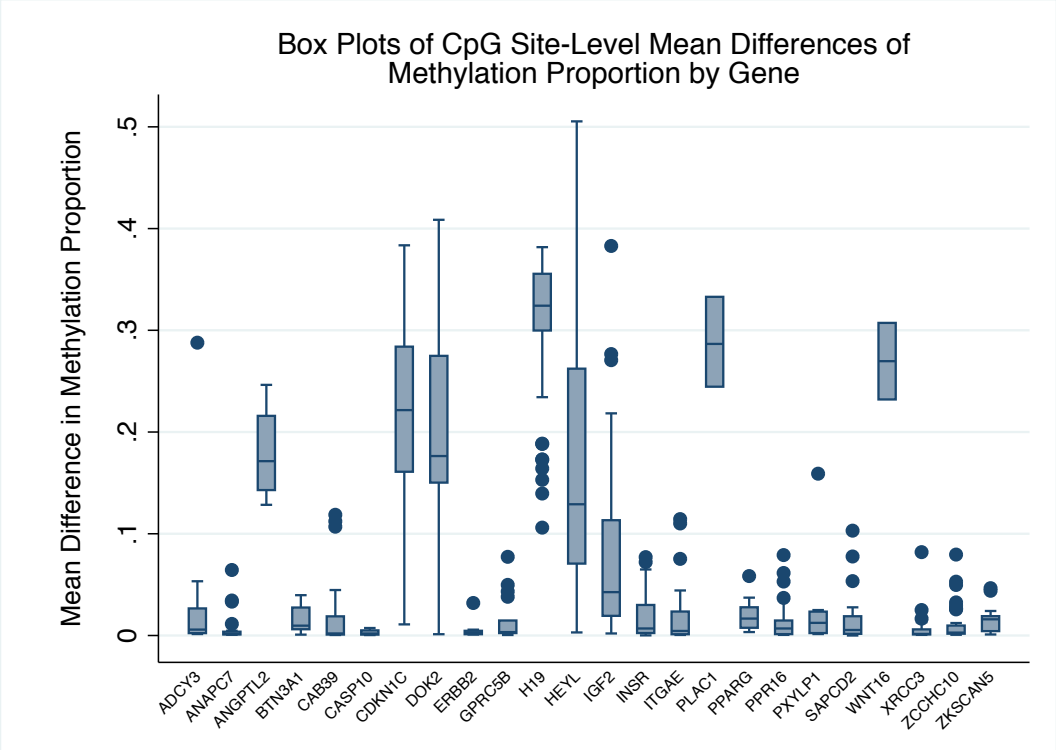

Figure A1. Box Plots of CpG Site-Level Mean Differences of Methylation Proportion by Gene. The absolute value of the mean difference was calculated between each pair of replicates for each CpG. This value was then averaged across all replicates to generate a mean difference (CpG site-level mean difference) for each CpG site.

Table A1. CpG Sites with Absolute Value Mean Differences that Exceed Gene-Level Mean Difference + 2SD

Gene Symbol	CpG Site Position[a]	CpG-Site Level Mean Difference[b]	Gene-Level Mean Difference (SD)[c]
ADCY3	24995895	0.2878	0.0449 (0.093)
ANAPC7	109325831	0.0643	0.0084 (0.017)
CAB39	231285851	0.1186	0.0175 (0.032)
	231285863	0.1122	0.0175 (0.032)
ERBB2	35098006	0.0319	0.0062 (0.011)
GPRC5B	19804118	0.0773	0.0149 (0.022)
HEYL	39877608	0.5054	0.1715 (0.133)
	39877610	0.4691	0.1715 (0.133)
	39877725	0.5045	0.1715 (0.133)
IGF2	2110623	0.2766	0.0769 (0.088)
	2110642	0.3829	0.0769 (0.088)
	2110657	0.2706	0.0769 (0.088)
INSR	7244450	0.0770	0.0196 (0.023)
	7244543	0.0721	0.0196 (0.023)
ITGAE	3573467	0.1100	0.0220 (0.035)
PPARG	12305326	0.0584	0.0198 (0.015)
PRR16	119827784	0.0614	0.0141 (0.020)
	119827801	0.0790	0.0141 (0.020)
PXYLP1	142432606	0.1590	0.0296 (0.053)
SAPCD2	139084682	0.0776	0.0152 (0.024)
	139084741	0.1030	0.0152 (0.024)
XRCC3	103251158	0.0819	0.0064 (0.014)
ZCCHC10	132390109	0.0529	0.0098 (0.016)
	132390152	0.0795	0.0098 (0.016)
	132390195	0.0496	0.0098 (0.016)
ZKSCAN5	98940131	0.0465	0.0138 (0.012)
	98940249	0.0439	0.0138 (0.012)

[a]Based on NCBI Build 36; [b]CpG site-level mean difference corresponds to the average mean difference across all replicates for a given CpG site; [c]Gene-level mean difference corresponds to the average of the CpG site-level mean differences within a given gene. Abbreviations: SD, Standard Deviation

Quality Control Samples: Laboratory Controls

Two types of laboratory control samples were also included for QC purposes, including water blanks containing no DNA (i.e., negative controls) and white blood cell DNA as a standard. Each sample type was included on all three Fluidigm Access Array chips. For the water blanks, we assessed mean coverage (i.e., the number of sequencing reads) levels for each gene and considered coverage of 50 reads or greater as contamination. The range of coverage averaged from 0 to 39.5 reads, with a majority of genes having a mean coverage of less than 10 reads across all three chips (data not shown).

Assessment of Chip ("Batch") Effects

Several visual tools were used to assess whether there were systematic differences in methylation proportions according to Fluidigm Access Array chip. First, a density plot was created to visualize the shape of the distribution of methylation values for each chip. In general, high overlap in the distribution of the curves did not suggest of batch effects by chip (Figure A2).

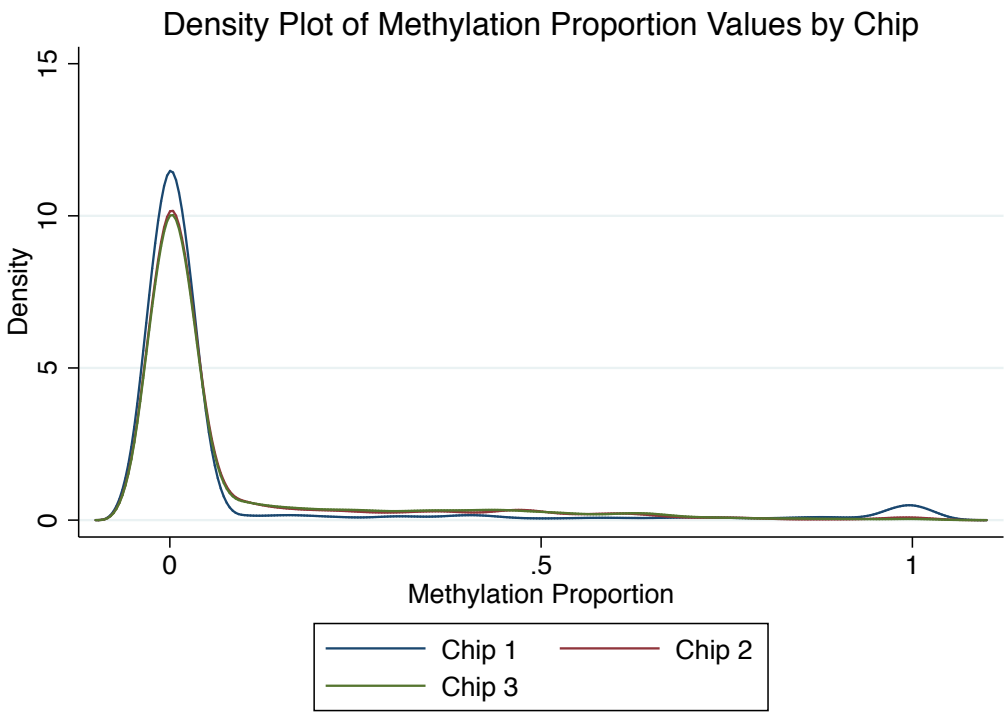

Figure A2. A density plot of the distribution of methylation values by chip. The distribution of methylation values for each Fluidigm Access Array chip is plotted using kernel density estimation. The overlap in the distribution curves is not suggestive of batch effects by chip.

Principle Component Analysis (PCA) was also employed to assess batch effects by chip. PCA aims to represent a large number of correlated measures (i.e., CpG sites) by a smaller number of uncorrelated variables to capture the majority of the variance in the data. The first two principle components were graphed using a scatterplot, with each sample colored according to chip. Separation of colors into distinct clusters would provide evidence for a batch effect, although this did not appear to be the case for our data (Figure A3).

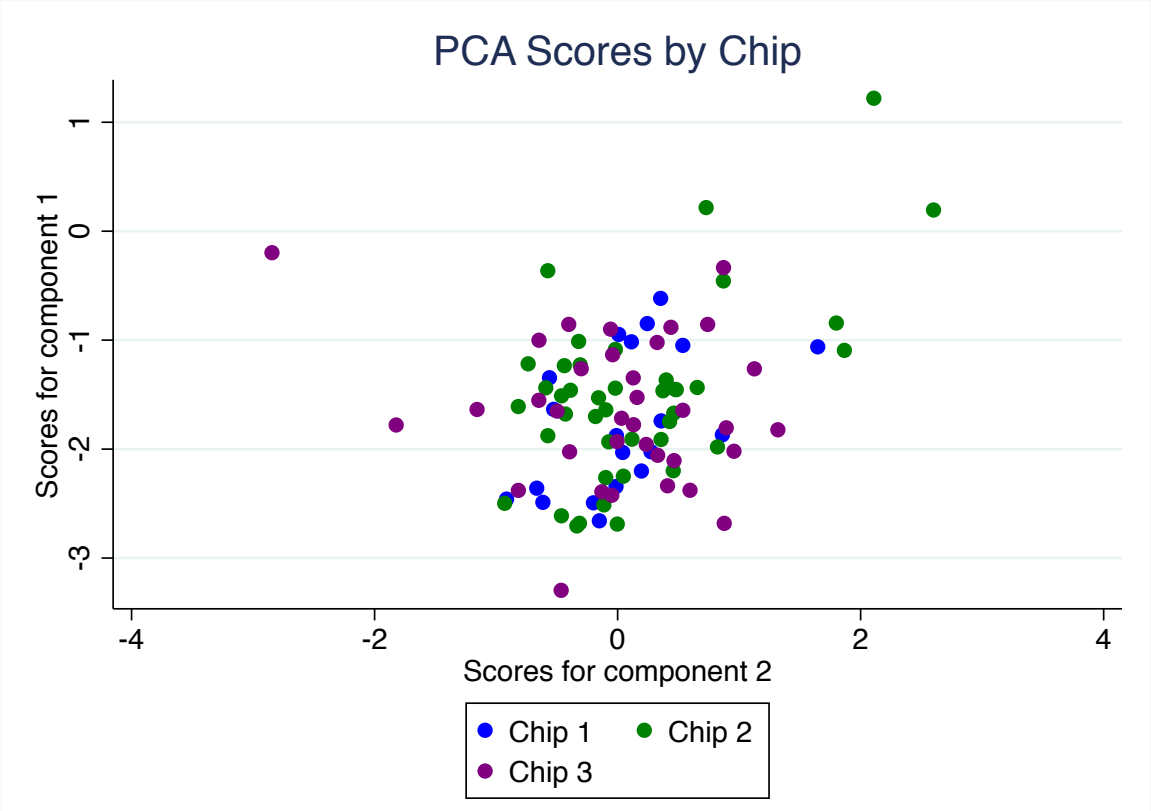

Figure A3. Principle Components Analysis by Chip. Principle components analysis was performed using the methylation values from all samples. Principle component score 1 is plotted against principle component score 2. Component scores are color-coded according to Fluidigm Access Array chip. This plot does not suggest systematic bias with respect to chip, although some outliers are observed. Abbreviations: PCA, Principle Components Analysis

Finally, we assessed sequencing performance according to chip by plotting the mean coverage values for each gene by chip in a histogram. As shown in Figure A4, coverage appeared to be consistent across chips, but did vary with respect to gene, ranging from 341.4 (ADCY3) to 38,188.8 reads (ZKSCAN5). In general, coverage was very high for all genes included in this study.

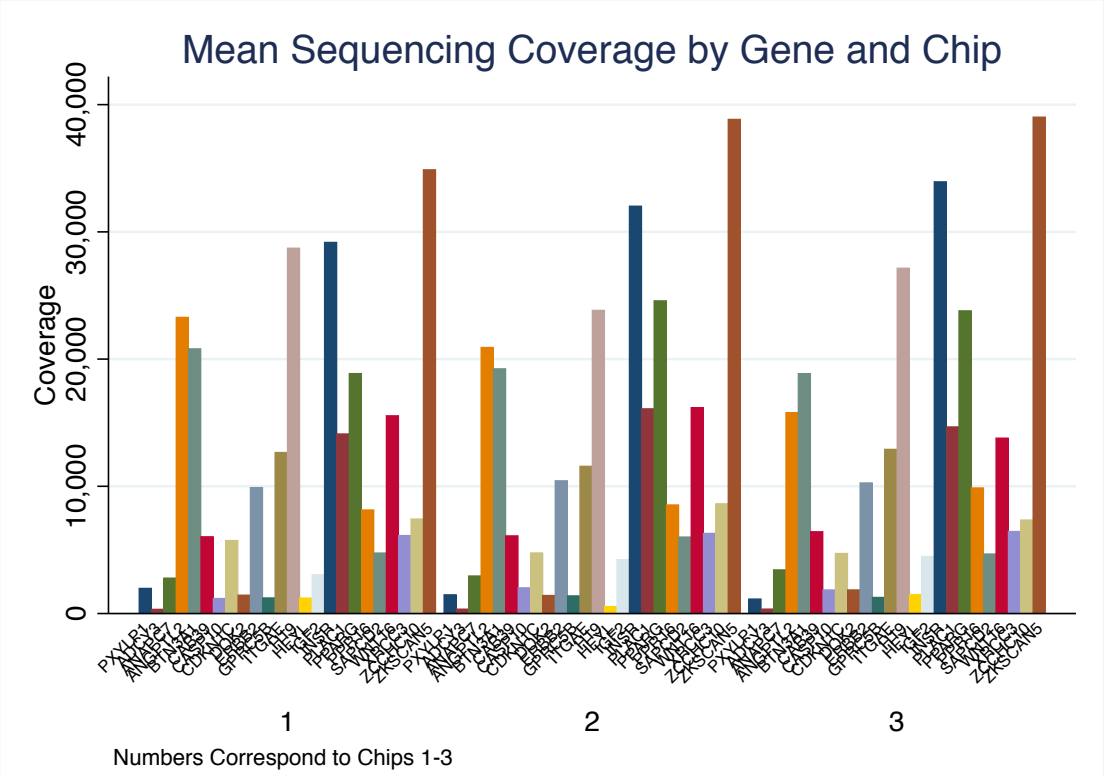

Figure A4. Plot of Coverage by Gene and Chip in all Samples. Sequencing coverage (i.e., the number of sequencing reads per amplicon) is plotted for each gene. Overall, coverage is high (>1,000 reads) and does not appear to be systematically biased according to chip; however coverage does appear to consistently vary with respect to gene across all three chips.

VALIDITY OF SELF-REPORTED PRE-PREGNANCY BMI IN THE EHUB STUDY

To assess the validity of self-reported pre-pregnancy BMI, we plotted self-reported pre-pregnancy BMI against BMI obtained at the first prenatal visit. In general, there was a high degree of correlation between pre-pregnancy BMI and BMI at the first prenatal visit (Figure A5, Pearson correlation coefficient = 0.93), suggesting high accuracy of self-

reported pre-pregnancy BMI. For our analyses, we chose to use pre-pregnancy BMI rather than first prenatal visit BMI to be in line with other studies, and because it most accurately reflects a woman's BMI prior to becoming pregnant.

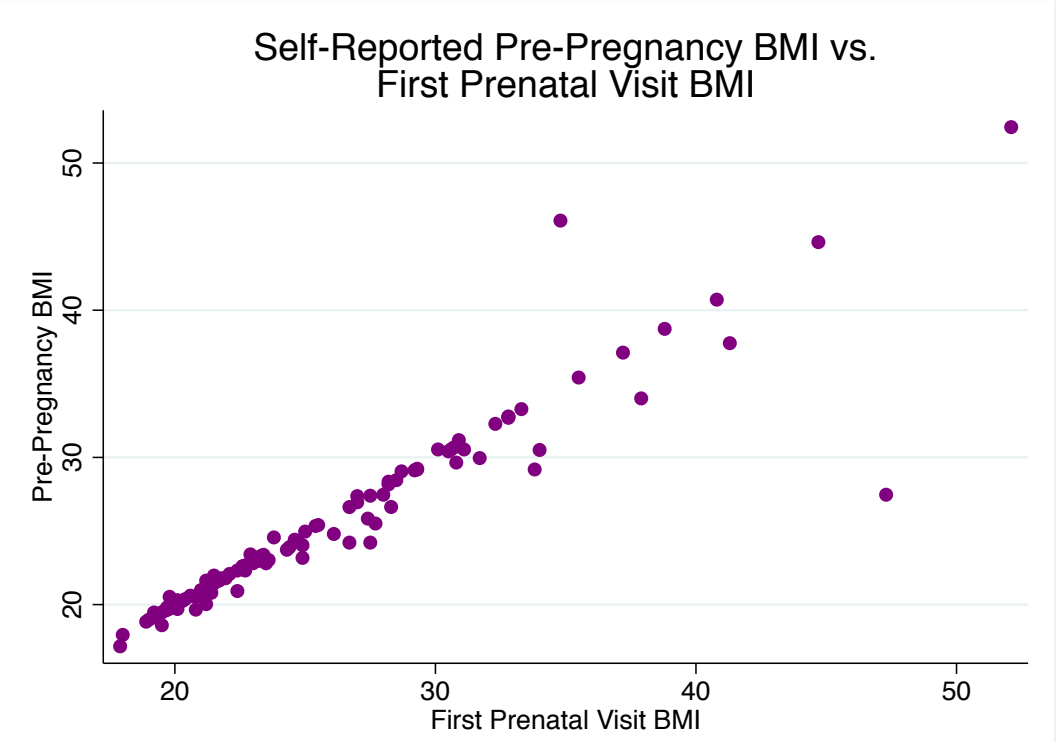

Figure A5. Self-Reported Pre-Pregnancy BMI vs. First Prenatal Visit BMI. A two-way scatterplot of self-reported pre-pregnancy BMI and first prenatal visit BMI shows a high degree of linear correlation between the two values (Pearson correlation coefficient = 0.93), suggesting high accuracy of self-reported pre-pregnancy BMI. Abbreviations: BMI, body mass index

References

1. American Cancer Society. *Cancer Prevention & Early Detection Facts & Figures 2015-2016.* (American Cancer Society, 2015).
2. Colditz, G. A. Carpe Diem: time to seize the opportunity for cancer prevention. *Am. Soc. Clin. Oncol. Educ. Book ASCO Am. Soc. Clin. Oncol. Meet.* 8–12 (2014). doi:10.14694/EdBook_AM.2014.34.8
3. Colditz, G. A. & Hunter, D. J. *Cancer Prevention: The Causes and Prevention of Cancer —*. (Springer Science & Business Media, 2006).
4. Markowitz, L. E. *et al.* Human papillomavirus vaccination: recommendations of the Advisory Committee on Immunization Practices (ACIP). *MMWR Recomm. Rep. Morb. Mortal. Wkly. Rep. Recomm. Rep. Cent. Dis. Control* **63,** 1–30 (2014).
5. Colditz, G. A., Wolin, K. Y. & Gehlert, S. Applying What We Know to Accelerate Cancer Prevention. *Sci. Transl. Med.* **4,** 127rv4 (2012).
6. Byers, T. *et al.* The American Cancer Society challenge goals. How far can cancer rates decline in the U.S. by the year 2015? *Cancer* **86,** 715–727 (1999).
7. Sedjo, R. L. *et al.* A midpoint assessment of the American Cancer Society challenge goal to decrease cancer incidence by 25% between 1992 and 2015. *CA. Cancer J. Clin.* **57,** 326–340 (2007).
8. Kohler, B. A. *et al.* Annual Report to the Nation on the Status of Cancer, 1975-2011, Featuring Incidence of Breast Cancer Subtypes by Race/Ethnicity, Poverty, and State. *J. Natl. Cancer Inst.* **107,** djv048 (2015).

9. Simard, E. P., Ward, E. M., Siegel, R. & Jemal, A. Cancers with increasing incidence trends in the United States: 1999 through 2008. *CA. Cancer J. Clin.* **62,** 118–128 (2012).

10. Edwards, B. K. *et al.* Annual Report to the Nation on the status of cancer, 1973–1999, featuring implications of age and aging on U.S. cancer burden. *Cancer* **94,** 2766–2792 (2002).

11. Mahabir, S. *et al.* Challenges and opportunities in research on early-life events/exposures and cancer development later in life. *Cancer Causes Control CCC* **23,** 983–990 (2012).

12. Ben-Shlomo, Y. & Kuh, D. A life course approach to chronic disease epidemiology: conceptual models, empirical challenges and interdisciplinary perspectives. *Int. J. Epidemiol.* **31,** 285–293 (2002).

13. Moolgavkar, S. H. The multistage theory of carcinogenesis and the age distribution of cancer in man. *J. Natl. Cancer Inst.* **61,** 49–52 (1978).

14. Wei, E. K., Wolin, K. Y. & Colditz, G. A. Time Course of Risk Factors in Cancer Etiology and Progression. *J. Clin. Oncol.* **28,** 4052–4057 (2010).

15. Calle, E. E., Rodriguez, C., Walker-Thurmond, K. & Thun, M. J. Overweight, obesity, and mortality from cancer in a prospectively studied cohort of U.S. adults. *N. Engl. J. Med.* **348,** 1625–1638 (2003).

16. Wolin, K. Y., Carson, K. & Colditz, G. A. Obesity and Cancer. *The Oncologist* **15,** 556–565 (2010).

17. American Institute for Cancer Research/World Cancer Research Fund. *Food, Nutrition, Physical Activity, and the Prevention of Cancer: A Global Perspective*. (AICR, 2007).

18. Freedland, S. J. & Platz, E. A. Obesity and prostate cancer: making sense out of apparently conflicting data. *Epidemiol. Rev.* **29,** 88–97 (2007).

19. Marengo, A., Rosso, C. & Bugianesi, E. Liver Cancer: Connections with Obesity, Fatty Liver, and Cirrhosis. *Annu. Rev. Med.* **67,** null (2016).

20. Flegal KM, Carroll MD, Ogden CL & Johnson CL. PRevalence and trends in obesity among us adults, 1999-2000. *JAMA* **288,** 1723–1727 (2002).

21. Ogden, C. L., Carroll, M. D. & Flegal, K. M. Prevalence of obesity in the United States. *JAMA* **312,** 189–190 (2014).

22. Calle, E. E. & Kaaks, R. Overweight, obesity and cancer: epidemiological evidence and proposed mechanisms. *Nat. Rev. Cancer* **4,** 579–591 (2004).

23. Renehan, A. G., Zwahlen, M. & Egger, M. Adiposity and cancer risk: new mechanistic insights from epidemiology. *Nat. Rev. Cancer* **15,** 484–498 (2015).

24. Roberts, D. L., Dive, C. & Renehan, A. G. Biological mechanisms linking obesity and cancer risk: new perspectives. *Annu. Rev. Med.* **61,** 301–316 (2010).

25. Vansaun, M. N. Molecular pathways: adiponectin and leptin signaling in cancer. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* **19,** 1926–1932 (2013).

26. Dalamaga, M., Diakopoulos, K. N. & Mantzoros, C. S. The role of adiponectin in cancer: a review of current evidence. *Endocr. Rev.* **33,** 547–594 (2012).

27. Pollak, M. The insulin and insulin-like growth factor receptor family in neoplasia: an update. *Nat. Rev. Cancer* **12,** 159–169 (2012).

28. Simpson, E. R. *et al.* Aromatase Cytochrome P450, The Enzyme Responsible for Estrogen Biosynthesis. *Endocr. Rev.* **15,** 342–355 (1994).

29. Kamat, A., Hinshelwood, M. M., Murry, B. A. & Mendelson, C. R. Mechanisms in tissue-specific regulation of estrogen biosynthesis in humans. *Trends Endocrinol. Metab.* **13,** 122–128 (2002).

30. Dickson, R. B. & Stancel, G. M. Estrogen receptor-mediated processes in normal and cancer cells. *J. Natl. Cancer Inst. Monogr.* 135–145 (2000).

31. Key, T. J. & Pike, M. C. The dose-effect relationship between 'unopposed' oestrogens and endometrial mitotic rate: its central role in explaining and predicting endometrial cancer risk. *Br. J. Cancer* **57,** 205–212 (1988).

32. Kaaks, R., Lukanova, A. & Kurzer, M. S. Obesity, endogenous hormones, and endometrial cancer risk: a synthetic review. *Cancer Epidemiol. Biomark. Prev. Publ. Am. Assoc. Cancer Res. Cosponsored Am. Soc. Prev. Oncol.* **11,** 1531–1543 (2002).

33. Ogden, C. L., Carroll, M. D., Kit, B. K. & Flegal, K. M. Prevalence of childhood and adult obesity in the United States, 2011-2012. *JAMA* **311,** 806–814 (2014).

34. Colditz, G. A., Bohlke, K. & Berkey, C. S. Breast cancer risk accumulation starts early: prevention must also. *Breast Cancer Res. Treat.* **145,** 567–579 (2014).

35. Grodstein, F., Goldman, M. B. & Cramer, D. W. Body mass index and ovulatory infertility. *Epidemiol. Camb. Mass* **5,** 247–250 (1994).

36. Stolzenberg-Solomon, R. Z., Schairer, C., Moore, S., Hollenbeck, A. & Silverman, D. T. Lifetime adiposity and risk of pancreatic cancer in the NIH-AARP Diet and Health Study cohort. *Am. J. Clin. Nutr.* **98,** 1057–1065 (2013).

37. Hofmann, J. N. *et al.* Body Mass Index and Physical Activity at Different Ages and Risk of Multiple Myeloma in the NIH-AARP Diet and Health Study. *Am. J. Epidemiol.* **177,** 776–786 (2013).

38. Dougan, M. M. *et al.* Prospective study of body size throughout the life-course and the incidence of endometrial cancer among premenopausal and postmenopausal women. *Int. J. Cancer* **137,** 625–637 (2015).

39. Chute, C. G. *et al.* A prospective study of body mass, height, and smoking on the risk of colorectal cancer in women. *Cancer Causes Control CCC* **2,** 117–124 (1991).

40. Oxentenko, A. S. *et al.* Body Size and Incident Colorectal Cancer: A Prospective Study of Older Women. *Cancer Prev. Res. (Phila. Pa.)* **3,** 1608–1620 (2010).

41. Han, X. *et al.* Body mass index at early adulthood, subsequent weight change and cancer incidence and mortality. *Int. J. Cancer* **135,** 2900–2909 (2014).

42. Lee, I. M. & Paffenbarger, R. S. Quetelet's index and risk of colon cancer in college alumni. *J. Natl. Cancer Inst.* **84,** 1326–1331 (1992).

43. Le Marchand, L., Wilkens, L. R. & Mi, M. P. Obesity in youth and middle age and risk of colorectal cancer in men. *Cancer Causes Control CCC* **3,** 349–354 (1992).

44. Levi, Z. *et al.* Measured Body Mass Index in Adolescence and the Incidence of Colorectal Cancer in a Cohort of 1.1 Million Males. *Cancer Epidemiol. Biomarkers Prev.* **20,** 2524–2531 (2011).

45. Cook, M. B., Freedman, N. D., Gamborg, M., Sørensen, T. I. A. & Baker, J. L. Childhood body mass index in relation to future risk of oesophageal adenocarcinoma. *Br. J. Cancer* **112,** 601–607 (2015).

46. Berentzen, T. L., Gamborg, M., Holst, C., Sørensen, T. I. A. & Baker, J. L. Body mass index in childhood and adult risk of primary liver cancer. *J. Hepatol.* **60,** 325–330 (2014).

47. Kitahara, C. M., Gamborg, M., Berrington de González, A., Sørensen, T. I. A. & Baker, J. L. Childhood height and body mass index were associated with risk of adult thyroid cancer in a large cohort study. *Cancer Res.* **74,** 235–242 (2014).

48. Cook, M. B., Gamborg, M., Aarestrup, J., Sørensen, T. I. A. & Baker, J. L. Childhood height and birth weight in relation to future prostate cancer risk: a cohort study based on the Copenhagen School Health Records Register. *Cancer Epidemiol. Biomark. Prev. Publ. Am. Assoc. Cancer Res. Cosponsored Am. Soc. Prev. Oncol.* **22,** (2013).

49. Aarestrup, J., Gamborg, M., Cook, M. B., Sørensen, T. I. A. & Baker, J. L. Childhood body mass index and the risk of prostate cancer in adult men. *Br. J. Cancer* **111,** 207–212 (2014).

50. Barker, D. J. Fetal origins of coronary heart disease. *BMJ* **311,** 171–174 (1995).

51. Barker, D. J. P., Eriksson, J. G., Forsén, T. & Osmond, C. Fetal origins of adult disease: strength of effects and biological basis. *Int. J. Epidemiol.* **31,** 1235–1239 (2002).

52. Roseboom, T., de Rooij, S. & Painter, R. The Dutch famine and its long-term consequences for adult health. *Early Hum. Dev.* **82,** 485–491 (2006).

53. Painter, R. C. *et al.* A possible link between prenatal exposure to famine and breast cancer: a preliminary study. *Am. J. Hum. Biol. Off. J. Hum. Biol. Counc.* **18,** 853–856 (2006).

54. Dabelea, D., Knowler, W. C. & Pettitt, D. J. Effect of diabetes in pregnancy on offspring: follow-up research in the Pima Indians. *J. Matern. Fetal Med.* **9,** 83–88 (2000).

55. Nelson, R. G., Morgenstern, H. & Bennett, P. H. Intrauterine diabetes exposure and the risk of renal disease in diabetic Pima Indians. *Diabetes* **47,** 1489–1493 (1998).

56. Boney, C. M., Verma, A., Tucker, R. & Vohr, B. R. Metabolic syndrome in childhood: association with birth weight, maternal obesity, and gestational diabetes mellitus. *Pediatrics* **115,** e290–296 (2005).

57. Smith, J. *et al.* Effects of maternal surgical weight loss in mothers on intergenerational transmission of obesity. *J. Clin. Endocrinol. Metab.* **94,** 4275–4283 (2009).

58. Guénard, F. *et al.* Differential methylation in glucoregulatory genes of offspring born before vs. after maternal gastrointestinal bypass surgery. *Proc. Natl. Acad. Sci. U. S. A.* **110,** 11439–11444 (2013).

59. Poston, L., Harthoorn, L. F. & van der Beek, E. M. Obesity in Pregnancy: Implications for the Mother and Lifelong Health of the Child. A Consensus Statement. *Pediatr. Res.* **69,** 175–180 (2011).

60. Oken, E. Maternal and child obesity: the causal link. *Obstet. Gynecol. Clin. North Am.* **36,** 361–377, ix–x (2009).

61. Catalano, P. M. & Ehrenberg, H. M. The short- and long-term implications of maternal obesity on the mother and her offspring. *BJOG Int. J. Obstet. Gynaecol.* **113,** 1126–1133 (2006).

62. Whincup, P. H. *et al.* Birth weight and risk of type 2 diabetes: a systematic review. *JAMA* **300,** 2886–2897 (2008).

63. Schellong, K., Schulz, S., Harder, T. & Plagemann, A. Birth weight and long-term overweight risk: systematic review and a meta-analysis including 643,902 persons from 66 studies and 26 countries globally. *PloS One* **7,** e47776 (2012).

64. Cook, M. B. *et al.* A systematic review and meta-analysis of perinatal variables in relation to the risk of testicular cancer--experiences of the son. *Int. J. Epidemiol.* **39,** 1605–1618 (2010).

65. Stephansson, O. *et al.* Perinatal risk factors for childhood testicular germ-cell cancer: a Nordic population-based study. *Cancer Epidemiol.* **35,** e100–104 (2011).

66. Heerwagen, M. J. R., Miller, M. R., Barbour, L. A. & Friedman, J. E. Maternal obesity and fetal metabolic programming: a fertile epigenetic soil. *Am. J. Physiol. Regul. Integr. Comp. Physiol.* **299,** R711–722 (2010).

67. Butte, N. F., Ellis, K. J., Wong, W. W., Hopkinson, J. M. & Smith, E. O. Composition of gestational weight gain impacts maternal fat retention and infant birth weight. *Am. J. Obstet. Gynecol.* **189,** 1423–1432 (2003).

68. Haggarty, P. Effect of placental function on fatty acid requirements during pregnancy. *Eur. J. Clin. Nutr.* **58,** 1559–1570 (2004).

69. Haggarty, P. Placental regulation of fatty acid delivery and its effect on fetal growth--a review. *Placenta* **23 Suppl A,** S28–38 (2002).

70. Catalano, P. M., Roman-Drago, N. M., Amini, S. B. & Sims, E. A. Longitudinal changes in body composition and energy balance in lean women with normal and

abnormal glucose tolerance during pregnancy. *Am. J. Obstet. Gynecol.* **179,** 156–165 (1998).

71. Di Cianni, G. *et al.* Maternal triglyceride levels and newborn weight in pregnant women with normal glucose tolerance. *Diabet. Med. J. Br. Diabet. Assoc.* **22,** 21–25 (2005).
72. Son, G. H., Kwon, J. Y., Kim, Y. H. & Park, Y. W. Maternal serum triglycerides as predictive factors for large-for-gestational age newborns in women with gestational diabetes mellitus. *Acta Obstet. Gynecol. Scand.* **89,** 700–704 (2010).
73. Radaelli, T. *et al.* Differential regulation of genes for fetoplacental lipid pathways in pregnancy with gestational and type 1 diabetes mellitus. *Am. J. Obstet. Gynecol.* **201,** 209.e1–209.e10 (2009).
74. Jaenisch, R. & Bird, A. Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nat. Genet.* **33 Suppl,** 245–254 (2003).
75. Hanley, B. *et al.* Metabolic imprinting, programming and epigenetics - a review of present priorities and future opportunities. *Br. J. Nutr.* **104 Suppl 1,** S1–25 (2010).
76. Illingworth, R. S. & Bird, A. P. CpG islands – 'A rough guide'. *FEBS Lett.* **583,** 1713–1720 (2009).
77. Waterland, R. A. & Michels, K. B. Epigenetic Epidemiology of the Developmental Origins Hypothesis. *Annu. Rev. Nutr.* **27,** 363–388 (2007).
78. Saffery, R. & Novakovic, B. Epigenetics as the mediator of fetal programming of adult onset disease: what is the evidence? *Acta Obstet. Gynecol. Scand.* **93,** 1090–1098 (2014).

79. Vidal, A. C. *et al.* Maternal stress, preterm birth, and DNA methylation at imprint regulatory sequences in humans. *Genet. Epigenetics* **6,** 37–44 (2014).

80. Joubert, B. R. *et al.* 450K epigenome-wide scan identifies differential DNA methylation in newborns related to maternal smoking during pregnancy. *Environ. Health Perspect.* **120,** 1425–1431 (2012).

81. Breton, C. V. *et al.* Prenatal tobacco smoke exposure is associated with childhood DNA CpG methylation. *PloS One* **9,** e99716 (2014).

82. Novakovic, B. *et al.* Postnatal stability, tissue, and time specific effects of AHRR methylation change in response to maternal smoking in pregnancy. *Epigenetics* **9,** 377–386 (2014).

83. Cardenas, A. *et al.* Differential DNA methylation in umbilical cord blood of infants exposed to mercury and arsenic in utero. *Epigenetics* **10,** 508–515 (2015).

84. Koestler, D. C., Avissar-Whiting, M., Houseman, E. A., Karagas, M. R. & Marsit, C. J. Differential DNA methylation in umbilical cord blood of infants exposed to low levels of arsenic in utero. *Environ. Health Perspect.* **121,** 971–977 (2013).

85. Pilsner, J. R. *et al.* Influence of prenatal arsenic exposure and newborn sex on global methylation of cord blood DNA. *PloS One* **7,** e37147 (2012).

86. Waterland, R. A. & Jirtle, R. L. Early nutrition, epigenetic changes at transposons and imprinted genes, and enhanced susceptibility to adult chronic diseases. *Nutrition* **20,** 63–68 (2004).

87. Dominguez-Salas, P. *et al.* Maternal nutrition at conception modulates DNA methylation of human metastable epialleles. *Nat. Commun.* **5,** 3746 (2014).

88. Waterland, R. A. *et al.* Season of conception in rural gambia affects DNA methylation at putative human metastable epialleles. *PLoS Genet.* **6,** e1001252 (2010).

89. Tobi, E. W. *et al.* DNA methylation differences after exposure to prenatal famine are common and timing- and sex-specific. *Hum. Mol. Genet.* **18,** 4046–4053 (2009).

90. Tobi, E. W. *et al.* DNA methylation signatures link prenatal famine exposure to growth and metabolism. *Nat. Commun.* **5,** 5592 (2014).

91. Heijmans, B. T. *et al.* Persistent epigenetic differences associated with prenatal exposure to famine in humans. *Proc. Natl. Acad. Sci.* **105,** 17046–17049 (2008).

92. Rayco-Solon, P., Fulford, A. J. & Prentice, A. M. Differential effects of seasonality on preterm birth and intrauterine growth restriction in rural Africans. *Am. J. Clin. Nutr.* **81,** 134–139 (2005).

93. Relton, C. L. *et al.* DNA methylation patterns in cord blood DNA and body size in childhood. *PloS One* **7,** e31821 (2012).

94. Morales, E., Groom, A., Lawlor, D. A. & Relton, C. L. DNA methylation signatures in cord blood associated with maternal gestational weight gain: results from the ALSPAC cohort. *BMC Res. Notes* **7,** 278 (2014).

95. Sharp, G. C. *et al.* Maternal pre-pregnancy BMI and gestational weight gain, offspring DNA methylation and later offspring adiposity: findings from the Avon Longitudinal Study of Parents and Children. *Int. J. Epidemiol.* (2015). doi:10.1093/ije/dyv042

96. Liu, X. *et al.* Maternal preconception body mass index and offspring cord blood DNA methylation: exploration of early life origins of disease. *Environ. Mol. Mutagen.* **55,** 223–230 (2014).

97. Godfrey, K. M. *et al.* Epigenetic gene promoter methylation at birth is associated with child's later adiposity. *Diabetes* **60,** 1528–1534 (2011).

98. Chesson, H. W., Dunne, E. F., Hariri, S. & Markowitz, L. E. The estimated lifetime probability of acquiring human papillomavirus in the United States. *Sex. Transm. Dis.* **41,** 660–664 (2014).

99. Moscicki, A.-B. *et al.* Updating the natural history of human papillomavirus and anogenital cancers. *Vaccine* **30 Suppl 5,** F24–33 (2012).

100. Kjaer, S. K. *et al.* High-risk human papillomavirus is sexually transmitted: evidence from a follow-up study of virgins starting sexual activity (intercourse). *Cancer Epidemiol. Biomark. Prev. Publ. Am. Assoc. Cancer Res. Cosponsored Am. Soc. Prev. Oncol.* **10,** 101–106 (2001).

101. Shew, M. L. *et al.* High frequency of human papillomavirus detection in the vagina before first vaginal intercourse among females enrolled in a longitudinal cohort study. *J. Infect. Dis.* **207,** 1012–1015 (2013).

102. Schiffman, M., Castle, P. E., Jeronimo, J., Rodriguez, A. C. & Wacholder, S. Human papillomavirus and cervical cancer. *The Lancet* **370,** 890–907 (2007).

103. Li, N., Franceschi, S., Howell-Jones, R., Snijders, P. J. F. & Clifford, G. M. Human papillomavirus type distribution in 30,848 invasive cervical cancers worldwide: Variation by geographical region, histological type and year of publication. *Int. J. Cancer J. Int. Cancer* **128,** 927–935 (2011).

104. Hariri, S. *et al.* Prevalence of genital human papillomavirus among females in the United States, the National Health And Nutrition Examination Survey, 2003-2006. *J. Infect. Dis.* **204,** 566–573 (2011).

105. Castle, P. E., Solomon, D., Schiffman, M. & Wheeler, C. M. Human papillomavirus type 16 infections and 2-year absolute risk of cervical precancer in women with equivocal or mild cytologic abnormalities. *J. Natl. Cancer Inst.* **97,** 1066–1071 (2005).

106. Schlecht, N. F. *et al.* Persistent human papillomavirus infection as a predictor of cervical intraepithelial neoplasia. *JAMA* **286,** 3106–3114 (2001).

107. Trottier, H. *et al.* Persistence of an incident human papillomavirus infection and timing of cervical lesions in previously unexposed young women. *Cancer Epidemiol. Biomark. Prev. Publ. Am. Assoc. Cancer Res. Cosponsored Am. Soc. Prev. Oncol.* **18,** 854–862 (2009).

108. International Collaboration of Epidemiological Studies of Cervical Cancer. Cervical carcinoma and sexual behavior: collaborative reanalysis of individual data on 15,461 women with cervical carcinoma and 29,164 women without cervical carcinoma from 21 epidemiological studies. *Cancer Epidemiol. Biomark. Prev. Publ. Am. Assoc. Cancer Res. Cosponsored Am. Soc. Prev. Oncol.* **18,** 1060–1069 (2009).

109. International Collaboration of Epidemiological Studies of Cervical Cancer. Cervical carcinoma and reproductive factors: collaborative reanalysis of individual data on 16,563 women with cervical carcinoma and 33,542 women without cervical carcinoma from 25 epidemiological studies. *Int. J. Cancer J. Int. Cancer* **119,** 1108–1124 (2006).

110. International Collaboration of Epidemiological Studies of Cervical Cancer *et al.* Carcinoma of the cervix and tobacco smoking: collaborative reanalysis of individual data on 13,541 women with carcinoma of the cervix and 23,017 women without carcinoma of the cervix from 23 epidemiological studies. *Int. J. Cancer J. Int. Cancer* **118,** 1481–1495 (2006).

111. Castle, P. E. *et al.* An association of cervical inflammation with high-grade cervical neoplasia in women infected with oncogenic human papillomavirus (HPV). *Cancer Epidemiol. Biomark. Prev. Publ. Am. Assoc. Cancer Res. Cosponsored Am. Soc. Prev. Oncol.* **10,** 1021–1027 (2001).

112. Palefsky, J. Human papillomavirus-related disease in people with HIV. *Curr. Opin. HIV AIDS* **4,** 52–56 (2009).

113. Saslow, D. *et al.* American Cancer Society, American Society for Colposcopy and Cervical Pathology, and American Society for Clinical Pathology screening guidelines for the prevention and early detection of cervical cancer. *J. Low. Genit. Tract Dis.* **16,** 175–204 (2012).

114. Gargano, J. W. *et al.* Prevalence of human papillomavirus types in invasive vulvar cancers and vulvar intraepithelial neoplasia 3 in the United States before vaccine introduction. *J. Low. Genit. Tract Dis.* **16,** 471–479 (2012).

115. Sinno, A. K. *et al.* Human papillomavirus genotype prevalence in invasive vaginal cancer from a registry-based population. *Obstet. Gynecol.* **123,** 817–821 (2014).

116. de Sanjosé, S., Bruni, L. & Alemany, L. HPV in genital cancers (at the exception of cervical cancer) and anal cancers. *Presse Médicale* **43,** e423–e428 (2014).

117. Hernandez, B. Y. *et al.* Human papillomavirus genotype prevalence in invasive penile cancers from a registry-based United States population. *Front. Oncol.* **4,** 9 (2014).

118. Castro, F. A. *et al.* Prevalence of and risk factors for anal human papillomavirus infection among young healthy women in Costa Rica. *J. Infect. Dis.* **206,** 1103–1110 (2012).

119. Machalek, D. A. *et al.* Anal human papillomavirus infection and associated neoplastic lesions in men who have sex with men: a systematic review and meta-analysis. *Lancet Oncol.* **13,** 487–500 (2012).

120. Steinau, M. *et al.* Human papillomavirus prevalence in invasive anal cancers in the United States before vaccine introduction. *J. Low. Genit. Tract Dis.* **17,** 397–403 (2013).

121. Silverberg, M. J. *et al.* Risk of anal cancer in HIV-infected and HIV-uninfected individuals in North America. *Clin. Infect. Dis. Off. Publ. Infect. Dis. Soc. Am.* **54,** 1026–1034 (2012).

122. Jemal, A. *et al.* Annual Report to the Nation on the Status of Cancer, 1975-2009, featuring the burden and trends in human papillomavirus(HPV)-associated cancers and HPV vaccination coverage levels. *J. Natl. Cancer Inst.* **105,** 175–201 (2013).

123. Kurdgelashvili, G. *et al.* Incidence of potentially HPV-related neoplasms in the United States, 1978–2007. *Cancer* **119,** 2291–2299 (2013).

124. Gillison, M. L., Chaturvedi, A. K., Anderson, W. F. & Fakhry, C. Epidemiology of Human Papillomavirus-Positive Head and Neck Squamous Cell Carcinoma. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* **33,** 3235–3242 (2015).

125. Giuliano, A. R. *et al.* EUROGIN 2014 roadmap: Differences in human papillomavirus infection natural history, transmission and human papillomavirus-related cancer incidence by gender and anatomic site of infection. *Int. J. Cancer* **136,** 2752–2760 (2015).

126. Beachler, D. C. *et al.* Risk Factors for Acquisition and Clearance of Oral Human Papillomavirus Infection Among HIV-Infected and HIV-Uninfected Adults. *Am. J. Epidemiol.* **181,** 40–53 (2015).

127. Kreimer, A. R., Clifford, G. M., Boyle, P. & Franceschi, S. Human Papillomavirus Types in Head and Neck Squamous Cell Carcinomas Worldwide: A Systematic Review. *Cancer Epidemiol. Biomarkers Prev.* **14,** 467–475 (2005).

128. D'Souza, G. *et al.* Six-month natural history of oral versus cervical human papillomavirus infection. *Int. J. Cancer J. Int. Cancer* **121,** 143–150 (2007).

129. Fakhry, C., Sugar, E., D'Souza, G. & Gillison, M. Two-week versus six-month sampling interval in a short-term natural history study of oral HPV infection in an HIV-positive cohort. *PloS One* **5,** e11918 (2010).

130. Kreimer, A. R. Prospects for prevention of HPV-driven oropharynx cancer. *Oral Oncol.* **50,** 555–559 (2014).

131. D'Souza, G. & Dempsey, A. The role of HPV in head and neck cancer and review of the HPV vaccine. *Prev. Med.* **53 Suppl 1,** S5–S11 (2011).

132. Bosch, F. X. *et al.* Comprehensive Control of Human Papillomavirus Infections and Related Diseases. *Vaccine* **31, Supplement 7,** H1–H31 (2013).

133. Garland, S. M. *et al.* Natural history of genital warts: analysis of the placebo arm of 2 randomized phase III trials of a quadrivalent human papillomavirus (types 6, 11, 16, and 18) vaccine. *J. Infect. Dis.* **199,** 805–814 (2009).

134. Lacey, C. J. N., Lowndes, C. M. & Shah, K. V. Chapter 4: Burden and management of non-cancerous HPV-related conditions: HPV-6/11 disease. *Vaccine* **24 Suppl 3,** S3/35–41 (2006).

135. Paavonen, J. *et al.* Efficacy of human papillomavirus (HPV)-16/18 AS04-adjuvanted vaccine against cervical infection and precancer caused by oncogenic HPV types (PATRICIA): final analysis of a double-blind, randomised study in young women. *Lancet Lond. Engl.* **374,** 301–314 (2009).

136. FUTURE II Study Group. Prophylactic efficacy of a quadrivalent human papillomavirus (HPV) vaccine in women with virological evidence of HPV infection. *J. Infect. Dis.* **196,** 1438–1446 (2007).

137. Herrero, R. *et al.* Prevention of persistent human papillomavirus infection by an HPV16/18 vaccine: a community-based randomized clinical trial in Guanacaste, Costa Rica. *Cancer Discov.* **1,** 408–419 (2011).

138. Muñoz, N. *et al.* Impact of human papillomavirus (HPV)-6/11/16/18 vaccine on all HPV-associated genital diseases in young women. *J. Natl. Cancer Inst.* **102,** 325–339 (2010).

139. Joura, E. A. *et al.* A 9-valent HPV vaccine against infection and intraepithelial neoplasia in women. *N. Engl. J. Med.* **372,** 711–723 (2015).

140. Nygård, M. *et al.* Evaluation of the Long-Term Anti-Human Papillomavirus 6 (HPV6), 11, 16, and 18 Immune Responses Generated by the Quadrivalent HPV Vaccine. *Clin. Vaccine Immunol. CVI* **22,** 943–948 (2015).

141. Kreimer, A. R. *et al.* Efficacy of fewer than three doses of an HPV-16/18 AS04-adjuvanted vaccine: combined analysis of data from the Costa Rica Vaccine and PATRICIA trials. *Lancet Oncol.* **16,** 775–786 (2015).

142. WHO | Human papillomavirus (HPV). *WHO* at <http://www.who.int/immunization/diseases/hpv/en/>

143. Jit, M., Brisson, M., Laprise, J.-F. & Choi, Y. H. Comparison of two dose and three dose human papillomavirus vaccine schedules: cost effectiveness analysis based on transmission model. *The BMJ* **350,** g7584 (2015).

144. Beachler, D. C. *et al.* Multisite HPV16/18 Vaccine Efficacy Against Cervical, Anal, and Oral HPV Infection. *J. Natl. Cancer Inst.* **108,** djv302 (2016).

145. Kreimer, A. R. *et al.* Efficacy of a bivalent HPV 16/18 vaccine against anal HPV 16/18 infection among young women: a nested analysis within the Costa Rica Vaccine Trial. *Lancet Oncol.* **12,** 862–870 (2011).

146. Herrero, R. *et al.* Reduced prevalence of oral human papillomavirus (HPV) 4 years after bivalent HPV vaccination in a randomized clinical trial in Costa Rica. *PloS One* **8,** e68329 (2013).

147. Petrosky, E. *et al.* Use of 9-valent human papillomavirus (HPV) vaccine: updated HPV vaccination recommendations of the advisory committee on immunization practices. *MMWR Morb. Mortal. Wkly. Rep.* **64,** 300–304 (2015).

148. Read the Law. *HHS.gov* (2015). at <http://www.hhs.gov/healthcare/about-the-law/read-the-law/index.html>

149. *HPV Vaccine: State Legislation and Statutes*. (National Conference of State Legislatures, 2015). at <http://www.ncsl.org/research/health/hpv-vaccine-state-legislation-and-statutes.aspx>

150. Reagan-Steiner, S. *et al.* National, Regional, State, and Selected Local Area Vaccination Coverage Among Adolescents Aged 13-17 Years--United States, 2014. *MMWR Morb. Mortal. Wkly. Rep.* **64,** 784–792 (2015).

151. U.S. Department of Health and Human Services Office of Disease Prevention and Health Promotion. Immunization and Infectious Diseases|Healthy People 2020. *HealthyPeople.gov* (2014). at <https://www.healthypeople.gov/2020/topics-objectives/topic/immunization-and-infectious-diseases/objectives>

152. Holman, D. M. *et al.* Barriers to human papillomavirus vaccination among US adolescents: a systematic review of the literature. *JAMA Pediatr.* **168,** 76–82 (2014).

153. Meites, E., Markowitz, L. E., Paz-Bailey, G., Oster, A. M. & NHBS Study Group. HPV vaccine coverage among men who have sex with men - National HIV Behavioral Surveillance System, United States, 2011. *Vaccine* **32,** 6356–6359 (2014).

154. *Accelerating HPV Vaccine Uptake: Urgency for Action to Prevent Cancer. A Report to the President of the United States from the President's Cancer Panel.* (U.S. Department of Health and Human Services: National Cancer Institute, 2014).

155. Maryland Department of Health and Mental Hygiene. Increasing the Uptake of the HPV Vaccine in Maryland. (2015). at <http://phpa.dhmh.maryland.gov/cancer/SitePages/HPV.aspx>

156. Freedman, L. S., Schatzkin, A., Midthune, D. & Kipnis, V. Dealing With Dietary Measurement Error in Nutritional Cohort Studies. *JNCI J. Natl. Cancer Inst.* **103,** 1086–1092 (2011).

157. Colditz, G. A. Overview of the epidemiology methods and applications: strengths and limitations of observational study designs. *Crit. Rev. Food Sci. Nutr.* **50 Suppl 1,** 10–12 (2010).

158. Giovannucci, E. *et al.* A comparison of prospective and retrospective assessments of diet in the study of breast cancer. *Am. J. Epidemiol.* **137,** 502–511 (1993).

159. Kushi, L. H. *et al.* American Cancer Society guidelines on nutrition and physical activity for cancer prevention. *CA. Cancer J. Clin.* **62,** 30–67 (2012).

160. American Institute for Cancer Research/World Cancer Research Fund. Continuous Update Project (CUP). (2015). at <http://www.wcrf.org/int/research-we-fund/continuous-update-project-cup>

161. American Institute for Cancer Research/World Cancer Research Fund. *Continuous Update Project Report. Food, Nutrition, Phyiscal Activity, and the Prevention of Colorectal Cancer.* (2011).

162. Aune, D. *et al.* Dairy products and colorectal cancer risk: a systematic review and meta-analysis of cohort studies. *Ann. Oncol. Off. J. Eur. Soc. Med. Oncol. ESMO* **23,** 37–45 (2012).

163. Bouvard, V. *et al.* Carcinogenicity of consumption of red and processed meat. *Lancet Oncol.* (2015). doi:10.1016/S1470-2045(15)00444-1

164. Vieira, A. R. *et al.* Fruits, vegetables and lung cancer risk: a systematic review and meta-analysis. *Ann. Oncol. Off. J. Eur. Soc. Med. Oncol. ESMO* (2015). doi:10.1093/annonc/mdv381

165. Aune, D. *et al.* Dietary fibre, whole grains, and risk of colorectal cancer: systematic review and dose-response meta-analysis of prospective studies. *BMJ* **343,** (2011).

166. Potischman, N. & Linet, M. S. Invited commentary: are dietary intakes and other exposures in childhood and adolescence important for adult cancers? *Am. J. Epidemiol.* **178,** 184–189 (2013).

167. Tseng, M., Olufade, T. O., Evers, K. A. & Byrne, C. Adolescent lifestyle factors and adult breast density in U.S. Chinese immigrant women. *Nutr. Cancer* **63,** 342–349 (2011).

168. Farvid, M. S., Cho, E., Chen, W. Y., Eliassen, A. H. & Willett, W. C. Adolescent meat intake and breast cancer risk. *Int. J. Cancer J. Int. Cancer* **136,** 1909–1920 (2015).

169. Linos, E., Willett, W. C., Cho, E., Colditz, G. & Frazier, L. A. Red meat consumption during adolescence among premenopausal women and risk of breast cancer. *Cancer Epidemiol. Biomark. Prev. Publ. Am. Assoc. Cancer Res. Cosponsored Am. Soc. Prev. Oncol.* **17,** 2146–2151 (2008).

170. Berkey, C. S. *et al.* Vegetable protein and vegetable fat intakes in pre-adolescent and adolescent girls, and risk for benign breast disease in young women. *Breast Cancer Res. Treat.* **141,** 299–306 (2013).

171. Carwile, J. L. *et al.* Sugar-sweetened beverage consumption and age at menarche in a prospective study of US girls. *Hum. Reprod.* **30,** 675–683 (2015).

172. Nimptsch, K. *et al.* Dietary patterns during high school and risk of colorectal adenoma in a cohort of middle-aged women. *Int. J. Cancer J. Int. Cancer* **134,** 2458–2467 (2014).

173. Ruder, E. H. *et al.* Adolescent and mid-life diet: risk of colorectal cancer in the NIH-AARP Diet and Health Study. *Am. J. Clin. Nutr.* **94,** 1607–1619 (2011).

174. Braganza, M. Z. *et al.* Adolescent and mid-life diet and subsequent risk of thyroid cancer in the NIH-AARP diet and health study. *Int. J. Cancer* **137,** 2413–2423 (2015).

175. Mikkilä, V., Räsänen, L., Raitakari, O. T., Pietinen, P. & Viikari, J. Consistent dietary patterns identified from childhood to adulthood: the cardiovascular risk in Young Finns Study. *Br. J. Nutr.* **93,** 923–931 (2005).

176. Holman, D. M. & White, M. C. Dietary behaviors related to cancer prevention among pre-adolescents and adolescents: the gap between recommendations and reality. *Nutr. J.* **10,** 60 (2011).

177. U.S. Department of Agriculture. *Dietary Guidelines for Americans, 2010: Executive Summary*. (2011). at <http://health.gov/dietaryguidelines/2010/>

178. Guenther, P. M. *et al.* Update of the Healthy Eating Index: HEI-2010. *J. Acad. Nutr. Diet.* **113,** 569–580 (2013).

179. Banfield, E. C., Liu, Y., Davis, J. S., Chang, S. & Frazier-Wood, A. C. Poor Adherence to US Dietary Guidelines for Children and Adolescents in the National

Health and Nutrition Examination Survey Population. *J. Acad. Nutr. Diet.* (2015). doi:10.1016/j.jand.2015.08.010

180. Rasmussen, M. *et al.* Determinants of fruit and vegetable consumption among children and adolescents: a review of the literature. Part I: Quantitative studies. *Int. J. Behav. Nutr. Phys. Act.* **3,** 22 (2006).
181. STORY, M., NEUMARK-SZTAINER, D. & FRENCH, S. Individual and Environmental Influences on Adolescent Eating Behaviors. *J. Am. Diet. Assoc.* **102,** S40–S51 (2002).
182. Centers for Disease Control and Prevention. *Trends in Prevalence of Obesity, Dietary Behaviors, and Weight Control Practices National YRBS:1991 - 2013*. (2014). at <http://www.cdc.gov/healthyschools/npao/pdf/us_obesity_trend_yrbs.pdf>
183. Dwyer, L., Oh, A., Patrick, H. & Hennessy, E. Promoting family meals: a review of existing interventions and opportunities for future research. *Adolesc. Health Med. Ther.* **6,** 115–131 (2015).
184. Hartman, T. J. *et al.* Dietary and Behavioral Factors Associated with Diet Quality among Low-income Overweight and Obese African American Women. *J. Am. Coll. Nutr.* **34,** 416–424 (2015).
185. Christiansen, K. M. H., Qureshi, F., Schaible, A., Park, S. & Gittelsohn, J. Environmental factors that impact the eating behaviors of low-income African American adolescents in Baltimore City. *J. Nutr. Educ. Behav.* **45,** 652–660 (2013).
186. Neumark-Sztainer, D., French, S. A., Hannan, P. J., Story, M. & Fulkerson, J. A. School lunch and snacking patterns among high school students: Associations with school food environment and policies. *Int. J. Behav. Nutr. Phys. Act.* **2,** 14 (2005).

187. Terry-McElrath, Y. M., O'Malley, P. M. & Johnston, L. D. Foods and beverages offered in US public secondary schools through the National School Lunch Program from 2011–2013: Early evidence of improved nutrition and reduced disparities. *Prev. Med.* **78,** 52–58 (2015).

188. Sanchez-Vaznaugh, E. V., Sánchez, B. N., Crawford, P. B. & Egerter, S. Association between competitive food and beverage policies in elementary schools and childhood overweight/obesity trends: differences by neighborhood socioeconomic resources. *JAMA Pediatr.* **169,** e150781 (2015).

189. United States Department of Agriculture. Smart Snacks in School Resources. (2015). at <http://healthymeals.nal.usda.gov/smartsnacks>

190. Terry-McElrath YM, O'Malley PM & Johnston LD. Potential impact of national school nutritional environment policies: Cross-sectional associations with us secondary student overweight/obesity, 2008-2012. *JAMA Pediatr.* **169,** 78–85 (2015).

191. Hamilton, B. E. & Ventura, S. J. Birth rates for U.S. teenagers reach historic lows for all age and ethnic groups. *NCHS Data Brief* 1–8 (2012).

192. Groth, S. W. The long-term impact of adolescent gestational weight gain. *Res. Nurs. Health* **31,** 108–118 (2008).

193. Joseph, N. P. *et al.* Pre-Pregnancy Body Mass Index among Pregnant Adolescents: Gestational Weight Gain and Long-Term Post Partum Weight Retention. *J. Pediatr. Adolesc. Gynecol.* **21,** 195–200 (2008).

194. Scholl, T. O., Hediger, M. L., Schall, J. I., Khoo, C. S. & Fischer, R. L. Maternal growth during pregnancy and the competition for nutrients. *Am. J. Clin. Nutr.* **60,** 183–188 (1994).

195. Scholl, T. O. & Hediger, M. L. Weight gain, nutrition, and pregnancy outcome: findings from the Camden study of teenage and minority gravidas. *Semin. Perinatol.* **19,** 171–181 (1995).

196. Howie, L. D., Parker, J. D. & Schoendorf, K. C. Excessive maternal weight gain patterns in adolescents. *J. Am. Diet. Assoc.* **103,** 1653–1657 (2003).

197. Chang, T., Choi, H., Richardson, C. R. & Davis, M. M. Implications of teen birth for overweight and obesity in adulthood. *Am. J. Obstet. Gynecol.* **209,** 110.e1–110.e7 (2013).

198. Nielsen, J. N., Gittelsohn, J., Anliker, J. & O'Brien, K. Interventions to improve diet and weight gain among pregnant adolescents and recommendations for future research. *J. Am. Diet. Assoc.* **106,** 1825–1840 (2006).

199. Black, M. M. *et al.* Overweight Adolescent African-American Mothers Gain Weight in Spite of Intentions to Lose Weight. *J. Am. Diet. Assoc.* **106,** 80–87 (2006).

200. Haire-Joshu, D. L. *et al.* A group randomized controlled trail integrating obesity prevention and control for postpartum adolescents in a home visiting program. *Int. J. Behav. Nutr. Phys. Act.* **12,** (2015).

201. Beauchamp, A., Backholer, K., Magliano, D. & Peeters, A. The effect of obesity prevention interventions according to socioeconomic position: a systematic review. *Obes. Rev.* **15,** 541–554 (2014).

202. Barker, D. J. P. The origins of the developmental origins theory. *J. Intern. Med.* **261,** 412–417 (2007).

203. Stein, Z. & Susser, M. The Dutch famine, 1944-1945, and the reproductive process. I. Effects on six indices at birth. *Pediatr. Res.* **9,** 70–76 (1975).

204. Painter, R. C., Roseboom, T. J. & Bleker, O. P. Prenatal exposure to the Dutch famine and disease in later life: an overview. *Reprod. Toxicol. Elmsford N* **20,** 345–352 (2005).

205. Lillycrop, K. A. & Burdge, G. C. Epigenetic changes in early life and future risk of obesity. *Int. J. Obes.* **35,** 72–83 (2011).

206. Robertson, K. D. DNA methylation and human disease. *Nat. Rev. Genet.* **6,** 597–610 (2005).

207. Morgan, H. D., Santos, F., Green, K., Dean, W. & Reik, W. Epigenetic reprogramming in mammals. *Hum. Mol. Genet.* **14,** R47–R58 (2005).

208. Perkins, E. *et al.* IGF2/H19 methylation at birth and risk of overweight and obesity in children. *J. Pediatr.* **161,** 31–39 (2012).

209. Gemma, C. *et al.* Maternal Pregestational BMI Is Associated With Methylation of the PPARGC1A Promoter in Newborns. *Obesity* **17,** 1032–1039 (2009).

210. Heijmans, B. T. & Mill, J. Commentary: The seven plagues of epigenetic epidemiology. *Int. J. Epidemiol.* **41,** 74–78 (2012).

211. Laird, P. W. Principles and challenges of genomewide DNA methylation analysis. *Nat. Rev. Genet.* **11,** 191–203 (2010).

212. Ku, C. S., Naidoo, N., Wu, M. & Soong, R. Studying the epigenome using next generation sequencing. *J. Med. Genet.* **48,** 721–730 (2011).

213. Rasmussen, K. M. & Yaktine, A. L. *Weight Gain During Pregnancy: Reexamining the Guidelines*. (Institute of Medicine, 2009).

214. Engel, S. M. *et al.* Neonatal genome-wide methylation patterns in relation to birth weight in the Norwegian Mother and Child Cohort. *Am. J. Epidemiol.* **179,** 834–842 (2014).

215. Dunn, G. A. & Bale, T. L. Maternal high-fat diet promotes body length increases and insulin insensitivity in second-generation mice. *Endocrinology* **150,** 4999–5009 (2009).

216. Dunn, G. A. & Bale, T. L. Maternal high-fat diet effects on third-generation female body size via the paternal lineage. *Endocrinology* **152,** 2228–2236 (2011).

217. Gong, L., Pan, Y.-X. & Chen, H. Gestational low protein diet in the rat mediates Igf2 gene expression in male offspring via altered hepatic DNA methylation. *Epigenetics* **5,** 619–626 (2010).

218. Muhlhausler, B. S., Duffield, J. A. & McMillen, I. C. Increased maternal nutrition stimulates peroxisome proliferator activated receptor-gamma, adiponectin, and leptin messenger ribonucleic acid expression in adipose tissue before birth. *Endocrinology* **148,** 878–885 (2007).

219. Paliwal, A. *et al.* Comparative Anatomy of Chromosomal Domains with Imprinted and Non-Imprinted Allele-Specific DNA Methylation. *PLoS Genet.* **9,** (2013).

220. Fernandez, A. F. *et al.* A DNA methylation fingerprint of 1628 human samples. *Genome Res.* **22,** 407–419 (2012).

221. Benjamini, Y., Drai, D., Elmer, G., Kafkafi, N. & Golani, I. Controlling the false discovery rate in behavior genetics research. *Behav. Brain Res.* **125,** 279–284 (2001).

222. Larsson, S. C. & Wolk, A. Obesity and colon and rectal cancer risk: a meta-analysis of prospective studies. *Am. J. Clin. Nutr.* **86,** 556–565 (2007).

223. Gabory, A., Ripoche, M.-A., Yoshimizu, T. & Dandolo, L. The H19 gene: regulation and function of a non-coding RNA. *Cytogenet. Genome Res.* **113,** 188–193 (2006).

224. Sun, Z., Cunningham, J., Slager, S. & Kocher, J.-P. Base resolution methylome profiling: considerations in platform selection, data preprocessing and analysis. *Epigenomics* **7,** 813–828 (2015).

225. Ziller, M. J., Hansen, K. D., Meissner, A. & Aryee, M. J. Coverage recommendations for methylation analysis by whole-genome bisulfite sequencing. *Nat. Methods* **12,** 230–232 (2015).

226. Centers for Disease Control and Prevention. *FDA licensure of quadrivalent human papillomavirus vaccine (HPV4, Gardasil) for use in males and guidance from the Advisory Committee on Immunization Practices (ACIP).* 630–2 (2010).

227. Centers for Disease Control and Prevention. *Recommendations on the use of quadrivalent human papillomavirus vaccine in males--Advisory Committee on Immunization Practices (ACIP)*. 1705–8 (2011).

228. Kjaer, S. K. *et al.* A pooled analysis of continued prophylactic efficacy of quadrivalent human papillomavirus (Types 6/11/16/18) vaccine against high-grade cervical and external genital lesions. *Cancer Prev. Res. Phila. Pa* **2,** 868–878 (2009).

229. de Sanjose, S. *et al.* Human papillomavirus genotype attribution in invasive cervical cancer: a retrospective cross-sectional worldwide study. *Lancet Oncol.* **11,** 1048–1056 (2010).

230. Elam-Evans, L. D. *et al.* National, regional, state, and selected local area vaccination coverage among adolescents aged 13-17 years--United States, 2013. *MMWR Morb. Mortal. Wkly. Rep.* **63,** 625–633 (2014).

231. Rand, C. M., Szilagyi, P. G., Albertin, C. & Auinger, P. Additional health care visits needed among adolescents for human papillomavirus vaccine delivery within medical homes: a national study. *Pediatrics* **120,** 461–466 (2007).

232. Perkins, R. B. *et al.* Factors affecting human papillomavirus vaccine use among White, Black and Latino parents of sons. *Pediatr. Infect. Dis. J.* **32,** e38–44 (2013).

233. Donahue, K. L., Stupiansky, N. W., Alexander, A. B. & Zimet, G. D. Acceptability of the human papillomavirus vaccine and reasons for non-vaccination among parents of adolescent sons. *Vaccine* **32,** 3883–3885 (2014).

234. Reiter, P. L., Oldach, B. R., Randle, K. E. & Katz, M. L. Acceptability of HPV vaccine for males and preferences for future education programs among Appalachian residents. *Am. J. Mens Health* **8,** 167–174 (2014).

235. Reiter, P. L., McRee, A.-L., Pepper, J. K., Chantala, K. & Brewer, N. T. Improving human papillomavirus vaccine delivery: a national study of parents and their adolescent sons. *J. Adolesc. Health Off. Publ. Soc. Adolesc. Med.* **51,** 32–37 (2012).

236. Reiter, P. L. *et al.* Longitudinal predictors of human papillomavirus vaccination among a national sample of adolescent males. *Am. J. Public Health* **103,** 1419–1427 (2013).

237. Reiter, P. L., Gilkey, M. B. & Brewer, N. T. HPV vaccination among adolescent males: results from the National Immunization Survey-Teen. *Vaccine* **31,** 2816–2821 (2013).

238. Du, P. *et al.* Human Papillomavirus Vaccination Among Adults and Children in 5 US States. *J. Public Health Manag. Pract. JPHMP* (2015). doi:10.1097/PHH.0000000000000271

239. Dempsey, A., Cohn, L., Dalton, V. & Ruffin, M. Patient and clinic factors associated with adolescent human papillomavirus vaccine utilization within a university-based health system. *Vaccine* **28,** 989–995 (2010).

240. Bednarczyk, R. A., Orenstein, W. A. & Omer, S. B. Impact of Gender-Specific Human Papillomavirus Vaccine Recommendations on Uptake of Other Adolescent Vaccines: Analysis of the NIS-Teen (2008-2012). *J. Public Health Manag. Pract. JPHMP* (2015). doi:10.1097/PHH.0000000000000335

241. Dobson, S. R. M. *et al.* Immunogenicity of 2 doses of HPV vaccine in younger adolescents vs 3 doses in young women: a randomized clinical trial. *JAMA* **309,** 1793–1802 (2013).

242. Romanowski, B. *et al.* Immunogenicity and safety of the HPV-16/18 AS04-adjuvanted vaccine administered as a 2-dose schedule compared with the licensed 3-dose schedule: results from a randomized study. *Hum. Vaccin.* **7,** 1374–1386 (2011).

243. Berenson, A. B. & Rahman, M. Gender differences among low income women in their intent to vaccinate their sons and daughters against human papillomavirus infection. *J. Pediatr. Adolesc. Gynecol.* **25,** 218–220 (2012).

244. Oldach, B. R. & Katz, M. L. Ohio Appalachia public health department personnel: human papillomavirus (HPV) vaccine availability, and acceptance and concerns among parents of male and female adolescents. *J. Community Health* **37,** 1157–1163 (2012).

245. Quinn, G. P., Murphy, D., Malo, T. L., Christie, J. & Vadaparampil, S. T. A national survey about human papillomavirus vaccination: what we didn't ask, but physicians wanted us to know. *J. Pediatr. Adolesc. Gynecol.* **25,** 254–258 (2012).

246. National Vaccine Advisory Committee. Financing vaccination of children and adolescents: National Vaccine Advisory Committee recommendations. *Pediatrics* **124 Suppl 5,** S558–562 (2009).

247. CDC. Vaccines for Children Program (VFC). *Vaccines for Children Program (VFC)* (2014). at <http://www.cdc.gov/vaccines/programs/vfc/index.html>

248. The Affordable Care Act and Immunization. *HHS.gov* (2013). at <http://www.hhs.gov/healthcare/facts-and-features/fact-sheets/aca-and-immunization/index.html>

249. Department of Defense Agency. TRICARE Retail Vaccination Program. (2015). at <https://www.express-scripts.com/TRICARE/news/Vaccine_List.pdf.>

250. Allison, M. A. *et al.* HPV vaccination of boys in primary care practices. *Acad. Pediatr.* **13,** 466–474 (2013).

251. Skinner, A. C. & Skelton, J. A. Prevalence and trends in obesity and severe obesity among children in the United States, 1999-2012. *JAMA Pediatr.* **168,** 561–566 (2014).

252. Flegal, K. M., Carroll, M. D., Ogden, C. L. & Curtin, L. R. Prevalence and trends in obesity among US adults, 1999-2008. *JAMA* **303,** 235–241 (2010).

253. Frieden, T. R., Dietz, W. & Collins, J. Reducing childhood obesity through policy change: acting now to prevent obesity. *Health Aff. Proj. Hope* **29,** 357–363 (2010).

254. Haire-Joshu, D. & Nanney, M. S. Prevention of overweight and obesity in children: influences on the food environment. *Diabetes Educ.* **28,** 415–423 (2002).

255. Showell, N. N. *et al.* A systematic review of home-based childhood obesity prevention studies. *Pediatrics* **132,** e193–200 (2013).

256. Story, M., Kaphingst, K. M., Robinson-O'Brien, R. & Glanz, K. Creating healthy food and eating environments: policy and environmental approaches. *Annu. Rev. Public Health* **29,** 253–272 (2008).

257. Briefel, R. R., Wilson, A. & Gleason, P. M. Consumption of low-nutrient, energy-dense foods and beverages at school, home, and other locations among school lunch participants and nonparticipants. *J. Am. Diet. Assoc.* **109,** S79–90 (2009).

258. Wu, Y. *et al. Future Research Needs for Childhood Obesity Prevention Programs: Identification of Future Research Needs From Comparative Effectiveness Review No. 115*. (Agency for Healthcare Research and Quality (US), 2013). at <http://www.ncbi.nlm.nih.gov/books/NBK154598/>

259. Institute of Medicine. *Early childhood obesity prevention policies.* (2011).

260. Velazquez, C. E., Pasch, K. E., Ranjit, N., Mirchandani, G. & Hoelscher, D. M. Are adolescents' perceptions of dietary practices associated with their dietary behaviors? *J. Am. Diet. Assoc.* **111,** 1735–1740 (2011).

261. Tabak, R. G., Joshu, C. E., Clarke, M. A., Schwarz, C. D. & Haire-Joshu, D. L. Postpartum Teens' Perception of the Food Environments at Home and School. *Health Educ. Behav. Off. Publ. Soc. Public Health Educ.* (2015). doi:10.1177/1090198115596734

262. Parents as Teachers. *Affiliate performance report 2012-2013*. (2013).

263. CDC National Health and Nutrition Examination Survey (NHANES). Anthropometry procedures manual. (2007). at <http:www.cdc.gov/nchs/data/nhanes/nhanes_07_08/manual_an.pdf.>

264. Ogden, C. L. & Flegal, K. M. Changes in terminology for childhood overweight and obesity. *Natl. Health Stat. Rep.* 1–5 (2010).

265. Echeverria, S. E., Diez-Roux, A. V. & Link, B. G. Reliability of self-reported neighborhood characteristics. *J. Urban Health Bull. N. Y. Acad. Med.* **81,** 682–701 (2004).

266. Glanz, K., Sallis, J. F., Saelens, B. E. & Frank, L. D. Healthy nutrition environments: concepts and measures. *Am. J. Health Promot. AJHP* **19,** 330–333, ii (2005).

267. Haire-Joshu, D. *et al.* High 5 for Kids: the impact of a home visiting program on fruit and vegetable intake of parents and their preschool children. *Prev. Med.* **47,** 77–82 (2008).

268. Haire-Joshu, D. *et al.* Improving dietary behavior in African Americans: the Parents As Teachers High 5, Low Fat Program. *Prev. Med.* **36,** 684–691 (2003).

269. Willett, W. C. *Overview of nutritional epidemiology*. (Oxford Scholarship Online, 1998).

270. Bevans, K. B., Sanchez, B., Teneralli, R. & Forrest, C. B. Children's eating behavior: the importance of nutrition standards for foods in schools. *J. Sch. Health* **81,** 424–429 (2011).

271. Cvjetan, B., Utter, J., Robinson, E. & Denny, S. The social environment of schools and adolescent nutrition: associations between the school nutrition climate and

adolescents' eating behaviors and body mass index. *J. Sch. Health* **84,** 677–682 (2014).

272. O'Toole, T. P., Anderson, S., Miller, C. & Guthrie, J. Nutrition services and foods and beverages available at school: results from the School Health Policies and Programs Study 2006. *J. Sch. Health* **77,** 500–521 (2007).

273. Haire-Joshu, D. *et al.* The Quality of School Wellness Policies and Energy-Balance Behaviors of Adolescent Mothers. *Prev. Chronic. Dis.* **8,** (2011).

274. Hood, N. E., Colabianchi, N., Terry-McElrath, Y. M., O'Malley, P. M. & Johnston, L. D. School wellness policies and foods and beverages available in schools. *Am. J. Prev. Med.* **45,** 143–149 (2013).

275. Robinson-O'Brien, R., Burgess-Champoux, T., Haines, J., Hannan, P. J. & Neumark-Sztainer, D. Associations between school meals offered through the National School Lunch Program and the School Breakfast Program and fruit and vegetable intake among ethnically diverse, low-income children. *J. Sch. Health* **80,** 487–492 (2010).

276. Jaime, P. C. & Lock, K. Do school based food and nutrition policies improve diet and reduce obesity? *Prev. Med.* **48,** 45–53 (2009).

277. Goh, Y.-Y. *et al.* Using community-based participatory research to identify potential interventions to overcome barriers to adolescents' healthy eating and physical activity. *J. Behav. Med.* **32,** 491–502 (2009).

278. Berge, J. M. *et al.* Youth dietary intake and weight status: healthful neighborhood food environments enhance the protective role of supportive family home environments. *Health Place* **26,** 69–77 (2014).

279. Leung, C. W. *et al.* Associations of food stamp participation with dietary quality and obesity in children. *Pediatrics* **131,** 463–472 (2013).

280. Waterland, R. A. Is epigenetics an important link between early life events and adult disease? *Horm. Res.* **71 Suppl 1,** 13–16 (2009).

281. Murphy, S. K. *et al.* Gender-Specific Methylation Differences in Relation to Prenatal Exposure to Cigarette Smoke. *Gene* **494,** 36–43 (2012).

282. Niccolai LM & Hansen CE. Practice- and community-based interventions to increase human papillomavirus vaccine coverage: A systematic review. *JAMA Pediatr.* **169,** 686–692 (2015).

283. Dietz, W. H. Overweight in Childhood and Adolescence. *N. Engl. J. Med.* **350,** 855–857 (2004).

284. Biro, F. M. & Deardorff, J. Identifying opportunities for cancer prevention during preadolescence and adolescence: puberty as a window of susceptibility. *J. Adolesc. Health Off. Publ. Soc. Adolesc. Med.* **52,** S15–20 (2013).

285. Thompson, A. L. Intergenerational impact of maternal obesity and postnatal feeding practices on pediatric obesity. *Nutr. Rev.* **71,** S55–S61 (2013).

286. Temel, S., Voorst, S. F. van, Jack, B. W., Denktaş, S. & Steegers, E. A. P. Evidence-Based Preconceptional Lifestyle Interventions. *Epidemiol. Rev.* **36,** 19–30 (2014).

287. The American Congress of Obstetrics and Gynecologists. *Committee Opinion: Obesity in Pregnancy*. (The American Congress of Obstetrics and Gynecologists, 2013). at <http://www.acog.org/Resources-And-Publications/Committee-Opinions/Committee-on-Obstetric-Practice/Obesity-in-Pregnancy>

288. The American Congress of Obstetrics and Gynecologists. *Committee Opinion: The Importance of Preconception Care in the Continuum of Women's Health Care*. (The American Congress of Obstetrics and Gynecologists, 2005). at <http://www.acog.org/Resources-And-Publications/Committee-Opinions/Committee-on-Gynecologic-Practice/The-Importance-of-Preconception-Care-in-the-Continuum-of-Womens-Health-Care>

289. Elster, A. Guidelines for adolescent preventive services. *UpToDate* (2015). at <http://www.uptodate.com/contents/guidelines-for-adolescent-preventive-services>

290. American Academy of Pediatrics. Bright Futures: Prevention and health promotion for infants, children, adolescents, and their families. *Bright Futures* (2015). at <https://brightfutures.aap.org/Pages/default.aspx>

291. American Academy of Pediatrics. HPV Champion Toolkit. *HPV Champion Toolkit* (2015). at <https://www.aap.org/en-us/advocacy-and-policy/aap-health-initiatives/Pages/HPV-Champion-Toolkit.aspx>

292. ASH. Code of Federal Regulations. at <http://www.hhs.gov/ohrp/humansubjects/guidance/45cfr46.html#subpartb>

293. Foulkes, M. A., Grady, C., Spong, C. Y., Bates, A. & Clayton, J. A. Clinical Research Enrolling Pregnant Women: A Workshop Summary. *J. Womens Health* **20,** 1429–1432 (2011).

294. Barnhart, H. X., Haber, M. & Song, J. Overall concordance correlation coefficient for evaluating agreement among multiple observers. *Biometrics* **58,** 1020–1027 (2002).

CURRICULUM VITAE

MEGAN A. CLARKE, MHS

CONTACT INFORMATION

School Address
Johns Hopkins Bloomberg School of Public Health
615 N. Wolfe Street
Baltimore, MD 21205

Home Address
2117 Bank Street
Baltimore, MD 21231
Email: mclark43@jhu.edu
Tel: (551) 427-5020

EDUCATION

Doctor of Philosophy	**Johns Hopkins Bloomberg School of Public Health** Epidemiology, Expected 2016 Track: Cancer Baltimore, MD Advisor: Corinne Joshu, PhD, MPH
Masters of Health Science	**Johns Hopkins Bloomberg School of Public Health** Biochemistry and Molecular Biology, May 2010 Concentration: Cancer and Reproductive Biology Baltimore, MD Master's Thesis: *Fetal Basis of Adult Disease: Endocrine Disruptors and Their Effect Throughout the Life Cycle*
Bachelors of Arts	**Johns Hopkins University** Natural Sciences, Behavioral Biology, May 2007 Baltimore, MD

PROFESSIONAL EXPERIENCE

2013 – Present	Research Assistant, the Center for Child and Community Health Research, Johns Hopkins Bayview, Baltimore, MD
2013 – Present	Research Assistant, the Johnson & Johnson Community Healthcare Scholars Program, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD

2010-2012	Pre-Doctoral Fellow, Clinical Genetics Branch, Division of Cancer Epidemiology and Genetics (DCEG), National Cancer Institute (NCI), National Institutes of Health (NIH), Rockville, MD Mentor: Mark Schiffman, MD, MPH
2007-2009	Clinical Research Assistant, Center for Genetic Disorders of Behavior and Cognition, Kennedy Krieger Institute, Baltimore, MD

PROFESSIONAL ACTIVITIES

2014-2015	TA Training Chair, Department of Epidemiology, JHSPH
2013-2014	Coordinator, Cancer Epidemiology, Prevention and Control Journal Club
2013-2014	Coordinator, Sexually Transmitted Infections Journal Club
2011-2012	Division of Cancer Epidemiology and Genetics Fellows' Colloquium Planning Committee, NCI
2011-2012	Member, DCEG Fellows' Committee, NCI
2005-2007	Secretary, Healthy Habits, Johns Hopkins University

EDITORIAL ACTIVITIES

Editor

2011-2012	DCEG Fellows' Editorial Board, NCI

Reviewer

Peer Reviewed Journals (cumulative)
American Journal of Obstetrics and Gynecology
BMC Cancer
European Journal of Obstetrics & Gynecology and Reproductive Biology
Preventing Chronic Diseases
Clinical Cancer Research

HONORS AND AWARDS

2015	Doctoral Thesis Research Fund Award, Department of Epidemiology
2015	American Association for Cancer Research-Millennium Pharmaceuticals, Inc. Scholar-in-Training Award
2014-present	T32 NRSA Training Grant, National Cancer Institute (T32CA009314)
2014	Charlotte Ferencz Scholarship Award
2014	Harvey M. Meyerhoff Fellowship in Cancer Prevention
2012-present	Johnson & Johnson Community Healthcare Scholars Program

2012-2014	T32 NRSA Training Grant, Sexually Transmitted Infections (5T32AI050056)
2010-2012	Cancer Research Training Award (CRTA), NCI (C3TR047321)
2007	Dean's List, Johns Hopkins University

PUBLICATIONS

Peer Reviewed Articles

1. Roberts JE, **Clarke MA**, Carter JC, Kaufmann WE. Autistic Behavior in Boys with Fragile X Syndrome: Social Approach and HPA-Axis Dysfunction. J Neurodevelop Disord 2009;1:283-291. PMCID: PMC3164009

2. Kaufmann WE, Tierney E, Rhode CA, Suarez-Pedraza MC, **Clarke MA**, Salorio CF, Bibat G, Bukelis I, Naram D, Lanham DC, Naidu S. Social impairments in Rett syndrome: Characteristics and relationship with clinical severity. J Intellect Disabil Res 2012;56:223-47.

3. **Clarke MA**, Gage JC, Olusegun AK, Wentzensen N, Akinfolarin AC, Burk RD, Schiffman M. A population-based cross-sectional study of age-specific risk factors for high risk human papillomavirus prevalence in rural Nigeria. Infect Agents Cancer. 2011;6:12. PMCID: PMC3162906

4. Schiffman M, Gage JC, **Clarke MA**. Accepting the universal truths of global HPV epidemiology in pursuit of the remaining mysteries. Sex Transm Dis 2011;38:907-8.

5. **Clarke MA**, Rodriguez AC, Gage JC, Herrero R, Hildesheim A, Wacholder S, Burk RD, Schiffman M. A large, population-based study of age-related associations between vaginal pH and human papillomavirus infection. BMC Infect Dis 2012;12:33. PMCID: PMC3292496

6. Wentzensen N, Sun C, Ghosh A, Kinney W, Mirabello L, Wacholder S, Shaber R, LaMere B, **Clarke M**, Lorincz A, Castle P, Schiffman M, Burk RD. Methylation of HPV18, HPV31, and HPV45 genomes is associated with cervical intraepithelial neoplasia grade 3. JNCI 2012;104:1738-49. PMCID: PMC3571257

7. **Clarke MA**, Wentzensen N, Mirabello L, Ghosh A, Wacholder S, Harari A, Lorincz A, Schiffman M, Burk RD. Human papillomavirus DNA methylation as a potential biomarker for cervical cancer. Cancer Epidemiol Biomarkers Prev 2012;21:2125-37. PMCID: PMC3664203

8. **Clarke MA**, Schiffman M, Wacholder S, Rodriguez AC, Hildesheim A, Quint W. A Prospective Study of Absolute Risk and Determinants of Human Papillomavirus Incidence among Young Women in Costa Rica. BMC Infect Dis 2013;13:308. PMCID: PMC3723935

9. **Clarke MA**, Haire-Joshu DL, Schwarz CD, Tabak RG, Joshu CE. Relative influence

of home and school environments on specific dietary behaviors among postpartum, high-risk teens, 2007 – 2009. 2015;12:140437. PMCID: PMC4436050

10. Tabak RG, Joshu CE, **Clarke MA**, Schwarz CD, Haire-Joshu DL. Postpartum Teens' Perception of the Food Environments at Home and School. Health Educ Behav. 2015; epub ahead of print.

11. **Clarke MA**, Coutinho F, Phelan-Emrick DF, Wilbur MA, Chou B, Joshu CE. Predictors of Human Papillomavirus Vaccination in a Large Clinical Population of Males Aged 11 to 26 years in Maryland, 2012 – 2013. Cancer Epidemiol Biomarkers Prev 2015; In Press.

Book Chapters

1. Kaufmann WE, Capone GT, **Clarke M**, Budimirovic DB (2008) Autism in genetic intellectual disability: Insights into idiopathic autism. In Zimmerman AW (Ed). Autism: Current Theories and Evidence. Totowa, NJ: The Humana Press Inc., pp. 81-108

PRESENTATIONS

Scientific Meetings

1. Performance of Cytology and HPV16/18 Genotyping Among a Large Cohort of HPV Positive Women Aged 30 Years and Older (Poster). The 2014 National STD Prevention Conference, Atlanta, GA.

2. Sex Partner Meeting Places Reported By Newly Diagnosed HIV-Infected MSM in Baltimore City: Exploring Individual Characteristics and Viral Loads By Meeting Place (Poster). The 2014 National STD Prevention Conference, Atlanta, GA.

3. Relative Influence of Home and School Environments on Dietary Behaviors Among Postpartum, High-Risk Teens (Poster). AACR 2014 Frontiers in Cancer Prevention Conference, New Orleans, LA.

4. Factors associated with HPV vaccine initiation among males aged 11-26 years attending outpatient clinics in the Baltimore Metro Area during 2012 – 2013 (Poster). Johns Hopkins Bloomberg School of Public Health's Delta Omega Poster Competition, Baltimore, MD.

5. Factors associated with HPV vaccine initiation among males aged 11-26 years attending outpatient clinics in the Baltimore Metro Area during 2012 – 2013 (Poster). AACR 2015 Annual Meeting, Philadelphia, PA.

6. Helping youth live healthy lives with character: Assessing the effectiveness of a multifaceted program designed for the prevention of childhood obesity in Charlotte, North Carolina (Poster). APHA 2015 Annual Meeting, Chicago, IL.

7. Farm to Family Obesity Initiative: A comprehensive prevention program involving physical activity, nutrition education and a food access program in Louisville, KY (Presentation). APHA 2015 Annual Meeting, Chicago, IL.

TEACHING AND MENTORING

Lead Teaching Assistant
Johns Hopkins School of Public Health

2014-present	Principles of Epidemiology (340.601.01) *290 graduate students*
	Etiology, Prevention, and Control of Cancer (340.624.01) *20 graduate students*

Teaching Assistant
Johns Hopkins School of Public Health

2013-2014	Epidemiologic Methods 3 (340.753.01) *60-75 graduate students (lab section)*
2013	Epidemiologic Methods 2 (340.752.01) *60-75 graduate students (lab section)*
2013	Principles of Epidemiology (340.601.01) *25 graduate students (lead, lab section)*

Johns Hopkins University

2013	Fundamentals of Epidemiology (280.350) *Grading TA*

Mentoring
Johns Hopkins School of Public Health

2013-2015 Peer Mentor, Department of Epidemiology Public Health

National Cancer Institute
2010-2011 High school student intern

Professional Societies

2014-present	American Association for Cancer Research
	Women in Cancer Research
	American Public Health Association